Impact of Violations of Longitudinal Measurement Invariance

in Latent Growth Models and Autoregressive Quasi-simplex Models

by

Margarita Olivera-Aguilar

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

Approved July 2013 by the
Graduate Supervisory Committee:

Roger Millsap, Chair
Roy Levy
David Mackinnon
Stephen G. West

ARIZONA STATE UNIVERSITY

August 2013

ABSTRACT

In order to analyze data from an instrument administered at multiple time points it is a common practice to form composites of the items at each wave and to fit a longitudinal model to the composites. The advantage of using composites of items is that smaller sample sizes are required in contrast to second order models that include the measurement and the structural relationships among the variables. However, the use of composites assumes that longitudinal measurement invariance holds; that is, it is assumed that that the relationships among the items and the latent variables remain constant over time. Previous studies conducted on latent growth models (LGM) have shown that when longitudinal metric invariance is violated, the parameter estimates are biased and that mistaken conclusions about growth can be made. The purpose of the current study was to examine the impact of non-invariant loadings and non-invariant intercepts on two longitudinal models: the LGM and the autoregressive quasi-simplex model (AR quasi-simplex). A second purpose was to determine if there are conditions in which researchers can reach adequate conclusions about stability and growth even in the presence of violations of invariance. A Monte Carlo simulation study was conducted to achieve the purposes. The method consisted of generating items under a linear curve of factors model (COFM) or under the AR quasi-simplex. Composites of the items were formed at each time point and analyzed with a linear LGM or an AR quasi-simplex model. The results showed that AR quasi-simplex model yielded biased path coefficients only in the conditions with large violations of invariance. The fit of the AR quasi-simplex was not affected by violations of invariance. In general, the growth parameter estimates of the LGM were biased under violations of invariance. Further, in the presence of non-invariant loadings the rejection rates of the hypothesis of linear growth increased as the proportion of non-invariant items and as the magnitude of violations of invariance increased. A discussion of the results and limitations of the study are provided as well as general recommendations.

To my mom and my sister

ACKNOWLEDGMENTS

First and foremost, I owe a debt of gratitude to all the professors in the Quantitative Psychology program at ASU: Leona Aiken, Craig Enders, David MacKinnon, Roger Millsap and Steve West for their expert training and for providing a very supportive learning environment. Thanks to my committee members David MacKinnon, Roy Levy and Steve West for their helpful suggestions at critical junctures in this project. I am especially grateful to my advisor, Roger Millsap, for his infinite patience, support and guidance throughout my studies at ASU.

I would like to thank my colleagues at Educational Testing Service for their support during the writing of my dissertation. In particular, I wish to thank Alina von Davier, who invited me to participate in the graduate fellowship program, and to Andreas Oranje, who supervised my participation in the program.

I would also like to thank my colleagues at Ceneval in Mexico for all the encouragement they gave me to pursue my PhD studies. I am especially grateful to Arturo Bouzas, Miguel Herrera and Lucía Monroy for all their support.

Thanks to my friends who always listened to my endless doubts and complaints with unconditional love and compassion, Daniel Albarrán, Ian Blood, Yasemin Kisbu, Yunuhen Nambo, Adriana Pichardo, Sandra Reyes and Victoria Serrano.

Last but not least, I want to express my deepest gratitude to my mom and my sister, who provided me with the strength and love that I needed in the most difficult times. I could not have done this without you.

TABLE OF CONTENTS

Page

LIST OF TABLES viii

LIST OF FIGURES xi

CHAPTER

1 INTRODUCTION 1

1.1 Measurement invariance 3

1.2 Longitudinal measurement invariance 7

1.3 The longitudinal common factor model 8

1.3.1 Longitudinal factorial invariance 12

1.3.2 Partial measurement invariance 15

1.3.3 The common factor model and factorial invariance using composites 17

1.3.4 Identification 19

1.3.5 Estimation 20

1.3.6 Model fit 22

1.4 Longitudinal models 25

1.4.1 Autoregressive simplex model 26

1.4.2 Latent growth models 30

1.5 Impact of violations of factorial invariance in longitudinal methods 35

1.6 Purpose of the study 45

CHAPTER Page

2 METHOD 47

2.1 Independent variables 48

2.1.1 Sample size...48

2.1.2 Total number of items and proportion of non-invariant items........49

2.1.3 Size of loading difference across measurement occasions 50

2.1.4 Size of intercept difference across measurement occasions 55

2.1.5 Summary of conditions 58

2.2 Generating models 60

2.2.1 Curve of factors model..60

2.2.2 Autoregressive quasi-simplex model 64

2.3 Data analysis 65

2.3.1 Convergence...66

2.3.2 Parameter estimation..66

2.3.3 Model fit 68

3 RESULTS 69

3.1 Latent growth model..70

3.1.1 Re-scaling of generating growth parameters...................................70

3.1.2 Non-convergence percentages 70

3.1.3 Parameter estimation..71

3.1.4 Model fit 93

3.2 Autoregressive quasi-simplex model 96

3.2.1 Change in identification constraints 96

CHAPTER Page

3.2.2 Non-convergence percentages..97

3.2.3 Parameter estimation..98

3.2.4 Model fit 112

4 DISCUSSION..114

4.1 Non-convergence rates 115

4.2 Bias in the parameter estimates 116

4.3 Model fit...121

4.4 Limitations 123

4.5 Recommendations..125

4.6 General conclusions...128

REFERENCES 129

APPENDIX

A Item variances and communalities 136

B Item means 139

C Re-scaling of growth parameter estimates 142

D Bias in the LGM growth parameter estimates 152

E Bias in the AR quasi-simplex parameter estimates.......................................155

LIST OF TABLES

Table Page

2.1 Items included in each condition 50

2.2 Generating item loadings per measurement occasion with small, medium and large violations of invariance 54

2.3 Generating item loadings per measurement occasion with small, medium and large violations of invariance 59

2.4 Generating means and variances for $\boldsymbol{\xi}_t$ in the curve of factors model 61

2.5 Generating item loadings, intercepts and unique variances in the invariant conditions 63

3.1 Acronyms for the conditions examined 69

3.2 Non-convergence percentages for the LGM conditions with N=100 71

3.3 Relative bias of LGM parameter estimates in the invariant conditions 73

3.4 Relative bias of LGM parameter estimates under violations of invariance 74

3.5 η^2 values from the ANOVAs on bias and relative bias of the LGM parameter estimates 75

3.6 η^2 values from the ANOVAs on the standard errors of the LGM parameter estimates 86

3.7 Standard errors of the LGM parameter estimates by the number of items and by sample size 87

3.8 η^2 values from the ANOVAs on the RMSE of the LGM parameter estimates 92

3.9 Rejection rates in the invariant conditions in LGM 93

3.10 Rejection rates in the conditions with violations of invariance in the LGM 95

Table Page

3.11 Non-convergence percentages in the invariant conditions of the AR quasi-simplex 97

3.12 Non-convergence percentages in the AR quasi-simplex conditions with violations of invariance 98

3.13 Relative bias of the AR quasi-simplex parameter estimates in the invariant conditions 99

3.14 Relative bias of the AR quasi-simplex parameter estimates in the conditions with violations of invariance 100

3.15 η^2 values from the ANOVAs on the relative bias of the AR quasi-simplex conditions with non-invariant loadings 106

3.16 η^2 values from the ANOVAs on the standard errors of the AR quasi-simplex parameter estimates 107

3.17 Standard errors of the AR quasi-simplex parameter estimates by the number of items and by sample size 107

3.18 η^2 values from the ANOVAs on the RMSE of the AR quasi-simplex parameter estimates 111

3.19 Rejection rates in the invariant conditions in the AR quasi-simplex model 112

3.20 Rejection rates in conditions with violations of invariance in the AR quasi-simplex model 113

5.1 Item variances and communalities in conditions with invariant loadings and intercepts 137

Table Page

5.2 Item variances and communalities in conditions with small violations of invariance in the loadings 137

5.3 Item variances and communalities in conditions with medium violations of invariance in the loadings 138

5.4 Item variances and communalities in conditions with large violations of invariance in the loadings 138

5.5 Item means in conditions with invariant loadings and intercepts 140

5.6 Item means in conditions with small violations of invariance in the intercepts...140

5.7 Item means in conditions with medium violations of invariance in the intercepts...141

5.8 Item means in conditions with large violations of invariance in the intercepts...141

5.9 Sum of item intercepts and loadings in the first measurement occasion...........145

5.10 Original and re-scaled true growth parameter values......................................150

5.11 Bias in the LGM parameter estimates in the invariant conditions...................153

5.12 Bias in the LGM parameter estimates in conditions with violations of invariance 154

5.13 Bias in the AR quasi-simplex parameter estimates in the invariant conditions...156

5.14 Bias in the AR quasi-simplex parameter estimates in conditions with violations of invariance 156

LIST OF FIGURES

Figure Page

1.1 Path diagram of an autoregressive simplex model with mean mean structure for composites Y measured at four time points 27

1.2 Path diagram of an autoregressive quasi-simplex model with meanstructure for latent variables ξ defined by multiple indicators **X** measured at four time points..28

1.3 Path diagram of an autoregressive quasi-simplex model with mean structure for composites Y measured at four time points..29

1.4 Latent growth model with a linear trajectory over four waves measured with composites Y 33

1.5 Curve of factors model with a linear trajectory over four waves measured latent variables ξ defined by multiple indicators..34

2.1 Curve of factors model with the generating parameter values 62

2.2 Autoregressive quasi-simplex model with the generating parameter values 64

3.1 Bias in the intercept factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn).................................76

3.2 Relative bias in the intercept factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn)...............77

3.3 Bias in the slope factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn).................................78

3.4 Relative bias in the slope factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 79

Figure Page

3.5 Bias in the slope factor variance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn)...............81

3.6 Relative bias in the slope factor variance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn)...............82

3.7 Bias in the intercept slope covariance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn)...............................84

3.8 Relative bias in the intercept slope covariance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn).............84

3.9 Standard errors of the intercept factor mean in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)...88

3.10 Standard errors of the slope factor mean in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)...88

3.11 Standard errors of the intercept factor variance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)...89

3.12 Standard errors of the slope factor variance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)...89

Figure Page

3.13 Standard errors of the intercept slope covariance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn) 90

3.14 Bias in ρ_{21} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 101

3.15 Relative bias in ρ_{21} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 102

3.16 Bias in ρ_{32} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 102

3.17 Relative bias in ρ_{32} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 103

3.18 Bias in ρ_{43} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 103

3.19 Relative bias in ρ_{43} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 104

3.20 Bias in ρ_{54} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn) 104

3.21 Relative bias in ρ_{54} in the non-invariant loading conditions (NiLd) and in non-invariant intercept conditions (NiIn) 105

3.22 Standard errors of ρ_{21} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)....108

Figure Page

3.23 Standard errors of ρ_{32} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)..........108

3.24 Standard errors of ρ_{43} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)..........109

3.25 Standard errors of ρ_{54} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn)..........109

Chapter 1

INTRODUCTION

Common longitudinal models for studying stability and change, such as latent growth models (LGM) and autoregressive models (AR) frequently use composites of items of a single instrument administered in repeated measurement occasions. One assumption made when using the same instrument in different time points, is that the meaning of the instrument used does not change over time. In other words, it is assumed that longitudinal measurement invariance holds. However, using the same instrument repeatedly does not guarantee that the relation between the instrument and the underlying latent variable remains the same over time. This relation might change if there has been an intervention between occasions or if the examinees have changed across time (McArdle, 2007). Longitudinal measurement invariance is fundamental to conclude that observed changes over time are due to changes in the target latent variable and not a consequence of the characteristics of the instruments (Chan, 1998; Khoo, West, Wu, & Kwok, 2005; Widaman, Ferrer, & Conger, 2010). Unfortunately, this assumption cannot be tested when using composites of items.

Measurement invariance across groups has been extensively studied (Borsboom, 2006; Byrne, Shavelson & Muthén, 1989; Cheung & Rensvold, 1999; Horn & McArdle, 1992; Johnson, Meade & DuVernet, 2009; Meade & Bauer, 2007; Meade & Lautenschlager, 2004; Meredith, 1993; Millsap, 2011; Schmitt & Kuljanin, 2008; Vandenberg & Lance, 2000; Widaman & Reise, 1997; Yoon & Millsap, 2007). In contrast, the research on measurement invariance over time has received less attention (Chan, 1998; Millsap & Cham, 2012; Tisak & Meredith, 1989; Widaman, Ferrer, &

Conger, 2010). Some studies have examined the impact of violations of longitudinal measurement invariance on the parameter estimates and model fit of the univariate LGM (Leite, 2007; Wirth, 2008). However, these studies did not systematically manipulate variables that have been shown to be relevant in the multiple group case. Further, the consequences of violations of invariance for other longitudinal models, such as AR models, are largely unknown.

The purpose of the present study was to examine the consequences of violations of longitudinal measurement invariance on the parameter estimates and model fit of the univariate LGM and the univariate AR quasi-simplex model when the analyses are conducted on composites of items. The univariate LGM and the univariate AR quasi-simplex model represent two of the most widely used models to analyze longitudinal data. While AR models have been one of the historically dominant approaches (Biesanz, 2012), the interest in LGM has increased during the past two decades (Ferrer, Balluerka, & Widaman, 2008; Leite, 2007). Since the use of composites in these models is a common practice, it is important to examine how the results from the univariate LGM and AR quasi-simplex model might change in the presence of violations of longitudinal measurement invariance.

The simulation study consisted of generating data for multiple indicators per measurement occasion with different levels of violations of invariance, forming composites of the items and analyzing the composites using a LGM or an AR quasi-simplex model. For the LGM, data were generated under a curve of factors model (COFM). Since the COFM can be considered an extension of the LGM that includes multiple indicators of the latent variable at each measurement occasion, it was a natural choice for generating data at the item level. For the AR quasi-simplex, the data were generated under an AR quasi-simplex model with multiple indicators. The AR quasi-simplex model is an extension of the AR simplex model that includes multiple indicators; hence, the AR quasi-simplex model was a natural choice to generate data.

After generating data for multiple indicators composites were formed and analyzed under a univariate LGM or an AR quasi-simplex model. The degree to which the parameter estimates recover the generating parameter values was examined by looking at the bias and relative bias of the parameter estimates, their stability across replications, and the root mean square error (RMSE). The fit of the models was examined by looking at the number of replications in which the χ^2 rejected the null hypothesis.

The document is organized as follows. First, the problem of measurement invariance in the multiple group case and in the longitudinal case are defined in a general way and discussed under the common factor model approach. Then, four longitudinal methods are described: the autoregressive simplex model, the autoregressive quasi-simplex model, the latent growth model and the curve of factors model. A description of previous studies that examined the impact of violations of longitudinal measurement invariance in latent growth models along with the general findings is provided. The simulation study is described along with the findings. Finally, the discussion of the results and the conclusions are presented.

1.1 Measurement invariance

Psychological tests are often used to compare groups with respect to some latent variable of interest. An important prerequisite for such comparisons is that the same construct is being measured across groups. When the measurement properties of the observed variables in relation to the target latent variable are the same across populations, we can say that measurement invariance holds. In other words, the knowledge about the group membership of the examinees should not alter the relationship between the observed and the latent variables (Millsap, 2011). As expressed by Horn and McArdle (1992),

> The general question of invariance of measurement is one of whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute. If there is no evidence indicating presence or absence of measurement invariance –the usual case- or there is evidence that such invariance does not obtain, then the basis for drawing scientific inferences is severely lacking: findings of differences between individuals and groups cannot be unambiguously interpreted. (pp. 117)

Mellenbergh (1989) provided a formal definition of an unbiased item as conditional independence,

$$P(\mathbf{X}|\mathbf{W},\mathbf{V}) = P(\mathbf{X}|\mathbf{W}) \qquad (1)$$

where **X** is a vector of observed variables, **W** is the vector of the target latent variables, and **V** contains indicators defining the groups assessed. Equation (1) indicates that the probability of the observed variables **X** given the latent variables **W** does not depend on

V. If measurement invariance holds, group membership should not affect the probability of the observed variables once the latent variables are taken into account. Another way of explaining Equation (1) is that under measurement invariance, two persons with the same values in **W** have the same probability of achieving a particular score on **X** regardless of their group membership.

It is important to note that the definition of measurement invariance does not require that the groups compared have the same distribution in the latent variables **W**. There could be population differences regarding **W** and measurement invariance can still hold. The key idea is that measurement invariance is studied in groups in which the values of **W** are matched. If individuals from different groups are matched in the latent variable of interest, there should no longer be differences in the probabilities of the observed values.

If Equation (1) does not hold, measurement bias is said to exist. Under measurement bias the scores in the observed variables **X** of two persons with the same values in **W** will depend on the groups they belong to. Measurement bias can be expressed as:

$$P(\mathbf{X}|\mathbf{W},\mathbf{V}) \neq P(\mathbf{X}|\mathbf{W}) \qquad (2)$$

Measurement bias implies that the distribution of the observed variables **X** conditional on the values of **W** will be different for at least one of the groups measured.

The conditions required by Equation (1) are stringent and often do not hold in practice. Weaker forms of measurement invariance are considered such as first-order and second-order measurement invariance (Millsap, 2011). *First-order measurement invariance* is defined as:

$$E(\mathbf{X}|\mathbf{W},\mathbf{V}) = E(\mathbf{X}|\mathbf{W}) \tag{3}$$

Equation (3) indicates that the conditional mean of the observed variables $\mathbf{X}$ is invariant across groups. In other words, two groups that are matched in the target latent variables $\mathbf{W}$ will have the same conditional expected value for the observed variables under first order measurement invariance. First-order measurement invariance is the minimum level of invariance that leads to meaningful comparisons across groups.

A stronger form of measurement invariance is *second-order measurement invariance*. In addition to the condition expressed in Equation (3), second-order measurement invariance requires that the conditional covariance structure is invariant across groups, as expressed in Equation (4),

$$\boldsymbol{\Sigma}(\mathbf{X}|\mathbf{W},\mathbf{V}) = \boldsymbol{\Sigma}(\mathbf{X}|\mathbf{W}) \tag{4}$$

Under the second order measurement invariance the covariance structure of $\mathbf{X}$ once the target latent variables $\mathbf{W}$ are taken into account, should be independent of group membership. This form of measurement invariance is also known as weak measurement invariance (Meredith, 1993; Meredith & Teresi, 2006).

1.2 Longitudinal measurement invariance

The above definition of measurement invariance specifies invariance in relation to group membership. Invariance can also be studied in relation to constructs measured in multiple occasions. In this case, the meaning of a construct measured with the same instrument over time, should be invariant regardless of the measurement occasion.

Millsap and Cham (2012) define longitudinal invariance in occasions t=1, 2,…T, if and only if for $\mathbf{W}_{\mathrm{t}} = \mathbf{W}_{\mathrm{t+n}}$ it is true that:

$$P(\mathbf{X}_{\mathrm{t}}|\mathbf{W}_{\mathrm{t}}) = P(\mathbf{X}_{\mathrm{t+n}}|\mathbf{W}_{\mathrm{t+n}}) \qquad (5)$$

Equation (5) is defined for all t such that $\mathrm{t} + \mathrm{n} \leq \mathrm{T}$. Equation (5) states that under longitudinal invariance, given the same values in a latent variable $\mathbf{W}$ measured in two or more occasions, the probability of getting some particular score in the measured variables **X** should be the same across occasions. In other words, if an instrument that exhibits longitudinal measurement invariance is used to measure a person that has the same value on a latent variable as another person measured at a subsequent point in time, both examinees will have the same probability of getting a particular score in the instrument regardless of the measurement occasion.

One assumption made to simplify the study of longitudinal measurement invariance is that once the latent variables at measurement occasion t are taken into account, the observed variables $\mathbf{X}_{\mathbf{t}}$ and earlier latent variables are no longer related. In other words, the effect of latent variables at previous occasions on the observed variables

$\mathbf{X_t}$ is completely mediated through the latent variables at occasion t (Millsap & Cham, 2012).

1.3 The longitudinal common factor model

Longitudinal measurement invariance can be studied under the common factor model, which is a widely used model to describe the relationship between the latent variables and the observed measures. The common factor model assumes that the measured variables are a linear combination of the underlying latent variables, or common factors, that influence the set of observed variables and the unique factors that are specific to each variable (MacCallum, 2009). It is expected that a number of common factors smaller than the number of variables will explain the associations between the observed variables.

The common factor model can be defined for occasion t as,

$$\mathbf{X}_{t} = \boldsymbol{\tau}_{t} + \boldsymbol{\Lambda}_{t}\boldsymbol{\xi}_{t} + \boldsymbol{\delta}_{t} \qquad (6)$$

where $\boldsymbol{\tau}_{t}$ is a p x 1 vector of latent measurement intercepts at time t, $\boldsymbol{\Lambda}_{t}$ is a p x r matrix of factor loadings at time t, $\boldsymbol{\xi}_{t}$ is a vector of r x 1 common factor scores at time t, and $\boldsymbol{\delta}_{t}$ is the p x 1 vector of unique factor scores at time t. The common factors $\boldsymbol{\xi}$ are the common dimensions that explain the correlations among the observed variables. It is important to mention that the unique factor scores $\boldsymbol{\delta}$ not only represent measurement error, they also contain reliable variance that is specific to an observed variable (Meredith & Horn, 2001; Meredith & Teresi, 2006; Millsap 2011).

In the longitudinal factor model, each of the elements of the common factor model expressed in Equation (6) are sub-matrices and sub-vectors contained in the super-matrices and super-vectors defined in this section (Corballis & Traub, 1970; MacCallum, 2009; McArdle, 2007; Millsap & Cham, 2012; Tisak & Meredith, 1989).

The measured variables, the measurement intercepts and the unique factor scores can be defined as a q x 1 super-vectors where q=pT, as

$$\mathbf{X} = \begin{vmatrix} \mathbf{X_1} \\ \mathbf{X_2} \\ \vdots \\ \mathbf{X_T} \end{vmatrix} \qquad \boldsymbol{\tau} = \begin{vmatrix} \boldsymbol{\tau}_\mathbf{1} \\ \boldsymbol{\tau}_\mathbf{2} \\ \vdots \\ \boldsymbol{\tau}_\mathbf{T} \end{vmatrix} \qquad \boldsymbol{\delta} = \begin{vmatrix} \boldsymbol{\delta}_\mathbf{1} \\ \boldsymbol{\delta}_\mathbf{2} \\ \vdots \\ \boldsymbol{\delta}_\mathbf{T} \end{vmatrix} \tag{7}$$

The loadings are defined as a q x s super-matrix, where s=rT. This super loading matrix contains the loadings of each variable in each factor at each measurement occasion,

$$\boldsymbol{\Lambda} = \begin{vmatrix} \boldsymbol{\Lambda}_\mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Lambda}_\mathbf{2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Lambda}_\mathbf{T} \end{vmatrix}. \tag{8}$$

The common factors are defined as a s x 1 super-vector,

$$\boldsymbol{\xi} = \begin{vmatrix} \boldsymbol{\xi}_\mathbf{1} \\ \boldsymbol{\xi}_\mathbf{2} \\ \vdots \\ \boldsymbol{\xi}_\mathbf{T} \end{vmatrix} \tag{9}$$

In Equation (6) it is assumed that the expected values of the common factor scores and the unique factor scores are,

$$E(\boldsymbol{\xi}_{\mathrm{t}}) = \boldsymbol{\kappa}_{\mathrm{t}} \qquad\qquad E(\boldsymbol{\delta}_{\mathrm{t}}) = 0 \tag{10}$$

where $\boldsymbol{\kappa}_{\mathrm{t}}$ represent the meansfor the common factors at time t. The factor means are expressed in a s x 1 super-vector,

$$\boldsymbol{\kappa} = \begin{vmatrix} \boldsymbol{\kappa}_1 \\ \boldsymbol{\kappa}_2 \\ \vdots \\ \boldsymbol{\kappa}_{\mathrm{T}} \end{vmatrix} \tag{11}$$

The covariance matrix of the common factors and the unique factor scores are assumed to be,

$$\mathrm{Cov}(\boldsymbol{\xi}_{\mathrm{t}}) = \boldsymbol{\phi}_{\mathrm{t}} \qquad\qquad \mathrm{Cov}(\boldsymbol{\delta}_{\mathrm{t}}) = \boldsymbol{\Theta}_{\mathrm{t}} \tag{12}$$

where $\boldsymbol{\Theta}_{\mathrm{t}}$ is a p x p diagonal matrix. It is assumed that the common factor scores and the unique factor scores at time t are uncorrelated $\mathrm{Cov}(\boldsymbol{\xi}_{\mathrm{t}}, \boldsymbol{\delta}_{\mathrm{t}}) = 0$. Lagged covariances over time between unique factor scores of the same variable are permitted, but not between different variables. The lagged covariances of the unique factor scores are expressed in Equation (13), where $\mathrm{t} + \mathrm{n} \leq \mathrm{T}$ and $\boldsymbol{\Theta}_{\mathrm{t,t+n}}$ is a diagonal covariance matrix,

$$\mathrm{Cov}(\boldsymbol{\delta}_{\mathrm{t}}, \boldsymbol{\delta}_{\mathrm{t+n}}) = \boldsymbol{\Theta}_{\mathrm{t,t+n}} \qquad (13)$$

Each of the covariance matrices $\boldsymbol{\Theta}_{\mathrm{t,t+n}}$ are included as sub-matrices of the q x q super matrix $\boldsymbol{\Theta}$. The super matrix $\boldsymbol{\Theta}$ is defined as a band diagonal matrix, since each sub-matrix $\boldsymbol{\Theta}_{\mathrm{t,t+n}}$ is a diagonal matrix,

$$\boldsymbol{\Theta} = \begin{vmatrix} \boldsymbol{\Theta}_{11} & \boldsymbol{\Theta}_{12} & \cdots & \boldsymbol{\Theta}_{1T} \\ \boldsymbol{\Theta}_{21} & \boldsymbol{\Theta}_{22} & \cdots & \boldsymbol{\Theta}_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Theta}_{T1} & \boldsymbol{\Theta}_{T2} & \cdots & \boldsymbol{\Theta}_{TT} \end{vmatrix}. \qquad (14)$$

Factor scores can freely correlate across time as indicated by Equation (15), where $\boldsymbol{\phi}_{\mathrm{t,t+n}}$ is an r x r covariance matrix.

$$\mathrm{Cov}(\boldsymbol{\xi}_{\mathrm{t}}, \boldsymbol{\xi}_{\mathrm{t+n}}) = \boldsymbol{\phi}_{\mathrm{t,t+n}} \qquad (15)$$

where each of the lagged covariance matrices $\boldsymbol{\phi}_{\mathrm{t,t+n}}$ are assembled in a s x s super-matrix $\boldsymbol{\Phi}$,

$$\boldsymbol{\Phi} = \begin{vmatrix} \boldsymbol{\Phi}_{11} & \boldsymbol{\Phi}_{12} & \cdots & \boldsymbol{\Phi}_{1T} \\ \boldsymbol{\Phi}_{21} & \boldsymbol{\Phi}_{22} & \cdots & \boldsymbol{\Phi}_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}_{T1} & \boldsymbol{\Phi}_{T2} & \cdots & \boldsymbol{\Phi}_{TT} \end{vmatrix} \qquad (16)$$

Under the common factor model, the first and second unconditional moments for the observed variables **X** at time tare expressed as:

$$E(\mathbf{X}_t) = \boldsymbol{\mu}_{\mathbf{x}_t} = \boldsymbol{\tau}_t + \boldsymbol{\Lambda}_t\boldsymbol{\kappa}_t \qquad \mathrm{Cov}(\mathbf{X}_t) = \boldsymbol{\Sigma}_{\mathbf{x}_t} = \boldsymbol{\Lambda}_t\boldsymbol{\Phi}_t\boldsymbol{\Lambda}_t' + \boldsymbol{\Theta}_t \qquad (17)$$

where all the elements are defined as before, and the means of the observed variables can be expressed in a q x 1 super-vector,

$$\boldsymbol{\mu} = \begin{vmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \\ \vdots \\ \boldsymbol{\mu}_T \end{vmatrix}. \qquad (18)$$

1.3.1 Longitudinal factorial invariance

Invariance within a factor model is denoted factorial invariance. An instrument exhibits longitudinal factorial invariance if the same factor structure relating the observed variables and the latent variables holds across measurement occasions. In other words, the factor structure expressed in Equation (6) should be invariant across measurement occasions for longitudinal factorial invariance to exist. It should be noted that longitudinal factorial invariance is concerned with second order measurement invariance expressed in Equations (3) and (4).

Different levels of factorial invariance can be defined by sequentially constraining parameters of the common factor model. Jöreskog (1971) initially proposed the sequential testing of models considering only the covariance structure. Sörbom (1974) extended the method proposed by Jöreskog (1971) to multiple group analysis with mean structures. The series of nested models used for testing invariance in multiple groups can also be used in the longitudinal case. The levels of factorial invariance are described next.

Configural invariance. This is the most basic form of factorial invariance. In the multiple group case this model holds when the number of factors and the pattern of zero and nonzero loadings is the same across groups (Horn & McArdle, 1992; Thurston, 1947). In the longitudinal case this baseline model holds when the same number of factors and the same pattern of zero and non-zero loadings are established across measurement occasions. If configural invariance holds, it can be concluded that each group has the same number of factors and that each factor is defined by the same variables (Millsap & Olivera-Aguilar, 2012). If the configural model shows a poor fit to the data because of a different number of factors across measurement occasions, no further invariance constraints should be imposed since the meaning of the target latent variables is changing across time. In this case, it would be reasonable to conduct further studies to clarify the nature of the target latent variable. In contrast, if the configural model does not fit the data because the pattern of loadings is changing for a fixed number of factors, further analysis should be undertaken to investigate these changes.

Metric invariance. Metric invariance (Horn & McArdle, 1992) is also called pattern invariance and weak measurement invariance (Widaman & Reise, 1997). If the configural model fits the data, the loadings can be evaluated for invariance over time. The $\mathbf{\Lambda}_t$ super-matrix is constrained such that each item has the same loading value in a factor across measurement occasions,

$$\mathbf{\Lambda_1} = \mathbf{\Lambda_2} = \cdots = \mathbf{\Lambda_T} \tag{19}$$

If metric invariance holds in the data, it can be concluded that the differences in the covariances between variables are due to the common factors. If metric invariance is rejected, one or more items have different loadings in one or more measurement occasions. In this case, the meaning of the latent factor may be changing across time. An important step if metric invariance is rejected is to find which items are violating invariance in the loadings.

Strong factorial invariance. If the hypothesis of pattern invariance is not rejected, invariance constraints in the latent intercepts are tested. Meredith (1993) named this form of invariance strong factorial invariance. It is also known as *scalar invariance* (Steenkamp & Baumgartner, 1998) The $\boldsymbol{\tau}_{\mathrm{t}}$ super vector is constrained so that the items have the same measurement intercepts across time as

$$\boldsymbol{\tau}_{\mathbf{1}} = \boldsymbol{\tau}_{\mathbf{2}} = \cdots = \boldsymbol{\tau}_{\mathbf{T}} \tag{20}$$

If strong factorial invariance holds in the data, systematic changes in the observed means are due to changes in the latent variables. On the other hand, if the hypothesis of strong factorial invariance is rejected, changes in the mean structure of the observed measures might just be reflecting differences in the measurement intercepts across time. Strong factorial invariance needs to be established in order to make clear interpretations of the change scores. Notice that the invariance constraints in the factor loadings and the latent intercepts ensure first-order measurement invariance as stated in Equation (3).

Strict factorial invariance. Strict factorial invariance holds when the unique factor variances are invariant across time,

$$\mathbf{\Theta_{11}} = \mathbf{\Theta_{22}} = \cdots = \mathbf{\Theta_{TT}}. \tag{21}$$

Under strict factorial invariance changes in the mean and covariance structures of the observed variables across time can be interpreted as changes in the latent variables. Second-order measurement invariance as defined in Equation (4) is accomplished when strong and strict factorial invariance hold. As mentioned previously, lagged covariances between unique factor scores of the same variable over time are allowed. Invariance in the lagged covariances is not a requirement for strict factorial invariance.

Strict factorial invariance is rarely studied in practice (Vandenberg & Lance, 2000; Schmitt & Kuljanin, 2008). However, it has been argued that strict factorial invariance is essential for group comparisons and should be investigated (Meredith, 1993; Meredith & Teresi, 2006; DeShon, 2004). DeShon (2004) argues that violations of strict factorial invariance may be due to unmodeled sources of systematic variances. Unmodeled variables affecting only one of the groups assessed might change the measurement process in that group, and these changes are only detected when examining strict factorial invariance.

1.3.2 Partial measurement invariance

If invariance cannot be established in the evaluation of metric, strong and strict factorial invariance, an alternative is to test a model in which some of the observed measures are constrained to invariance while others are allowed to vary between groups. Partial invariance is the term used to denote invariance in only a subset of parameters (Byrne, Shavelson & Muthén, 1989). Partial invariance can be found at different levels of factorial invariance; partial metric invariance denotes invariance in only some of the loadings, partial strong factorial invariance denotes invariance in some of the measurement intercepts, and partial strict factorial invariance refers to invariance in some of the unique variances. Models with partial invariance are found to fit the data more often than invariance in the entire matrices evaluated. However, there are important unsolved issues in partial invariance: the specification and the meaningfulness problems (Millsap & Meredith, 2007). The specification problem deals with modifying an initial model with lack of fit to the data, until a good fitting model is found. The problem is that model re-specifications frequently do not lead to the true model, are data driven and often do not generalize to other samples (MacCallum, 1986; MacCallum, Roznowski & Necowitz, 1992). In the context of measurement invariance model modifications involve allowing the non-invariant observed measures to have different loadings, intercepts or unique variances across groups until a partial invariant model provides a good fit to the data. The issue to solve is how to locate the items that should have different parameters across time. Several methods have been proposed to locate the items that violate invariance (Byrne et al., 1989; Cheung & Rensvold, 1999, Yoon & Millsap, 2007; Woods, 2009).

Once non-invariant items are detected, a second issue to consider is the meaningfulness problem or the impact that partial invariance has on the practical conclusions made from the instrument. Unfortunately, there are no clear guidelines that indicate how large the violation of invariance must be to be meaningful for practical decisions. Further, non-invariance at the item level does not necessarily mean violations of invariance at the scale level (Stark, Chernyshenko & Drasgow, 2004) which makes it difficult to judge the impact of partial invariance. In the context of selecting individuals from their results in an instrument, Millsap and Kwok (2004) proposed a method to evaluate the consequences of partial invariance by looking at the decisions made about the examinees in the minority or low scoring group (focal group) in comparison to the majority or high scoring group (reference group). Measures such as sensitivity and specificity for the focal and reference groups are evaluated to determine the impact of partial invariance in selecting individuals from both groups.

The consequences of partial invariance in longitudinal studies have also been studied. For example, a study conducted by Ferrer, Balluerka and Widaman (2008) and the study by Wirth (2008) show that the conclusions about the growth trajectory in LGM change when longitudinal measurement invariance fails to hold. However, the question about how large the violation of invariance must be to change the conclusions of longitudinal studies has not been answered.

1.3.3 The common factor model and factorial invariance using composites

Frequently, composites of items or indicators are formed and analyzed instead of the individual items or indicators. For example, it is a common practice to fit latent growth models and autoregressive models to composites of items formed at each measurement occasion. The characteristics of the items will be reflected in the composites, such that if the items can be modeled by a common factor model, the composite can also be expressed as following a common factor model. The relationship between the common factor model at the item level and at the composite level is relevant for the purposes of the present document in which the consequences of violations of longitudinal invariance at the item level are examined on longitudinal models fitted to item composites. In this section, the common factor model and factorial invariance at the item level are explained in relation to the composites.

For a longitudinal common factor model with one factor per measurement occasion, composites of the sums of the items at time t can be formed for each individual as,

$$\mathbf{Y}_t = \mathbf{1'X}_t \tag{22}$$

where **1** is a p x 1 unit vector. If a single factor model fits the items, there is a relationship between the common factor model at the item level and at the composite level as can be observed in,

$$\mu_{Y_t} = \tau_t^* + \lambda_t^* \kappa_t \qquad \sigma_{Y_t}^2 = \lambda_t^{*2} \phi_t + \theta_t^* \tag{23}$$

where μ_{Y_t} is the mean over individuals of the composite Y at time t, $\sigma_{Y_t}^2$ is the variance of composite Y at time t, τ_t^* is the sum of the measurement intercepts of all the observed items at time point t, λ_t^* is the sum of the factor loadings of all the items at time point t, θ_t^* is the sum of the unique variances of all the items at time point t; κ_t and ϕ_t correspond to the factor mean and variance respectively, at time point t. In other words, the mean of the composite at time t is a function of the sum of the item intercepts, the sum of the item loadings and the latent mean, while its variance is a function of the sum of the item loadings, the sum of the item unique variances and the factor variance.

Strong factorial invariance at the item level $\mathbf{X}_t$, implies strong invariance in the composite Y_t. However, it should be noted that strong invariance in Y_t does not imply strong invariance in $\mathbf{X}_t$. It could be the case that differences in $\boldsymbol{\Lambda}_t$ and $\boldsymbol{\tau}_t$ across items cancel out when forming the sums of the loadings λ_t^* and the sums of the intercepts τ_t^*. In other words, if strong factorial invariance holds at the item level, strong factorial invariance will hold in the composites formed with those items, but the reverse need not be true. This relationship between invariance at the item and at the composite level exists for metric, strong and strict factorial invariance.

1.3.4 Identification

There are an infinite number of values that the $\boldsymbol{\Lambda}_t$, $\boldsymbol{\Theta}_t$ and $\boldsymbol{\phi}_t$ matrices, and the $\boldsymbol{\tau}_t$ and $\boldsymbol{\kappa}_t$ vectors can adopt that will reproduce the same mean and covariance structures of the measured variables at each time point. In order to obtain a unique solution for the factor model described in Equation (6) identification constraints are required.

Two requisites that will be assumed and that greatly simplify the identification of the longitudinal common factor model is to have each factor defined by at least three measured variables and that each measured variable loads on only one factor. Other models are possible but a set of identification constraints different from the ones to be presented are needed.

In order to identify the covariance structure it is necessary to constrain some factor loadings and/or the factor variances to nonzero values. One option is to fix the loading of one chosen measured variable per factor to one at each measurement occasion. The chosen variables are known as referent indicators. Another option to identify the covariance structure relies on loadings that are invariant across time. If the loadings are invariant, the variances of all factors at one measurement occasion can be fixed to one. That is, if metric invariance holds the loadings and the covariance structure can be identified by fixing the factors variances to one in a single measurement occasion and freely estimating the factor variances at other measurement occasions. These identification constraints are useful since it is not necessary to select an item as a referent indicator.

To identify the mean structure, the measurement intercepts and/or the factor means must be constrained. One option is to constrain the measurement intercepts of one measured variable per factor to zero in each measurement occasion. Usually, this constraint is imposed in the referent indicator. If the intercepts are found to be invariant across time another option for identifying the mean structure is to fix the factor means to zero at one measurement occasion.

Special attention is needed when choosing the referent indicator. There is evidence that choosing a non-invariant item as a referent item leads to a distorted factor solution (Johnson, Meade & DuVernet, 2009; Yoon & Millsap, 2007).

1.3.5 Estimation

The most common estimation technique used for continuous observed measures is maximum likelihood. Maximum likelihood estimates have characteristics that make them desirable. At large sample sizes the estimates are consistent, normally distributed, and efficient. Maximum likelihood estimation across measurement occasions typically makes the assumption that the measured variables $\mathbf{X}$ at each time point have a multivariate normal distribution. Under MVN the discrepancy function to be minimized that includes the mean and the covariance structure is:

$$F_{\mathrm{ML}} = (\overline{\mathbf{X}} - \boldsymbol{\mu}_{0\mathrm{X}})'\boldsymbol{\Sigma}_{0\mathrm{X}}^{-1}(\overline{\mathbf{X}} - \boldsymbol{\mu}_{0\mathrm{X}}) + \ln\left|\frac{\boldsymbol{\Sigma}_{0\mathrm{X}}}{\mathbf{S}_{\mathrm{X}}}\right| + \mathrm{tr}[\boldsymbol{\Sigma}_{0\mathrm{X}}^{-1}\mathbf{S}_{\mathrm{X}}] - p \qquad (24)$$

where $\boldsymbol{\mu}_{0\mathrm{X}}$ and $\boldsymbol{\Sigma}_{0\mathrm{X}}$ are the population values for the means and the covariance matrix, while $\overline{\mathbf{X}}$ and $\mathbf{S}_{\mathrm{X}}$ are the sample estimators of $\boldsymbol{\mu}_{\mathrm{X}}$ and $\boldsymbol{\Sigma}_{\mathrm{X}}$ calculated as,

$$\overline{\mathbf{X}} = \mathrm{N}^{-1}\sum_{i=1}^{\mathrm{N}}\mathbf{x}_{\mathrm{i}}\ \mathbf{S}_{\mathrm{X}} = \mathrm{N}^{-1}\sum_{i=1}^{\mathrm{N}}(\mathbf{x}_{\mathrm{i}} - \overline{\mathbf{X}})(\mathbf{x}_{\mathrm{i}} - \overline{\mathbf{X}})' \qquad (25)$$

The fitted mean and covariance structures are expressed as $\boldsymbol{\mu}_{\mathrm{X}}$ and $\boldsymbol{\Sigma}_{\mathrm{X}}$ and are defined as a function of the parameters $(\boldsymbol{\tau}, \boldsymbol{\Lambda}, \boldsymbol{\theta}, \boldsymbol{\kappa}, \boldsymbol{\phi})$ as shown in Equation (17). Maximum likelihood estimation looks for the set of parameters that will minimize the discrepancy function in Equation (24) after the proper constraints for identification and for invariance are imposed.

It is important to mention that the discrepancy function expressed in Equation (24) assumes complete data. Although missing data and attrition are common problems in longitudinal studies, complete data is assumed in the study proposed in the present document.

1.3.5 Model fit

In order to test the fit of invariance constraints, global fit indices such as the chi-square fit statistic, the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) are used (Bollen, 1989). A brief description of the three global fit indices considered is provided next.

Chi-square fit statistic. The chi-square goodness-of-fit statistic is used as a global measure of exact fit. The null hypothesis that is tested is,

$$H_0\colon \boldsymbol{\Sigma}_{\mathrm{X}} = \boldsymbol{\Sigma}_{0\mathrm{X}}, \boldsymbol{\mu}_{\mathrm{X}} = \boldsymbol{\mu}_{0\mathrm{X}} \tag{26}$$

The chi-square goodness-of-fit statistic is defined as,

$$\chi^2 = (N-1)\hat{F}_{ML} \tag{27}$$

where $\hat{F}_{ML}$ is the sample discrepancy function value, with $df = \frac{p(p+3)}{2} - c$ where p is the number of measured variables and c is the number of independent parameters to be estimated.

The difference-in-chi-square test can be used to compare nested models, such as the models for the different levels of factorial invariance. This test is used to determine the fit of the more constrained model in comparison with the less restricted one, assuming the latter fits well. Suppose that model B is nested in model A. The difference in chi-square can be calculated as:

$$\chi^2_{\mathrm{D}} = \chi^2_{\mathrm{B}} - \chi^2_{\mathrm{A}} \text{ with } df_{\mathrm{D}} = df_{\mathrm{B}} - df_{\mathrm{A}} \tag{28}$$

where $\chi^2_{\mathrm{A}}(df_{\mathrm{A}})$ and $\chi^2_{\mathrm{B}}(df_{\mathrm{B}})$ are the chi-square values and the degrees of freedom for models A and B respectively. The difference in the chi-square values is compared to the chi-square distribution with degrees of freedom equal to the difference in degrees of freedom in the two models. In order to conduct this difference-in-chi-square test model A must fit the data as indicated by the chi-square goodness-of-fit statistic in Equation (26). When testing the levels of factorial invariance this means that the configural invariance, model which is the less restrictive model, must hold in order to test for metric factorial invariance. Another assumption made is that the data are normally distributed.

Root mean square error of approximation (RMSEA). The root mean square error of approximation is a summary index indicates the model lack of fit per degree of freedom (Browne & Cudeck, 1993; Steiger, 1990). The RMSEA can be expressed in terms of the discrepancy function as,

$$\mathrm{RMSEA} = \left[\max\left\{\frac{\hat{\mathrm{F}}_{\mathrm{ML}}}{\mathrm{df}} - \frac{1}{\mathrm{N}-1}, 0\right\}\right]^{1/2} \tag{29}$$

where df is the degrees of freedom in the specified model and $\hat{\mathrm{F}}_{\mathrm{ML}}$ is the discrepancy function in the sample. As long as the discrepancy function incorporates the mean and the covariance structure as specified in Equation (24), RMSEA will evaluate the error in using $(\boldsymbol{\Sigma}_{0\mathrm{X}}, \boldsymbol{\mu}_{0\mathrm{X}})$ as an approximation of $(\boldsymbol{\Sigma}_{\mathrm{X}}, \boldsymbol{\mu}_{\mathrm{X}})$. RMSEA introduces a penalty for model complexity by dividing the discrepancy function by the degrees of freedom.

RMSEA indicates how bad the fit of the model is; smaller values are better. The minimum value is bounded at zero and there is no theoretical maximum. Values below .05 are considered to indicate "good fit" while values between .05 and .08 indicate "fair fit" (Browne & Cudeck, 1993).

RMSEA makes the assumption that the discrepancy function is adequate for the data, and that the error of approximation is not too large. It should also be considered that RMSEA is a large sample index in the sense that it has large standard errors at small sample sizes.

Standardized root mean square residual (SRMR). Another fit index frequently used is the SRMR that is calculated as the square root of the average squared standardized residual. The values of SRMR are non-negative, and lower values indicate better fit. Values below .05 are considered good fit. One problem with the SRMR is that it only considers the covariance structure, so it does not provide any direct information regarding the misfit of the means. The SMSR is calculated as,

$$\mathrm{SMSR} = \left[\frac{\mathbf{e}'_{\mathrm{C}}\mathbf{e}_{\mathrm{C}}}{p(p-1)/2}\right]^{1/2} \tag{30}$$

where $\mathbf{e}_{\mathrm{C}}$ is a vector of standardized residuals of the covariance matrix computed as the difference between the sample covariance matrix and the model implied covariance matrix (Hu & Bentler, 1999; West, Taylor & Wu, 2012).

1.4 Longitudinal models

Two of the most common statistical approaches for studying longitudinal data are the univariate AR simplex models and univariate LGM (Khoo, et al. 2006). In these models, the analyzed variables are observed measured variables and frequently correspond to composites of items. The use of composites of items assumes that longitudinal measurement invariance holds. However, this assumption cannot be tested at the composite level.

In the present section a description of the univariate AR simplex model and univariate LGM used to analyze composites is provided. Extensions of the univariate AR simplex models and univariate LGM that include multiple indicators per measurement occasion are also described, such as the AR quasi-simplex model and the curve of factors model (COFM). The advantage of these models is that longitudinal measurement invariance can be tested and not only assumed. In a subsequent section, studies conducted to assess the impact of violations of longitudinal measurement invariance are described in detail.

The standard notation of these models is slightly modified for ease of presentation and to avoid defining matrices with changing meaning across models. Although the AR model and LGM are described as different methods, it should be noted that recent research has shown that they can be considered special cases of a more general model, the AR latent trajectory (ALT) model (Bollen & Curran, 2004).

1.4.1 Autoregressive simplex model (AR)

In AR models, variables measured across time are modeled as a direct function of the same variable observed at an earlier measurement occasion (Heise, 1969; Jöreskog, 1970a, 1979a; Werts, Jöreskog & Linn, 1971; Wiley & Wiley, 1970).

Jöreskog (1970b) distinguished the simplex and the quasi-simplex models. While a perfect simplex assumes that measurement errors are negligible, quasi-simplex models allow for measurement error. The univariate AR simplex model is shown in Figure 1.1 and is formally expressed as,

$$\mathbf{Y}_t = \boldsymbol{\alpha}_t + \boldsymbol{\rho}_{t,t-1}\mathbf{Y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{31}$$

where **Y** are observed measures here defined as composites of items, $\boldsymbol{\alpha}_t$ is a fixed intercept for time t; $\boldsymbol{\rho}_{t,t-1}$ represents the autoregressive parameter and indicates the impact of **Y** in the time point t-1 on the value of **Y** at the current measurement occasion t. The value of $\boldsymbol{\rho}_{t,t-1}$ is frequently constrained to be the same from one measurement occasion to the next for ease of interpretation (Biesanz, 2012). It is important to note that AR models assume that all individuals can be represented by the autoregressive parameter $\boldsymbol{\rho}_{t,t-1}$. In other words, individual differences in trajectories across time are not modeled. The variable $\boldsymbol{\varepsilon}_t$ is the time specific error. It is assumed that the time specific errors are distributed as $\boldsymbol{\varepsilon}_t \sim \mathrm{N}(0, \sigma_t^2)$. Further assumptions made in this model are that the residuals $\boldsymbol{\varepsilon}_t$ are uncorrelated with $\mathbf{Y}_{t-1}$, and that the residuals are uncorrelated across individuals and across measurement occasions. The first measurement of **Y** is treated as predetermined such that,

$$Y_1 = \alpha_1 + \varepsilon_1 \tag{32}$$

In the basic AR model it is also assumed that the observed measure **Y** at time point t is only affected by **Y** at a previous time point; this model is denoted as AR(1). This assumption is relaxed in other auto-regressive models it which earlier lagged values of **Y** affect its current value.

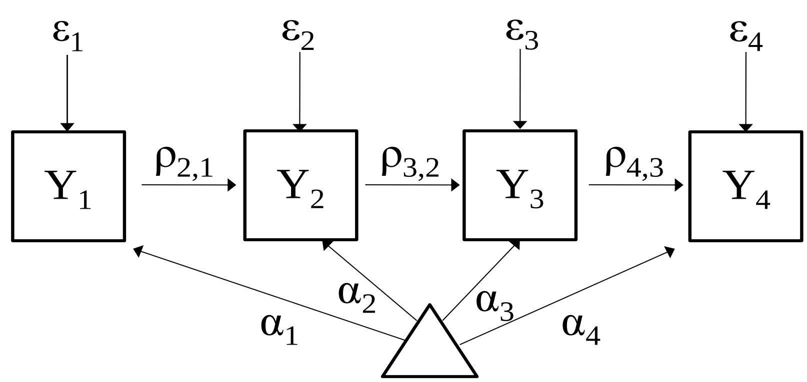

Figure 1.1 Path diagram of an autoregressive simplex model with mean structure for composites **Y** measured at four time points.

The main focus of AR models is to determine the stability of the relative standing of individuals over measurement occasions (Khoo, et al. 2006). Perfect stability is expressed by a correlation of 1.0 in the **Y** measures across time points, and it indicates stability in the rank order of individuals from one time period to the next.

The AR simplex model stated in Equation (31) can be modified so that latent variables are the focus of analysis. The AR model with latent variables is called autoregressive quasi-simplex model. The advantage of the AR quasi-simplex model is that the observed variance can be partitioned into variance due to the latent variable of interest and residual variance. By removing the influence of measurement error it is possible to obtain stability coefficients that are not attenuated (Khoo, et al. 2006; Jöreskog, 1979b).

The AR quasi-simplex model is expressed as,

$$\boldsymbol{\xi}_t = \boldsymbol{\alpha}_t + \boldsymbol{\rho}_{t,t-1}\boldsymbol{\xi}_{t-1} + \boldsymbol{\zeta}_{\xi_t} \qquad (33)$$

where $\boldsymbol{\xi}_t$ are the latent variables formed by multiple indicators **X** at time t, and $\boldsymbol{\zeta}_{\xi_t}$ is the time specific error for the latent variables $\boldsymbol{\xi}$ at time point t. The parameters $\boldsymbol{\alpha}_t$ and $\boldsymbol{\rho}_{t,t-1}$ are defined in the same way as in Equation (31). The measurement part of the model is defined as in Equation (6). The AR model with latent variables is shown in Figure 1.2.

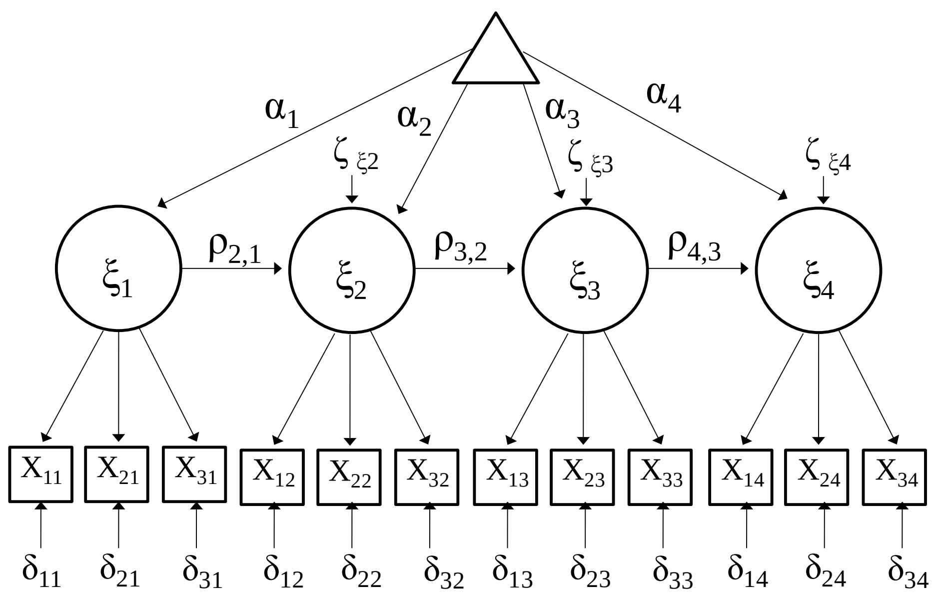

Figure 1.2 Path diagram of an autoregressive quasi-simplex model with mean structure for latent variables $\boldsymbol{\xi}$ defined by multiple indicators **X** measured at four time points

The autoregressive quasi-simplex model can also be expressed for the case in which composites are formed from the multiple indicators **X**, as shown in Figure 1.3. The equations that define the model are,

$$\mathbf{Y}_t = \boldsymbol{\xi}_t + \varepsilon_t \qquad \text{for t=1, 2, ..., T} \tag{34}$$

$$\boldsymbol{\xi}_t = \boldsymbol{\alpha}_t + \boldsymbol{\rho}_{t,t-1}\boldsymbol{\xi}_{t-1} + \boldsymbol{\zeta}_{\xi_t} \qquad \text{for t=2, ..., T} \tag{35}$$

The model in Figure 1.3 is not identified and the source of the indeterminacy is in the outer variables, that is, in Y_1 and Y_4. Therefore, in order to identify the model ϕ_1, θ_{11} or $\rho_{2,1}$ must be specified, and ϕ_4 or θ_{44} must also be specified (Jöreskog, 1979b).

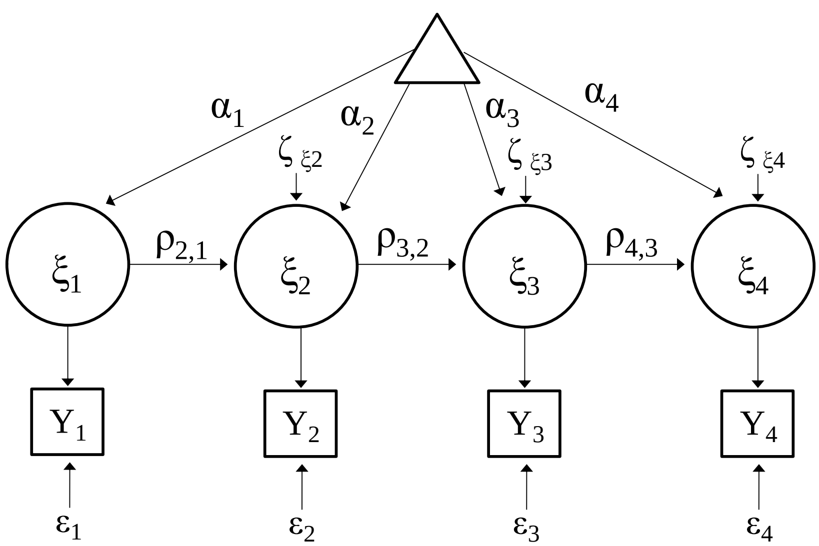

Figure 1.3 Path diagram of an autoregressive quasi-simplex model with mean structure for composites **Y** measured at four time points.

1.4.2 Latent growth models

A second common set of models used for analyzing repeated measures with three or more measurement waves are latent growth models (LGM), often also called latent curve analysis (Meredith & Tisak, 1990) or growth curve models. In contrast to AR models, in LGM it is stated that latent trajectories that are directly unobservable and that underlie the repeated measures govern the observed changes across time (Bollen & Curran, 2006; Muthén & Curran, 1997). The focus of analysis in LGM is the implications of the latent trajectories for the measured variables.

Univariate LGMs are appropriate models to study repeated measures of one target latent variable when it is believed that change is related to the passage of time (Duncan, Duncan, Strycker, Li & Alpert, 1999). One of the advantages of LGM is that it models group trajectories over time but also models individual differences in growth trajectories. In other words, LGMs incorporate information of the groups but also model individual differences. Other advantages of LGM are that it is possible to test for linear and quadratic growth curve trajectories, and that predictors of growth could be included in the model (Duncan, et al. 1999; Bollen & Curran, 2006). For the purposes of the present study the description of LGM will be restricted to the basic model without predictors.

LGM can be viewed as a common factor model defined as (Meredith & Tisak, 1990; Bollen & Curran, 2006),

$$\mathbf{Y} = \boldsymbol{\Gamma\eta} + \boldsymbol{\varepsilon} \tag{36}$$

where **Y** is the t x 1 vector of observed variables here defined as composites of items; **Γ** is a t x m matrix of factor loadings; **η** is an m x 1 vector of m latent factors; and **ε** is t x 1 vector of individual time specific residuals distributed as $\boldsymbol{\varepsilon} \sim N(0, \sigma_t^2)$ in which the variance of the residuals may vary over t. It is also assumed that the covariance between the individual time specific residuals across individuals is zero.

In a linear LGM there are two latent factors which correspond to an intercept factor η_1 and a slope factorη_2 as shown in Figure 1.4. The intercept factor η_1 refers the level of the composite at the measurement occasion defined as 0, and the slope factor η_2 represents the linear rate at which the **Y** measures change (Muthén & Khoo, 1998; Preacher, Wichman, MacCallum, Briggs, 2008). In quadratic models an additional latent variable η_3 representing a quadratic slope trajectory is included.

The η_1 and η_2 variables are random coefficients in the sense that they can be modeled as deviations from the population model as shown in,

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \mu_{\eta_1} \\ \mu_{\eta_2} \end{bmatrix} + \begin{bmatrix} \zeta_{\eta_1} \\ \zeta_{\eta_2} \end{bmatrix} \tag{37}$$

where μ_{η_1} represent the population mean for the intercept and μ_{η_2} the population mean for the slope factor, and the **ζ** residuals represent individual's deviations from the population means. It is assumed that the residual terms **ζ** have zero means, $E(\zeta_{\eta i}) = 0$ and that $Cov(\varepsilon_{it}, \zeta_{\eta i}) = 0$.

The residual terms $\boldsymbol{\zeta}$ in Equation (37) are called random effects, and are assumed to have mean 0 and a covariance matrix among the latent intercept and slope factors expressed as,

$$\boldsymbol{\Psi} = \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix} \quad (38)$$

The elements of the **Γ** matrix in Equation (36) are called basis functions (Meredith & Tisak, 1990), basic curves or latent growth vectors (Singer & Willet, 2003). In contrast to the traditional CFA, the loadings in the **Γ** matrix are not estimated but instead are fixed values. Loadings relating the intercept factor to the **Y** repeated measures are fixed to 1.0 indicating that the intercept factor equally influences all the **Y** repeated measures. The slope loadings are chosen as fixed values that adequately represent the scaling of time. Depending on the research question of interest the origin of the time scale can be defined at different measurement occasions, usually at the first time point, in which the intercept is interpreted as the initial status, or at the last time point in which the intercept is then interpreted as the final status. The origin of the scale is defined by setting the loading of the slope factor of a specific measurement occasion at 0.

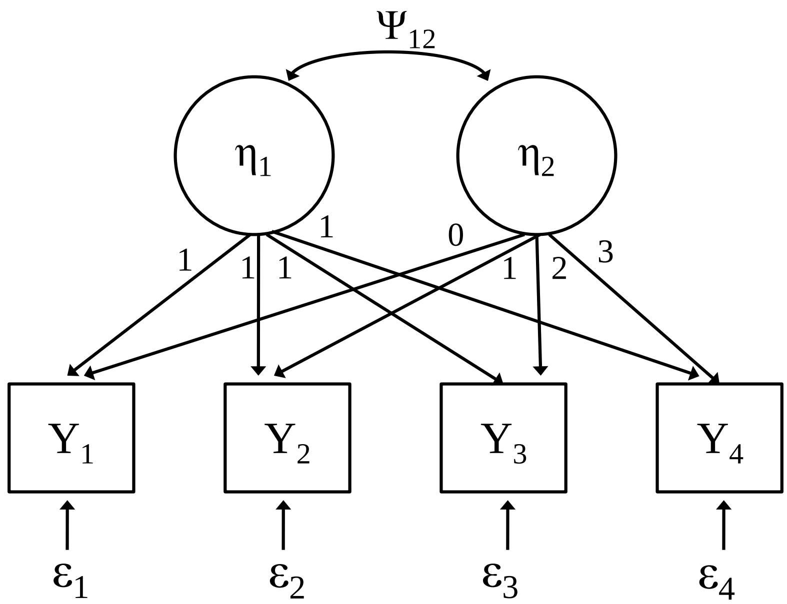

Figure 1.4 Latent growth model with a linear trajectory over four waves measured with composites **Y**.

It should be noted that models in which loadings are freely estimated have been proposed (Meredith & Tisak, 1990; McArdle, 1988; Preacher et al. 2008). In these models, the shape of the growth function is unknown and must be estimated from the data. The specific form of growth is not tested. These models are exploratory in the sense that their purpose is to gain insight about the appropriate form of the growth trajectory. These models are not further described since they are not the main focus of the present research.

Latent variables measured by multiple indicators can also be modeled across time with latent variables as in the univariate LGM. Curve of factors model consist of second-order latent growth model that includes the measurement model relating the individual items with the underlying latent construct (first-order factors) and the growth model in which the intercept and slope latent variables correspond to the second-order factors (McArdle, 1988). The curve of factors model is shown in Figure 1.5

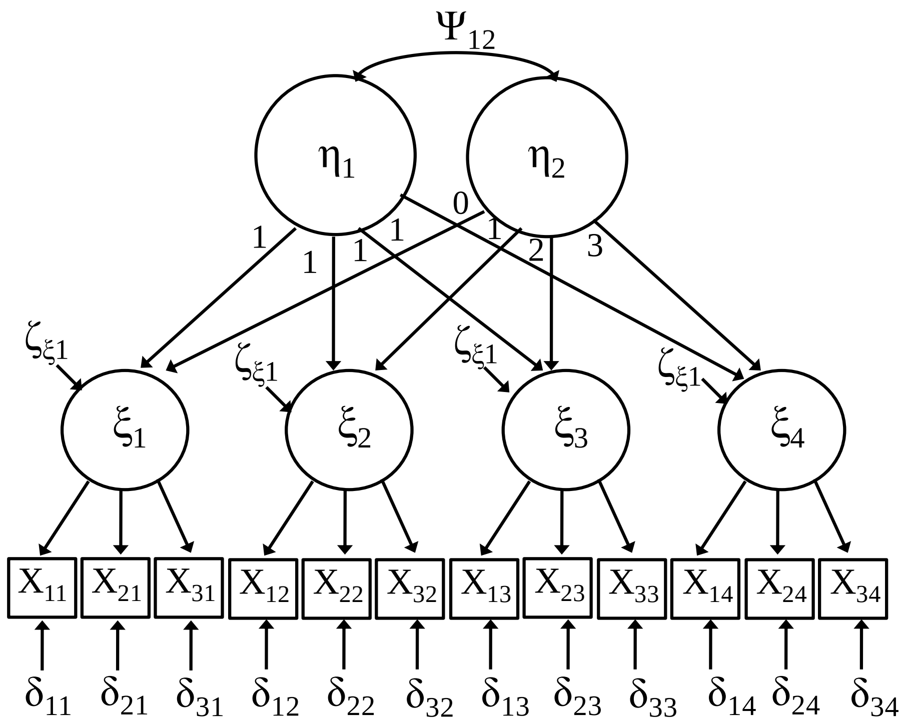

Figure 1.5 Curve of factors model with a linear trajectory over four waves measured with latent variables ξ defined by multiple indicators.

The curve of factors model is expressed as

$$\boldsymbol{\xi} = \boldsymbol{\Gamma}\boldsymbol{\eta} + \boldsymbol{\zeta}_{\xi} \qquad (39)$$

where $\boldsymbol{\xi}$ is the first order latent variable formed by multiple indicators **X**, $\boldsymbol{\zeta}_{\xi}$ is the time specific error for the latent variable $\boldsymbol{\xi}$, and $\boldsymbol{\Gamma}$ and $\boldsymbol{\eta}$ are defined as in Equation (36). The relationship between the items and the first order latent variables is defined as in Equation (6).

1.5 Impact of violations of factorial invariance in longitudinal methods

Most research done regarding measurement invariance has been developed in the multiple group case. Although studies about the impact of violations of factorial invariance in the longitudinal case have received less attention, some research has been conducted (Ferrer, Balluerka & Widaman, 2008; Leite, 2007; Wirth, 2008).

Ferrer, Balluerka, and Widaman (2008) studied the impact of measurement noninvariance in a second-order latent growth model using real data from an alcohol prevention program. An instrument assessing alcohol expectancy using 3 items was administered to 610 children measured for the first time in Grade 5, and followed through Grades 6, 7, 9 and 10. A confirmatory factor analysis showed that the model of metric invariance did not fit the data, indicating the possibility of partial invariance, but this hypothesis was not further explored. The authors fitted two curve of factors model that only differed in the item chosen to have a loading fixed to one; i. e. the models compared were the same except for the item used as the reference indicator. The results showed completely different growth trajectories obtained from the two models; using one item as a referent indicator a significant linear growth trajectory was found, while no significant growth was detected when using a different item as the reference indicator. Although this study exemplifies how the results of a longitudinal study can change when partial invariance is present, no general conclusions can be made since the study was conducted with real data and no simulation study was performed.

Leite (2005) demonstrated why fitting a latent growth model to composites formed by items with violations of metric longitudinal measurement invariance can yield to biased parameter estimates and a poor fitting model. Wirth (2008) included violations of strong factorial invariance over time in the demonstration provided by Leite (2005).

Consider the composites **Y** at four measurement occasions with a linear growth trajectory,

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{40}$$

where all the elements are defined as in Equation (34), and the composites $\mathbf{Y_t}$ are formedas sums of four items **X**. If a single factor model fits the items **X,** the composites $\mathbf{Y_t}$ can also be expressed in terms of the common factor model as $\mathbf{Y}_\mathrm{t} = \boldsymbol{\tau}_\mathrm{t}^* + \boldsymbol{\Lambda}_\mathrm{t}^*\boldsymbol{\xi}_\mathrm{t} + \boldsymbol{\delta}_\mathrm{t}^*$. Rewriting the LGM model shown in Equation (40),

$$\begin{bmatrix} \tau_1^* + \Lambda_1^*\xi_1 + \delta_1^* \\ \tau_2^* + \Lambda_2^*\xi_2 + \delta_2^* \\ \tau_3^* + \Lambda_3^*\xi_3 + \delta_3^* \\ \tau_4^* + \Lambda_4^*\xi_4 + \delta_4^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{41}$$

For the following explanation, it will be assumed that all the items have the same intercept and loading values. This is a restrictive assumption and in practice it is more common to find items that measure one latent factor but that have different loadings, intercepts and unique variances. However, for ease of presentation for the next explanation it will be assumed that all the items share the same parameters at each time point. In addition, strict factorial invariance over time will be assumed.

Setting all intercept values equal to 0, and all loadings equal to 1, Equation (41) can be written as,

$$\begin{bmatrix} 0 + 4\xi_1 + \delta_1^* \\ 0 + 4\xi_2 + \delta_2^* \\ 0 + 4\xi_3 + \delta_3^* \\ 0 + 4\xi_4 + \delta_4^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{42}$$

and thus,

$$\begin{aligned} \xi_1 &= .25(\eta_1 + \varepsilon_1 - \delta_1^*) \\ \xi_2 &= .25(\eta_1 + \eta_2 + \varepsilon_2 - \delta_2^*) \\ \xi_3 &= .25(\eta_1 + 2\eta_2 + \varepsilon_3 - \delta_3^*) \\ \xi_4 &= .25(\eta_1 + 3\eta_2 + \varepsilon_4 - \delta_4^*) \end{aligned} \tag{43}$$

It can be seen in Equation (43) that since there is strict factorial invariance, the measurement loadings are multiplicative constants that do not affect the estimation of growth parameters. Now suppose that while the unique variances are still invariant over time, there are violations of strong factorial invariance such that,

$$\begin{aligned} \tau' &= [0 \quad 0 \quad 0 \quad 0 \quad .25 \quad .25 \quad .25 \quad .25 \quad .5 \quad .5 \quad .5 \quad .5 \quad .75 \quad .75 \quad .75 \quad .75] \\ \Lambda' &= \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.5 & 1.5 & 1.5 & 1.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2.5 & 2.5 & 2.5 & 2.5 \end{bmatrix} \end{aligned} \tag{44}$$

Using the values of the intercepts and loadings in (44), Equation (40) can be written as,

$$\begin{bmatrix} 0+4\xi_1+\delta_1^* \\ 1+6\xi_2+\delta_2^* \\ 2+8\xi_3+\delta_3^* \\ 3+10\xi_4+\delta_4^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{45}$$

Equation (45) can be solved for the latent factors $\boldsymbol{\xi}_t$ as,

$$\begin{aligned} \xi_1 &= .25(\eta_1+\varepsilon_1-\delta_1^*) \\ \xi_2 &= .17(\eta_1+\eta_{2_1}-1+\varepsilon_2-\delta_2^*) \\ \xi_3 &= .125(\eta_1+2\eta_2-2+\varepsilon_3-\delta_3^*) \\ \xi_4 &= .1(\eta_1+3\eta_2-3+\varepsilon_4-\delta_4^*) \end{aligned} \tag{46}$$

Equation (46) shows that in the presence of violations of metric and strong factorial invariance over time, the measurement loadings and intercepts have different effects at each measurement occasion, altering the linear trajectory. Although there is a linear growth trajectory in the latent factors over time, $\xi_1 ... \xi_4$, the model would show a poor fit to the data and bias in the parameter estimates.

Wirth (2008) conducted a simulation study with the purpose of examining the impact of violations of factorial invariance in the growth parameter estimates of a univariate LGM and in its fit to the data using different ways of compositing items. Composites were defined as item means or as factor scores. Violations of factorial invariance were simulated in the intercepts only, or in the intercepts and loadings. The results indicated that mean and factor scores in the presence of violations of invariance resulted in biased growth estimates, and biased χ^2 statistics. Further, when examining a free loading LGM, nonlinear trend estimates over time were found even though the data were simulated to follow a linear growth trajectory. As expected, the results showed that violations of measurement invariance over time can alter the conclusions about growth trajectories. However, two important variables that might affect the results of the simulation study by Wirth (2008) and that were not systematically examined were the size of the violations of invariance and the number of non-invariant items.

Leite (2007) conducted a simulation study to examine the lack of invariance under different methods for fitting latent growth models. The two methods compared were the univariate latent growth model (LGM) of composites of multiple items and the curve of factors model (McArdle, 1988). Although the author generated the data with different levels of invariance (configural, metric or strict factorial invariance), the actual values of the loadings and intercepts were randomly selected without a manipulation of the size of the parameter difference across measurement occasions. No partial invariance conditions were included; either all the parameters were invariant or all items reflected violations of invariance. One more difficulty with the study of Leite is that when analyzing the data using the curve of factor models all of the parameters were allowed to be freely estimated across measurement occasions. A more appropriate way to conduct the curve of factors model is to constrain the parameters to invariance across measurement occasions to ensure that the same construct is being measured across time.

The results indicated that the curve of factors model fitted the data better than the univariate LGM and that the growth estimates were less biased. However, these findings are due to the fact that no invariance constraints were imposed in the curve of factors model while using composites in the univariate LGM assumes that the items are invariant. In other words, it was expected that the curve of factors model would fit better since it imposes less restrictions that the univariate LGM.

Studies in which violations of measurement invariance over time are examined in models other than LGM have not been reported. However, it can be shown that the results of an autoregressive quasi-simplex model when fitting composites will be affected by lack of invariance of the items forming the composites.

To demonstrate how the results of an autoregressive quasi-simplex model can be altered when there are violations of measurement invariance, consider composites of items measured in 4 measurement occasions. Consider the AR quasi-simplex model using composites shown in Figure 1.3, where composites Y_t are formed as the sum of 4 items at each time point.

Following Equations (34) and (35), the composites for the second, third and fourth measurement occasions can be expressed as,

$$Y_t = \alpha_t + \rho_{t,t-1}\xi_{t-1} + \zeta_{\xi_t} + \varepsilon_t \quad (47)$$

Specifically, the composite at the second, third and fourth measurement occasion are expressed as,

$$\begin{aligned} Y_2 &= \alpha_2 + \rho_{2,1}\xi_1 + \zeta_{\xi_2} + \varepsilon_2 \\ Y_3 &= \alpha_3 + \rho_{3,2}\xi_2 + \zeta_{\xi_3} + \varepsilon_3 \\ Y_4 &= \alpha_4 + \rho_{4,3}\xi_3 + \zeta_{\xi_4} + \varepsilon_4 \end{aligned} \quad (48)$$

Composites Y_2, Y_3, and Y_4, can be expressed in terms of the common factor model as $\mathbf{Y}_t = \boldsymbol{\tau}_t^* + \boldsymbol{\Lambda}_t^* \boldsymbol{\xi}_t + \boldsymbol{\delta}_t^*$ such that,

$$\begin{aligned} \tau_2^* + \Lambda_2^* \xi_2 + \delta_2^* &= \alpha_2 + \rho_{2,1}\xi_1 + \zeta_{\xi_2} + \varepsilon_2 \\ \tau_3^* + \Lambda_3^* \xi_3 + \delta_3^* &= \alpha_3 + \rho_{3,2}\xi_2 + \zeta_{\xi_3} + \varepsilon_3 \\ \tau_4^* + \Lambda_4^* \xi_4 + \delta_4^* &= \alpha_4 + \rho_{4,3}\xi_3 + \zeta_{\xi_4} + \varepsilon_4 \end{aligned} \tag{49}$$

Solving for $\boldsymbol{\xi}_t$,

$$\begin{aligned} \xi_2 &= \frac{1}{\Lambda_2^*}(\alpha_2 - \tau_2^* + \rho_{2,1}\xi_1 + \zeta_{\xi_2} + \varepsilon_2 - \delta_2^*) \\ \xi_3 &= \frac{1}{\Lambda_3^*}\left(\alpha_3 - \tau_3^* + \rho_{3,2}\xi_2 + \zeta_{\xi_3} + \varepsilon_3 - \delta_3^*\right) \\ \xi_4 &= \frac{1}{\Lambda_4^*}(\alpha_4 - \tau_4^* + \rho_{4,3}\xi_3 + \zeta_{\xi_4} + \varepsilon_4 - \delta_4^*) \end{aligned} \tag{50}$$

Composite Y_1 can also be expressed in terms of the factor model, such that $Y_1 = \tau_1^* + \Lambda_1^* \xi_1 + \delta_1^*$. Solving for ξ_1,

$$\xi_1 = \frac{1}{\Lambda_1^*}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1) \tag{51}$$

Substituting $\boldsymbol{\xi}_1$ in the right hand side of the equations in (50),

$$\xi_2 = \frac{1}{\Lambda_2^*}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{\Lambda_1^*}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)$$

$$\xi_3 = \frac{1}{\Lambda_3^*}\left(\alpha_3 - \tau_3^* + \rho_{3,2}\left(\frac{1}{\Lambda_2^*}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{\Lambda_1^*}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\right)$$

$$\xi_4 = \frac{1}{\Lambda_4^*}\left(\alpha_4 - \tau_4^* + \rho_{4,3}\left(\frac{1}{\Lambda_3^*}\left(\alpha_3 - \tau_3^* + \rho_{3,2}\left(\frac{1}{\Lambda_2^*}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{\Lambda_1^*}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\right)\right) + \zeta_{\xi_4} - \delta_4^* + \varepsilon_4\right) \tag{52}$$

If strict factorial invariance over time holds, and all intercepts equal 0 and all loadings equal 1, then $\tau_t^* = 0$ and $\Lambda_t^* = 4$,

$$\xi_2 = \frac{1}{4}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)$$

$$\begin{aligned}\xi_3 = \frac{1}{4}\Bigg(&\alpha_3 - \tau_3^* \\ &+ \rho_{3,2}\left(\frac{1}{4}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2}\right.\right. \\ &\left.\left. - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\Bigg)\end{aligned}$$

(53)

$$\begin{aligned}\xi_4 = \frac{1}{4}\Bigg(&\alpha_4 - \tau_4^* \\ &+ \rho_{4,3}\Bigg(\frac{1}{4}\Bigg(\alpha_3 - \tau_3^* \\ &+ \rho_{3,2}\left(\frac{1}{4}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2}\right.\right. \\ &\left.\left. - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\Bigg)\Bigg) + \zeta_{\xi_4} - \delta_4^* + \varepsilon_4\Bigg)\end{aligned}$$

If there are violations of invariance such that the values of the loadings and the intercepts change over time as shown in (44), Equations in (53) can be written as,

$$\xi_2 = \frac{1}{6}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)$$

$$\xi_3 = \frac{1}{8}\left(\alpha_3 - \tau_3^* + \rho_{3,2}\left(\frac{1}{6}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\right)$$

$$\xi_4 = \frac{1}{10}\left(\alpha_4 - \tau_4^* + \rho_{4,3}\left(\frac{1}{8}\left(\alpha_3 - \tau_3^* + \rho_{3,2}\left(\frac{1}{6}\left(\alpha_2 - \tau_2^* + \rho_{2,1}\left(\frac{1}{4}(Y_1 - \tau_1^* - \delta_1^* - \varepsilon_1)\right) + \zeta_{\xi_2} - \delta_2^* + \varepsilon_2\right)\right) + \zeta_{\xi_3} - \delta_3^* + \varepsilon_3\right)\right) + \zeta_{\xi_4} - \delta_4^* + \varepsilon_4\right) \quad (54)$$

Equations in (54) show that in the presence of violations of metric and strong invariance the measurement loadings and intercepts have different effects at each measurement occasion, which will bias the parameter estimates in the AR quasi-simplex model.

1.6 Purpose of the study

The studies of Ferrer et al. (2008) and Wirth (2008) showed some of the consequences of partial invariance in conclusions of growth estimates. However, the former was not a simulation study and no generalizations can be made, while the latter did not include variables that have been shown to be important in invariance studies. Simulation studies regarding the methods to identify non-invariant items (French & Finch, 2008; Johnson, Meade & DuVernet, 2009; Woods, 2009; Yoon & Millsap, 2007), and studies about the power to identify violations of invariance (Meade & Bauer, 2007; Meade & Lautenschlager, 2004) have consistently found that relevant variables in the study of invariance are the total number of items, the proportion of items violating invariance, and the size of the parameter difference across groups.

The study of Leite (2007) compared the violations of invariance in two different methods for studying growth: univariate LGM and curve of factors model. However, while in LGM using composites of items it was assumed that the items were invariant across time, in the curve of factors model no invariance constrains were imposed in the items across time. The better fit of the curve of factors model and the less biased growth estimates can be explained as a consequence of the lack of invariance constraints across measurement occasions. Further, there was no systematic manipulation of the differences in parameter estimates across time and no partial invariance conditions were examined.

Some questions that remain unanswered regarding the impact of partial invariance in longitudinal methods are: How many items should be invariant so that the conclusions about growth would not change? How different the measurement parameters across time need to be to distort the growth estimates? Is the latent growth model affected in the same way by violations of invariance as other models to study longitudinal data?

The purpose of the present study was to examine the impact of partial invariance in a univariate LGM and in an AR quasi-simplex model. The manipulated variables were sample size, total number of items, proportion of items with violations of invariance in the loadings or in the intercepts, and size difference across time in the loadings or in the intercepts. The impact of partial invariance was examined by looking at the bias in the parameter estimates, the stability of the parameter estimates over replications, RMSE, and by the rejection rates as indicated by the χ^2.

The following hypotheses were tested:

1. Larger changes over time in the item loadings will increase the bias in the growth parameter estimates and in the autoregressive parameters.
2. Larger changes over time in item intercepts will increase bias in the growth parameter estimates and in the autoregressive parameters.
3. Larger proportions of items with violations of invariance will increase bias in the parameter estimates in LGM and in AR quasi-simplex model.
4. Larger changes over time in the size of the loadings will lead to higher rates of rejection of the LGM and AR quasi-simplex models by the χ^2 test.
5. Larger changes over time in the size of the intercepts will lead to higher rates of rejection of the LGM and AR quasi-simplex models by the χ^2 test.
6. Larger proportions of items that violate invariance will lead to higher rates of rejection of the LGM and AR quasi-simplex models by the χ^2 test.

Chapter 2

METHOD

A Monte Carlo simulation study was conducted to examine the impact of violations of longitudinal measurement invariance in LGM and in AR quasi-simplex models. In general, the method consisted of generating data for five measurement occasions under the COFM (see Figure 1.5) and under the AR quasi-simplex model (see Figure 1.2) in which there are multiple indicators defining the latent variable at each measurement occasion. The multiple indicators were generated with different levels of violations of longitudinal measurement invariance. After generating the data, composites of the items at each measurement occasion were formed. The univariate LGM shown in Figure 1.4 was fit to the composites of the items that were generated from a COFM, and the AR quasi-simplex model depicted in Figure 1.3 was fit to composites of items generated under the AR quasi-simplex model with multiple indicators. The results were evaluated by examining the extent to which the parameter estimates recover the generating parameter values, as indicated by measures of bias and variability of the parameter estimates, and by examining the fit of the models.

The data were generated in M*plus* version 6.1 (Muthén & Muthén, 2010) under multivariate normality via Monte Carlo simulations. Composites of item sums at each measurement occasion were computed in SAS version 9.2 and the composites were analyzed under a LGM or AR quasi-simplex using M*plus* version 6.1. The different evaluation criteria measures, such as bias and relative bias of the estimated parameter values, the standard errors of the parameter estimates, and the root mean square error were computed in SAS. The graphs shown in the results section were obtained using the free software R (R Core Team, 2013).

In this section, the independent variables, the generating models and parameters, and the dependent variables are described.

2.1 Independent variables

The manipulated variables were sample size, total number of items per measurement occasion, proportion of non-invariant items, size of the difference in the loadings across time, and size of the difference in intercepts across time. In order to decide the conditions to be studied for each independent variable, previous studies with similar manipulations were considered.

2.1.1 Sample size

Hamilton, Gagne and Hancock (2003) showed in a simulation study that as sample size increased the percentage of replications that converged to a solution improved as well as the model fit, and suggested a minimum sample size of 100. Fan and Fan (2005) studied the power of LGM in detecting linear growth in a single group and found that for a small effect a sample size of 200 is needed, while for a medium effect size a sample size of 100 is enough.

Previous studies about the impact of partial invariance in LGM were also used as a reference to decide the sample size conditions. Leite (2007) included sample sizes of 100, 200, 500 and 1000, while Wirth (2008) studied sample sizes of 250 and 750.

In the present simulation study, sample sizes of 100, 200, 500 and 1,000 were examined.

2.1.2 Total number of items and proportion of non-invariant items

Wirth (2008) simulated eight items per measurement occasion but did not include conditions in which the number of items was manipulated. Leite (2007) included conditions for 5, 10 and 15 items per measurement occasion and the results indicated that the bias in the slope mean in LGM decreased as the number of items increased.

Studies that examine measurement invariance in the multiple group case have included conditions with 6 and 12 items (Yoon & Millsap, 2007; Meade & Lautenschlager, 2004). Yoon and Millsap (2007) proposed a method for identifying non-invariant items using modification indices and conducted a simulation study in which the total number of items was 6 or 12. The results showed that the 6 item condition yielded a higher percentage of samples that recovered the generating model and that had no false detections in contrast with the 12 item condition. However, when 2/3 of the items were non-invariant, the number of total items no longer influenced the number of samples recovering the generating model and the number of false detections. These results indicate that the decision about the number of total items per measurement occasion should consider the number of items that will be generated with violations of invariance.

For the present study conditions with 6, 9 and 15 items per measurement occasion were examined. These numbers were chosen to be comparable to the conditions in the studies by Wirth (2008) and by Leite (2007), and also to be able to manipulate the number of items violating invariance as 1/3 or 2/3 as the simulation study conducted by

Yoon and Millsap (2007). Another reason to select these quantities is to target social psychological scales often used in longitudinal tests in which the number of items tend to be relatively short.

The total item pool consisted of nine items such that for the 6-item condition, six items were selected from the item pool, in the 9-item condition all the items in pool will be used, and in the 15-item condition all the items were used and additionally six items were selected twice. Table 2.1 shows the items that were included in the 6, 9 and 15-item condition.

Table 2.1
Items included in each condition

6-item condition	9-item condition	15-items condition	
Item 1*	Item 1*	Item 1*	Item 5***
Item 2*	Item 2*	Item 2*	Item 6***
Item 3**	Item 3*	Item 3*	Item 7***
Item 4**	Item 4**	Item 4*	Item 8***
Item 7***	Item 5**	Item 5*	Item 9***
Item 8***	Item 6**	Item 6**	
	Item 7***	Item 1**	
	Item 8***	Item 2**	
	Item 9***	Item 3**	
		Item 4**	

* Non-invariant items in the 1/3 and 2/3 conditions, ** non-invariant items in the 2/3 conditions,
*** invariant items across all conditions.

2.1.3 Size of loading difference across measurement occasions

The item loadings were generated to be either invariant across time or to have small, medium or large violations of metric longitudinal invariance. In these conditions, the intercepts and unique variances were invariant over time.

Under violations of measurement invariance item loadings in a test can decrease, or increase over time, or there could be a mixed pattern in which some items increase while others decrease. For example, Obradovic, Pardini, Long and Loeber (2007) analyzed responses of parents and teachers to an instrument that assessed interpersonal callousness of children initially interviewed at 8 years old and measured annually until they were 16 years old. The authors examined the items for longitudinal invariance and found that, some item loadings decreased over time and others increased. In general, the item loadings decreased in the first three measurement occasions and then remained invariant from the fourth to the ninth wave when the teachers answered the items. When the answers from the parents were analyzed it was found that 3 item loadings increased over time while 3 item loadings decreased over time. It was concluded that the items were not equally representative of interpersonal callousness over time.

In another study Willoughby, Wirth and Blair (2012) examined the longitudinal invariance of a battery of six tests that were administered repeatedly over time to assess executive function (defined as cognitive abilities in the control and coordination of information in the service of goal directed actions). The authors found that two tests were invariant over time, while the other four tests were non-invariant, with loadings that increased at some measurement occasions and decreased at others.

While the studies described indicate that it is plausible to find item loadings that increase or decrease over time, in the present study all the item loadings in the non-invariant conditions decreased over time. No manipulations in which the loadings increase or have a mixed pattern of increasing and decreasing values over time were included to keep the total number of conditions within manageable limits.

To create the values of the loadings over time, a three step procedure similar to the one used by Yoon and Millsap (2007) was followed. The item loadings in the first measurement occasion were the same in all conditions and the loadings in the fifth measurement occasion were selected to represent small, medium and large violations of metric longitudinal invariance. As indicated by Yoon and Millsap (2007) imposing a fixed change in all item loadings might have a different meaning across items depending on the magnitude of the initial item loading; for example, the impact of a 0.1 change might be different in an item changing from a loadings of 0.9 to a loading of 0.8 than in an item changing from a loading of 0.3 to a loading of 0.2. For this reason, the first step to create the item loadings for the fifth measurement occasion was to define effect sizes for violations of metric invariance over time with respect to one specific item. The effect size was defined as the change in the loading of item 1 from the first measurement occasion, with a loading value of 0.7, to the fifth measurement occasion. A small violation to longitudinal metric invariance was defined as a change of .1 from the first to the fifth measurement occasion (a change from a loading of 0.7 to a loading of 0.6), a medium violation corresponded to a change of 0.2 (from 0.7 to 0.5), and a change of 0.3 defined a large violation of longitudinal metric invariance (from 0.7 to 0.4).

The second step was to define the loading values at the fifth measurement occasion for the rest of the non-invariant items by subtracting a proportional amount to the change in the loadings of item 1. For example, when there are small violations to measurement invariance, item 1 changes from a loading value of 0.7 in the first measurement occasion to 0.6 in the fifth measurement occasion, which corresponds to a proportional drop of $0.1/0.7 = 1/7$. In order to determine the loadings of the items at the fifth measurement occasion, the loadings of the non-invariant items were multiplied by 6/7=0.857. In the same way, to create medium and large violations of invariance the loadings were multiplied by 0.714 (5/7) and 0.571 (4/7), respectively.

Once the value of the loadings at the fifth measurement occasion is defined the loadings for the second, third and fourth measurement occasions were defined. The total change from the first to the fifth measurement occasion in the item loadings was be divided in equal parts so that there is a constant change from one measurement occasion to the next. For example, the change of 0.1 in item 1 was divided so that the change from one measurement occasion to the next was .025.

Table 2.2 shows the item loadings at each measurement occasion for each size of violation in measurement invariance. It should be noted that six of the nine items were generated to show violations of metric invariance over time, while three items were invariant over time. Appendix A shows the item variances at each time point and at each condition.

Table 2.2

Generating item loadings per measurement occasion with small, medium and large violations of invariance

	Item	Time1	Small violations				Medium violations				Large violations			
			Time 2	Time 3	Time 4	Time 5	Time 2	Time 3	Time 4	Time 5	Time 2	Time 3	Time 4	Time 5
Non-invariant	1	0.7	0.675	0.650	0.625	0.600	0.650	0.600	0.550	0.500	0.625	0.550	0.475	0.400
	2	0.9	0.868	0.836	0.804	0.771	0.836	0.771	0.707	0.643	0.804	0.707	0.611	0.514
	3	0.5	0.482	0.464	0.446	0.429	0.464	0.429	0.393	0.357	0.446	0.393	0.339	0.286
	4	0.6	0.579	0.557	0.536	0.514	0.557	0.514	0.471	0.429	0.536	0.471	0.407	0.343
	5	0.8	0.771	0.743	0.714	0.686	0.743	0.686	0.629	0.571	0.714	0.629	0.543	0.457
	6	0.4	0.386	0.371	0.357	0.343	0.371	0.343	0.314	0.286	0.357	0.314	0.271	0.229
Invariant	7	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
	8	1	1	1	1	1	1	1	1	1	1	1	1	1
	9	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

2.1.4 Size of intercept difference across measurement occasions

The item intercepts were created to be either invariant or to have small, medium or large violations of strong factorial invariance over time. It should be noted that the item loadings were invariant over time under this condition.

As in the case of non-invariant loadings, the intercepts can increase, decrease or have a mixed pattern over time. For example, Millsap and Cham (2012) analyzed data from children assessed from 8 to 16 years of age that participated in an intervention to help them cope with the death of a parent. The authors found that two of the subscales used to assess the acting coping dimension violated strong factorial invariance: the intercept of the Optimism subscale decreased from 2.803 in wave 1 to 2.648 in wave 2, while the intercept of the Direct Problem Solving subscale decreased from 2.657 in wave 2 to 2.607 in wave 3. In another study, Willoughby et al. (2011) found violations of strong longitudinal invariance in 4 of the 6 battery tests to assess executive function. Although the intercepts were not reported, the means of the subscales were reported to increase at each of the three measurement waves. Since the hypothesis of strong factorial invariance was rejected, part of the increase in the test means was due to violations of strong longitudinal invariance. In the present study it was decided to examine only conditions in which the intercepts increase over time. No conditions in which the item intercepts decrease or have a mixed pattern were considered to keep the total number of conditions under manageable limits.

The size of the violations in strong longitudinal invariance was defined as the change in intercepts from the first to the fifth measurement occasion. To define small, medium and large violations of strong longitudinal invariance the ratio of the difference in intercepts to the difference in the item means at two time points was considered (Equation 55). This measure was proposed by Millsap and Olivera-Aguilar (2012) in the multiple group case but can be used for studying invariance over time,

$$\mathrm{d} = \frac{\boldsymbol{\tau}_{\mathrm{p}1} - \boldsymbol{\tau}_{\mathrm{p}5}}{\boldsymbol{\mu}_{\mathrm{p}1} - \boldsymbol{\mu}_{\mathrm{p}5}} \tag{55}$$

where $\boldsymbol{\tau}_{\mathrm{p}1} - \boldsymbol{\tau}_{\mathrm{p}5}$ corresponds to the difference in intercepts for item p from time 1 to time 5, $\boldsymbol{\mu}_{\mathrm{p}1} - \boldsymbol{\mu}_{\mathrm{p}5}$corresponds to the difference in means for item p from time 1 to time 5, and d represents the proportion of the difference in means across time that is due to differences in intercepts over time. A *d* value of 0.2 was considered a small effect size (Millsap & Olivera-Aguilar, 2012), 0.5 represented a medium effect size, and 0.8 corresponded to a large effect size.

The item means at the first and fifth measurement occasions needed in the denominator of Equation (55) are substituted in Equation (17). From Equation (17) it can be seen that $\boldsymbol{\mu}_{\mathrm{p}_\mathbf{1}} = \boldsymbol{\tau}_{\mathrm{p}1} + \boldsymbol{\Lambda}_{\mathrm{p}1}\boldsymbol{\kappa}_1$, and that $\boldsymbol{\mu}_{\mathrm{p}_\mathbf{5}} = \boldsymbol{\tau}_{\mathrm{p}5} + \boldsymbol{\Lambda}_{\mathrm{p}5}\boldsymbol{\kappa}_5$**,** so that Equation (55) can be written as,

$$\mathrm{d} = \frac{\boldsymbol{\tau}_{\mathrm{p}1} - \boldsymbol{\tau}_{\mathrm{p}5}}{(\boldsymbol{\tau}_{\mathrm{p}1} + \boldsymbol{\Lambda}_{\mathrm{p}1}\boldsymbol{\kappa}_1) - (\boldsymbol{\tau}_{\mathrm{p}5} + \boldsymbol{\Lambda}_{\mathrm{p}5}\boldsymbol{\kappa}_5)} \tag{56}$$

Since the loadings are invariant $\boldsymbol{\Lambda}_{\mathrm{p}1} = \boldsymbol{\Lambda}_{\mathrm{p}5}$, they can be expressed as $\boldsymbol{\Lambda}_{\mathrm{p}}$. Replacing $\boldsymbol{\Lambda}_{\mathrm{p}1}$ and $\boldsymbol{\Lambda}_{\mathrm{p}5}$ with $\boldsymbol{\Lambda}_{\mathrm{p}}$ and solving Equation (56) for $\boldsymbol{\tau}_{\mathrm{p}5}$**,**

$$\boldsymbol{\tau}_{p5} = \frac{d\boldsymbol{\tau}_{p1} + d\boldsymbol{\Lambda}_p\boldsymbol{\kappa}_1 - d\boldsymbol{\Lambda}_p\boldsymbol{\kappa}_5 - \boldsymbol{\tau}_{p1}}{d - 1} \tag{57}$$

Equation (57) was used to determine the item intercepts at time 5. The item loadings corresponded to the loadings of the first measurement occasion shown in Table 2.2; the factor mean at the first time point was 0 and at the fifth measurement occasion it was 0.8, as determined from the curve of factors model (see below); the values for the intercepts in the first measurement occasion were held constant across conditions, and correspond to the values for the first measurement occasion in Table 2.3. The values of *d* were 0.2, 0.5 and 0.8.

For example, substituting the loading for the first item ($\lambda_1 = 0.7$), its intercept at the first measurement occasion ($\tau_{11} = 0.5$), the factor means at the first ($\kappa_1 = 0$) and fifth ($\kappa_5 = 0.8$) measurement occasions, for a small difference in intercepts (d=0.2), the intercept at the fifth measurement occasion is,

$$0.640 = \frac{(0.2^{*}0.5) + (0.2 * 0.7 * 0) - (0.20^{*}0.7^{*}0.8) - 0.5}{0.20 - 1} \tag{58}$$

The item intercepts at the fifth measurement occasion were determined for all items using Equation (57). To determine the item intercepts at the second, third and fourth measurement occasions, the total difference in intercepts from time 1 to 5 was divided so that there is a constant change in intercepts over time. For example, the total change of 0.140 in the intercept of item 1 from the first to the fifth measurement occasion was divided so that the change from one measurement occasion to the next was .035.

Table 2.3 shows the resulting item intercepts across measurement occasions. It should be noted that six of the nine items in the item pool show violations of invariance while three items remain invariant over time. For the 6-item condition, 6 items were selected from Table 2.3; for the 15-item condition, all the items from the item pool were used and six items were selected twice as indicated in Table 2.1. Appendix B shows the item means at each time point.

2.1.5 Summary of conditions

A total of 312 conditions were examined; 156 conditions in each of the generating models. For each of the generating models, the conditions examined corresponded to four sample sizes, three total numbers of items per measurement occasion, two proportions of items violating invariance, and six sizes of violation of invariance (small, medium and large violations of metric measurement invariance, and small, medium and large violations of strong factorial invariance). Additionally, in the conditions in which measurement invariance holds, four sample sizes and three total numbers of items were examined.

Table 2.3
Generating item intercepts per measurement occasion with small, medium and large violations of invariance.

	Item	Time1	Small violations				Medium violations				Large violations			
			Time 2	Time 3	Time 4	Time 5	Time 2	Time 3	Time 4	Time 5	Time 2	Time 3	Time 4	Time 5
Non-invariant	1	0.5	0.535	0.570	0.605	0.640	0.640	0.780	0.920	1.060	1.060	1.620	2.180	2.740
	2	0.6	0.645	0.690	0.735	0.780	0.780	0.960	1.140	1.320	1.320	2.040	2.760	3.480
	3	0.3	0.325	0.350	0.375	0.400	0.400	0.500	0.600	0.700	0.700	1.100	1.500	1.900
	4	0.4	0.430	0.460	0.490	0.520	0.520	0.640	0.760	0.880	0.880	1.360	1.840	2.320
	5	0.6	0.640	0.680	0.720	0.760	0.760	0.920	1.080	1.240	1.240	1.880	2.520	3.160
	6	0.4	0.420	0.440	0.460	0.480	0.480	0.560	0.640	0.720	0.720	1.040	1.360	1.680
Invariant	7	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
	8	0	0	0	0	0	0	0	0	0	0	0	0	0
	9	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3

2.2 Generating models

The data were generated under the COFM or under the AR quasi-simplex model. Data were generated for five measurement occasions, since Leite (2007) and Fan and Fan (2005) reported no convergence problems in LGM with five measurement occasions.

2.2.1 Curve of factors model

The mean and variance for the growth latent variables were set as the same values used in the simulation study by Muthén and Muthén (2002). The generating parameter values for the growth latent variables were,

$$\begin{bmatrix} \mu_{\eta_1} \\ \mu_{\eta_2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0.2 \end{bmatrix} \qquad \boldsymbol{\Psi} = \begin{bmatrix} 0.5 & 0.044721 \\ 0.044721 & 0.1 \end{bmatrix} \tag{59}$$

The covariance between the intercept and the slope latent variables were set such that the correlation corresponds to a value of 0.2. The loadings from the intercept latent variable η_1 to the first order latent variables $\boldsymbol{\xi}$ were set to 1, and the loadings from the slope latent factor η_2 to the latent variable $\boldsymbol{\xi}$ were set to 0, 1, 2, 3 and 4 at each measurement occasion.

The residual variances for the latent variables $\boldsymbol{\xi}$ were chosen such that the proportion of variance in $\boldsymbol{\xi}$ explained by the latent growth factors $\boldsymbol{\eta}$ correspond to values of 0.80 in all measurement occasions. The variance of $\boldsymbol{\xi}$ is defined as,

$$\sigma^2_{\xi_t} = \gamma_t \Psi \gamma_t' + \theta_{\xi_t} \tag{60}$$

$$\sigma^2_{\xi_t} = \Psi_{11} + \gamma_t^2\Psi_{22} + 2\gamma_t\Psi_{21} + \theta_{\xi_t}$$

The proportion of variance in $\boldsymbol{\xi}$ explained by the latent growth factors $\boldsymbol{\eta}$ is defined as,

$$R^2(\xi_t) = \frac{\Psi_{11} + \gamma_t^2\Psi_{22} + 2\gamma_t\Psi_{21}}{\Psi_{11} + \gamma_t^2\Psi_{22} + 2\gamma_t\Psi_{21} + \theta_{\xi_t}} \tag{61}$$

The resulting variances for $\boldsymbol{\xi}_t$ and its residual variances $\boldsymbol{\theta}_{\xi_t}$ that yield R^2 values of 0.80 are shown in Table 2.4. Table 2.4 also shows the means of $\boldsymbol{\xi}_t$ calculated as,

$$E(\boldsymbol{\xi}_t) = \boldsymbol{\mu}_{\xi_t} = \boldsymbol{\Gamma}\boldsymbol{\mu}_\eta \tag{62}$$

Table 2.4
Generating means and variances for $\boldsymbol{\xi}_t$ in the curve of factors model

	Mean μ_{ξ_t}	Variance $\sigma^2_{\xi_t}$	Residual var. θ_{ξ_t}
ξ_1	0	0.62	0.12
ξ_2	0.2	0.86	0.17
ξ_3	0.4	1.34	0.26
ξ_4	0.6	2.07	0.40
ξ_5	0.8	3.06	0.60

Figure 2.1 shows the curve of factors model with the generating parameters for the structural part of the model. The generating parameters for the measurement part of the model are not shown.

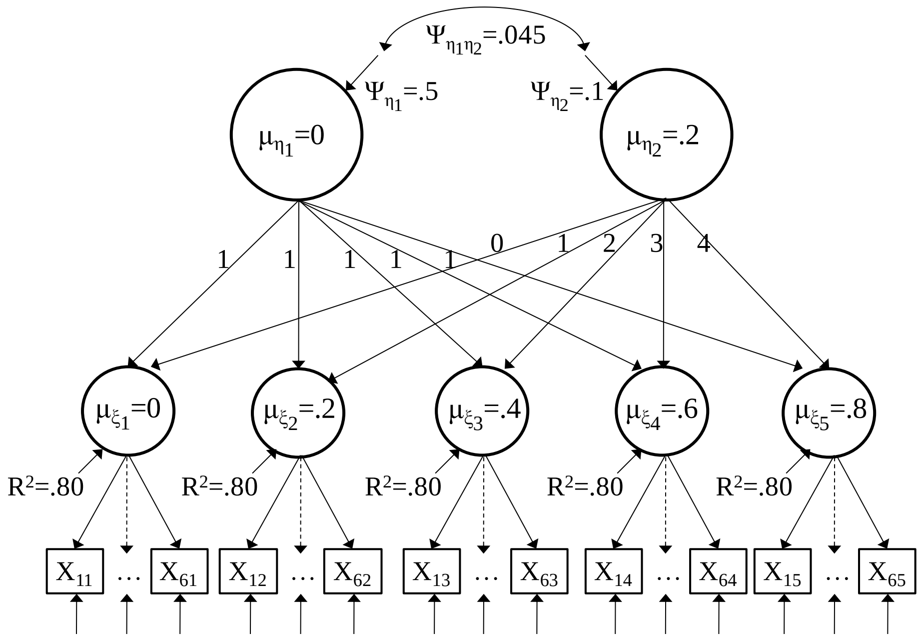

Figure 2.1 Curve of factors model with the generating parameter values

The generating parameter values for the measurement part of the model were determined for a pool of 9 items. Table 2.5 shows the generating parameter values of the intercepts, loadings and unique variance for the condition in which longitudinal measurement invariance holds. In the conditions with small, medium and large violations of metric longitudinal invariance, the parameter values for the loadings changed over time as shown in Table 2.2, while the intercepts and unique variances were invariant over time. In the same way, in the conditions with violations of strong factorial invariance the values of the intercepts changed over time as indicated in Table 2.3, while the loadings and unique variances were invariant over time.

Table 2.5
Generating item loadings, intercepts and unique variances used in the invariant conditions

	Intercept τ	Loading λ	Unique variance θ
Item 1	0.5	0.7	0.7
Item 2	0.6	0.9	1.3
Item 3	0.3	0.5	0.6
Item 4	0.4	0.6	0.8
Item 5	0.6	0.8	1.3
Item 6	0.4	0.4	0.4
Item 7	0.3	0.5	0.7
Item 8	0	1	1.9
Item 9	0.3	0.6	1

The values of the unique variances were selected such that the item communalities across all measurement occasions were between 0.1 and 0.68 based on the Equation (57). Appendix A contains the communality values for each item in each condition.

$$h_j^2 = \frac{\lambda_j^2 \phi_{\xi_j}}{\lambda_j^2 \phi_{\xi_j} + \theta_\delta} \tag{63}$$

One thing to notice is that no lagged covariances between unique factor scores of the same variable over time were allowed. Although it is reasonable to assume that the unique factor scores are correlated over time, when composites of items are formed at each time point the covariances between unique factor scores are ignored. Hence, simulating data with lagged covariances and then compositing the items would introduce a source of bias. Wirth (1998) found that mean scores at each measurement occasion were less biased when there were no lagged covariances between the same item over time than when the items were generated to have lagged covariances over time. On the other hand, Biesanz (2012) found that when using composites of items with correlated unique variances over time, the autoregressive parameters are inflated since the correlations among the unique variances are ignored when using composites.

2.2.2 Autoregressive quasi-simplex model

To select the generating values of an autoregressive quasi-simplex model the study of Morera, et al. (1998) was considered in which six waves of data from a smoking intervention study were analyzed using a quasi-simplex model. For the purposes of the present study only estimates from five of the six waves were considered. Figure 2.2 shows the unstandardized parameter estimates obtained by Morera et al. for waves 1, 2, 3, 4 and 5, and used in the present study as generating values for the structural part of the autoregressive model. The one-lagged unstandardized autoregressive path coefficients were constrained to the same value across measurement occasions for ease of interpretation (Biesanz, 2012). Path coefficients for lags greater than one were set to zero.

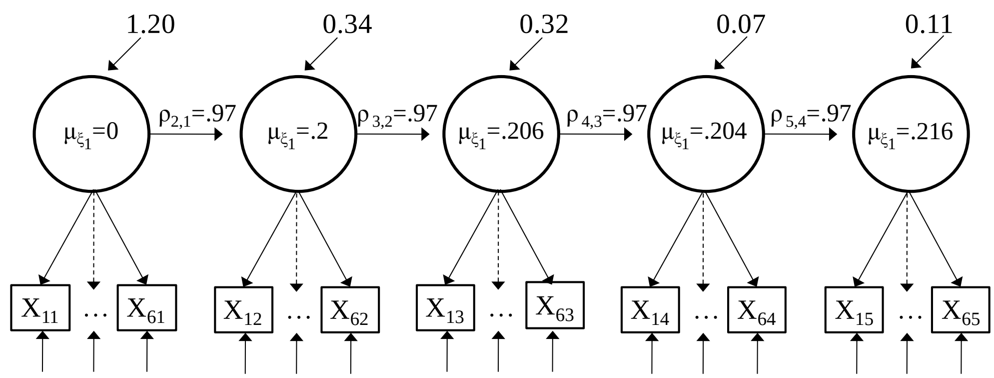

Figure 2.2. Autoregressive quasi-simplex model with the generating parameter values

Since the effect sizes defined for the violations in strong measurement invariance take into account the factor means as indicated in Equation (57), the factor means in the autoregressive model were generated with the same values as the factor means in the LGM and shown in Table 2.4. The factor means from the first to the fifth measurement occasions were 0, 0.2, 0.4, 0.6, and 0.8.

The generating parameter values for the measurement part of the model, not shown in Figure 2.2, were determined as described above and using the values from Tables 2.2, and 2.3 for the non-invariant conditions, and with the values from Table 2.5 for the condition in which longitudinal measurement invariance holds.

2.3 Data analysis

After generating the data under the curve of factors model and the AR quasi-simplex model, composites were generated by summing the items at each measurement occasion. Either a univariate LGM or an AR quasi-simplex model were fit to the composites of the items.

When fitting the univariate LGM the factor loadings of the intercept factor η_1 were fixed to one and the factor loadings of the slope factorη_2 were fixed to 0, 1, 2, 3, and 4 for time points 1 to 5. The means of the intercept and the slope factors were freely estimated, as well as their variances and covariances.

When fitting the AR quasi-simplex model the autoregressive paths were freely estimated. Also, the residual variances of the composites Y at the measurement occasions one and two were constrained to have the same values $\theta_{11} = \theta_{22}$, as well as the residual variances of the composites Y at the measurement occasions four and five $\theta_{44} = \theta_{55}$. The intercepts of the composites Y were fixed to 0. The means and variances of the latent variables were freely estimated at each time point.

2.3.1 Convergence

The first step in the analysis of the results was to examine the number of replications needed to obtain 1000 converged solutions. Solutions with convergence problems or with improper solutions were not included in subsequent analyses.

2.3.2 Parameter estimation

The ability to recover the generating parameter values in the presence of violations of invariance was evaluated in the univariate LGM and the AR quasi-simplex model. A raw bias statistic was computed for each of the estimated parameters denoted $\hat{\theta}_c$ as shown in Equation (64). In the latent curve model the estimated parameters correspond to the means, variances and covariances of the intercept and slope factors, and in the AR the estimated parameters are the path coefficients.

$$B(\hat{\theta}_c) = R^{-1} \sum_{r=1}^{R} (\hat{\theta}_{rc} - \theta_c) \qquad (64)$$

where R refers to the total number of replications that converged to a solution, θ_c refers to the generating parameter value in LGM and in the AR quasi-simplex model; $\hat{\theta}_{rc}$ refers to the parameter estimate for replication r in condition c.

Another criteria for assessing the impact of violations of invariance was relative bias, in which the mean difference of the parameter estimates at each condition and the generating parameter value is divided by the generating parameter value,

$$RB(\hat{\theta}_c) = R^{-1} \sum_{r=1}^{R} \frac{(\hat{\theta}_{rc} - \theta_c)}{\theta_c} \tag{65}$$

Hoogland and Boomsma (1998) indicated that the relative bias of parameter estimates is considered acceptable when its absolute value is less than .05.

The stability of the parameter estimates in LGM and in AR quasi-simplex model was evaluated with the standard error of $\hat{\theta}_c$. The standard error of the estimates is defined as,

$$SE(\hat{\theta}_c) = \sqrt{R^{-1} \sum_{r=1}^{R} (\hat{\theta}_{rc} - \bar{\bar{\theta}})^2} \tag{66}$$

where $\bar{\bar{\theta}}$ is the mean of the parameter estimates across conditions and replications.

To have an overall measure of the accuracy of the parameter estimates that considered both the bias in the parameter estimates and their stability, the root mean square error (RMSE) was also calculated,

$$\text{RMSE}_{\hat{\theta}_c} = \sqrt{R^{-1}\sum_{r=1}^{R}(\hat{\theta}_{rc} - \theta_c)^2} \tag{67}$$

To determine the effect of the independent variables (sample size, number of items, proportion of non-invariant items and magnitude of the violations of invariance) several ANOVAs were conducted on the bias, relative bias, standard errors and RMSE of each of the parameter estimates of the LGM and AR-simplex. The meaningfulness of the ANOVA results were determined by η^2 and Cohen's (1988) values were used to judge small 0.01, medium 0.06 and large 0.14 associations between the variables. Graphic devices were used to compare bias, relative bias and standard errors across conditions.

2.3.3 Model fit

The χ^2 value of model fit in each replication was compared to the χ^2 critical value that would be expected from a correctly specified model that is a function of the degrees of freedom of the LGM or the AR quasi-simplex model and an α=.05. The rejection rates were computed in each condition as the proportion of replications in which the χ^2 value of model fit indicates lack of fit.

Chapter 3

RESULTS

This chapter presents the results of the simulation study in which a LGM or an AR quasi-simplex model were fit to composites formed from items with different levels of violations to measurement invariance. First, the results of the LGM are described followed by the description of the AR quasi-simplex results. Some issues were encountered when analyzing the results of the LGM and the AR quasi-simplex; in the LGM it was found that an adjustment of the true growth parameter values was necessary, while in the AR quasi-simplex the identification constraints initially proposed were modified. The sections of each of the models start with a description of these issues.

The presentation of the results of the LGM and the AR quasi-simplex is divided into three sections. First, the non-convergence percentages are presented. Then, the ability to recover the generating parameter values is described in terms of the bias, relative bias, standard errors and RMSE of each parameter estimate. Finally, the fit of the models are examined in terms of the rejection rates.

For ease of presentation, graphs are shown when possible. To facilitate the display of the information the name of the conditions with invariant loadings and intercepts, and with non-invariance in the intercepts or loadings are abbreviated as shown in Table 3.1.

Table 3.1
Acronyms for the conditions examined

Acronym	Condition
InvLI	Invariant loadings and invariant intercepts
NiLd	Non-invariant loadings, invariant intercepts
NiIn	Non-invariant intercepts, invariant loadings

3.1 Latent growth model

3.1.1 Re-scaling of the generating growth parameters

As indicated in the method section, multiple indicators were generated under the curve of factors models. The multiple indicators were summed at each time point, and an LGM was used to analyze the composites. It was found that by substituting the latent factors ξ_t of the curve of factors model, by the composites Y_t a scaling factor was introduced that changed the growth parameter estimates even in the conditions with invariant loadings and intercepts. Since this change was systematic, it was possible to develop a re-scaling of the growth parameters that corrected for the change. The re-scaling depended on the number of items per measurement occasion and the values of the item intercept sums and item loadings sums (See Appendix C). The following results were obtained after the re-scaling of the growth parameters.

3.1.2 Non-convergence percentages

The first criterion used to evaluate the results was the number of replications per condition with convergence problems. As expected, in the conditions with invariant loadings and intercepts there were no replications with convergence problems.

Overall, in the conditions with violations of invariance over time the total number of replications with convergence problems or improper solutions was small. While all the replications with sample sizes of 200, 500 and 1000 reached convergence, the conditions with a sample size of 100 had some replications with non-convergence, which suggests that convergence problems were a consequence of small sample sizes and not due to the violations of invariance. Table 3.2 shows the percentage of replications with convergence problems for the conditions with a sample size of 100. It can be observed that among those conditions, only a small percentage of replications (0.1 to 0.3%) resulted in non-convergence.

The replications with non-convergence were replaced so that the computations of the bias, relative bias, standard errors and RMSE of the growth parameter estimates were based in a total of 1000 converged replications.

Table 3.2
Non-convergence percentages for the LGM conditions with N=100

Number of items	Proportion of non-inv. items	Effect size	Non-invariant loadings	Non-invariant intercepts
6	1/3	Small	0	0
		Medium	0	0.1
		Large	0.1	0
	2/3	Small	0.1	0
		Medium	0.1	0
		Large	0.1	0
9	1/3	Small	0.1	0
		Medium	0	0
		Large	0	0
	2/3	Small	0	0
		Medium	0	0
		Large	0.1	0.1
15	1/3	Small	0.1	0
		Medium	0	0.1
		Large	0.3	0
	2/3	Small	0.3	0.1
		Medium	0.1	0.1
		Large	0.1	0.2

3.1.3 Parameter estimation

Parameter estimation under violations of invariance was evaluated by examining at the bias, relative bias, standard errors and RMSE for each growth parameter estimate.

The results are divided in three sections: bias and relative bias, standard errors and RMSE.

Bias and relative bias

Since the results of bias and relative bias are comparable, only the tables with the relative bias results are presented. It was decided to present the tables with the relative bias results since there is a clear cutoff for judging the magnitude of relative bias (Hoogland & Boomsma, 1998). The tables with the bias results across conditions can be consulted in Appendix D. To further simplify the relative bias tables the results for the different sample sizes were collapsed, since the ANOVA results described below showed that the sample size did not have an effect on the relative bias of the growth parameter estimates.

As expected, the bias and relative bias of the growth parameter estimates were acceptable in the conditions with invariant loadings and intercepts. A series of ANOVAs were conducted to examine the effect of the independent variables on the bias and relative bias of the parameter estimates. The results of the ANOVAs were judged by the overall η^2 and by the η^2 values of the interactions and main effects of the independent variables. The ANOVA results indicated that none of the conditions had a η^2 value above 0.01, which is a small effect size following Cohen's suggestion (Cohen, 1988). Table 3.3 shows the relative bias values of the parameter estimates in the conditions with invariant loadings and intercepts.

The relative bias values for each growth parameter estimate in the non-invariant conditions are presented in Table 3.4. In general, it can be observed that with non-invariant loadings the absolute values of the relative bias of the slope factor mean, the slope factor variance and the intercept-slope covariance are larger than in the invariant conditions and larger than the suggested cutoff of 0.05. In contrast, with non-invariant intercepts the only parameter that showed relative bias values larger than 0.05 was the slope factor mean.

Table 3.3
Relative bias of LGM parameter estimates in the invariant conditions

Num. items	Intercept factor mean	Intercept factor variance	Slope factor mean	Slope factor variance	Intercept-slope covariance
6	0.001	0.001	-0.003	-0.004	0.006
9	0.001	0.000	-0.005	-0.005	0.031
15	-0.001	0.002	-0.007	-0.006	0.014

Table 3.4

Relative bias of LGM parameter estimates under violations of invariance

Num. items	Effect size	Prop. non-inv.	Non-invariant loadings					Non-invariant intercepts				
			Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.
6	Small	1/3	0.009	**-0.053**	-0.001	**-0.122**	**-0.126**	0.000	**0.101**	-0.007	-0.009	0.013
		2/3	0.017	**-0.092**	0.001	**-0.192**	**-0.238**	0.001	**0.159**	-0.008	-0.003	0.012
	Medium	1/3	0.023	**-0.112**	0.002	**-0.228**	**-0.295**	0.002	**0.383**	-0.002	-0.010	0.015
		2/3	0.037	**-0.183**	0.001	**-0.355**	**-0.508**	-0.001	**0.643**	-0.008	-0.005	0.009
	Large	1/3	0.036	**-0.167**	0.005	**-0.324**	**-0.446**	0.000	**1.527**	-0.011	-0.005	0.007
		2/3	**0.063**	**-0.280**	0.013	**-0.500**	**-0.803**	0.002	**2.572**	-0.005	-0.004	0.017
9	Small	1/3	0.010	**-0.050**	-0.004	**-0.111**	**-0.124**	-0.001	**0.087**	-0.005	-0.003	0.010
		2/3	0.016	**-0.095**	0.004	**-0.198**	**-0.236**	0.000	**0.160**	-0.004	-0.007	0.016
	Medium	1/3	0.019	**-0.101**	0.001	**-0.213**	**-0.270**	0.000	**0.351**	-0.005	-0.007	0.004
		2/3	0.036	**-0.189**	0.005	**-0.361**	**-0.514**	0.000	**0.650**	-0.003	-0.009	0.007
	Large	1/3	0.029	**-0.153**	0.004	**-0.301**	**-0.417**	0.000	**1.403**	-0.006	-0.008	0.014
		2/3	**0.053**	**-0.285**	0.012	**-0.507**	**-0.817**	0.000	**2.601**	-0.003	-0.003	0.021
15	Small	1/3	0.008	**-0.052**	-0.001	**-0.115**	**-0.115**	-0.002	**0.087**	-0.004	-0.006	0.014
		2/3	0.014	**-0.095**	-0.001	**-0.202**	**-0.245**	-0.003	**0.164**	-0.005	-0.004	0.006
	Medium	1/3	0.018	**-0.105**	-0.001	**-0.212**	**-0.259**	-0.001	**0.354**	-0.007	-0.005	0.012
		2/3	0.031	**-0.196**	0.006	**-0.376**	**-0.526**	-0.001	**0.669**	-0.004	-0.003	0.007
	Large	1/3	0.025	**-0.153**	0.005	**-0.310**	**-0.406**	0.001	**1.415**	-0.006	-0.005	0.015
		2/3	**0.052**	**-0.299**	0.018	**-0.523**	**-0.833**	-0.001	**2.668**	-0.006	-0.004	0.018

Note: The bolded numbers correspond to relative bias absolute values larger than 0.05.

Next, a detailed description of the bias, relative bias and ANOVA results for each growth parameter estimates is provided. The η^2 values for the conditions with at least a small effect size are shown in Table 3.5.

Table 3.5
η^2values from the ANOVAs on bias and relative bias of the LGM parameter estimates

	Non-invariant loadings								Non-invariant intercepts	
	Int. mean		Slope mean		Slope var.		I-S covar.		Slope mean	
	Bias	Rel. bias	Bias	Rel. bias	Bias	Rel. bias	Bias	Rel. bias	Bias	Rel. bias
Overall effect	0.04	0.03	0.47	0.36	0.86	0.65	0.44	0.24	0.99	0.98
N. Items	--	--	0.14	--	0.52	--	0.21	--	0.10	--
Prop. Non-inv	--	--	0.09	0.11	0.07	0.20	0.05	0.07	0.07	0.08
Magnitude	0.02	0.02	0.18	0.23	0.15	0.42	0.09	0.15	0.66	0.83
N. items x Prop.	--	--	0.02	--	0.04	--	0.03	--	0.01	--
N. items x Mag.	--	--	0.02	--	0.07	--	0.04	--	0.09	--
Prop. x Mag.	--	--	0.02	0.02	--	0.02	0.01	0.02	0.06	0.07

Intercept factor mean

As indicated in Table 3.4, the intercept factor mean estimates showed relative bias values larger or at the cutoff of 0.05 only in the conditions with non-invariant loadings with large violations of invariance and with 2/3 of non-invariant items. In the rest of the non-invariant loading conditions and in all the non-invariant intercept conditions the relative bias values were below 0.05.

Figures 3.1 and 3.2 show the bias and relative bias of the intercept factor mean across conditions. The results by the number of items were collapsed since no differences were found. It can be seen that the values are very similar under violations of invariance and in the invariant conditions. As expected, the ANOVA results (Table 3.5) showed only a small effect size for the magnitude of the violations under the conditions with non-invariant loadings.

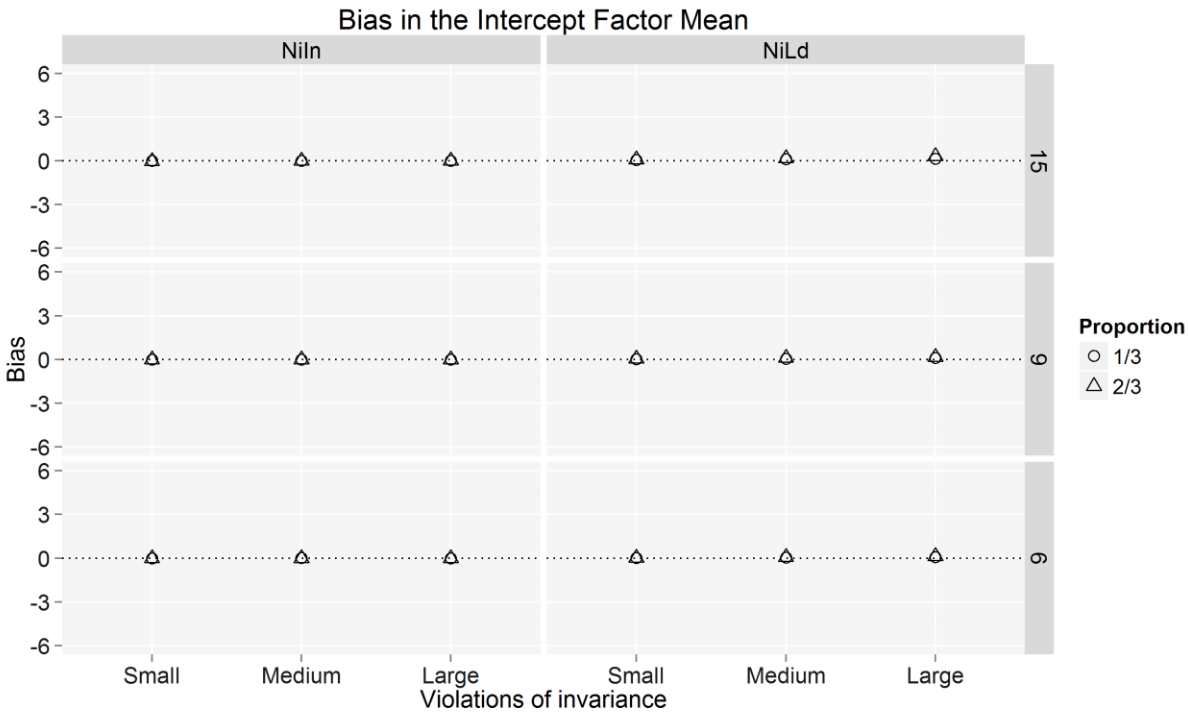

Figure 3.1 Bias in the intercept factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the intercept factor mean in the conditions with invariant loadings and invariant intercepts.

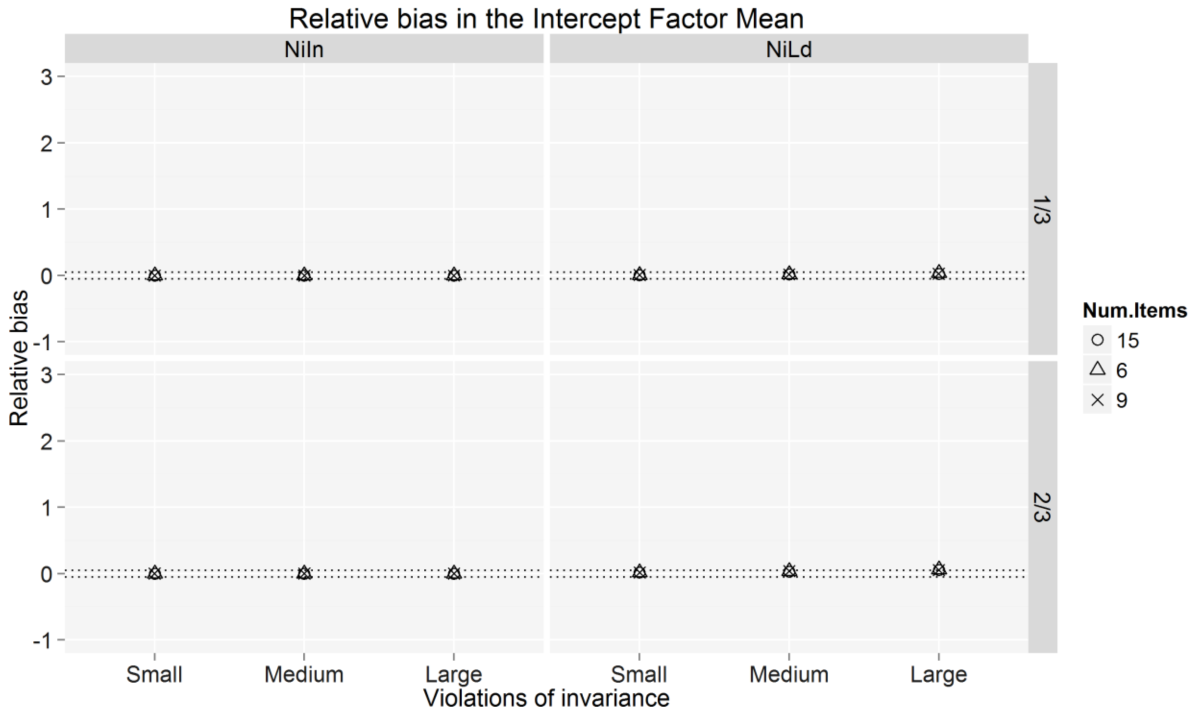

Figure 3.2 Relative bias in the intercept factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

Slope factor mean

As indicated in Table 3.4, across all conditions with non-invariant loadings the parameter estimates of the slope factor mean were underestimated while they were overestimated in the conditions with non-invariant intercepts. Only the relative bias absolute values in the non-invariant loading conditions with small violations of invariance and 1/3 of non-invariant items were close to cutoff of 0.05. The relative bias absolute values for the rest of the conditions showed absolute values above the recommended cutoff.

It should be noticed in Table 3.4 that the relative bias values were larger in the non-invariant intercept conditions than under the non-invariant loading conditions. While the relative bias absolute values in the non-invariant loading conditions ranged from 0.05 to 0.30, the absolute values under the non-invariant intercepts ranged from .09 to 2.67. This finding can be observed in Figures 3.3 and 3.4. It can also be seen that as the magnitude of violations increase and as the proportion of non-invariant items increase, the bias and relative bias absolute values increased.

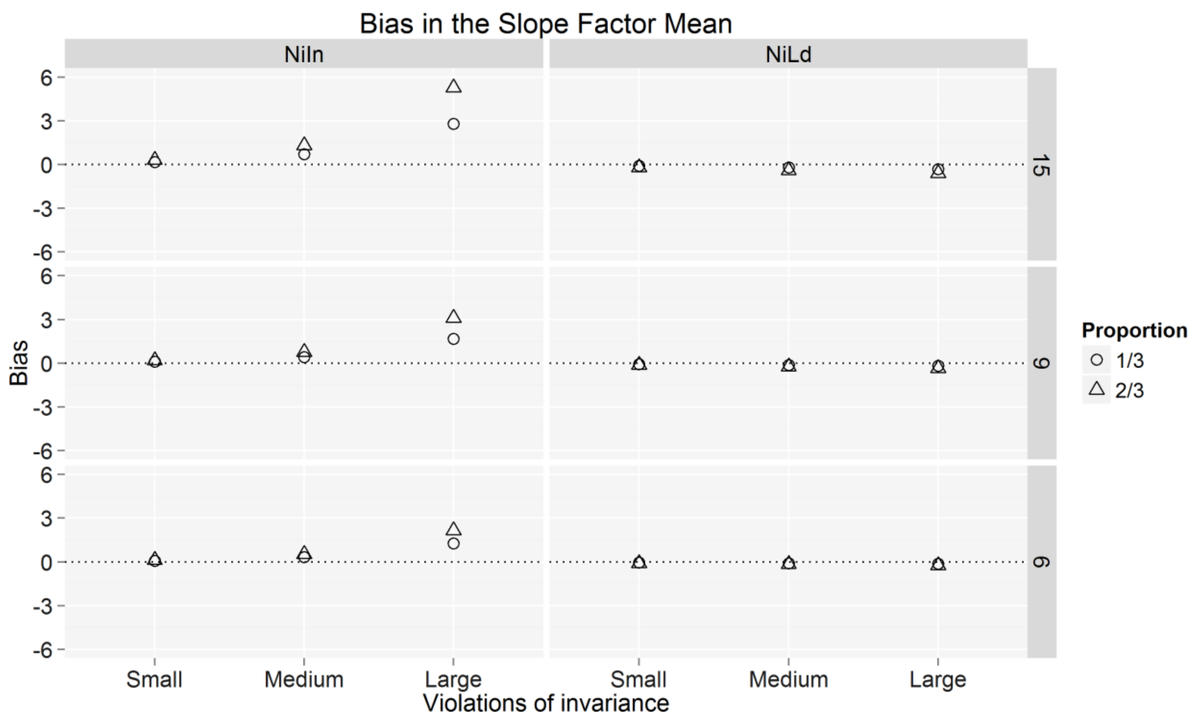

Figure 3.3 Bias in the slope factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the slope factor mean in the conditions with invariant loadings and invariant intercepts (InLI).

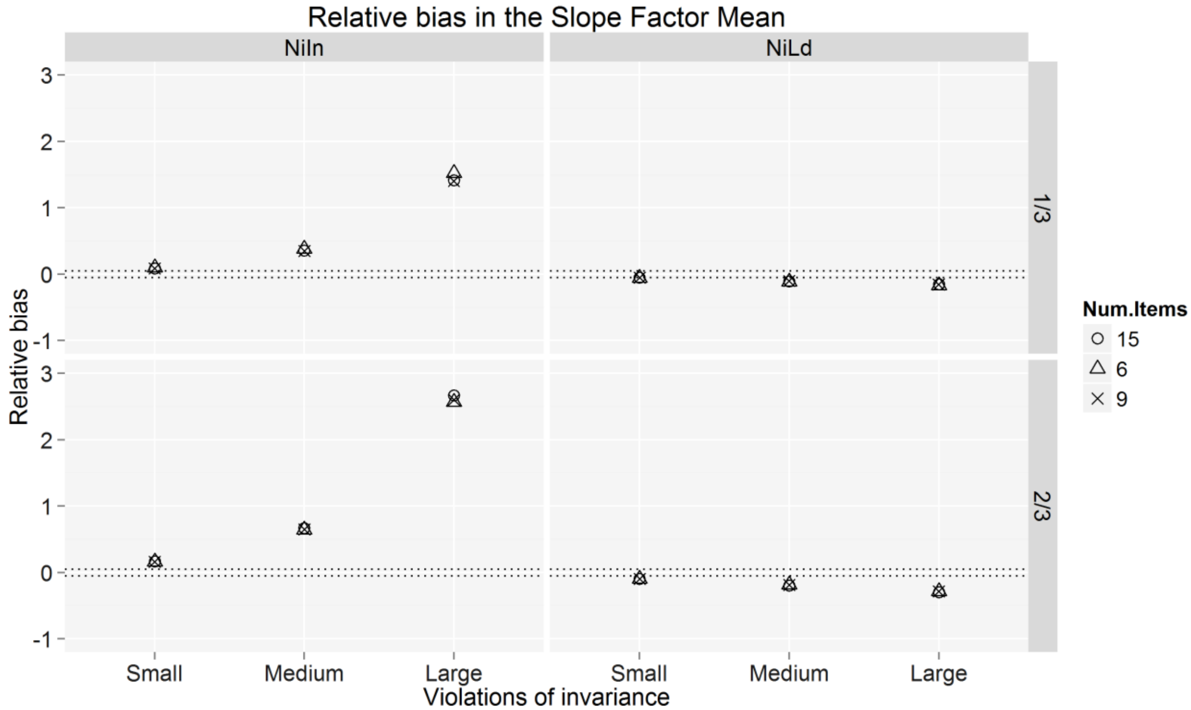

Figure 3.4 Relative bias in the slope factor mean in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

The ANOVA results on the bias of the slope factor mean indicated that in the non-invariant loading conditions the magnitude of violations explained 18% of the variance, while the number of items explained 14% and the proportion of non-invariant items explained 9%. Regarding relative bias, the ANOVA results revealed a medium effect size for the proportion of non-invariant items ($\eta^2 = .11$) and a large effect size ($\eta^2 = .23$) for the magnitude of violations.

Under the non-invariant intercept conditions, the ANOVA results on the bias of the slope factor mean indicated a medium effect size for the interaction between the number of items and the magnitude of violations of invariance ($\eta^2 = 0.09$), and for the interaction between the proportion of non-invariant items and the magnitude of the violations ($\eta^2 = 0.06$). The main effect of the magnitude of violations, the number of items, and the proportion of non-invariant items explained 66, 10 and 7% of the variance respectively.

The ANOVA on the relative bias also showed a medium effect size for the interaction between the proportion of non-invariant items and the magnitude of the violations ($\eta^2 = 0.07$). The proportion of non-invariant items explained 8% of the variance, and the proportion of variance in the relative bias explained by the magnitude of violations increased to 83%.

Intercept factor variance

Table 3.4 shows that the relative bias absolute values of the intercept factor variance were smaller than the cutoff of 0.05 in all conditions with violations of invariance. It should be noted that the intercept factor variance was the only growth parameter estimate that was unbiased in the non-invariant loading conditions. The ANOVA results showed that the independent variables did not have an effect on the bias and relative bias of the intercept factor variance estimates under violations of invariance.

Slope factor variance

In the conditions with non-invariant loadings, the parameter estimates of the slope variance underestimated the true value in all conditions. The relative bias absolute values were larger than 0.05, ranging from 0.11 to 0.52, as shown in Table 3.4. In contrast, in all the conditions with non-invariant intercepts the relative bias absolute values were lower than 0.05.

Figures 3.5 and 3.6 show the bias and relative bias for the slope factor variance across conditions. It should be noticed that in the non-invariant loading conditions, as the magnitude of violations and the proportion of non-invariant items increased, the bias and relative bias absolute values increased. These results were confirmed by the ANOVAs that showed large effects for the magnitude of violations and the proportion of non-invariant items (Table 3.5).

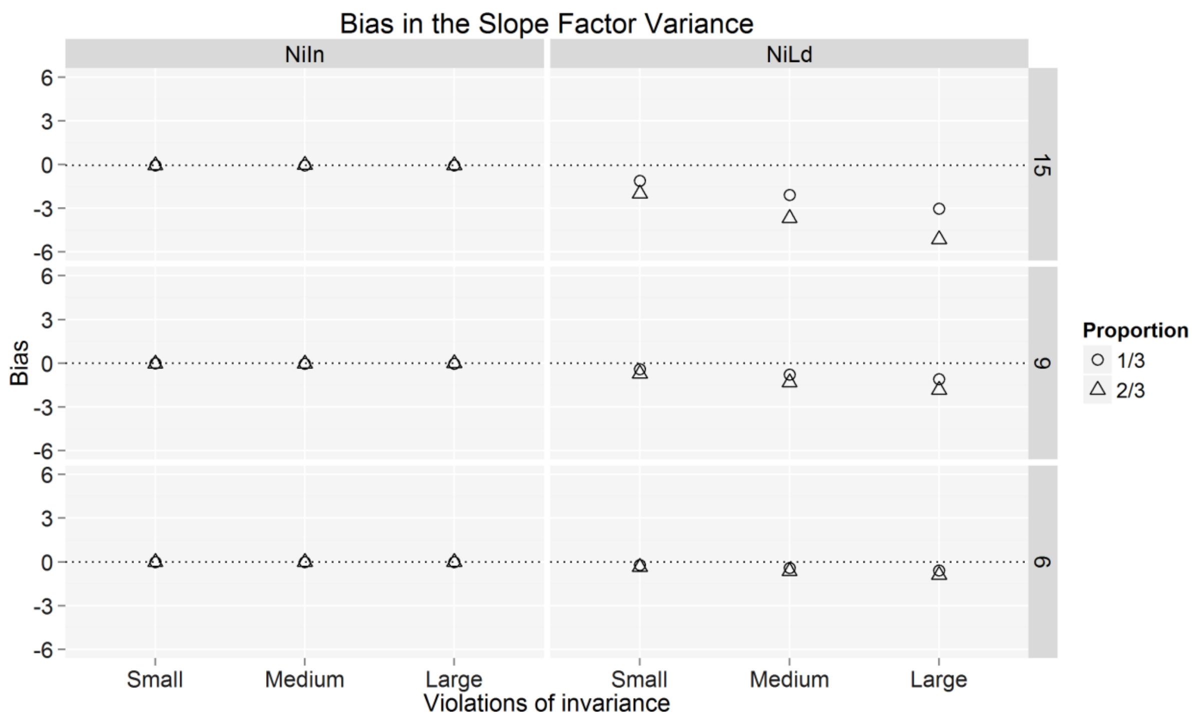

Figure 3.5 Bias in the slope factor variance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the slope factor variance in the conditions with invariant loadings and invariant intercepts (InLI).

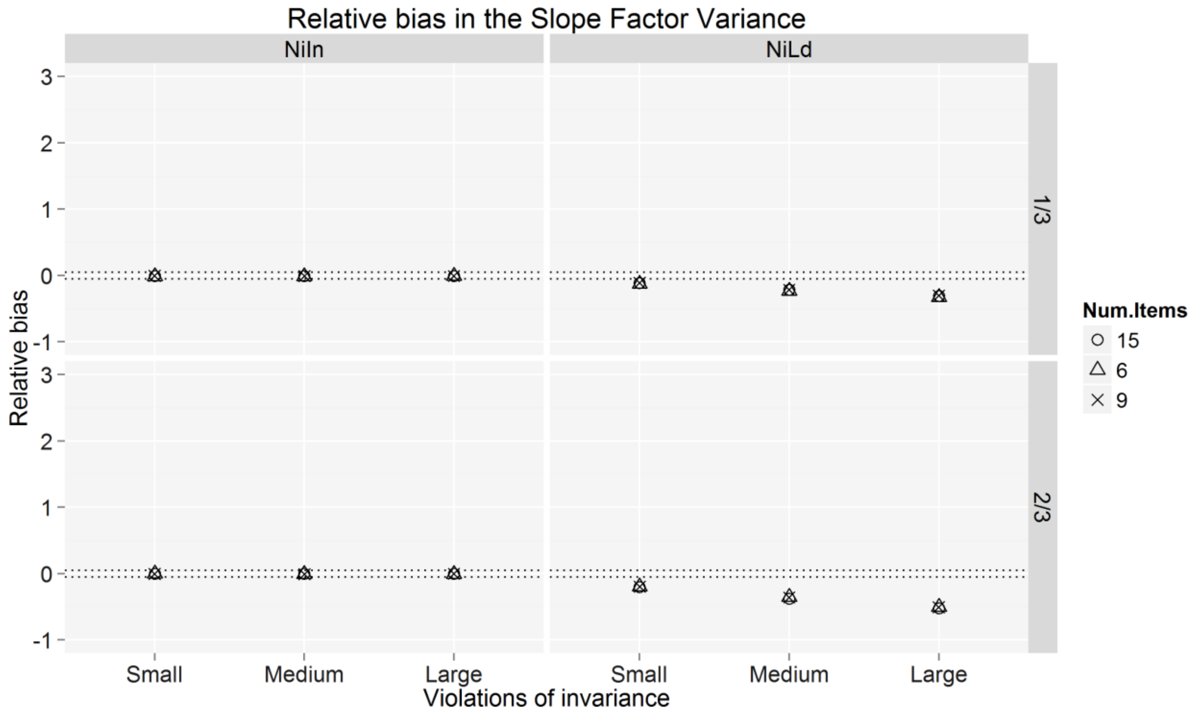

Figure 3.6 Relative bias in the slope factor variance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

More specifically, the ANOVA conducted in the non-invariant loading conditions indicated that the number of items explained 52% of the variance on the bias of the slope factor variance, while the magnitude of violations explained 15% and the proportion of non-invariant items explained 7%. A medium effect size was also found for the interaction between the number of items and the magnitude of the violations to invariance ($\eta^2 = 0.07$). The ANOVA on the relative bias showed that the magnitude of violations explained 42% of the variance and the proportion of non-invariant items explained 20%.

In the non-invariant intercept conditions, the ANOVA on the bias and relative bias did not show a η^2 value larger than 0.01.

Intercept-slope covariance

The covariance between the intercept and slope factors was underestimated across all the non-invariant loading conditions as shown in Figure 3.7. The relative bias absolute values were above 0.05 even in the conditions with small violations of invariance and 1/3 of non-invariant items. Figures 3.7 and 3.8, as well as Table 3.4, show that as the magnitude of violations of invariance and as the proportion of non-invariant items increase, the bias and relative bias absolute values also increased.

The ANOVA conducted in the non-invariant loading conditions showed a large effect size for the number of items (η^2=.21), and a medium effect size for the magnitude of violations of invariance (η^2= .09) on the bias of the intercept-slope covariance. The ANOVA on the relative bias showed a large effect size for the magnitude of violations (η^2 = 0.15) and a medium effect size for the proportion of non-invariant items (η^2 = 0.07).

The ANOVAs conducted on the bias and relative bias in the conditions with non-invariant intercepts did not show η^2 values larger than 0.01.

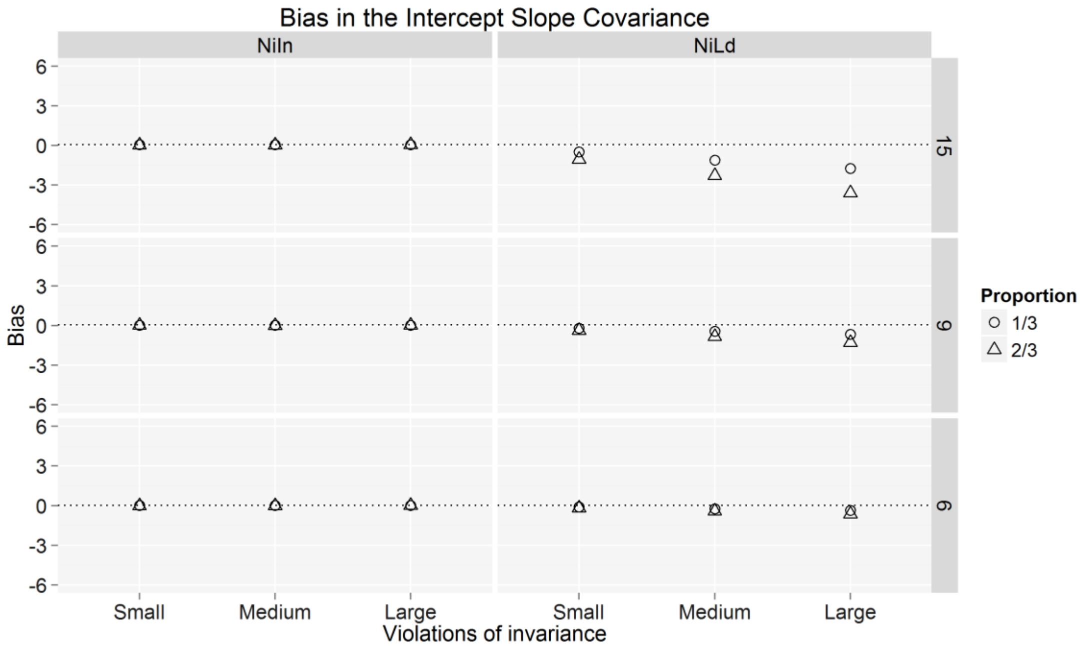

Figure 3.7 Bias in the intercept-slope covariance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the slope factor variance in the conditions with invariant loadings and invariant intercepts (InLI)

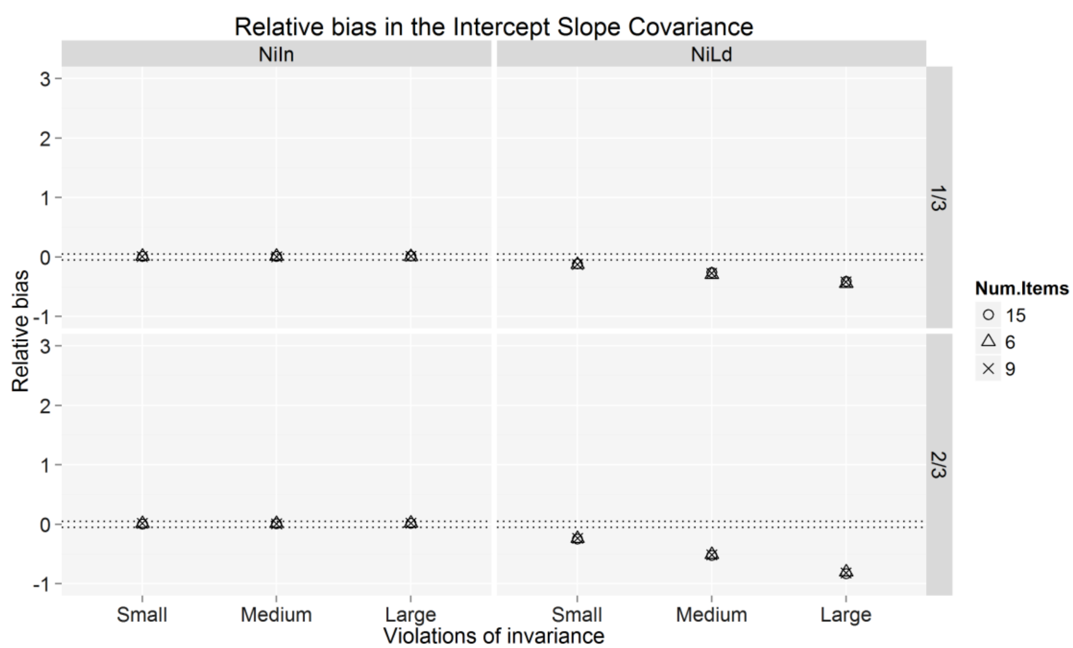

Figure 3.8 Relative bias in the intercept-slope covariance in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

Standard errors

The ANOVAs conducted on the standard errors of the growth parameter estimates showed that the number of items and the sample size had a large effect in all conditions. No other independent variable had medium or large effect sizes. Table 3.6 shows η^2 values only for the conditions with at least small effect sizes.

The standard errors of the growth parameter estimates are shown in Figures 3.9 to 3.13 and are also presented in Table 3.7. Since the magnitude of violations and the proportion of non-invariant items did not have effects on the standard errors, the results shown are averaged over sample size and the number of items. It should be noticed that in order to accommodate the large standard errors observed for the intercept factor variance, the y axis of the graph of the intercept factor variance (Figure 3.11) is in a different scale than the graphs of the other growth parameter estimates.

In general, it can be observed that the standard errors of growth parameter estimates were very similar in the invariant conditions and in the conditions with violations of invariance. In all conditions, the standard errors decreased as the sample size increase and as the number of items decreased. The effect of the number of items can be seen very clearly in Figure 3.11 that corresponds to the standard errors of the intercept variance. The influence of the number of items is also observed in the standard errors of the slope factor variance (Figure 3.12) and in the standard errors of the intercept-slope covariance (Figure 3.13).

Table 3.6

η^2 values from the ANOVAs on the standard errors of the LGM parameter estimates

	Invariant loadings and intercepts					Non-invariant loadings					Non-invariant intercepts				
	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.
Overall	0.94	0.94	0.89	0.89	0.89	0.94	0.93	0.89	0.87	0.89	0.94	0.94	0.89	0.89	0.89
N. Items	0.33	0.35	0.59	0.60	0.59	0.34	0.32	0.60	0.56	0.58	0.33	0.35	0.60	0.60	0.61
Sample size	0.62	0.59	0.31	0.29	0.30	0.60	0.59	0.29	0.28	0.30	0.60	0.59	0.29	0.29	0.29
Magnitude	--	--	--	--	--	--	0.01	--	0.02	0.01	--	--	--	--	--
Proportion	--	--	--	--	--	--	0.01	--	0.01	--	--	--	--	--	--

Table 3.7

Standard errors of the LGM parameter estimates by the number of items and by sample size

Num. items	Sample size	Invariant loadings and intercepts					Non-invariant loadings					Non-invariant intercepts				
		Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.
6	100	0.36	0.16	1.89	0.36	0.58	0.36	0.14	1.88	0.27	0.52	0.35	0.15	1.85	0.35	0.57
	200	0.26	0.11	1.27	0.25	0.42	0.25	0.10	1.31	0.20	0.37	0.25	0.11	1.29	0.24	0.41
	500	0.16	0.07	0.85	0.15	0.26	0.16	0.06	0.81	0.13	0.23	0.16	0.07	0.82	0.15	0.26
	1000	0.11	0.05	0.59	0.11	0.18	0.11	0.04	0.59	0.09	0.17	0.11	0.05	0.58	0.11	0.18
9	100	0.47	0.21	3.49	0.62	1.08	0.48	0.19	3.41	0.51	0.96	0.48	0.21	3.41	0.65	1.05
	200	0.34	0.15	2.42	0.45	0.75	0.34	0.13	2.43	0.35	0.66	0.34	0.15	2.36	0.46	0.76
	500	0.22	0.10	1.50	0.29	0.48	0.22	0.08	1.51	0.23	0.42	0.21	0.10	1.50	0.29	0.48
	1000	0.15	0.07	1.09	0.20	0.33	0.15	0.06	1.09	0.16	0.30	0.15	0.07	1.05	0.21	0.33
15	100	0.74	0.33	8.24	1.51	2.58	0.76	0.29	8.38	1.19	2.26	0.76	0.34	8.20	1.61	2.55
	200	0.54	0.24	5.72	1.14	1.82	0.54	0.20	5.87	0.85	1.60	0.53	0.24	5.79	1.14	1.82
	500	0.32	0.15	3.69	0.66	1.13	0.34	0.13	3.76	0.54	1.01	0.34	0.15	3.66	0.71	1.15
	1000	0.24	0.11	2.54	0.49	0.82	0.24	0.09	2.65	0.38	0.72	0.24	0.11	2.62	0.51	0.83

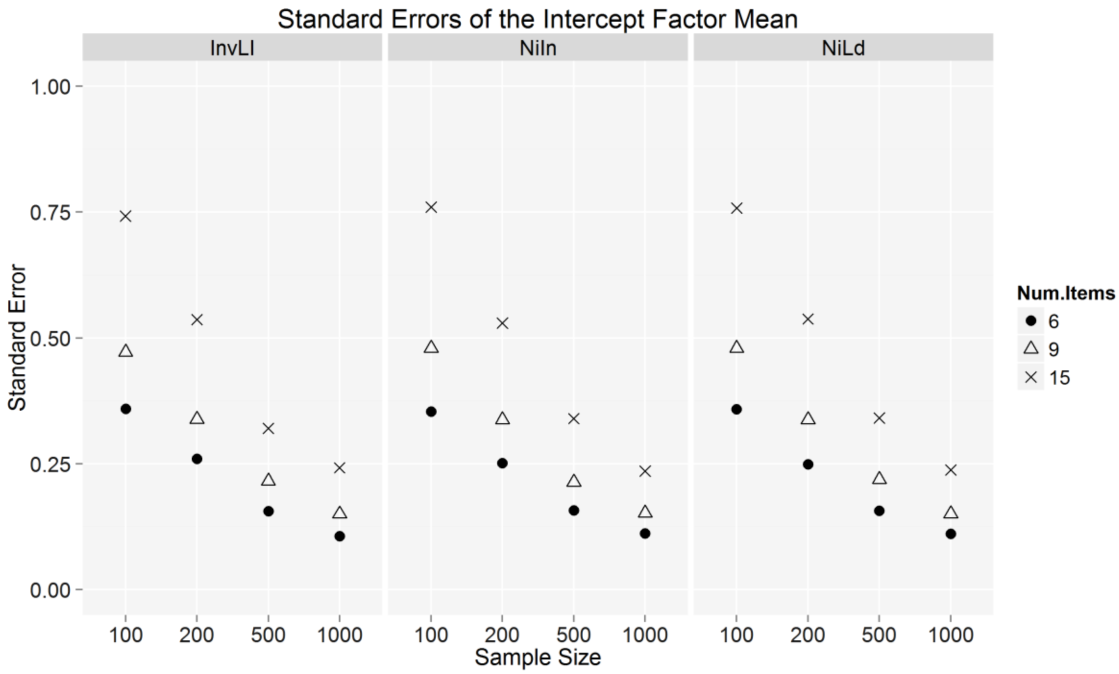

Figure 3.9 Standard errors of the intercept factor mean in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

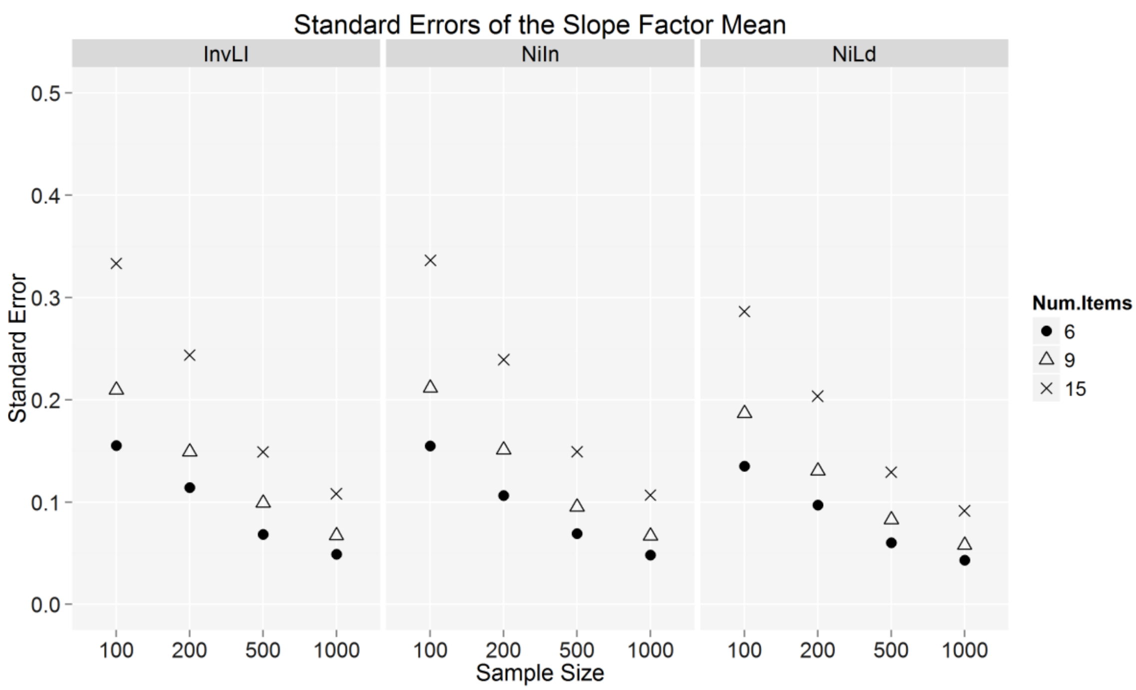

Figure 3.10 Standard errors of the slope factor mean in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

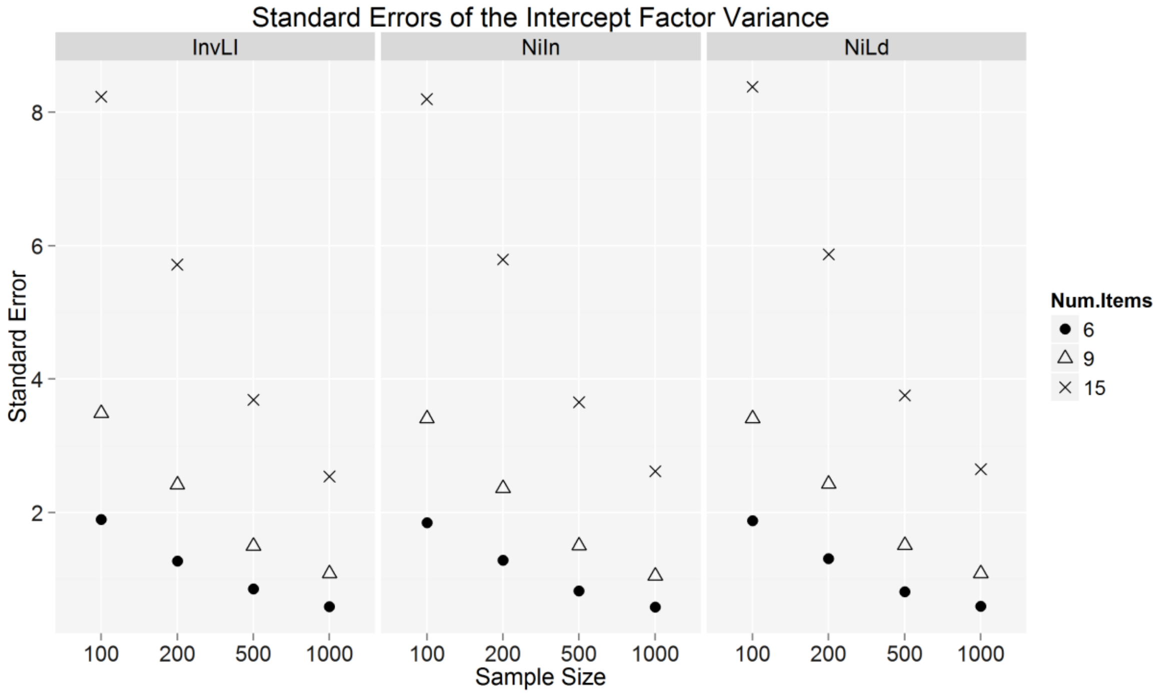

Figure 3.11 Standard errors of the intercept factor variance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

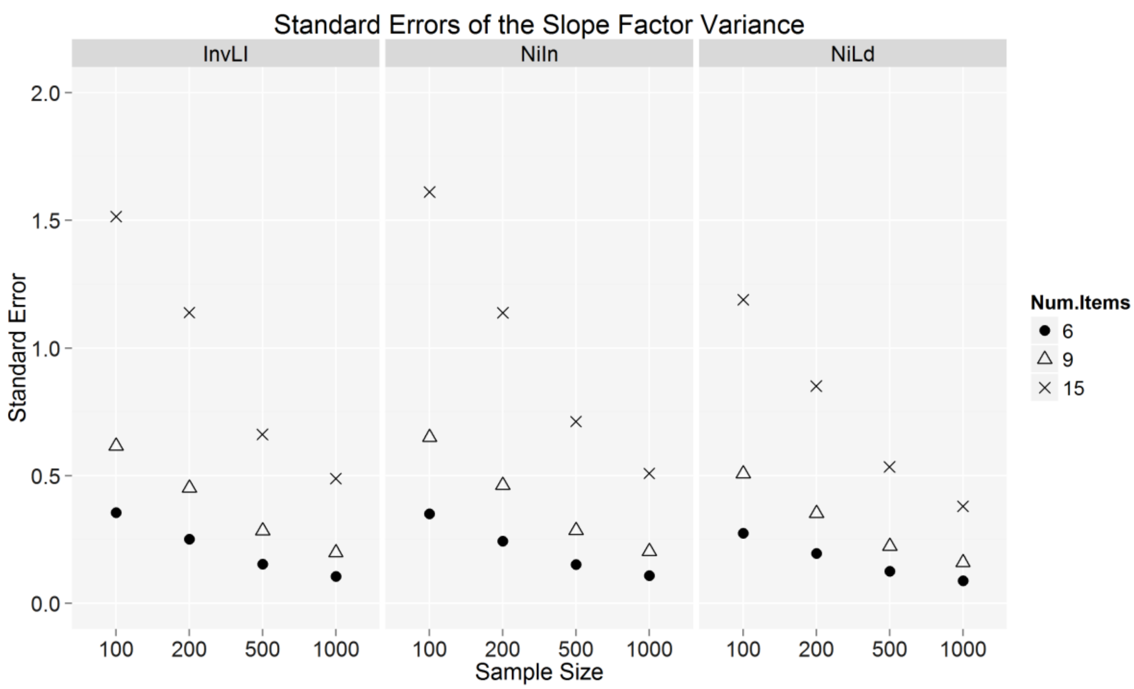

Figure 3.12 Standard errors of the slope factor variance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

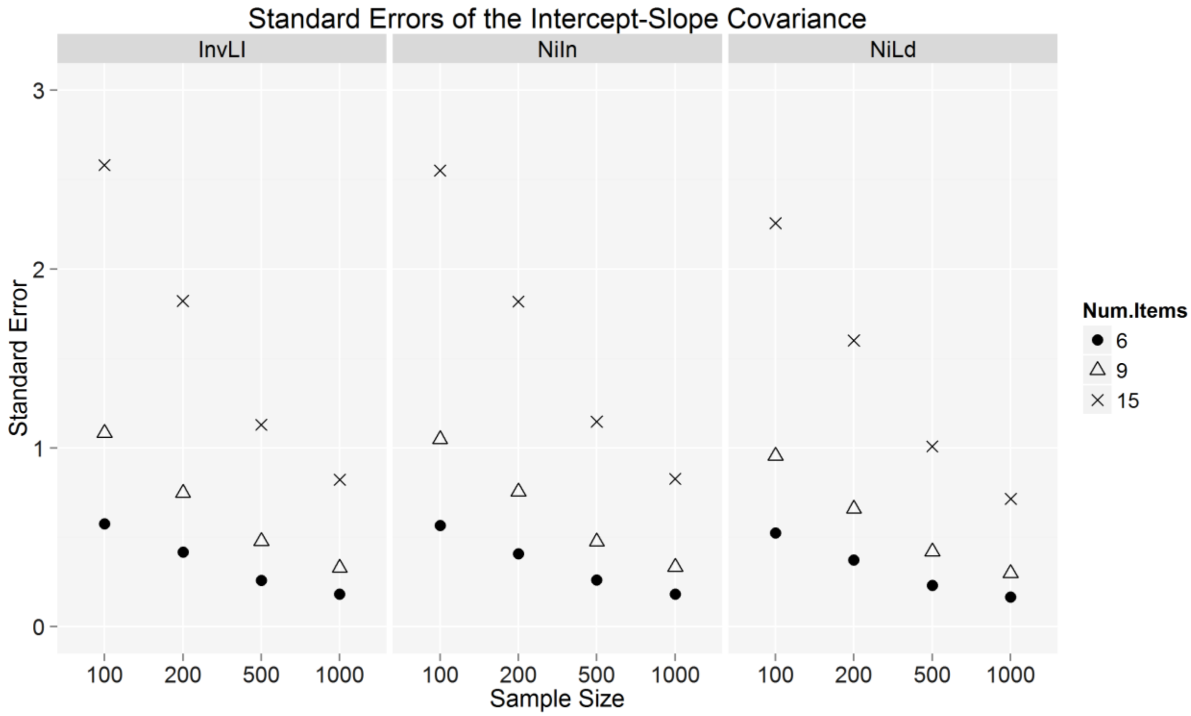

Figure 3.13 Standard errors of the interceptslope covariance in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

Root mean square error (RMSE)

Table 3.8 contains the η^2 values obtained from the ANOVAs on the RMSE of the growth parameter estimates. Table 3.8 indicates that the number of items and the sample size had a large effect on the RMSE in almost all conditions, explaining in some growth parameter estimates up to 66% and 62% of the total variance respectively.

In the non-invariant loading conditions a large effect size was observed for the magnitude of violations of invariance on the slope factor mean (η^2=.24) and on the slope factor variance (η^2=.14), with RMSE values increasing as the magnitude of violations increased. The RMSE values also increased with increases in the proportion of non-invariant items, that showed a large effect on the slope factor mean (η^2=.13) and a medium effect on the slope factor variance (η^2=.07).

In the non-invariant intercept conditions, the magnitude of violations had a large effect size (η^2=.66) on the intercept factor mean, while the proportion of non-invariant items showed a medium effect size (η^2=.07). The direction of the effects was the same as in the non-invariant loading conditions.

Table 3.8

η^2 values from the ANOVAs on the RMSE of the LGM parameter estimates

	Invariant loadings and intercepts					Non-invariant loadings					Non-invariant intercepts				
	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.
Overall	0.94	0.94	0.89	0.89	0.89	0.94	0.88	0.89	0.87	0.94	0.83	0.89	0.89	0.89	0.94
N. Items	0.33	0.35	0.59	0.60	0.59	0.40	0.44	0.61	0.66	0.33	0.11	0.60	0.60	0.61	0.33
Magnitude	--	--	--	--	--	0.01	0.24	--	0.14	--	0.66	--	--	--	--
Proportion	--	--	--	--	--	--	0.13	--	0.07	--	0.07	--	--	--	--
SS	0.62	0.59	0.31	0.29	0.30	0.53	0.06	0.29	--	0.60	--	0.30	0.29	0.29	0.60

3.1.4 Model fit

The last criterion examined was the fit of the LGM under violations of invariance. In general, in the invariant conditions the percentage of replications that incorrectly rejected the hypothesis of linear growth remained close to the nominal level as shown in Table 3.9. In the conditions with just 6 items the rejection rates were below the nominal level for sample sizes larger than 100. In the conditions with 9 items the rejection rates were larger than expected (6%) but as the sample size increased it got closer to the nominal 5%. In the conditions with 15 items the rejection rates were on average close to the nominal level.

Table 3.9
Rejection rates in the invariant conditions in LGM

Number of items	Sample size	Rejection rates
6	100	4.7
	200	3.2
	500	3.6
	1000	3.9
9	100	5.2
	200	5.6
	500	5.6
	1000	4.3
15	100	4.5
	200	5.3
	500	4.3
	1000	4.5

In the non-invariant loading conditions the rejection rates were larger than in the invariant conditions as shown in Table 3.10. As the number of items, the magnitude of the violations of invariance, the proportion of non-invariant items, and the sample size increased, the percentage of replications in which the hypothesis of linear growth was incorrectly rejected increased. With only 6 items, the rejection rates ranged from 5.4 up to 85 in the condition with large violations to invariance, 2/3 of non-invariant items, and a sample size of 1000.

When the total number of items was 9, only in the conditions with small violations to invariance and 1/3 of non-invariant items the rejection rates were close to the nominal level. In all the other conditions the rejection rates were higher. For example, in the condition with large violations to invariance, 2/3 of non-invariant items and a sample size of 1000, 96.4% of the replications rejected the null hypothesis.

In the conditions with 15 items, the rejection error rates were never lower than 7.8%, and increased up to 100% when 2/3 of the items had large violations of invariance and a sample size of 1000.

In contrast, in the conditions in which composites were formed from items with non-invariant intercepts the rejection rates remained close to 5%. Although there were some conditions in which the rejection rates were slightly inflated, the rates are comparable to the invariant conditions.

Table 3.10
Rejection rates in the conditions with violations of invariance in the LGM

Num. Items	Magnitude of violations	Sample size	Non-invariant loadings		Non-invariant intercepts	
			1/3 non-inv.	2/3 non-inv.	1/3 non-inv.	2/3 non-inv.
6	Small	100	6.4	6.2	5.0	5.1
		200	7.0	5.0	5.3	4.7
		500	4.6	7.9	3.9	4.7
		1000	6.7	10.8	4.4	5.1
	Medium	100	5.2	7.8	4.1	6.1
		200	4.7	10.3	5.6	4.6
		500	8.3	21.2	3.6	4.6
		1000	15.5	42.9	4.4	4.2
	Large	100	7.4	11.2	4.2	5.1
		200	8.6	18.0	6.9	4.3
		500	15.7	49.7	5.1	3.9
		1000	30.8	84.4	5.3	4.9
9	Small	100	3.9	5.1	5.5	6.2
		200	4.9	6.8	4.4	5.8
		500	6.4	9.2	3.3	4.3
		1000	5.7	14.4	4.3	4.8
	Medium	100	5.5	10.0	4.9	4.8
		200	8.0	12.3	3.9	5.1
		500	10.6	32.2	3.9	5.7
		1000	17.2	63.6	5.1	4.6
	Large	100	8.8	14.8	6.6	5.8
		200	9.3	26.3	4.7	4.3
		500	20.6	69.5	4.4	5.1
		1000	40.3	95.7	3.9	3.9
15	Small	100	6.7	7.6	6.1	5.9
		200	5.9	9.4	5.2	5.3
		500	7.9	16.8	5.2	3.5
		1000	9.3	30.0	3.8	6.1
	Medium	100	6.2	12.0	5.4	5.6
		200	10.0	25.5	4.2	4.4
		500	16.8	61.6	5.4	4.1
		1000	33.1	92.0	4.0	5.4
	Large	100	9.1	27.5	5.3	5.3
		200	14.5	55.0	5.5	4.7
		500	40.0	95.4	4.8	4.6
		1000	73.4	100.0	4.1	4.5

3.2 Autoregressive quasi-simplex model

3.2.1 Change in identification constraints

The set of constraints initially proposed to identify the AR quasi-simplex model consisted of fixing the intercepts of the composites to zero, constraining to equal values the Autoregressive quasi-simplex path coefficients across waves $\rho_{21} = \rho_{32} = \rho_{43} = \rho_{54}$, and constraining to equal values the unique variances of the Y composites of the first and second measurement occasions $\theta_{11} = \theta_{22}$ and at the fourth and fifth measurement occasions $\theta_{44} = \theta_{55}$. This set of identification constraints resulted in large non-convergence rates in most conditions. In the invariant conditions the percentage of replications with convergence problems was between 0.6 and 32, while in the conditions with violations of invariance the percentage was between 0.3 and 42.

In order to reduce the non-convergence rates, the identification constraints were changed. As indicated by Jöreskog (1979b) and Biesanz (2012) the first and last latent variables in the AR quasi-simplex are not identified and constraints in the residual variances are needed by constraining their values to zero, or equal to the values of the adjacent waves. The new set of constraints consisted of constraining the residual variances of the Y composites at first and second measurement occasions to equal values $\theta_{11} = \theta_{22}$ (as in the initial set of constraints proposed) and to constrain the residual variance of the fifth measurement occasion to zero. It should be noted that although this constraint identifies the model, it implies that the assessment in the last wave contains no error which is an unrealistic assumption in practice. The intercepts of the composites were fixed to zero to identify the mean structure of the latent variables. The path coefficient estimates were allowed to vary across waves. This new set of constraints yielded lower non-convergence rates as described in the next section.

3.2.2 Non-convergence percentages

The convergence percentages for the invariant conditions under the AR quasi-simplex model are shown in Table 3.11. It can be observed that the invariant conditions with sample sizes of 100 and 200 presented between 3 and 8% of replications with non-convergence. With sample sizes of 500, only one replication had convergence problems when the number of items was 6, and no convergence problems were found for conditions with sample sizes of 1000.

Table 3.11
Non-convergence percentages in the invariant conditions of the AR quasi-simplex model

Num. Items	Sample size	% of replications
6	100	7.9
	200	0.9
9	100	3.1
	200	0.2
15	100	7.0
	200	0.8

The non-convergence percentages for the conditions with violations of longitudinal invariance are shown in Table 3.12. With sample sizes of 100 the non-convergence percentages ranged from 3.9 to 8.7, while for conditions with sample sizes of 200 the largest non-convergence rate was 1.5.

Table 3.12 shows that the sample size is the only variable that had a clear impact on the non-convergence percentages. It was concluded that the convergence problems were related to the small sample sizes and the particular set of constraints that were chosen for the AR quasi-simplex model.

Table 3.12
Non-convergence percentages in the AR quasi-simplex conditions with violations of invariance

Num. Items	Proportion non-inv.	Magnitude of violations	Non-invariant loadings		Non-invariant intercepts	
			N=100	N=200	N=100	N-200
6	1/3	Small	6.5	1.2	6.7	0.8
		Medium	5.7	1.1	8.5	1.4
		Large	6.2	1.5	5.6	1.1
	2/3	Small	7.1	1.1	6.1	1.0
		Medium	7.9	1.0	6.7	1.3
		Large	7.5	0.9	6.5	1.1
9	1/3	Small	4.3	0.1	4.1	0.1
		Medium	3.4	0	4.0	0.4
		Large	4.8	0.1	4.7	0.3
	2/3	Small	4.0	0.1	4.8	0.2
		Medium	4.5	0.3	6.3	0.3
		Large	4.4	0.4	6.4	0.3
15	1/3	Small	6.3	0.3	8.0	0.8
		Medium	5.6	1.2	6.9	0.7
		Large	5.7	0.8	8.0	1.0
	2/3	Small	6.9	0.4	7.8	0.7
		Medium	3.9	0.5	8.2	1.3
		Large	4.0	0.2	8.7	0.9

3.2.3 Parameter estimation

In this section, the bias, relative bias, standard errors and RMSE of the four autoregressive coefficients are described.

Bias and relative bias

As in the result section under LGM, only the tables with the relative bias values are presented in this section, but a graphical representation of the bias and relative bias are included. (Appendix E presents the tables with the mean bias across conditions). To further simplify the presentation of the findings, the results are averaged over sample size, since the ANOVA results did not show an effect of sample size on the relative bias.

The relative bias values of the parameter estimates for the invariant conditions and the conditions with violations to invariance are shown in Tables 3.13 and 3.14, respectively. As expected, the relative bias values were close to zero in all the invariant conditions (Table 3.13). In the non-invariant intercept conditions the relative bias values were also close to zero (Table 3.14).

Table 3.13
Relative bias of the AR quasi-simplex parameter estimates in the invariant conditions

Num. items	$\rho 21$	$\rho 32$	$\rho 43$	$\rho 54$
6	0.001	0.000	0.000	0.001
9	0.002	0.002	0.001	0.001
15	0.002	0.002	0.001	0.002

In contrast to the non-invariant intercept conditions, in the conditions with non-invariant loadings the AR quasi-simplex parameter estimates had relative bias values above 0.05 in some conditions. As shown in Table 3.14, the conditions with 2/3 of non-invariant items had relative bias absolute values between .047 and .087. It can be observed that as the magnitude of violations and the proportion of non-invariant items increased, the relative bias absolute values also increased.

Table 3.14.
Relative bias of the AR quasi-simplex parameter estimates in the conditions with violations of invariance

Num. Items	Effect size	Prop. non-inv.	Non-invariant loadings				Non-invariant intercepts			
			ρ21	ρ32	ρ43	ρ54	ρ21	ρ32	ρ43	ρ54
6	Small	1/3	-0.013	-0.013	-0.013	-0.013	-0.001	-0.001	0.002	0.002
		2/3	-0.022	-0.022	-0.024	-0.023	0.000	-0.001	0.002	0.003
	Medium	1/3	-0.029	-0.028	-0.028	-0.028	0.000	0.000	0.000	-0.002
		2/3	-0.047	-0.047	-0.049	**-0.051**	0.000	-0.002	0.003	0.000
	Large	1/3	-0.041	-0.043	-0.042	-0.045	-0.001	-0.001	0.001	0.003
		2/3	**-0.069**	**-0.073**	**-0.079**	**-0.086**	-0.002	0.001	0.000	0.000
9	Small	1/3	-0.011	-0.011	-0.012	-0.013	0.001	0.002	-0.001	0.002
		2/3	-0.023	-0.023	-0.022	-0.024	0.002	0.001	0.001	0.001
	Medium	1/3	-0.023	-0.026	-0.026	-0.027	-0.001	-0.001	0.001	0.001
		2/3	-0.045	-0.048	-0.049	**-0.053**	0.002	0.001	0.001	0.000
	Large	1/3	-0.037	-0.038	-0.041	-0.042	0.003	0.000	0.001	0.001
		2/3	**-0.069**	**-0.074**	**-0.079**	**-0.088**	0.004	0.001	0.000	0.000
15	Small	1/3	-0.011	-0.012	-0.011	-0.010	0.001	0.001	0.001	0.000
		2/3	-0.023	-0.024	-0.022	-0.024	0.003	0.001	0.001	0.001
	Medium	1/3	-0.024	-0.025	-0.025	-0.024	0.001	0.001	0.001	0.001
		2/3	-0.047	-0.049	**-0.052**	**-0.053**	0.001	0.000	0.000	0.001
	Large	1/3	-0.037	-0.037	-0.041	-0.038	0.002	0.002	0.001	0.000
		2/3	**-0.071**	**-0.076**	**-0.083**	**-0.087**	0.002	0.003	0.001	0.001

Note: The bolded numbers correspond to relative bias absolute values at the cutoff of 0.05 or larger.

The general pattern described in Table 3.14 can also be observed in Figures 3.14 to 3.21 that show the bias and relative bias results averaged over sample size. In the non-invariant intercept conditions the bias and relative bias of the parameter estimates are similar to the results obtained in the invariant conditions. In contrast, with non-invariant loadings the parameter estimates were underestimated in some conditions. It can be seen that as the magnitude of the violations increased, the bias and relative bias absolute values also increased.

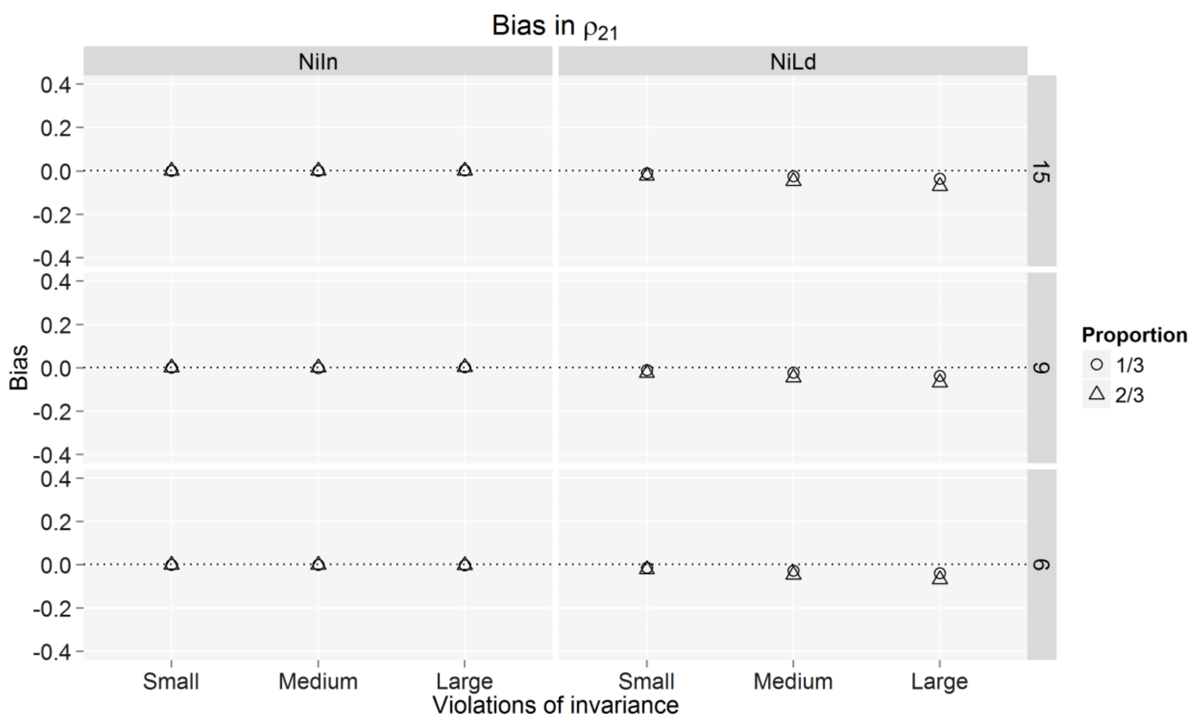

Figure 3.14 Bias in ρ_{21} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the path coefficient ρ_{21} in the conditions with invariant loadings and invariant intercepts (InLI).

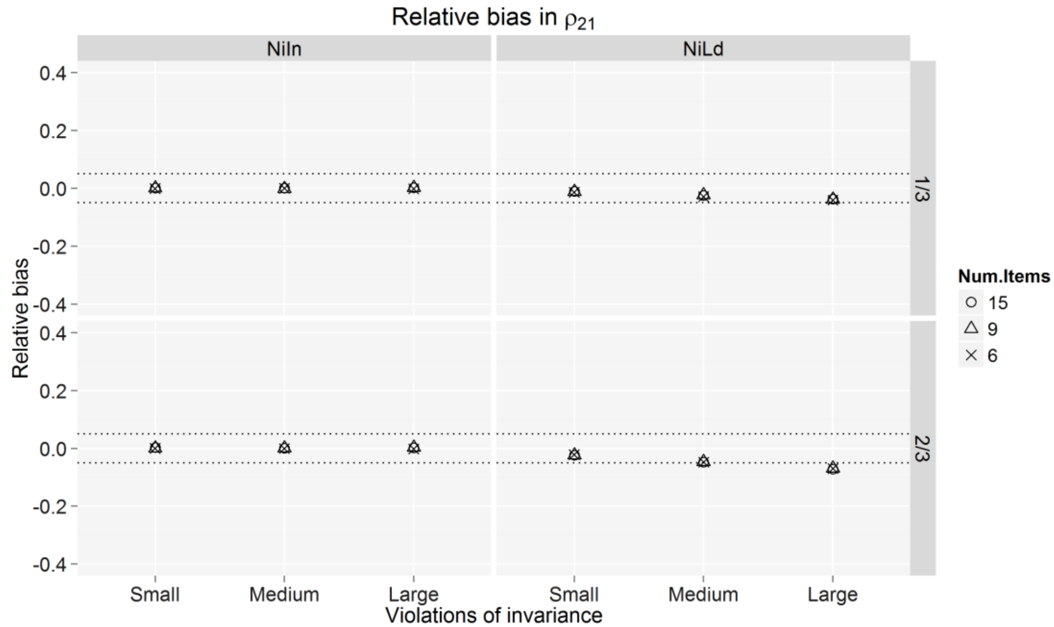

Figure 3.15 Relative bias in ρ_{21} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

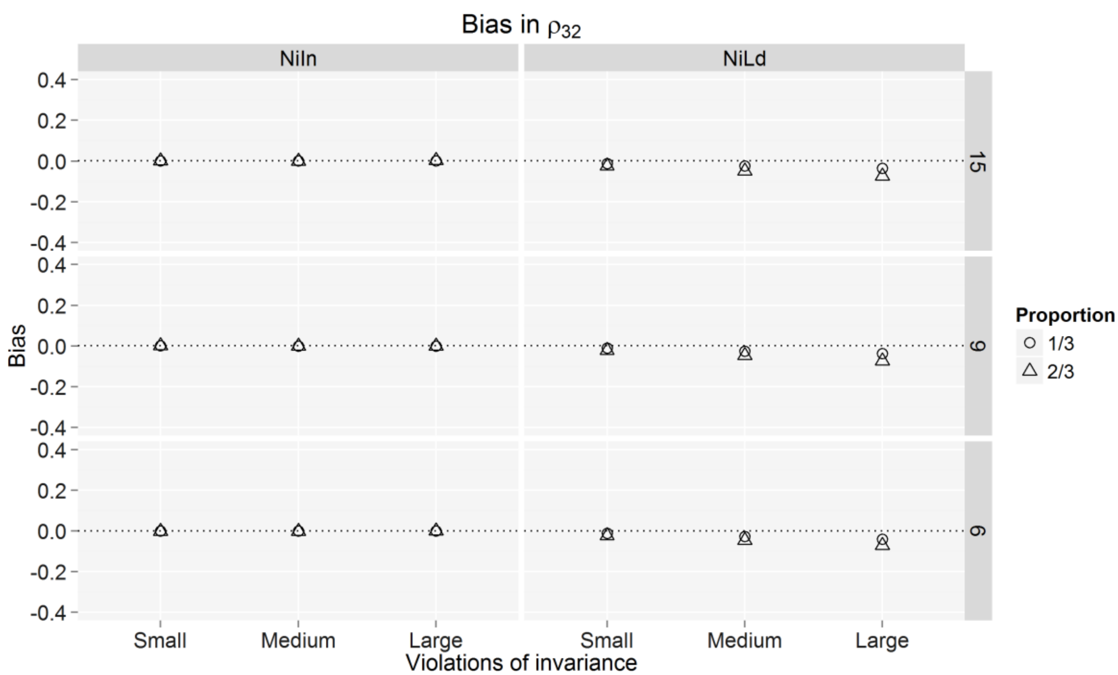

Figure 3.16 Bias in ρ_{32} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the path coefficient ρ_{32} in the conditions with invariant loadings and invariant intercepts (InLI).

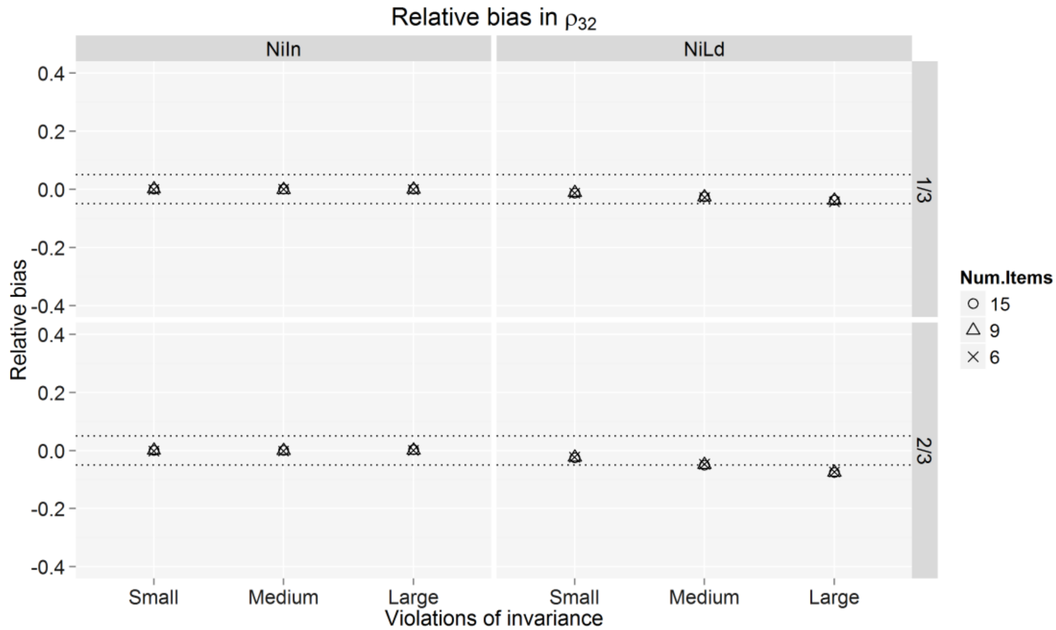

Figure 3.17 Relative bias in ρ_{32} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn).The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

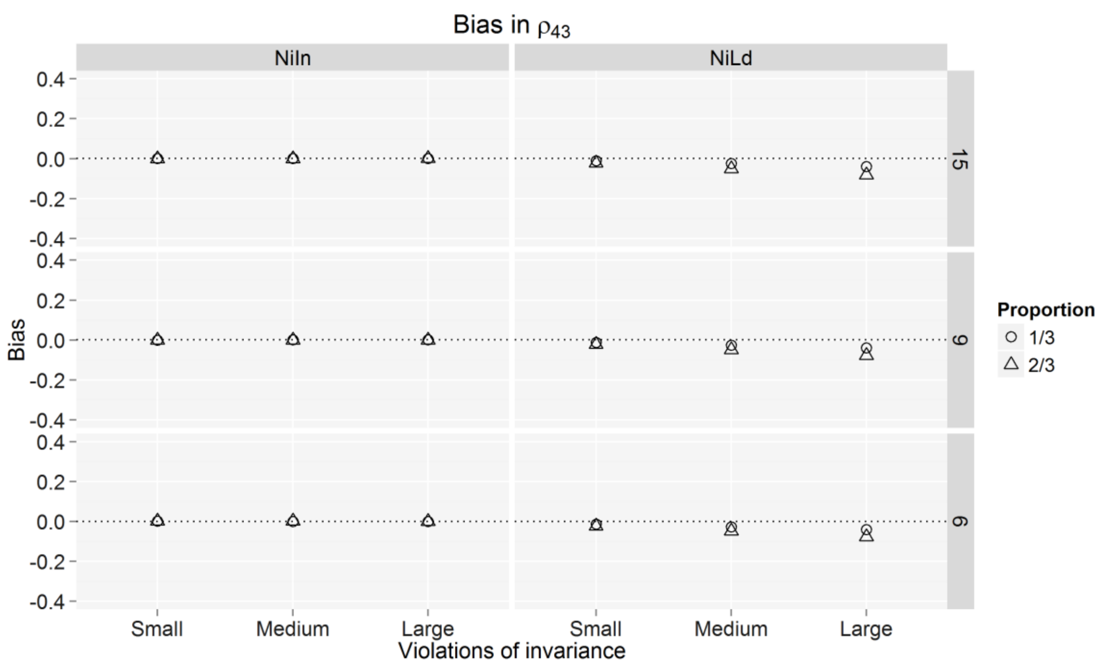

Figure 3.18 Bias in ρ_{43} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the path coefficient ρ_{43} in the conditions with invariant loadings and invariant intercepts (InLI).

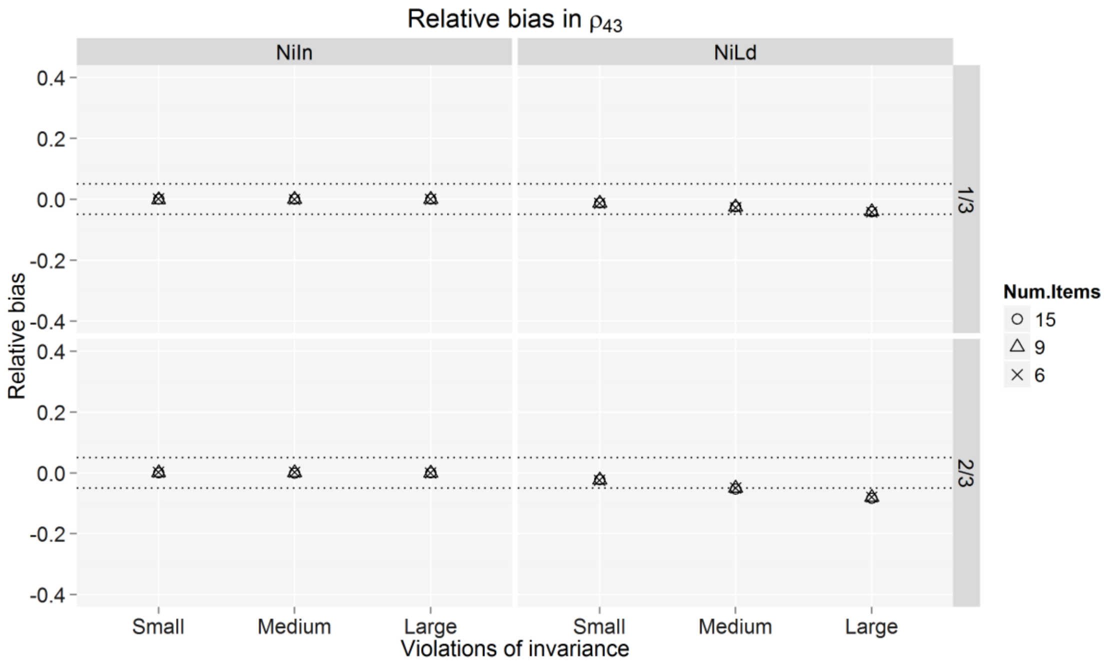

Figure 3.19 Relative bias in ρ_{43} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

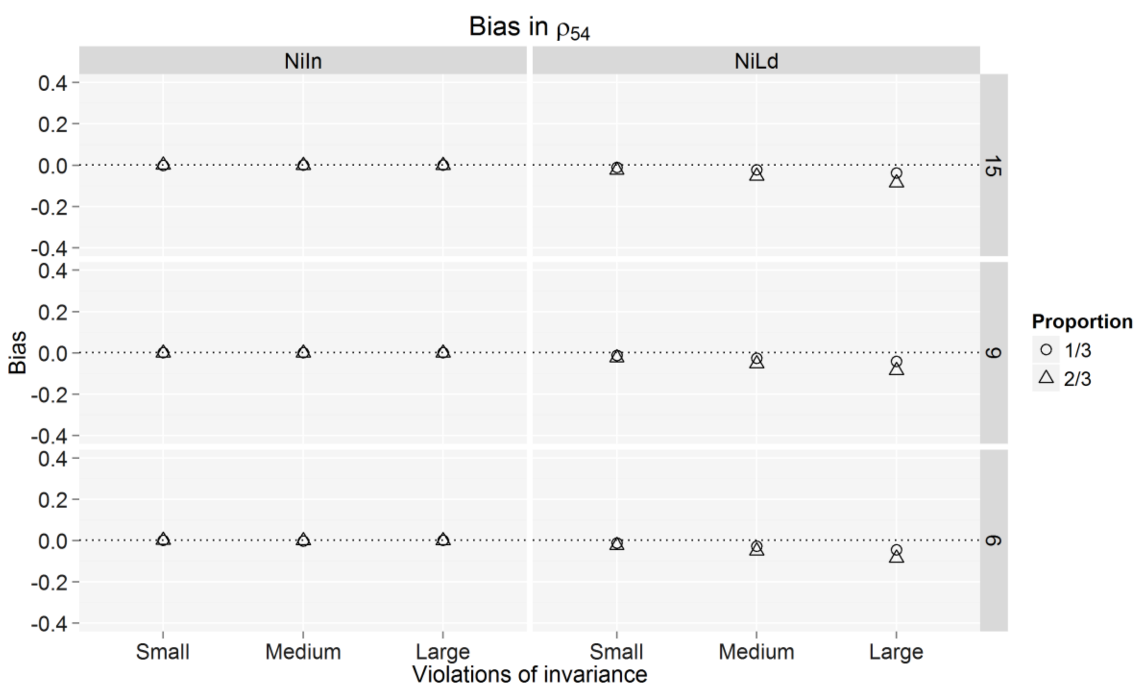

Figure 3.20 Bias in ρ_{54} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn). The horizontal lines show the bias in the path coefficient ρ_{54} in the conditions with invariant loadings and invariant intercepts (InLI).

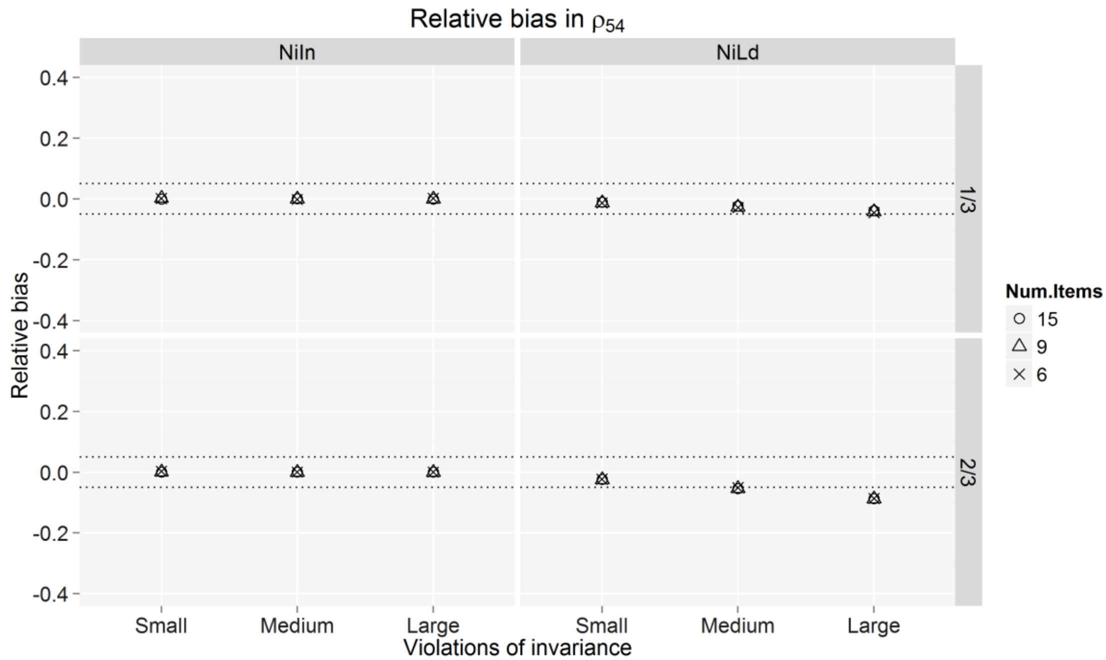

Figure 3.21 Relative bias in ρ_{54} in the non-invariant loading conditions (NiLd) and in the non-invariant intercept conditions (NiIn).The horizontal lines show the cutoff values of 0.05 and -0.05. Relative bias values inside the lines were considered acceptable values.

The ANOVAs conducted on the bias and relative bias of each of the AR quasi-simplex parameter estimates confirmed the patterns observed in Tables 3.13 and 3.14 and Figures 3.14 to 3.21. In the invariant and in the non-invariant intercept conditions no overall η^2 values larger than 0.01 were found, indicating that the independent variables did not have an effect on the bias and relative bias of the parameter estimates.

In the non-invariant loading conditions overall medium and large effect sizes were found. Table 3.15 presents the η^2 values for at least small effect sizes in the relative bias values. The η^2 values obtained for the bias were the same as for relative bias. The results confirmed that the proportion of non-invariant items and the magnitude of the violations of invariance had medium and large effects on the bias and relative bias of the AR quasi-simplex parameter estimates.

Table 3.15
η^2 values from the ANOVAs on the relative bias of the AR quasi-simplex conditions with non-invariant loadings

	Relative bias			
	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
Overall	.07	.09	.13	.17
Prop non-inv.	.02	.03	.04	.05
Magnitude	.04	.06	.08	.10

Standard errors

The ANOVAs conducted on the standard errors of the AR quasi-simplex parameter estimates revealed that in the invariant conditions and the conditions with violations of factorial invariance, the sample size explained 90% of the variance and the number of items explained between 7 to 10%. Neither the magnitude of the violations nor the proportion of non-invariant items had an effect on the standard errors of the AR quasi-simplex parameter estimates. Table 3.16 shows the η^2 values of the conditions with at least a small effect size. Table 3.17 shows the standard errors of the AR quasi-simplex parameter estimates. Since the magnitude of violations and the proportion of non-invariant items did not have an effect on the standard errors, the results are averaged over sample size and the number of items. Figures 3.22 to 3.25 show the standard errors of the AR quasi-simplex parameter estimates. The figures show that there are no differences between the invariant conditions and the conditions with violations to invariance. It can also be seen that the standard errors decreased as the sample size and as the number of items increased.

Table 3.16

η^2 values from the ANOVAs on the standard errors of the AR quasi-simplex parameter estimates

	Invariant loadings and intercepts				Non-invariant loadings				Non-invariant intercepts			
	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
Overall	0.99	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
N. Items	0.09	0.08	0.08	0.08	0.10	0.08	0.08	0.07	0.09	0.09	0.07	0.07
Sample size	0.90	0.90	0.91	0.91	0.88	0.90	0.91	0.91	0.89	0.90	0.91	0.91
N. Items x SS	--	--	--	--	0.01	0.02	0.01	0.01	0.01	0.02	0.01	0.01

Table 3.17

Standard errors of the AR quasi-simplex parameter estimates by the number of items and by sample size

Num. items	Sample size	Invariant loadings and intercepts				Non-invariant loadings				Non-invariant intercepts			
		ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
6	100	0.12	0.11	0.09	0.08	0.12	0.11	0.10	0.09	0.12	0.10	0.09	0.08
	200	0.09	0.07	0.07	0.06	0.09	0.07	0.07	0.06	0.09	0.07	0.07	0.06
	500	0.06	0.05	0.04	0.04	0.06	0.05	0.04	0.04	0.06	0.05	0.04	0.04
	1000	0.04	0.03	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.03	0.03
9	100	0.11	0.09	0.08	0.07	0.10	0.09	0.08	0.08	0.11	0.09	0.08	0.07
	200	0.08	0.07	0.06	0.05	0.08	0.07	0.06	0.05	0.08	0.07	0.06	0.05
	500	0.05	0.04	0.04	0.03	0.05	0.04	0.04	0.03	0.05	0.04	0.04	0.03
	1000	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.02	0.04	0.03	0.03	0.02
15	100	0.09	0.07	0.07	0.06	0.09	0.08	0.07	0.07	0.09	0.08	0.07	0.06
	200	0.06	0.06	0.05	0.05	0.06	0.05	0.05	0.05	0.06	0.05	0.05	0.04
	500	0.04	0.03	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.03	0.03
	1000	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.03	0.02	0.02	0.02

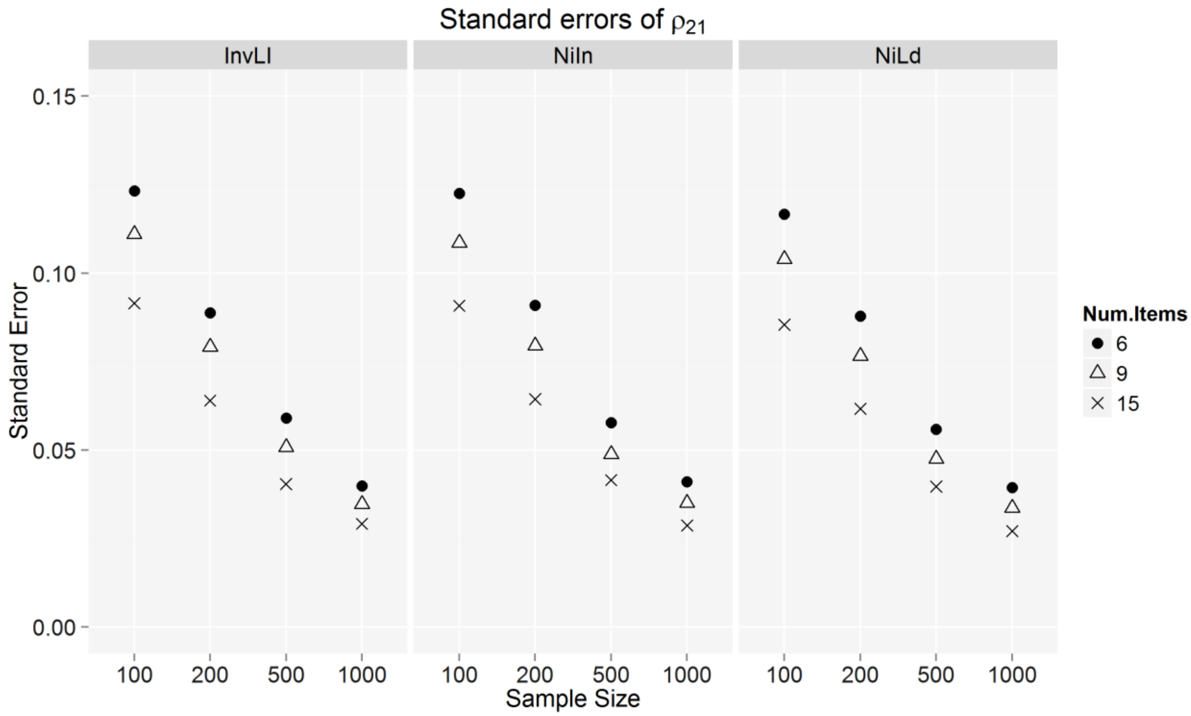

Figure 3.22 Standard errors of ρ_{21} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

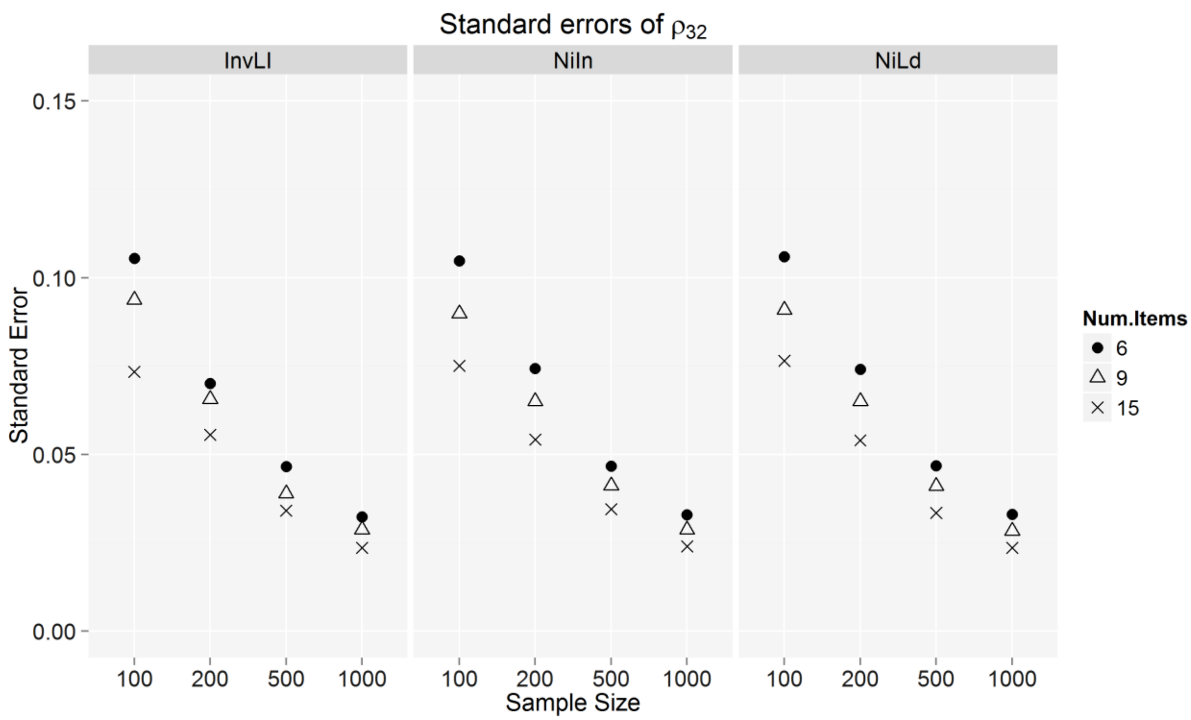

Figure 3.23 Standard errors of ρ_{32} errors of ρ_{21} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

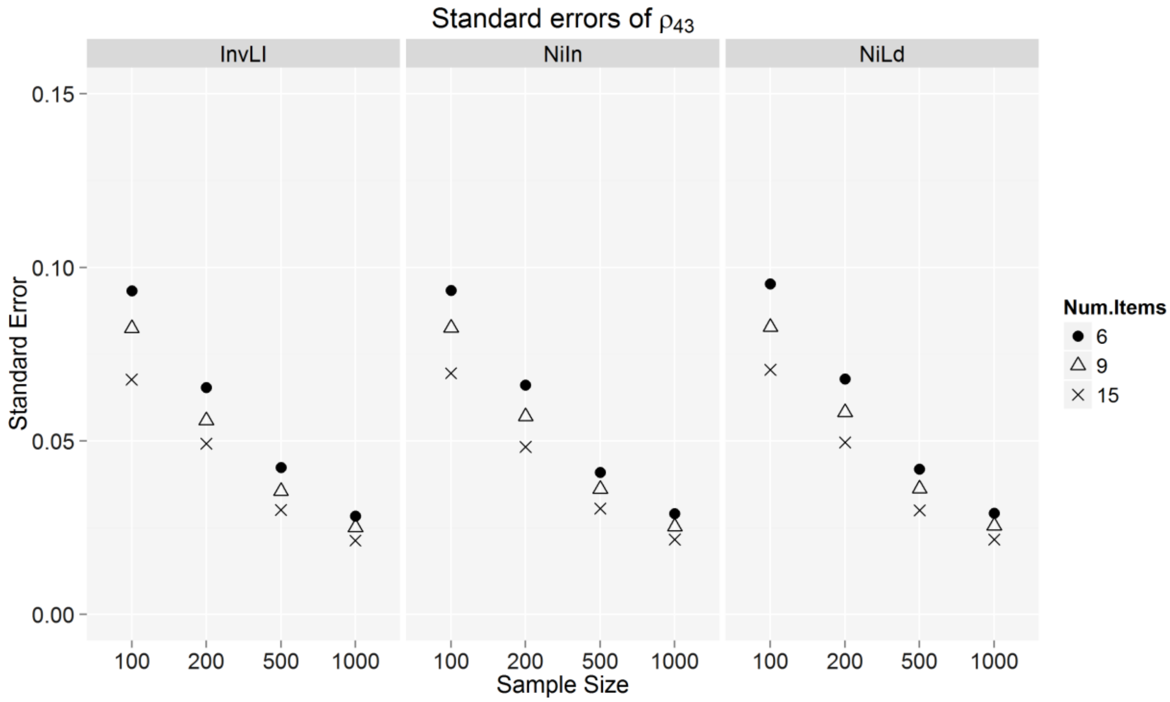

Figure 3.24 Standard errors of ρ_{43} errors of ρ_{21} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

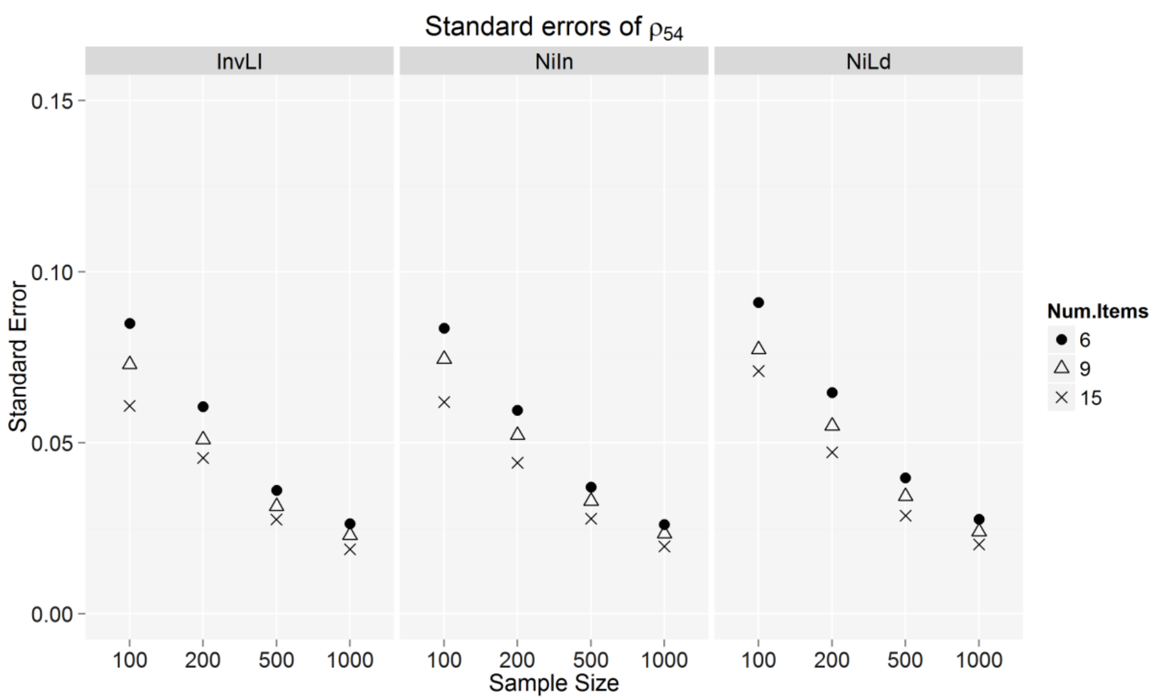

Figure 3.25 Standard errors of ρ_{54} errors of ρ_{21} in the invariant conditions (InLI), non-invariant loading conditions (NiLd) and non-invariant intercept conditions (NiIn).

Root mean square error (RMSE)

Table 3.18 presents the η^2 values obtained from the ANOVAs conducted on the RMSE of the AR quasi-simplex parameter estimates. Table 3.18 shows that in the invariant conditions and in the non-invariant intercept conditions the sample size explained 90% of the variance and the number of items explained 8%. In the non-invariant intercept conditions neither the magnitude of the violations nor the proportion of non-invariant items showed effect sizes larger than 0.01.

In contrast, in the non-invariant loading conditions all independent variables had at least a small effect size on the RMSE of the AR quasi-simplex parameter estimates. Sample size was still the independent variable that explained the larger amount of variance, between 42 and 76%, but the number of items, the magnitude of violations to invariance, and the proportion of non-invariant items had a larger effect than in the invariant and in the non-invariant intercept conditions. The number of items explained between 4 and 8% of the variance, the magnitude of violations explained between 6 and 30%, and the proportion of non-invariant items explained between 4 and 16%. The interactions between the number of items and the sample size, and between the magnitude of violations and the proportion of non-invariant items showed small and medium effect sizes.

Table 3.18

η^2 values from the ANOVAs on the RMSE of the AR quasi-simplex parameter estimates

	Invariant loadings and intercepts				Non-invariant loadings				Non-invariant intercepts			
	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
Overall	0.99	0.98	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
N. Items	0.09	0.08	0.08	0.08	0.08	0.06	0.04	0.04	0.09	0.09	0.07	0.07
Sample size	0.90	0.90	0.91	0.91	0.76	0.66	0.54	0.42	0.89	0.90	0.91	0.91
Magnitude	--	--	--	--	0.06	0.14	0.21	0.30	--	--	--	--
Proportion	--	--	--	--	0.04	0.07	0.12	0.16	--	--	--	--
N. Items x SS	--	--	--	--	0.02	0.02	0.01	0.01	0.01	0.02	0.01	0.01
Mag. X Prop.	--	--	--	--	0.02	0.03	0.05	0.06	--	--	--	--

3.2.4 Model fit

Tables 3.19 and 3.20 show the rejection rates for the invariant conditions and the conditions with violations of invariance, respectively. It can be observed that the rejection rates in all the conditions are close to 5%.

Table 3.19
Rejection rates in the invariant conditions in the AR quasi-simplex model

Number of items	Sample size	Rejection rates
6	100	5
	200	5.7
	500	4.3
	1000	4.1
9	100	5.7
	200	5.6
	500	4.7
	1000	5.6
15	100	5.2
	200	6.3
	500	5.4
	1000	4.6

Table 3.20 shows the rejection rates for the conditions with violations of invariance. It can be seen that although there are some conditions with rejection rates smaller and others larger than the nominal level, the results under violations to invariance (Table 3.20) are comparable to the rates in the invariant conditions (Table 3.19). No pattern was detected in terms of the number of items, the sample size, the proportion of non-invariant items, or the magnitude of the violations to invariance. These findings indicate that the manipulation of the violations of invariance did not affect the rejection rates in the AR quasi-simplex model.

Table 3.20
Rejection rates in the conditions with violations of invariance in the AR quasi-simplex model

Num. Items	Magnitude of violations	Sample size	Non-invariant loadings		Non-invariant intercepts	
			1/3 non-inv.	2/3 non-inv.	1/3 non-inv.	2/3 non-inv.
6	Small	100	6.4	3.9	4.6	5.9
		200	4.8	5.1	5.0	5.1
		500	5.5	5.6	5.2	4.6
		1000	5.7	3.6	5.8	5.5
	Medium	100	3.7	5.3	5.7	4.3
		200	4.1	5.1	5.9	3.8
		500	5.4	6.0	5.3	5.4
		1000	5.3	4.6	4.3	5.5
	Large	100	5.5	4.7	5.4	6.1
		200	6.3	5.9	4.6	5.0
		500	5.5	6.0	4.6	3.9
		1000	7.0	5.7	6.2	4.2
9	Small	100	4.5	6.3	6.3	5.5
		200	5.5	6.1	4.3	5.6
		500	5.7	5.4	5.0	5.4
		1000	4.1	6.6	3.3	4.6
	Medium	100	5.2	6.1	5.4	5.6
		200	5.2	5.9	4.2	5.8
		500	4.9	5.3	3.6	5.5
		1000	4.5	4.2	5.6	5.6
	Large	100	4.4	5.9	6.6	6.5
		200	4.9	5.9	5.5	4.7
		500	4.8	6.3	4.6	6.1
		1000	5.1	5.2	4.5	5.6
15	Small	100	6.7	6.5	6.2	4.8
		200	4.8	4.5	5.8	6.1
		500	4.4	6.5	5.9	4.2
		1000	5.1	5.8	4.5	4.2
	Medium	100	6.8	4.7	5.8	6.0
		200	6.0	5.6	4.7	5.2
		500	4.6	6.0	5.7	4.7
		1000	5.4	5.2	4.2	5.2
	Large	100	6.3	6.4	5.3	6.7
		200	5.6	4.7	4.6	6.3
		500	5.5	5.0	5.6	4.2
		1000	5.9	4.9	5.6	5.1

Chapter 4

DISCUSSION

The impact of analyzing composites formed by items that violate longitudinal measurement invariance has been explored in latent growth models. Leite (2007) and Wirth (2008) showed that wrong conclusions about growth can be formulated: biased growth parameter estimates as well as biased χ^2 fit indices can be obtained under non-invariance. However, no published research was found that examined the impact of violations of invariance in other longitudinal models.

Several questions guided the present research: *How many items should be invariant so that the conclusions about growth would not change? How different do the measurement parameters across time need to be to distort the growth estimates? Is the latent growth model affected in the same way by violations of invariance as are other models for longitudinal data?* By answering these questions, the present study aimed to inform researchers about when they could have confidence in growth conclusions even in the presence of composites formed from items violating invariance. It was also of interest to study the violations of invariance in another method used to analyze longitudinal data, the AR quasi-simplex model. To answer these questions, a simulation study was conducted where the number of items violating invariance was manipulated as well as the magnitude of the violations over time. Composites were formed from the simulated items and were analyzed either by the LGM or the AR quasi-simplex model. Bias in the parameter estimates and the fit of the model were examined.

A different pattern of results was found for the LGM and for the AR quasi-simplex model. While the LGM parameter estimates were biased and the model fit was severely affected by violations of invariance, the AR quasi-simplex parameters were usually unbiased and the model fit under violations of invariance were comparable to the invariant conditions. In this chapter the results are discussed, along with the limitations of the study and the conclusions.

4.1 Non-convergence rates

The percentage of samples with non-convergence was higher in the AR quasi-simplex model than in the LGM. The invariant and the non-invariant conditions had similar rates of non-convergence which indicates that the convergence problems were related to the models themselves and not to the invariance issues.

Similar results were found by the simulation study by Stockdale (2007). An AR quasi-simplex model and a LGM were fit to data simulated under the linear LGM and to data generated under an AR quasi-simplex model. He found higher convergence problems and inadmissible solutions when the data were analyzed with AR quasi-simplex than when the LGM was fit to the data, regardless of the model used to generate the data. Further, Stockdale found higher convergence problems and inadmissible solutions when the AR quasi-simplex was fit to data generated under the same model than under the LGM. He found that low path coefficients (ρ=.30), small sample sizes (N=100) and large residual variance ($\theta = 1.11$) were associated with non-convergence and inadmissible solutions.

Stockdale originally identified the AR quasi-simplex model by constraining all residual variances to equality across time. In order to explain the high non-convergence rates he compared models with different constraints but the non-convergence rates did not change drastically.

The results of the present study are congruent with the results of Stockdale (2007). More studies should be conducted to explain the non-convergence rates in the AR quasi-simplex model. Since this was not purpose of the current study this issue was not further explored.

4.2 Bias in the parameter estimates

In general, it was found that as the magnitude of the violations increased and as the proportion of non-invariant items increased, the bias in the parameter estimates also increased, bringing support to Hypotheses 1 to 3. However, the effects of the independent variables were different for each parameter estimate and for each model. In the LGM, the magnitude of violations and proportion of non-invariant item intercepts only affected the slope factor mean, while the effect of the non-invariant loadings was mostly in the slope factor mean, variance, and the covariance between the intercept and the slope. In contrast, in the AR quasi-simplex model, the non-invariant loadings equally affected all the path coefficients while non-invariant intercepts had no effect. In this section, an explanation for the different pattern of results is provided.

In the LGM the slope factor was most affected by the violations of invariance in contrast to the intercept factor. To understand why the intercept factor mean was not affected, it should be noticed that the loadings in the LGM were chosen such that the intercept factor was defined by the composite of the first measurement occasion. Hence, the mean of the intercept factor mean was defined by the mean of the composite at the first measurement occasion as shown in Equation (68),

$$\hat{\mu}_{\eta_1} = \mu_{Y_1} \tag{68}$$

The mean of the composite can be expressed in terms of the common factor model as in Equation (69),

$$\hat{\mu}_{\eta_1} = \tau_1^* + \lambda_1^* \mu_{\xi_1} \tag{69}$$

The first order factor mean μ_{ξ_1} at time 1 was generated with a value of 0, and as a consequence the loadings did not have an impact in the intercept factor mean. The intercept factor mean adopted the value of the sum of the intercepts at the first measurement occasion, τ_1^*. Since the violations of invariance in the intercepts were only shown from the second to the fifth waves, the item intercepts at the first measurement occasion were not affected by the lack of invariance. Hence, the intercept factor mean adopted the value that was expected.

It should be noted that the intercept factor can be defined by any wave and not only by the first measurement occasion as was the case in the present analysis. If the intercept factor mean were defined by a composite from the second to the fifth measurement occasion, greater bias in the intercept factor mean would have been observed. After wave one, the true values of μ_{ξ} were larger than zero, so that the effect of the loadings would not be cancelled. Also, the violations of invariance in the intercepts would be shown after wave one, changing the value of the sum of the item intercepts.

To explain why the slope factor mean was affected, two cases will be considered, one in which the loadings are changing over time, and the second case in which intercepts are changing over time.

In case 1, there are non-invariant loadings, but invariant intercepts as expressed in Equation (70)

$$\begin{matrix} \tau^* + \lambda_1^* \mu_{\xi_1} \\ \tau^* + \lambda_2^* \mu_{\xi_2} \\ \tau^* + \lambda_3^* \mu_{\xi_3} \\ \tau^* + \lambda_4^* \mu_{\xi_4} \\ \tau^* + \lambda_5^* \mu_{\xi_5} \end{matrix} = \begin{matrix} \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \end{matrix} + \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{70}$$

Since, as explained before, $\hat{\mu}_{\eta_1} = \tau^*$, Equation (70) can be rewritten as,

$$\begin{matrix} \lambda_1^* \mu_{\xi_1} \\ \lambda_2^* \mu_{\xi_2} \\ \lambda_3^* \mu_{\xi_3} \\ \lambda_4^* \mu_{\xi_4} \\ \lambda_5^* \mu_{\xi_5} \end{matrix} = \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{71}$$

When the loadings are non-invariant over time, each first order latent mean factor is changed by a different amount. For example, using the generating values for the conditions with 6 items, large violations of invariance and 2/3 of non-invariant items, the loading sums at each time point are 4.20, 3.91, 3.62, 3.33 and 3.04, such that,

$$\begin{matrix} 4.20\mu_{\xi_1} \\ 3.91\mu_{\xi_2} \\ 3.62\mu_{\xi_3} \\ 3.33\mu_{\xi_4} \\ 3.04\mu_{\xi_5} \end{matrix} = \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{72}$$

Equation (72) shows that the means of the first order latent factors were re-scaled at each measurement occasion by a different amount. Since the item loadings sums were decreasing, the change over time in the first order latent factors was smaller than it should be considering the true values, and hence the slope factor mean was underestimated. It should be noted that if the item loadings were generated to have increasing values over time, the opposite pattern of results would have been observed. That is, with increasing loadings over time it would be expected that the estimated slope factor mean would overestimate the true value.

In the second case the loadings are invariant over time but the intercepts are non-invariant. As a consequence, in each first order latent factor a different amount is added. For example, for the conditions with 6 items, large violations of invariance and 2/3 of non-invariant items, the loading sums at each time point are 2.1, 4.26, 6.42, 8.58 and 10.74 at each time point, such that,

$$\begin{matrix} 2.1 + \lambda^{*}\mu_{\xi_1} \\ 4.26 + \lambda^{*}\mu_{\xi_2} \\ 6.42 + \lambda^{*}\mu_{\xi_3} \\ 8.58 + \lambda^{*}\mu_{\xi_4} \\ 10.74 + \lambda^{*}\mu_{\xi_5} \end{matrix} = \begin{matrix} \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \end{matrix} + \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{73}$$

Equation (73) shows that at each time point, the mean of the composites were changed by a different amount. Since this amount was increasing over time, the composites seemed to change at a higher rate than what the true growth parameter values were generated to be. As a consequence, the slope factor mean was overestimated. It should be noted that if the item intercepts were generated to decrease over time, then the slope factor mean would have been underestimated.

Equation (73) shows that the amount by which the composites were changed due to non-invariant intercepts is much larger than the amount they changed due to non-invariant loadings (Equation 72). Hence, the larger bias in the slope factor mean is observed with non-invariant intercepts.

Regarding the variances of the growth factors and the covariance, it should be noted that while the intercepts have no impact in the covariance structure the loadings do have an impact. It was expected that non-invariant loadings would affect the growth factor variances. However, the intercept factor variance was unbiased in the presence of violations of invariance. The reason is that the intercept factor was defined by the first composite and the violations of invariance change the values of the loadings only after the second measurement occasion. If the intercept factor was defined by the composite at a different wave, larger bias would have been observed in the intercept factor variance.

It should be emphasized that the same pattern of results was found in the simulation study conducted by Wirth (2008). He found that the slope factor mean showed the largest degree of bias, and that the intercept factor mean and variance resulted in the least amount of bias. These results correspond to the conditions in which no correlations over time in the unique factors were simulated, which is the way the data were simulated in the present study. Wirth found that in the conditions in which the items were generated to have correlated unique factors the bias in the variances and covariances of the growth factors increased even in the invariant conditions.

The path coefficients of the AR quasi-simplex model showed small bias only in the conditions of non-invariant loadings. In general, the path coefficients are dependent on the correlations among the measures, and the correlations are affected by the item loadings. Since the item intercepts do not impact the correlations among the items, the AR quasi-simplex coefficients were unbiased regardless of the invariance in the item intercepts. In this study, bias in the means of the latent variables of the AR quasi-simplex was not examined.

4.3 Model fit

Hypotheses 4 to 6 concern the impact of the violations of invariance on the fit of the LGM and the AR quasi-simplex model. The results differed in the two models examined. While in the AR quasi-simplex the rejection rates can be interpreted as Tupe I error rates, in the LGM there was a different pattern of results in the conditions with non-invariant loadings and in the conditions with non-invariant intercepts. In the LGM conditions with non-invariant intercepts the rejection remained close to the nominal level. However, with non-invariant loadings, the percentage of replications in which the χ^2 rejected the null hypothesis was larger than 5% in most conditions.

Even though the rejection rates were initially conceptualized as Type I error rates, an alternative explanation is that the non-invariant loadings changed the functional form of the growth trajectory, and as a consequence a misspecified model was fit to the data. If this was the case, the high rejection rates shown in Table 3.10 could be interpreted as statistical power. Although a LGM with alternative growth trajectories was not examined in the present study, Wirth (2008) simulation study indicates that it is possible that violations of invariance changed the true structural model. Wirth compared the fit of two different LGM models under violations of invariance: a model in which the basis functions (the loadings relating the growth factors to the composites) were fixed to reflect a linear trajectory, and a model in which the basis functions were freely estimated so that no specific trajectory form was imposed. It was found that a model with freely estimated basis functions was accepted over a linear LGM with non-invariant item loadings over time, which indicated the existence of non-linear trajectories. It was argued that the freely estimated basis functions absorbed the non-invariance in the item loadings which changed the functional form of the trajectories. Further, Wirth found that non-invariant intercepts did not affect the fit of the model as long as the loadings were invariant. These results are consistent with what was found in the present study. The non-invariant intercepts did not change the functional form of the growth trajectory; however, non-invariant loadings affected the fit of the LGM. It could be the case that non-invariant loadings changed the functional form of the growth trajectory and that as the sample size increased, the power of the LGM to correctly reject the misspecified model increased. To examine the change in the functional form of the growth trajectories it would be necessary to compare the fit of a quadratic LGM in comparison to the linear LGM.

It should be noted that in the present simulation study the generating growth trajectory was fit to the data. However, in practice the true model is unknown. The simulation results suggest that if a researcher were to fit a quadratic LGM to composites formed by sums of items with non-invariant loadings, it might be mistakenly concluded that the data follows a quadratic trajectory. This hypothesis should be tested in a simulation study.

4.4 Limitations

As in any simulation study, the present research has a number of limitations that need to be addressed. The two major limitations of the study concern the extent to which the results can be generalized to real situations and the comparability of results in the LGM and in the AR quasi-simplex model.

The extent to which the results can be generalized to situations encountered in practice is related to the selection of the parameter values used to generate the data. In the AR quasi-simplex model the generating parameter values were chosen from a published paper in which real data was analyzed using the AR quasi-simplex (Morera, et al., 1998). In the case of the LGM model the growth parameter values were chosen based on previous simulation studies (Muthén & Muthén, 2002) that in turn chose the values based in results found in practice. Muthén and Muthén (2002) simulated the data such that the R^2 values of the analyzed composites over time ranged from .50 to .74. To avoid another source of variability, in the present study it was chosen to maintain the R^2 values constant over time and R^2 values of .80 were chosen. Although it could be argued that this value is higher than what is frequently found, studies have reported R^2 values between 83 and 84 (Bollen & Curran, 2006). Wirth (2008) used a constant R^2 value of .70, and obtained the same pattern of results as the ones reported in the present study. In general, when the growth factors can explain proportions of variance in these ranges, it is expected that the results can be generalized. In the cases in which the growth factors explain a smaller proportion of variance, it is expected that the lack of invariance will have a lower impact in the growth estimates.

Another limitation of the present study concerns the choice of identification constraints in the AR quasi-simplex model. In the present study, the latent variable of the fifth measurement occasion was identified by setting its unique variance to zero as suggested by Biesanz (2012) and Jöreskog (1979b). The implication of this identification constraint is that the last composite is measured without error. In practice, this may be an unrealistic assumption that might lead researchers to choose a different set of identification constraints, such as constraining the unique variances of the last two measurement occasions to equality. This more realistic constraint was initially proposed but high non-convergence rates were obtained. Although the change in identification constraints should not affect the fit of the model, previous studies have shown that in certain models the change in identification constraints altered the model fit (Millsap, 2001). For these reasons, the results of the AR quasi-simplex model under violations of invariance should be studied with a different set of constraints.

The extent to which the results of the LGM can be compared to the results of the AR quasi-simplex model should be examined. Although it is tempting to conclude that the impact of violations of invariance is larger in the LGM than in the AR quasi-simplex model, more studies should be conducted before this statement can be made. Marsh, Hau & Wen (2004) showed that conclusions made from models with different levels of misspecification can be misleading. It could be the case that the level of misspecification in the AR quasi-simplex model lead to an acceptable misspecified model, while the level of misspecification was larger in the LGM. One way to explore this would be to conduct a simulation study with the same conditions explored in the present study but increasing the sample size to a larger N, (e.g., 500,000). By doing this, it could be determined if the two models show the same levels of misspecification in the population.

The last limitation identified is that in the present study conditions with violations of invariance in the loadings or in the intercepts were simulated, but no conditions in which both parameters violated invariance were examined. In practice, it is frequently the case that if an item has a non-invariant loading its intercept will also be non-invariant. From the results of the simulation study it can be inferred that when the items have violations of invariance in loadings and intercepts, the same pattern of results observed in the conditions with non-invariant loadings would be found but it would be expected that the bias in the slope factor mean would increase.

4.5 Recommendations

Based on the results of the present study several recommendations can be offered to researchers interested in making conclusions from longitudinal data. The first recommendation is that longitudinal invariance should be routinely tested, rather than assumed. The results of the simulation study add to the existing literature showing that when there are violations of invariance, wrong conclusions can be made, especially when analyzing the data with the LGM (Ferrer, Balluerka, & Widaman, 2008; Leite, 2007; Wirth, 2008). The longitudinal confirmatory factor analysis described in Chapter 1 should be used to test for invariance by sequentially constraining item parameters as suggested by Jöreskog (1971). If it is found that the items have invariant loadings and intercepts, the items can be summarized in composites and analyzed using a LGM or an AR quasi-simplex model. However, if some of the items are found to have violations of invariance, models that incorporate both the measurement and the structural relations should be used, such as the curve of factors model and the AR quasi-simplex with items defining each latent factor. Leite (2007) showed that the curve of factors model can yield unbiased estimates of growth under violations of invariance when item parameters are allowed to be freely estimated over time.

An alternative is to test for invariance in a model that incorporates the measurement and the structural relations, such as in the curve of factors model. However, if a model such as the curve of factors model shows poor fit to the data when testing for invariance, the lack of fit can be due to violations of measurement invariance or due to a misspecified structural model. For example, a source of structural misfit can come from fitting a non-linear trajectory to the data that follows a linear trajectory. If the structural model is misspecified it can alter the measurement model which may have consequences in the conclusions about measurement invariance. In order to avoid the confounding of the sources of misfit, the approach suggested by Anderson and Gerbing (1988) and by Mulaik and Millsap (2000) should be followed. This approach consists of testing a series of nested models in which first, the fit of the measurement model is evaluated by saturating the structural relations between the latent variables. If the measurement model fits the data, the structural relations are examined. This approach would permit distinguishing among the sources of lack of fit, if any.

Researchers interested in examining invariance using models that incorporate the measurement and the structural relations should be cautious about the consequences of selecting a non-invariant item as the referent indicator to identify the model. Ferrer, Balluerka, and Widaman (2008) studied the impact of measurement non-invariance in a curve of factors model using real data from an alcohol prevention program. A series of confirmatory factor analysis were conducted and the hypothesis of metric invariance was rejected indicating that some items were non-invariant. Two different second-order latent growth curve models were fit to the data that only differed in the item chosen as the referent indicator. The results showed completely different growth trajectories obtained from the two models; using one item as a referent indicator yielded a significant linear growth trajectory, while no significant growth was detected when using a different item as the reference indicator. These results indicate that partial invariance can have a drastic impact in the conclusions made regarding growth, depending on the choice of referent indicator.

The results of the AR quasi-simplex model suggest that in general, researchers can obtain unbiased path coefficients with small and medium violations of invariance and with 1/3 of non-invariant items. This does not suggest that researchers should stop testing for invariance in the AR quasi-simplex model. The researchers need to determine the extent of the violations of invariance. If the magnitude of the violations is comparable to the conditions of the present simulation study and if there are only 1/3 of non-invariant items, the researchers could use item composites if the sample sizes do not permit the use of a full SEM model that incorporates the measurement and the structural relations.

The last recommendation is that whenever possible the use of composites should be avoided if longitudinal invariance is unexplored. As shown in the present results, forming composites when there are violations of invariance can yield biased conclusions.

There are other disadvantages related to the use of composites. The first one is that if the items have correlated unique variances over time, these are ignored when forming composites. Wirth (2008) showed that when the correlated unique variances are ignored the growth estimates can be biased even with invariant items.

4.6 General conclusion

Analytic results presented in Chapter 1 showed that the violations of longitudinal invariance can bias the parameter estimates of models such as the LGM and the AR quasi-simplex model. The present simulation study further showed that the impact of non-invariance can vary by the longitudinal model used. In general, researchers should expect that violations of metric and strong factorial invariance would bias the parameter estimates of the LGM as well as the fit of the model. Violations of metric and strong longitudinal invariance would yield unbiased AR quasi-simplex path coefficients, and adequate rejection rates.

The present study emphasizes the importance of examining longitudinal measurement invariance before forming composites of the items to obtain adequate conclusions from longitudinal studies. Special caution is advised when using the LGM since biased estimates and a different growth trajectory can be found under non-invariance.

Finally, it is advised to avoid the use of item composites if longitudinal measurement invariance has not been investigated.

REFERENCES

Anderson, J. C. & Gerbin, D. W. (1988). Structural Equation Modeling in Practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411-423.

Biesanz, J. C. (2012). Autoregressive longitudinal models. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (pp. 459-471). New York: Guilford Press.

Bollen, K. A. (1989). *Structural equations with latent variables*. Oxford, England: John Wiley & Sons, Oxford.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods and Research, 32,* 336-383.

Bollen, K. A., & Curran, P. J. (2006). Latent curve models: A structural equation perspective. New Jersey: John Wiley & Sons.

Borsboom, D. (2006). When does measurement invariance matter? *Medical care, 44,* 176-181.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage Publications.

Byrne, B.M., Shavelson, R.J., & Muthén, B. (1989). Testing for equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456-466.

Chan, D. (1998). The conceptualization and analysis of change over time: An integrative approach incorporating longitudinal mean and covariance structures (LMACS) and multiple indicator latent growth modeling (MLGM). *Organizational Research Methods, 1*, 421-483.

Cheung, G. W. & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25,* 1-27.

Cohen, J. (1988). Statistical power analysis of behavioral sciences (2nded.) New York: Academic Press.

Corballis, M. C. & Traub, R. E. (1970). Longitudinal factor analysis. *Psychometrika, 35,* 79-95.

DeShon, R. P. (2004). Measures are not invariant across groups without error variance homogeneity. *Psychology Science, 46*, 137-149.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). An introduction to latent variable growth curve modeling: Concepts, issues, and applications. New Jersey: Lawrence Erlbaum Associates.

Fan, X., & Fan, X. (2005). Power of latent growth modeling for detecting linear growth: number of measurements and comparison with other analytic approaches. *The Journal of Experimental Education, 73*, 121-139.

Ferrer, E., Balluerka, N., & Widaman, K. (2008). Factorial Invariance and the specification of second order latent growth models. *Methodology, 4*, 22-36.

Hamilton, J., Gagne, P. E., & Hancock, G. R. (2003, April). *The effect of sample size on latent growth models*. Paper presented at the Annual Meeting of the American Education Research Association, Chicago, IL.

Heise, D. R. (1969). Separating reliability and stability in test-retest correlation. *American Sociological Review, 34*, 93-101.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: an overview and a meta-analysis. *Sociological Methods & Research, 26*, 329-367.

Horn, J.L., & McArdle, J.J. (1992). A practical guide to measurement invariance in research on aging. *Experimental Aging Research, 18*, 117-144.

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

Johnson, E.C., Meade, A.W., & DuVernet, A.M. (2009). The role of referent indicators in tests of measurement invariance. *Structural Equation Modeling, 16*, 642-657.

Jöreskog, K. G. (1970a). A general method for analysis of covariance structures. *Biometrika, 57*, 239-251.

Jöreskog, K. G. (1970b). Estimation and testing of simplex models. *British Journal of Mathematical and Statistical Psychology, 23*, 121-145.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychomerika, 36*, 409-426.

Jöreskog, K. G. (1979a). Statistical models and methods for analysis of longitudinal data. In K. G. Jöreskog & D. Sörbom (Eds.) *Advances in factor analysis and structural equation models*. New York: University Press of America.

Jöreskog, K. G. (1979b). Statistical estimation of structural models in longitudinal-developmental investigations. In J. R. Nesselroade & P. B. Baltes (Eds.) *Longitudinal research in the study of behavior and development*. New York: Academic Press.

Khoo, S.-T., West, S. G., Wu, W. & Kwok, Oi-Man (2005). Longitudinal methods. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 301-317), Washington DC: American Psychological Association.

Leite, W. L. (2005). *A comparison of latent growth models for constructs measured by multiple indicators* (Doctoral dissertation). Available from ProQuest Dissertations and Thesesdatabase. (UMI No. 3175266).

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

MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*, 107-120.

MacCallum, R. C. (2009). Factor Analysis. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE Handbook of Quantitative Methods in Psychology* (pp.123-147), Thousand Oaks, CA: Sage Publications.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*, 490-504.

Marsh, H, W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis testing approaches to setting cutoff values for fit indices and dangers 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 J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561-614). New York: Plenum.

McArdle, J. J. (2007). Five steps in the structural factor analysis of longitudinal data. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 99-130), Mahwah, NJ: Lawrence Erlbaum Associates.

Meade, A. W. & Bauer, D. J. (2007). Power and precision in confirmatory factor analytic tests of metric invariance. *Structural Equation Modeling, 14*, 611-635.

Meade, A., W. & Lautenschlager, G. J. (2004). A Monte-Carlo study of confirmatory factor analytic tests of measurement equivalence/invariance. *Structural Equation Modeling, 11*, 60-72.

Mellenbergh, G. J. (1989). Item bias and Item Response Theory. *International Journal of Educational Research, 13*, 127-143.

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, *30*, 419-440.

Meredith, W. & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L.M. Collins & A.G. Sayer (Eds.). *New methods for the analysis of change* (pp. 203-240). Washington D.C.: American Psychological Association.

Meredith, W. & Teresi, J. E. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*, 69-77.

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

Millsap, R. E. (2001). When trivial constraints are not trivial: The choice of uniqueness constraints in confirmatory factor analysis. *Structural Equation Modeling, 8*, 1-17.

Millsap, R. E. (2011). *Statistical Approaches to Measurement Invariance*. New York: Routledge.

Millsap. R. E. & Cham, H. (2012). Investigating factorial invariance in longitudinal data. In B. Laursen, T. D., Little, & N. A. Card. (Ed.), *Handbook of Developmental Research Methods*. (pp. 109-126) NY: Guilford.

Millsap, R.E. & Kwok, Oi-Man (2004). Evaluating the impact of partial factorial invariance on selection in two populations. *Psychological Methods, 9*, 93-115.

Millsap, R. E. & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. In R. Cudeck & R. MacCallum (Eds.) *Factor Analysis at 100*. (pp. 131-152) Mahwah, NJ: Lawrence Erlbaum.

Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (pp. 380-392). New York: Guilford Press.

Morera, O. F., Johnson, T. P., Freels, S., Parsons, J., Crittenden, K. S., Flay, B. R., & Warnecke, R. B. (1998). The measure of stage of readiness to change: Some psychometric considerations. *Psychological Assessment, 10*, 182-186.

Mulaik, S. A. & Millsap, R. E. (2000). Doing the four step right. *Structural Equation Modeling 7*, 36-73.

Muthén, B. O. & Curran, P. J. (1997). Experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2*, 371-402.

Muthén, B. O. & Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10*, 73-101.

Muthén, L. K. & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599-620.

Muthén, L. K. & Muthén, B. O. (2010). M*plus* (Version 6.1), Los Angeles: Muthén & Muthén.

Obradovic, J., Pardini, D. A., Long, J. D., & Loeber, R. (2007). Measuring interpersonal callousness in boys from adolescence: An examination of longitudinal invariance and temporal stability. *Journal of Clinical Child and Adolescent Psychology, 36*, 276-292.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Thousand Oaks, CA: Sage Publications.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

SAS (Version 9.2) [Computer software] (2008). Cary, NC: SAS Institute.

Schmitt, N. & Kuljanin, G. (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*, 210-222.

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

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27*, 229-239.

Stark, S., Chernyshenko, O.S., & Drasgow, F. (2004). Examining the effects of differential item (functioning and differential) test functioning on selection decisions: When are statistically significant effects practically important? *Journal of Applied Psychology, 89,* 497-508.

Steiger, J. H. (1990). Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research, 25,* 173-180.

Steenkamp, J.E.M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. Journal of Consumer Research, 25, 78-90.

Stockdale, G. D. (2007). *Factors affecting goodness of fit of the quasi-simplex, linear growth curve, and latent difference score models to opposite data structures: A simulation study.* (Doctoral dissertation) Available from ProQuest Dissertations and Theses database (UMI No. 3303209)

Thurston, L. L. (1947). *Multiple Factor Analysis.* Chicago: University of Chicago Press.

Tisak, J. & Meredith, W. (1989). Exploratory longitudinal factor analysis in multiple populations. *Psychometrika, 54,* 261-281.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4-70.

Werts, C. E., Jöreskog, K. G., & Linn, R. L. (1971). Comment on the estimation of measurement error in panel data. *American Sociological Review, 36,* 110-112.

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.) *Handbook of Structural Equation Modeling* (pp. 380-392). New York: Guilford Press.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4,* 10-18.

Widaman K. F. & Reise S. P. (1997) Exploring the measurement invariance of psychological instruments: applications in the substance use domain. In K. J. Bryant, M., Windle & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse* (pp. 281-324). Washington: American Psychological Association.

Wiley, D. E., & Wiley, J. A. (1970). The estimation of measurement error in panel data. *American Sociological Review, 35,* 112-117.

Willoughby, M. T., Wirth, R. J., & Blair, C. B. (2012). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment*, *24*, 418-431.

Wirth, R. J. (2008). *The effects of measurement non-invariance on parameter estimation in latent growth models* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3331053).

Woods, C. M. (2008). Empirical selection of anchors for tests of differential item functioning. *Applied Psychological Measurement*, *33*, 42-57.

Yoon, M. & Millsap, R.E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling*, *14*, 435-463.

APPENDIX A

ITEM VARIANCES AND COMMUNALITIES

In this appendix the item variances and the communalities are shown for conditions with invariant loadings over time, and with small, mediums, and large violations of invariance over time.

Table 5.1
Item variances and communalities in conditions with invariant loadings and intercepts

Item	Time1		Time2		Time3		Time4		Time5	
	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2
1	1.00	0.30	1.12	0.38	1.36	0.48	1.71	0.59	2.20	0.68
2	1.80	0.28	2.00	0.35	2.39	0.46	2.98	0.56	3.78	0.66
3	0.76	0.21	0.82	0.26	0.94	0.36	1.12	0.46	1.37	0.56
4	1.02	0.22	1.11	0.28	1.28	0.38	1.55	0.48	1.90	0.58
5	1.70	0.23	1.85	0.30	2.16	0.40	2.62	0.50	3.26	0.60
6	0.50	0.20	0.54	0.26	0.61	0.35	0.73	0.45	0.89	0.55
7	0.86	0.18	0.92	0.23	1.04	0.32	1.22	0.43	1.47	0.52
8	2.52	0.25	2.76	0.31	3.24	0.41	3.97	0.52	4.96	0.62
9	1.22	0.18	1.31	0.24	1.48	0.33	1.75	0.43	2.10	0.52

Table 5.2
Item variances and communalities in conditions with small violations of invariance in the loadings

Item	Time1		Time2		Time3		Time4		Time5	
	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2
1	1.00	0.30	1.09	0.36	1.27	0.45	1.51	0.54	1.80	0.61
2	1.80	0.28	1.95	0.33	2.24	0.42	2.64	0.51	3.12	0.58
3	0.76	0.21	0.80	0.25	0.89	0.32	1.01	0.41	1.16	0.48
4	1.02	0.22	1.09	0.26	1.22	0.34	1.39	0.43	1.61	0.50
5	1.70	0.23	1.81	0.28	2.04	0.36	2.36	0.45	2.74	0.53
6	0.50	0.20	0.53	0.24	0.58	0.32	0.66	0.40	0.76	0.47
7	0.86	0.18	0.92	0.23	1.04	0.32	1.22	0.43	1.47	0.52
8	2.52	0.25	2.76	0.31	3.24	0.41	3.97	0.52	4.96	0.62
9	1.22	0.18	1.31	0.24	1.48	0.33	1.75	0.43	2.10	0.52

Table 5.3
Item variances and communalities in conditions with medium violations of invariance in the loadings

Item	Time1		Time2		Time3		Time4		Time5	
	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2
1	1.00	0.30	1.06	0.34	1.18	0.41	1.33	0.47	1.47	0.52
2	1.80	0.28	1.90	0.32	2.10	0.38	2.34	0.44	2.56	0.49
3	0.76	0.21	0.79	0.24	0.85	0.29	0.92	0.35	0.99	0.39
4	1.02	0.22	1.07	0.25	1.15	0.31	1.26	0.37	1.36	0.41
5	1.70	0.23	1.77	0.27	1.93	0.33	2.12	0.39	2.30	0.43
6	0.50	0.20	0.52	0.23	0.56	0.28	0.60	0.34	0.65	0.38
7	0.86	0.18	0.92	0.23	1.04	0.32	1.22	0.43	1.47	0.52
8	2.52	0.25	2.76	0.31	3.24	0.41	3.97	0.52	4.96	0.62
9	1.22	0.18	1.31	0.24	1.48	0.33	1.75	0.43	2.10	0.52

Table 5.4
Item variances and communalities in conditions with large violations of invariance in the loadings

Item	Time1		Time2		Time3		Time4		Time5	
	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2	σ^2	h^2
1	1.00	0.30	1.04	0.32	1.11	0.37	1.17	0.40	1.19	0.41
2	1.80	0.28	1.86	0.30	1.97	0.34	2.07	0.37	2.11	0.38
3	0.76	0.21	0.77	0.22	0.81	0.26	0.84	0.28	0.85	0.29
4	1.02	0.22	1.05	0.24	1.10	0.27	1.14	0.30	1.16	0.31
5	1.70	0.23	1.74	0.25	1.83	0.29	1.91	0.32	1.94	0.33
6	0.50	0.20	0.51	0.22	0.53	0.25	0.55	0.28	0.56	0.29
7	0.86	0.18	0.92	0.23	1.04	0.32	1.22	0.43	1.47	0.52
8	2.52	0.25	2.76	0.31	3.24	0.41	3.97	0.52	4.96	0.62
9	1.22	0.18	1.31	0.24	1.48	0.33	1.75	0.43	2.10	0.52

APPENDIX B

ITEM MEANS

In this appendix the item means are shown for conditions with invariant intercepts over time, and with small, mediums, and large violations of invariance over time.

Table 5.5
Item means in conditions with invariant loadings and intercepts

Item	Time1	Time2	Time3	Time4	Time5
1	0.5	0.68	0.85	1.03	1.20
2	0.6	0.83	1.05	1.28	1.50
3	0.3	0.43	0.55	0.68	0.80
4	0.4	0.55	0.70	0.85	1.00
5	0.6	0.80	1.00	1.20	1.40
6	0.4	0.50	0.60	0.70	0.80
7	0.3	0.40	0.50	0.60	0.70
8	0	0.20	0.40	0.60	0.80
9	0.3	0.42	0.54	0.66	0.78

Table 5.6
Item means in conditions with small violations of invariance in the intercepts

Item	Time1	Time2	Time3	Time4	Time5
1	0.5	0.68	0.85	1.03	1.20
2	0.6	0.83	1.05	1.28	1.50
3	0.3	0.43	0.55	0.68	0.80
4	0.4	0.55	0.70	0.85	1.00
5	0.6	0.80	1.00	1.20	1.40
6	0.4	0.50	0.60	0.70	0.80
7	0.3	0.40	0.50	0.60	0.70
8	0	0.20	0.40	0.60	0.80
9	0.3	0.42	0.54	0.66	0.78

Table 5.7
Item means in conditions with medium violations of invariance in the intercepts

Item	Time1	Time2	Time3	Time4	Time5
1	0.5	0.78	1.06	1.34	1.62
2	0.6	0.96	1.32	1.68	2.04
3	0.3	0.5	0.7	0.9	1.1
4	0.4	0.64	0.88	1.12	1.36
5	0.6	0.92	1.24	1.56	1.88
6	0.4	0.56	0.72	0.88	1.04
7	0.3	0.4	0.5	0.6	0.7
8	0	0.2	0.4	0.6	0.8
9	0.3	0.42	0.54	0.66	0.78

Table 5.8
Item means in conditions with large violations of invariance in the intercepts

Item	Time1	Time2	Time3	Time4	Time5
1	0.5	1.2	1.9	2.6	3.3
2	0.6	1.5	2.4	3.3	4.2
3	0.3	0.8	1.3	1.8	2.3
4	0.4	1	1.6	2.2	2.8
5	0.6	1.4	2.2	3	3.8
6	0.4	0.8	1.2	1.6	2
7	0.3	0.4	0.5	0.6	0.7
8	0	0.2	0.4	0.6	0.8
9	0.3	0.42	0.54	0.66	0.78

APPENDIX C

RE-SCALING OF THE GROWTH PARAMETER ESTIMATES

In order to study violations of invariance in the LGM, item level data were generated from a curve of factors model (COFM). Item composites were formed at each time point by summing the items, and a LGM was used to analyze the composites as planned. However unexpected results were obtained. The bias and relative bias in the control conditions were computed using the generating parameter values shown in Equation 60. Since in the control conditions all the items were invariant over time, bias and relative bias values near zero were expected, but this was not the case. For example, for the control condition with 6 items and a sample size of 1000, relative bias values as large as 16 were found for the intercept factor variance, the slope factor variance and the covariance between the intercept and the slope factors. It was found that a re-scaling of the true growth parameter values was needed.

In this appendix the re-scaling of the growth parameter values that occurred by modeling composites of items instead of the first order latent factors is explained.

Intercept factor mean

In the LGM and the COFM the intercept factor mean (μ_{η_1}) represents the mean of the composite (μ_{Y_1}) or the mean of the first order latent factor (μ_{ξ_1}) in which the slope factor loading is set to zero. The data were generated under a COFM in which the first order factor at wave 1 had a zero slope factor loading (See Figure 2.1). In other words, the intercept factor mean was defined as the mean of first order factor at wave 1:

$$\mu_{\eta_1} = \mu_{\xi_1} \tag{74}$$

The data were generated with $\mu_{\xi_1} = 0$ and $\mu_{\eta_1} = 0$. However, the data were not analyzed using a COFM. Instead, a LGM was fitted to composites. As explained in Chapter 1, when the items used to generate the composites can be modeled by the common factor model, the mean of the composite can be expressed as the mean of the latent factor weighted by the sum of the item loadings plus the sum of the intercepts. The composite of the first measurement occasion can be expressed as,

$$\mu_{Y_1} = \tau_1^* + \lambda_1^* \mu_{\xi_1} \tag{75}$$

where τ_1^* and λ_1^* are the sum of the item intercepts and the item loadings at time 1, respectively.

Since a LGM was fitted to the composites of the items, the estimated intercept factor mean $\hat{\mu}_{\eta_1}$, was defined as the mean of the composite at the first wave:

$$\hat{\mu}_{\eta_1} = \mu_{Y_1} \tag{76}$$

Substituting Equation (75) in (76),

$$\hat{\mu}_{\eta_1} = \tau_1^* + \lambda_1^* \mu_{\xi_1} \tag{77}$$

Since the mean of the latent factor at time 1 was generated as zero $\mu_{\xi_1} = 0$ (Table 2.4 contains the generating parameter values), the estimated latent factor mean adopted the value of the sum of the intercepts at wave 1,

$$\hat{\mu}_{\eta_1} = \tau_1^* \tag{78}$$

Table 5.9 shows the sums of the generating item intercepts and loadings at wave 1. It should be noted that this sum was the same in all time points in the conditions in which there were invariant loadings and invariant intercepts. It can be seen that the sum of the intercepts and loadings varied by the number of items.

Table 5.9
Sum of item intercepts and loadings in the first measurement occasion

Num. Items	Intercept sum	Loading sum
6	2.1	4.2
9	3.4	6.0
15	6.2	9.9

In other words, Equation (78) shows that although the intercept factor mean was generated to be zero in the COFM, by analyzing composites of the items, the estimated intercept factor mean adopted the value of the sum of the item intercepts at wave 1.

Slope factor mean

In the COFM, the mean of the first order latent factors can be expressed as a function of the intercept and slope factors as,

$$\begin{matrix} \mu_{\xi_1} \\ \mu_{\xi_2} \\ \mu_{\xi_3} \\ \mu_{\xi_4} \\ \mu_{\xi_5} \end{matrix} = \begin{matrix} \mu_{\eta_1} \\ \mu_{\eta_1} \\ \mu_{\eta_1} \\ \mu_{\eta_1} \\ \mu_{\eta_1} \end{matrix} + \begin{matrix} 0\mu_{\eta_2} \\ 1\mu_{\eta_2} \\ 2\mu_{\eta_2} \\ 3\mu_{\eta_2} \\ 4\mu_{\eta_2} \end{matrix} \tag{79}$$

The mean of the slope factor mean can be expressed as,

$$\begin{aligned} \mu_{\xi_1} - \mu_{\eta_1} &= 0\mu_{\eta_2} \\ \mu_{\xi_2} - \mu_{\eta_1} &= 1\mu_{\eta_2} \\ \mu_{\xi_3} - \mu_{\eta_1} &= 2\mu_{\eta_2} \\ \mu_{\xi_4} - \mu_{\eta_1} &= 3\mu_{\eta_2} \\ \mu_{\xi_5} - \mu_{\eta_1} &= 4\mu_{\eta_2} \end{aligned} \tag{80}$$

Since the intercept factor mean was generated to be zero $\mu_{\eta_1} = 0$, then,

$$\begin{aligned} \mu_{\xi_2} &= \mu_{\eta_2} \\ \frac{\mu_{\xi_3}}{2} &= \mu_{\eta_2} \\ \frac{\mu_{\xi_4}}{3} &= \mu_{\eta_2} \\ \frac{\mu_{\xi_5}}{4} &= \mu_{\eta_2} \end{aligned} \tag{81}$$

However, the data were not analyzed using the COFM. Instead, item composites were analyzed using a LGM. Substituting the mean of the first order latent factors μ_{ξ_t} in Equation (79) with the mean of the item composites expressed as in Equation (75), yields,

$$\begin{matrix} \tau^* + \lambda^*\mu_{\xi_1} \\ \tau^* + \lambda^*\mu_{\xi_2} \\ \tau^* + \lambda^*\mu_{\xi_3} \\ \tau^* + \lambda^*\mu_{\xi_4} \\ \tau^* + \lambda^*\mu_{\xi_5} \end{matrix} = \begin{matrix} \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \\ \hat{\mu}_{\eta_1} \end{matrix} + \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{82}$$

where $\hat{\mu}_{\eta_1}$ and $\hat{\mu}_{\eta_2}$ represent the estimated intercept and slope factor means. In the conditions in which longitudinal invariance holds, the item intercepts and loadings sums did not change over time so the subindices denoting time were dropped. Since, $\hat{\mu}_{\eta_1} = \tau_1^*$ asdetermined in Equation (78), it follows that,

$$\begin{matrix} \lambda^*\mu_{\xi_1} \\ \lambda^*\mu_{\xi_2} \\ \lambda^*\mu_{\xi_3} \\ \lambda^*\mu_{\xi_4} \\ \lambda^*\mu_{\xi_5} \end{matrix} = \begin{matrix} 0\hat{\mu}_{\eta_2} \\ 1\hat{\mu}_{\eta_2} \\ 2\hat{\mu}_{\eta_2} \\ 3\hat{\mu}_{\eta_2} \\ 4\hat{\mu}_{\eta_2} \end{matrix} \tag{83}$$

Since the mean of the first order latent factor at wave 1 was generated to be zero, $\mu_{\xi_1} = 0$, then,

$$\begin{gathered} \lambda^*\mu_{\xi_2} = \hat{\mu}_{\eta_2} \\ \frac{\lambda^*\mu_{\xi_3}}{2} = \hat{\mu}_{\eta_2} \\ \frac{\lambda^*\mu_{\xi_4}}{3} = \hat{\mu}_{\eta_2} \\ \frac{\lambda^*\mu_{\xi_5}}{4} = \hat{\mu}_{\eta_2} \end{gathered} \tag{84}$$

When comparing Equation (84) to Equation (81), it is observed that the latent factor means were re-scaled by multiplying the sum of the loadings at each time point. As a consequence, the estimated slope factor mean represented the change in the re-scaled latent factor means.

Covariance structure

In the COFM, the variance of the first order latent factors, ϕ_{ξ_t}, can be expressed as,

$$\phi_{\xi_t} = \Psi_{11} + \Psi_{22} + 2\Psi_{12} \quad (85)$$

However, composites were analyzed instead of the first order latent factors, as

$$\sigma^2_{Y_t} = \hat{\Psi}_{11} + \hat{\Psi}_{22} + 2\hat{\Psi}_{12} \quad (86)$$

where $\hat{\Psi}_{11}$ is the estimated variance of the intercept factor, $\hat{\Psi}_{22}$ is the estimated variance of the slope factor, and $\hat{\Psi}_{12}$ is the estimated covariance of the intercept and slope factors. Expressing the composites as a function of the common factor model,

$$\lambda_t^{*2}\phi_{\xi_t} + \theta_t^* = \hat{\Psi}_{11} + \hat{\Psi}_{22} + 2\hat{\Psi}_{12} \quad (87)$$

where θ_t^* is the sum of the unique variances at wave 1. Since the growth latent factors are modeling only the variance of the first order factor and not the unique variance, the unique variance term can be dropped,

$$\lambda_t^{*\mathbf{2}} \phi_{\xi_t} = \Psi_{11} + \Psi_{22} + 2\Psi_{12} \quad (88)$$

Equation (88) shows that the variance of the first order latent factor was re-scaled by multiplying the square of the sum of the intercept loadings. Hence, the estimated intercept factor variance, the estimated slope factor variance, and the estimated covariance between the intercept and the slope factors modeled the re-scaled first order latent factor.

Re-scaling the growth parameter values

One of the purposes of the present research was to determine the bias, relative bias, standard errors and RMSE of the estimated parameter values of the LGM when composites were formed by items that violated longitudinal factorial invariance. In order to calculate bias, relative bias and the RMSE of the parameter estimates, the true parameter values must be known. Since this research used a simulation study, the true parameter values were known. However, as shown above, by using composites instead of the latent factors there was a re-scaling that affected the estimated growth parameters even in the conditions in which the items were invariant over time. If the re-scaling were not corrected, inflated bias, relative bias and RMSE values would have been obtained even in the invariant conditions.

To correct for the re-scaling that occurred by using composites of items instead of the first order latent factors there were two options: to re-scale the estimated growth parameter values or to re-scale the true growth parameter values. In order to re-scale the estimated growth parameter values, it was necessary to change the estimated growth parameter values in each replication in each condition. Instead, the re-scale in the true parameter values was a one-time change, so this was the approached followed.

To re-scale the intercept factor mean, the sum of the item intercepts were added to the true value. The slope factor mean was re-scaled by multiplying the generating value by the sum of the item loadings. Finally, the intercept factor variance, slope factor variance and intercept-slope covariance were re-scaled by multiplying their generating parameter values by the square of the sum of the loadings. Since the sum of the intercepts and loadings varied depending on the number of items (Table 5.9), a different set of re-scaled true values were obtained for conditions with 6, 9 and 15 items as shown in Table 5.10.

Table 5.10
Original and re-scaled true growth parameter values

	Original	Re-scaled values		
		6 items	9 items	15 items
Intercept mean (μ_{η_1})	0	2.10	3.40	6.20
Slope mean (μ_{η_2})	0.20	0.84	1.20	1.98
Intercept variance (Ψ_{11})	0.50	8.82	18	49.01
Slope variance (Ψ_{22})	0.10	1.76	3.60	9.80
Intercept-slope covariance (Ψ_{12})	0.044	0.78	1.58	4.31

After implementing the re-scaling of the growth parameter estimates, the large bias and relative bias values previously observed in the conditions with invariant loadings and invariant intercepts decreased substantially. Tables 3.3 and Table 5.11 in Appendix D show that the relative bias and bias values of the growth parameter estimates under longitudinal invariance is zero, as expected.

APPENDIX D

BIAS IN THE LGM GROWTH PARAMETER ESTIMATES

In this appendix the bias of the growth parameter estimates by number of items, can be found for the invariant conditions. The bias results for the conditions with non-invariant loadings and with non-invariant intercepts are shown by number of items, magnitude of the violations, and the proportion of non-invariant items.

Table 5.11
Bias in the LGM parameter estimates in the invariant conditions

Num. items	Intercept factor mean	Intercept factor variance	Slope factor mean	Slope factor variance	Intercept-slope covariance
6	0.001	0.001	-0.023	-0.007	0.005
9	0.002	0.000	-0.092	-0.016	0.049
15	-0.007	0.003	-0.362	-0.063	0.059

Table 5.12
Bias in the LGM parameter estimates in conditions with violations of invariance

Num. items	Effect size	Prop. invariant	Non-invariant loadings					Non-invariant intercepts				
			Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.	Int. mean	Slope mean	Int. var.	Slope var.	I-S covar.
6	Small	1/3	0.02	-0.04	-0.01	-0.21	-0.10	0.00	0.09	-0.07	-0.02	0.01
		2/3	0.04	-0.08	0.01	-0.34	-0.18	0.00	0.13	-0.07	-0.01	0.01
	Medium	1/3	0.05	-0.09	0.02	-0.40	-0.23	0.00	0.32	-0.01	-0.02	0.01
		2/3	0.08	-0.15	0.01	-0.63	-0.39	0.00	0.54	-0.07	-0.01	0.01
	Large	1/3	0.08	-0.14	0.04	-0.57	-0.35	0.00	1.28	-0.09	-0.01	0.01
		2/3	0.13	-0.24	0.12	-0.88	-0.62	0.00	2.16	-0.04	-0.01	0.01
9	Small	1/3	0.03	-0.06	-0.07	-0.40	-0.20	0.00	0.10	-0.10	-0.01	0.02
		2/3	0.05	-0.11	0.07	-0.71	-0.37	0.00	0.19	-0.07	-0.02	0.03
	Medium	1/3	0.06	-0.12	0.02	-0.77	-0.43	0.00	0.42	-0.09	-0.03	0.01
		2/3	0.12	-0.23	0.09	-1.30	-0.81	0.00	0.78	-0.06	-0.03	0.01
	Large	1/3	0.10	-0.18	0.08	-1.08	-0.66	0.00	1.68	-0.10	-0.03	0.02
		2/3	0.18	-0.34	0.22	-1.83	-1.29	0.00	3.12	-0.05	-0.01	0.03
15	Small	1/3	0.05	-0.10	-0.07	-1.12	-0.50	-0.01	0.17	-0.20	-0.06	0.06
		2/3	0.09	-0.19	-0.04	-1.98	-1.06	-0.02	0.33	-0.25	-0.04	0.03
	Medium	1/3	0.11	-0.21	-0.04	-2.08	-1.12	0.00	0.70	-0.33	-0.05	0.05
		2/3	0.19	-0.39	0.30	-3.68	-2.27	-0.01	1.32	-0.19	-0.03	0.03
	Large	1/3	0.15	-0.30	0.24	-3.04	-1.75	0.01	2.80	-0.30	-0.05	0.06
		2/3	0.32	-0.59	0.89	-5.12	-3.59	0.00	5.28	-0.29	-0.04	0.08

APPENDIX E

BIAS IN THE AR QUASI-SIMPLEX PARAMETER ESTIMATES

In this appendix the bias of the AR quasi-simplex parameter estimates by number of items, can be found for the invariant conditions. The bias results for the conditions with non-invariant loadings and with non-invariant intercepts are shown by number of items, magnitude of the violations, and the proportion of non-invariant items.

Table 5.13
Bias in the AR quasi-simplex parameter estimates in the invariant conditions

Num. items	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
6	0.00	0.00	0.00	0.00
9	0.04	0.00	0.00	0.00
15	0.04	0.00	0.00	0.00

Table 5.14
Bias in the AR quasi-simplex parameter estimates in conditions with violations of invariance

Num. Items	Effect size	Prop. non-inv.	Non-invariant loadings				Non-invariant intercepts			
			ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}	ρ_{21}	ρ_{32}	ρ_{43}	ρ_{54}
6	Small	1/3	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
		2/3	-0.02	-0.02	-0.02	-0.02	0.00	0.00	0.00	0.00
	Medium	1/3	-0.03	-0.03	-0.03	-0.03	0.00	0.00	0.00	0.00
		2/3	-0.05	-0.05	-0.05	-0.05	0.00	0.00	0.00	0.00
	Large	1/3	-0.04	-0.04	-0.04	-0.04	0.00	0.00	0.00	0.00
		2/3	-0.07	-0.07	-0.08	-0.08	0.00	0.00	0.00	0.00
9	Small	1/3	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
		2/3	-0.02	-0.02	-0.02	-0.02	0.00	0.00	0.00	0.00
	Medium	1/3	-0.02	-0.03	-0.02	-0.03	0.00	0.00	0.00	0.00
		2/3	-0.04	-0.05	-0.05	-0.05	0.00	0.00	0.00	0.00
	Large	1/3	-0.04	-0.04	-0.04	-0.04	0.00	0.00	0.00	0.00
		2/3	-0.07	-0.07	-0.08	-0.09	0.00	0.00	0.00	0.00
15	Small	1/3	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00
		2/3	-0.02	-0.02	-0.02	-0.02	0.00	0.00	0.00	0.00
	Medium	1/3	-0.02	-0.02	-0.02	-0.02	0.00	0.00	0.00	0.00
		2/3	-0.05	-0.05	-0.05	-0.05	0.00	0.00	0.00	0.00
	Large	1/3	-0.04	-0.04	-0.04	-0.04	0.00	0.00	0.00	0.00
		2/3	-0.07	-0.07	-0.08	-0.08	0.00	0.00	0.00	0.00