



Departamento de Métodos Quantitativos

LATENT GROWTH MIXTURE MODELING:
AN APPLICATION IN THE AERONAUTIC TRAINING
ENVIRONMENT

Ana Patrícia Correia Gomes

Thesis submitted as partial requirement for the conferral of

Master in Data Mining and Data Analysis

Supervisor:
Prof. Doutor José Gonçalves Dias

Setembro 2009

Departamento de Métodos Quantitativos

**LATENT GROWTH MIXTURE MODELING:
AN APPLICATION IN THE AERONAUTIC TRAINING
ENVIRONMENT**

Ana Patrícia Correia Gomes

Thesis submitted as partial requirement for the conferral of

Master in Data Mining and Data Analysis

Supervisor:
Prof. Doutor José Gonçalves Dias

Setembro 2009

RESUMO

A aplicação de modelo de mistura com crescimento latente a dados longitudinais oferece uma generalização importante dos modelos de crescimento convencionais, permitindo a identificação de diferentes padrões de crescimento, tendo em conta a heterogeneidade da população.

O principal objectivo deste estudo consiste em analisar o processo de aprendizagem no treino de pilotos (modelos com crescimento latente), identificar diferentes padrões de crescimento resultantes da heterogeneidade existente (modelo de mistura com crescimento latente), e identificar as variáveis explicativas da variabilidade e do padrão de crescimento. O objecto de estudo é o desempenho no treino de pilotos *ab-initio* (n=297), candidatos à Academia da Força Aérea Portuguesa (avaliados em seis medidas repetidas).

Os resultados obtidos demonstram que existe heterogeneidade não observada na população e que o modelo mais adequado é um modelo de mistura com crescimento latente de duas classes.

A coordenação motora (SMA) demonstrou um efeito significativo no intercepto (estado inicial) e o prognóstico de *Adaptabilidade Geral* (dimensão Personalidade/Motivacional) demonstrou um efeito significativo quer no intercepto (estado inicial) quer no declive (aprendizagem). A classe latente 1 (66% da amostra) caracteriza-se por apresentar uma performance em voo superior no estado inicial (intercepto), um efeito significativo da *Adaptabilidade Geral* no intercepto, e melhores resultados nos testes realizados na fase de avaliação psicológica. Por sua vez, a classe latente 2 (34% da amostra) apresenta piores resultados relativos ao estado inicial da performance em voo, e um efeito significativo da *Adaptabilidade Geral* na aprendizagem (declive).

Palavras-chave: modelos com mudança latente, modelo de mistura com crescimento latente, treino aeronáutico, heterogeneidade não observada.

Classifications from JEL Classification System: C23, M53.

ABSTRACT

The application of growth mixture modeling to longitudinal data offers an important extension of conventional modeling tools, enabling the identification of different patterns in growth, by accounting for population heterogeneity.

The main goal of this study is to analyze the shape of the learning process in pilot training (Latent Growth Modeling), as well as to recognize different patterns in growth due to population heterogeneity (Growth Mixture Modeling). Moreover, the research intends to identify predictors that explain that variability and the pattern of growth. The object of study is the performance in flight training of ab-initio pilot applicants ($n=297$) to the Portuguese Air Force Academy (evaluated through six repeated measures).

The results showed the existence of unobserved heterogeneity in the population and the best fitting model is a 2-class mixture model.

Psychomotor coordination (SMA) showed a significant effect on the intercept (initial status) and the prognostic of *General Adaptability* (Personality and Motivation dimension) depicted a significant effect on the intercept (initial status) and on slope (development). The latent class 1 (66% of the sample) presents the highest initial flight performance, a positive significant effect of the General Adaptation on the intercept and the best results in the tests performed in the psychological phase. The latent class 2 (34% of the sample) presents the worst initial flight performance, and a positive significant effect of General Adaptation on the slope.

Keywords: latent change models, growth mixture modeling, aeronautic training, unobserved heterogeneity.

Classifications from JEL Classification System: C23, M53.

SUMÁRIO EXECUTIVO

Os modelos com crescimento latente (MCL) apresentam-se como uma valiosa ferramenta para modelar dados longitudinais. Nos últimos anos verificou-se um interesse crescente pela utilização desta técnica, devido em parte à sua grande aplicabilidade às ciências económicas, ciências psicológicas e ao marketing.

Os modelos com crescimento latente (MCL) focam-se essencialmente na modelação de curvas de crescimento, modelando simultaneamente diferenças intraindividuais (um indivíduo ao longo do tempo) e inter-individuais (diferença entre as curvas para os vários indivíduos). É ainda assumido que a amostra provém de uma única população homogénea. Contudo, o pressuposto de homogeneidade nos parâmetros de mudança nem sempre é válido.

O modelo de mistura com crescimento latente (MMCL) é apresentado como uma extensão ao MCL, permitindo uma estrutura de clusters na modelação. Este tipo de modelos é considerado como uma segunda geração de modelos de equações estruturais com variáveis latentes, ao permitir além de variáveis latentes contínuas, também variáveis latentes discretas que definem uma tipologia de grupos e permite modelar populações heterogéneas.

Este trabalho foi aplicado ao meio aeronáutico, nomeadamente ao processo de aprendizagem no treino de pilotos militares, tendo como principal objectivo identificar diferentes padrões de crescimento resultantes da heterogeneidade existente, e variáveis explicativas da variabilidade e do padrão de crescimento. O objecto de estudo é o desempenho no treino de pilotos ab-initio ($n = 297$), candidatos à Academia da Força Aérea Portuguesa (avaliados em seis medidas repetidas). A amostra é constituída pela totalidade de candidatos, que entre os anos de 2001 a 2008 completaram com sucesso as provas psicológicas, médicas e físicas. As seis medidas repetidas são operacionalizadas pelos desempenhos obtidos nos seis voos do Estágio de Selecção de Voo (ESV). O Estágio de Selecção de Voo constitui a última etapa do processo de selecção e tem como principal objectivo a eliminação dos candidatos que não se consigam adaptar às exigências do meio aeronáutico. São também incluídas neste estudo, variáveis preditoras do cluster e do crescimento latente, nomeadamente a coordenação motora, a aptidão espacial e o prognóstico de adaptabilidade.

Os resultados obtidos demonstram que existe heterogeneidade não observada na população e que o modelo mais adequado é um modelo de mistura com duas classes. A coordenação

motora demonstrou um efeito significativo no estado inicial e o prognóstico de adaptabilidade demonstrou um efeito significativo quer estado inicial quer na aprendizagem. A classe latente 1 (66% da amostra) caracteriza-se por apresentar uma performance em voo superior no estado inicial e um efeito significativo do prognóstico de adaptabilidade no estado inicial. Por sua vez, a classe latente 2 (34% da amostra) apresenta piores resultados relativos ao estado inicial da performance em voo, e um efeito significativo do prognóstico de adaptabilidade na aprendizagem.

O MMCL com variáveis ordinais está ainda em desenvolvimento e encontram-se ainda muito poucos trabalhos publicados com aplicações, e mesmo os próprios pacotes estatísticos estão ainda em fase de aperfeiçoamento. O presente trabalho apresenta-se como um contributo para a exploração desta abordagem como também uma primeira tentativa de modelação longitudinal da performance aeronáutica.

ACKNOWLEDGMENTS

I would like to thank some people, who directly or not contributed to the completion of this master's thesis:

First, I wish to thank my supervisor, Professor José Gonçalves Dias, about the way he supervised me through this study. His endless enthusiasm, dedication and guidance have encouraged me to bring this research into completion. I am deeply indebted to him for his help.

I would like to thank to the Portuguese Air Force Psychology Center for giving me the opportunity of developing this research, and to all pilot applicants that made this work possible.

Special thanks to my dear friends for their support and encouragement.

I would also like to express thanks to my family for their love and understanding.

I also wish to thank my fiancé's family for their friendship.

Finally, I would like to thank my fiancé Tó, for the unconditional support throughout this process, companionship, concern, and helpful advices in the decisive moments, which made it possible to achieve this goal.

INDEX

RESUMO	I
ABSTRACT	III
SUMÁRIO EXECUTIVO	V
ACKNOWLEDGMENTS	VII
INDEX	IX
ACRONYMOUS	XI
1 INTRODUCTION	1
2 LATENT GROWTH MIXTURE MODELING	5
2.1 INTRODUCTION	5
2.2 LATENT GROWTH CURVE MODELS	6
2.2.1 <i>Model Specification</i>	6
2.2.2 <i>Model Estimation</i>	8
2.2.3 <i>LGM with Covariates</i>	9
2.2.4 <i>Measurement Model for Ordinal Variables</i>	9
2.3 LATENT GROWTH MIXTURE MODELING	11
2.3.1 <i>Background</i>	11
2.3.2 <i>Model Specification</i>	13
2.3.3 <i>Model Estimation</i>	16
2.4 MODEL ASSESSMENT	17
2.4.1 <i>Akaike information criterion (AIC)</i>	18
2.4.2 <i>Bayesian Information Criterion (BIC)</i>	18
2.4.3 <i>Adjusted Bayesian Information Criterion (aBIC)</i>	18
2.4.4 <i>Lo-Mendell-Rubin likelihood ratio test – LMR LRT</i>	19
2.5 SYNTHETIC DATA EXAMPLE	20
3 EMPIRICAL STUDY	23
3.1 INTRODUCTION	23
3.2 HYPOTHESES	24
3.3 DATA	25
3.3.1 <i>Sample</i>	25
3.3.2 <i>Flight Screening (FS)</i>	25
3.3.3 <i>Covariates</i>	29
3.3.3.1 <i>Personality and Motivation Dimension</i>	29
3.3.3.2 <i>Perceptive-cognitive and Psychomotor Dimension</i>	31
3.4 EMPIRICAL RESULTS	34
3.4.1 <i>Single-class growth curve model – Assuming homogeneity</i>	34
3.4.2 <i>Checking population heterogeneity</i>	36
3.4.3 <i>Selecting the appropriate number of latent classes</i>	37
3.4.4 <i>Two-class Mixture Model</i>	39
3.4.5 <i>LGMM with Covariates</i>	43
3.5 DISCUSSION AND IMPLICATIONS OF THE RESULTS	48

4	CONCLUSIONS.....	51
4.1	EMPIRICAL FINDINGS	51
4.2	LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH.....	52
	REFERENCES.....	55
	APPENDICES	59

ACRONYMOUS

ABIC	Adjusted Bayesian Information Criterion
AIC	Akaike Information Criterion
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
BIC	Bayesian Information Criterion
CAIC	Consistent Akaike Information Criterion
CFI	Comparative Fit Index
EM	Expectation Maximization
FS	Flight Screening
GGMM	General Growth Mixture Model
GMM	Growth Mixture Modeling
HLM	Hierarchical Linear Modeling
INS2	Instruments 2
LCA	Latent Class Analysis
LCM	Latent Curve Model
LGM	Latent Growth Model
LGCM	Latent Growth Curve Model
LMR – LRT	Lo-Mendell-Rubin likelihood ratio test
MANOVA	Multivariate Analysis of Variance
ML	Maximum Likelihood
NNFI	Non-Normed Fit Index
PILAV	Pilot Aviator
PoAFA	Portuguese Air Force Academy
RMSEA	Root Mean Square Error of Approximation
SEM	Structural Equation Model
SMA	Sensory Motor Apparatus
VLMR	Vung-Lo-Mendell-Rubin statistic

LIST OF FIGURES

Figure 1 – Representation of the growth mixture model.....	14
Figure 2 – Representation of the model being simulated.....	20
Figure 3 – Scores obtained in the selection flights.....	26
Figure 4 – Recoded scores obtained in the selection flights.....	27
Figure 5 – Hands (spatial aptitude test).....	32
Figure 6 – Instruments interpretation.....	32
Figure 7 – Sensory Motor Apparatus.....	33
Figure 8 – Representation of the single class mixture model.....	35
Figure 9 – Growth curve for a single latent class model.....	35
Figure 10 – Model assessment.....	38
Figure 11 – Two-class mixture model.....	40
Figure 12 – Growth curve for each of the two latent class membership.....	41
Figure 13 – Two-class mixture model with covariates.....	43
Figure 14 – Latent Class profiles.....	49

LIST OF TABLES

Table 1 – Parameter estimates for latent class 1 (simulated data set).	21
Table 2 – Parameter estimates for latent class 2 (simulated data set).	22
Table 3 – Applicants per year and results of the FS.....	26
Table 4 – Flight scores frequencies.....	27
Table 5 – Recoded flight scores frequencies.....	28
Table 6 – Summary of categorical data proportions.	28
Table 7 – Personality and Motivation Dimensions.	30
Table 8 – Descriptive statistics for General Adaptability.	31
Table 9 – Descriptive statistics.	34
Table 10 – Parameter estimates – single latent class.....	36
Table 11 – Vuong-Lo-Mendell-Rubin likelihood ratio test for 1 (H0) versus 2 classes	37
Table 12 – Lo-Mendell-Rubin Adjusted LRT test.	37
Table 13 – Model selection criteria.....	38
Table 14 – Vuong-Lo-Mendell-Rubin likelihood ratio test for 2 (H0) versus 3 classes	39
Table 15 – Lo-Mendell-Rubin Adjusted LRT test (2 versus 3 classes)	39
Table 16 – Final latent class proportions and classification.....	41
Table 17 – Parameter estimates (unconditional model).	42
Table 18 – Final latent class proportions and classification (conditional model).	44
Table 19 – Parameter estimates for latent class 1 (conditional model).	44
Table 20 – Parameter estimates for latent class 2 (conditional model).	45
Table 21 – Parameter estimates (conditional model).	46
Table 22 – Parameter estimates for the effects of the covariates on <i>C</i>	47

1 Introduction

Recent developments in latent growth modeling (LGM) provide a better tool to model longitudinal data than traditional statistic methods. Longitudinal modeling is frequently encountered in behavioral research (Muthén and Curran, 1997). No matter the subject area or the time interval, social and behavioral scientists have a strong interest in describing and explaining the time trajectories of their variables (Bollen and Curran, 2006). Latent growth modeling postulates the existence of *latent trajectories*. The term *latent* stands for a process that is not directly observed. The trajectory process is observed only indirectly using manifest repeated measures. However, this trajectory can differ at individual level (Bollen and Curran, 2006). In the recent years there has been a growing interest among researchers in the use of latent growth modeling (Jung and Wickrama, 2008). These techniques have become popular in applied social and psychological sciences, in part due to its software implementation (e.g., Mplus, Gllamm).

The statistical analysis of data over time appears in the nineteenth century. The studies were oriented taking into account the change in groups rather than in individuals. The focus on aggregate change continued in the twentieth century. These early applications utilized mostly complex functional forms of growth (e.g., nonlinear polynomials and logistic curves) to examine change in an entire group of individuals (Bollen and Curran, 2006). The theory of LGM appeared two decades ago with McArdle and Epstein (1987) and Meredith and Tisak (1990). This model estimates growth function with fixed and random parts, which describes the average development in the population as well as the variation of individual development, respectively (McArdle and Epstein, 1987; Meredith and Tisak, 1990; Willett and Sayer, 1994). These methods allow researchers to move beyond the use of ad-hoc categorization procedures for constructing developmental trajectories.

This modeling approach has been extended to include the estimation of the impact of covariates on individual growth as well as (besides observed variables with continuous scale) categorical variables (Muthén, 2004a). The estimation of model parameters can be handled using structural equation modeling (SEM) programs, such as LISREL (Jöreskog *et al.*, 1999), EQS (Bentler, 1995), Amos (Arbuckle, 2006) and Mplus (Muthén and Muthén, 1998-2007).

In recent years, the extensiveness and intensity of longitudinal studies have raised, and a particularly attention has been given to “inter-individual variation” (variation between individuals) and “intra-individual variation” (variation within individuals). The approaches focusing on inter-individual variation point out to the establishment of general development principles that apply to all individuals. On the other hand, approaches focusing on intra-individual variation emphasize understanding change within individuals, viewing the establishment of general principles as a secondary aim. Growth is a phenomenon that occurs within the individual, and therefore intra-individual variability is a primary interest in statistical modeling of longitudinal data (McGrath and Tschan, 2004).

Latent growth models can analyze the development of individuals in one or more outcome variables over time. Observed outcome variables can be continuous, censored, binary, ordered categorical (ordinal), counts, or combinations of these variable types if more than one growth process is being modeled. In a latent variable modeling framework, the random effects are conceptualized as continuous latent variables, that is, growth factors are used to capture individual differences in development (Muthén and Muthén, 1998-2007).

Another important recent latent variable modeling extension connected with the SEM context is mixture modeling. This modeling is based on the idea that population contains subpopulations, i.e. latent classes. These classes can be identified and their parameters estimated (Muthén and Shedden, 1999; Muthén, 2001). The distribution of observed variables is over a mixture distribution, so that each subpopulation has its own model parameters values (Lubke and Muthén, 2005).

Different terminologies are used in this context. For example, Muthén (2004a) uses the term “latent growth curve model”, whereas Meredith and Tisak (1990) uses the term “latent curve analysis” to refer the modeling of growth in the context of SEM. When viewed from a SEM perspective, the model is called latent growth model (Hoeksma and Kelderman, 2007). On the other hand, the term “latent growth mixture modeling” (LGMM or LGM modeling; Muthén, 2001; Muthén, 2004a), in turn, is used to refer to the combination of latent growth curve modeling and mixture modeling.

Despite twenty years old, the LGM and LGMM are novel methods in practice. As these methodologies become more common and important to model development, the functionality of the models needs more research. Because most of inferential analysis is based on asymptotic results, researchers can trust the results when sample size is large, at least 1000 observations. However, in many empirical studies, sample size is limited, say, to 100-500 observations. With simulated data is possible to examine the performance of these methods with different sample sizes (Bauer and Curran, 2003; Muthén, 2004a).

The main goal of this thesis consists in modeling and understanding the nature and heterogeneity in pilot performance development. This study focuses on two different types of data: simulated data sets ($n=500$, $n=1000$, $n=5000$, $n=10000$), and an empirical example ($n=297$) from the aeronautic training context. The simulated data makes it possible to draw conclusions concerning the performance of LGMM with different sample sizes, while the application of LGMM in empirical data shows that this approach is able to capture the underlying process being modeled. Methods for latent growth modeling do allow incorporating attributes that can help improve predictions and, hence, design more effective interventions. Indeed, at the moment the performance in flight has been studied in the perspective of trying to find predictors of flight performance (Burke and Hunter, 1995). However, the dichotomous split of training success into *pass* versus *fail* dominates the pilot selection literature (Roe, 2008).

Martinussen (2003) pointed out that the supported interest in the theme of pilot selection drifts essentially of the high cost associated to pilot training mainly due to the use of aircrafts. The search for better predictors of success in pilot training is connected to high costs related to personnel training, and the need for recruiting competent and well-suited people. According to the same author, the cognitive and the psychomotor measures were pointed out as the best predictors of flight performance.

Previous studies concerning pilot performance in PoAF (Portuguese Air Force) identified key aptitudes as predictors of flight performance (Bártolo-Ribeiro *et al.*, 2004). The tests applied in the psychological phase, concerning the assessment of the psychomotor coordination (Sensory Motor Apparatus - SMA) and spatial aptitude (INSB2 and Hands) are predictors of pilot training successful performance. This fact is not surprising as the spatial aptitude and motor coordination are two factors strongly related to the flight adaptation.

Concerning the training of military pilots, the candidates have little or even no experience in the field of the aviation. After initial screening (psychological, medical and physical phases), the pilot candidates accomplish a selection training program, the FS (Flight Screening), composed by seven flights, in which they have to demonstrate their learning capabilities. This research analyzes the shape of the learning process in pilot training (LGM), allowing for heterogeneous patterns of growth due to sample heterogeneity (GMM).

Next, we present the structure of this thesis. Chapter 2 gives an overview of LGM and LGMM frameworks. Section 2.1 consists of an introduction to the setting of LGM. In Section 2.2, the LGM is discussed in four stages: model specification, inclusion of covariates, model estimation, and the application to ordinal variables. The LGMM framework is explored in Section 2.3, and Section 2.4 discusses model assessment. Section 2.5 illustrates these approaches using simulated data sets. Chapter 3 deals with the empirical application. Section 3.1 gives an overview to the pilot evaluation framework, Section 3.2 presents the study hypotheses, Section 3.3 describes the data set, Section 3.4 provides the results and Section 3.5 provides a discussion of the empirical results. The thesis ends with a final chapter in which a summary of main conclusions, major limitations of this research, and suggestions for further research are provided.

2 Latent Growth Mixture Modeling

2.1 *Introduction*

The most applied approach to modeling change with continuous variables is *growth curve models*. They fit growth trajectories for individuals and relate characteristics of these individual growth trajectories (e.g., slope) to covariates (Collins, 2005).

One type of analytic technique designed to address questions of individual differences in change over time is referred to as random coefficients modeling. The translation into latent variable models can be made using a simple random coefficient growth curve model as a starting point that illustrates the key features of the latent variable longitudinal model approach (Curran and Muthén, 1999).

The latent growth model (LGM) is commonly composed of two factors. The first and second factors relate to the level (i.e., intercept) and slope of growth, respectively. These two factors correspond to the mean value of the intercept and slope and individual random variation of these two latent components. Slope is either fixed to describe linear change or, alternatively, the pattern of slope is estimated (Aunola *et al.*, 2004).

2.2 *Latent Growth Curve Models*

2.2.1 **Model Specification**

In LGM the focus of interest is on the average development over time and individual variation around this average. The pattern of slope component is the same for each individual, but the strength of this pattern may vary individually (Collins, 2005). It is possible to define the latent growth model in two parts:

a) **Measurement Model**

$$y_{it} = \eta_{0i} + \eta_{1i}(t - 1) + \varepsilon_{it}, i = 1, 2, \dots, n \text{ and } t = 1, 2, \dots, T, \quad (1)$$

where

y_{it} is the observation of individual i at the time point t ,

η_{0i} is the intercept component of individual i ,

η_{1i} is the linear slope component for individual i ,

ε_{it} is the measurement error for individual i at the time point t ,

n is the number of observations and

T is the number of time points.

The intercept and slope parameters are random effects; in other words, they may vary across individuals, as reflected in the need for the subscript denoting individual.

b) **Latent Model**

$$\begin{aligned} \eta_{0i} &= \alpha_0 + \zeta_{0i}, \\ \eta_{1i} &= \alpha_1 + \zeta_{1i}, i = 1, 2, \dots, n, \end{aligned} \quad (2)$$

where α_0 and α_1 are expected values (fixed part of the model) of latent growth factors η_{0i} and η_{1i} , respectively; and ζ_{0i} and ζ_{1i} are random variables that configure individual growth. This model describes linear growth over time, so that intercept (η_{0i}) and slope (η_{1i}) represent

the initial status and linear growth rate of the growth trajectory of individual i , respectively (Curran and Muthén, 1999).

The standard latent curve can be presented by using a general SEM framework (Hipp and Bauer, 2006)

$$\begin{aligned} \mathbf{y}_i &= \mathbf{\Lambda}\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i \\ \boldsymbol{\eta}_i &= \boldsymbol{\alpha} + \boldsymbol{\zeta}_i \end{aligned} \quad (3)$$

where \mathbf{y}_i is a $T \times 1$ vector of observed repeated measures for individual i , where T is the number of waves of data, $\boldsymbol{\eta}_i = (\eta_{0i}, \eta_{1i})^T$ is a 2×1 vector of latent growth factors (or random coefficients; i.e., intercept and slope), $\boldsymbol{\varepsilon}_i$ is a $T \times 1$ a vector of measurement errors, and $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)^T$ is a 2×1 vector of expected values of $\boldsymbol{\eta}_i$.

We assume further that $E(\boldsymbol{\varepsilon}_i) = \mathbf{0}$ and $E(\boldsymbol{\zeta}_i) = \mathbf{0}$, and

$$\begin{aligned} \text{COV}(\boldsymbol{\varepsilon}_i) &= \boldsymbol{\Theta} = \text{diag}(\theta_1, \theta_2, \dots, \theta_T) \\ \text{COV}(\boldsymbol{\zeta}_i) &= \boldsymbol{\Psi} = \begin{bmatrix} \psi_{00} & \psi_{01} \\ \psi_{10} & \psi_{11} \end{bmatrix}, \end{aligned} \quad (4)$$

where $\boldsymbol{\Theta}$ is the $T \times T$ covariance matrix of the time-specific residuals and $\boldsymbol{\Psi}$ is covariance matrix of the growth factors.

The mean vector and covariance matrix of $\boldsymbol{\eta}$ are

$$\begin{aligned} E(\boldsymbol{\eta}) &= \boldsymbol{\mu}_\eta = \boldsymbol{\alpha}, \\ \text{COV}(\boldsymbol{\eta}) &= \boldsymbol{\Psi}. \end{aligned} \quad (5)$$

The mean vector and the covariance matrix of the observed variables \mathbf{y} are

$$\begin{aligned} E(\mathbf{y}) &= \boldsymbol{\mu}_y = \mathbf{\Lambda}E(\boldsymbol{\eta}) = \mathbf{\Lambda}\boldsymbol{\alpha}, \\ \text{COV}(\mathbf{y}) &= \boldsymbol{\Sigma}_y = \mathbf{\Lambda}\boldsymbol{\Psi}\mathbf{\Lambda}^T + \boldsymbol{\Theta}. \end{aligned} \quad (6)$$

where $\boldsymbol{\Sigma}_y$ is the variance-covariance matrix of y , $\boldsymbol{\Theta}$ and $\boldsymbol{\Psi}$ represent the covariance structure of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\zeta}$, respectively. The functional form of the individual trajectories is defined by fixing

the coefficients in the $T \times 2$ factor-loading matrix Λ to predetermined values, as we have seen before. It is traditionally assumed that the growth factors and time-specific residuals are multivariate normally distributed as

$$\begin{bmatrix} \eta_i \\ \varepsilon_i \end{bmatrix} \sim N \left(\begin{bmatrix} \alpha \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Psi & \mathbf{0} \\ \mathbf{0} & \Theta \end{bmatrix} \right). \quad (7)$$

According to equation (5), the $T \times 1$ vector \mathbf{y} follows the multivariate normal distribution

$$\mathbf{y}_i \sim N_T(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y), i = 1, 2, \dots, n, \quad (8)$$

where $\boldsymbol{\mu}_y$ and $\boldsymbol{\Sigma}_y$ are given in equation (6).

2.2.2 Model Estimation

Maximum-likelihood (ML) estimation is carried out by maximizing the log-likelihood function (Muthén and Khoo, 1998). Raykov (2005) refers that this is a very popular estimation and testing framework. Its essence is the estimation of unknown parameters that underlie an assumed model by values that maximize the probability of observing the data at hand (probability density). ML is employed for parameter estimation, model fit evaluation, and hypothesis testing purposes, in particular when examining hypotheses of parameter restrictions via likelihood-ratio tests. The application of the ML method for parameter estimation assumes here the multivariate normality of the observed data. The distribution of the observed (independent) individual is given by the product of “probabilities”. This product is viewed as dependent on the unknown characteristics of the underlying distribution, such as means, variances, and covariances.

Let $\boldsymbol{\varphi} = \text{vec}(\boldsymbol{\alpha}, \boldsymbol{\Psi}, \boldsymbol{\Theta})$ be a $k \times 1$ vector, which consists of all free parameters in $\boldsymbol{\alpha}$, $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}$. The parameters $\boldsymbol{\varphi}$ can be estimated using the maximum likelihood estimation (ML).

The likelihood function is the *probability* density of the observed data considered as a function of the model parameters. For a set of n independent observations (given), the likelihood function can be expressed as

$$L(\boldsymbol{\varphi}; \mathbf{y}) = \prod_{i=1}^n f(\mathbf{y}_i; \boldsymbol{\varphi}), \quad (9)$$

and the log-likelihood can be expressed as

$$l(\boldsymbol{\varphi}; \mathbf{y}) = \log L(\boldsymbol{\varphi}; \mathbf{y}) = \sum_{i=1}^n \log f(\mathbf{y}_i; \boldsymbol{\varphi}). \quad (10)$$

This log-likelihood function can be maximized iteratively by the Newton-Raphson algorithm.

2.2.3 LGM with Covariates

It is possible and desirable to include predictors in addition to time. Let the latent growth for a single outcome variable y observed for individual i at time point t be related to a covariate x_i . In this case, considering only one covariate, the idea is that each individual has his or her own growth trajectory. Growth will be expressed in terms of an LGM, with the inclusion of predictors:

$$\begin{aligned} \mathbf{y}_i &= \Lambda \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i \\ \boldsymbol{\eta}_i &= \boldsymbol{\alpha} + \Gamma \mathbf{x}_i + \boldsymbol{\zeta}_i \end{aligned} \quad (11)$$

The log-likelihood function of the conditional model turns out to be more difficult to maximize and time consuming as it demands for numerical integration.

2.2.4 Measurement Model for Ordinal Variables

Modeling ordinal variables is relatively new in the applied SEM literature. In ordinal repeated measures designs, the data consist of a multidimensional contingency table. A model for ordinal repeated measures must predict the proportion of individuals within each response pattern in terms of hypothesized model parameters. A Latent Growth Curve Model for ordinal panel data is a model of multivariate proportions. An LGC model, however, imposes restrictions on sample means and covariances rather than on sample proportions. Therefore, we must somehow extract means and covariances from the observed ordinal proportions (Mehta *et al.*, 2004).

A growth model for ordinal measures predicts joint ordinal proportions. The LGC model hypothesizes that the observed response (\mathbf{y}_t) at each time point t is related to the

corresponding unobserved or latent continuous variable (y_t^*). Unknown cutpoints or thresholds are hypothesized to relate the unobserved scores on the latent variable, with the observed ordinal responses at each time point t . The underlying latent variables (y_t^*) are assumed to follow a multivariate normal distribution with unknown means and covariance. In the analysis of categorical variables, thresholds are provided for each increased scale level, or we can say that each threshold represents a portion of the underlying continuous scale. For example, the observed variable is represented by y , with three categories, and y^* represents the underlying latent or unobserved continuous variable. An ordinal variable does not have a metric scale, so the latent response variable y^* is used to describe the relationship between that categorical variable and other variables in the model.

The relationship between an observed ordinal measure (y), with R response categories ($r = 0, \dots, R - 1$) and underlying continuous variable (y^*), can be defined by means of $R - 1$ thresholds (τ_r) as (Mehta *et al.*, 2004):

$$y = r, \text{ if } \tau_r < y^* \leq \tau_{r+1}, \quad (12)$$

where $\tau_0 = -\infty$ and $\tau_R = +\infty$. If y^* is assumed to be normally distributed the probability that y^* is less than or equal to τ_r is:

$$P(y^* \leq \tau_r) = \Phi\left(\frac{\tau_r - E(y^*)}{SD(y^*)}\right), \quad (13)$$

where Φ is the cumulative standard normal distribution and $E(y^*)$ and $SD(y^*)$ are the mean and standard deviation of the latent variable y^* , respectively. Equation (13) gives the cumulative probability that the latent continuous y^* is less than or equal to the r^{th} threshold τ_r .

The probability that the observed ordinal Y is equal to a given category r is:

$$P(y = r) = P(y^* \leq \tau_{r+1}) - P(y^* \leq \tau_r). \quad (14)$$

The thresholds define the cutoff points on a normal distribution. For a given mean and standard deviation of y^* and the two thresholds (τ_r and τ_{r+1}), the probability that y is in category r (i.e., the probability that y is between the two thresholds) is the area under the standard normal curve bounded by the two thresholds on the z scale (τ_r^z and τ_{r+1}^z).

The correlation between pairs of outcomes is called the *polychoric correlation coefficient* when ordinal variables have a bivariate normal distribution. The correlation between an ordered categorical outcome and a variable measured on a continuous interval scale is called the *polyserial correlation coefficient*. Correlations such as the polychoric correlation are not computed from actual observed variables but rather from theoretical correlations of the underlying y^* variables. The use of these variables allows the focus of growth to be on changes in the continuous variables y^* . In spite of this, the y^* formulation can also be used to focus on changes in the probabilities of y (Duncan *et al.*, 2006).

2.3 *Latent Growth Mixture Modeling*

2.3.1 *Background*

Growth curve models, as have been seen in the previous section, focused on growth curve modeling that considers both intra-individual change and inter-individual differences in such change, but treats observed data as if collected from a single homogeneous population. Conventional latent growth modeling can be used to estimate the amount of variation across individuals in the latent growth factors (random intercepts and slopes) as well as the average growth. In other words, in a conventional latent growth model the individual variation around the estimated average trajectory is expressed as growth factors that are allowed to vary across individuals (Kreuter and Muthén, 2008). This assumption of homogeneity in the growth parameters does not always correspond to the reality, and if heterogeneity is ignored, statistical analyses and their effects can be seriously biased.

Latent growth mixture modeling assumes that the population of interest is not homogeneous (as measured by response probabilities) but consists of heterogeneous subpopulations with varying parameters, allowing for within-class variation (Muthén, 2006). The common theme in mixture modeling is the partition of the population into an unknown number of latent classes or subpopulations (Duncan *et al.*, 2002).

Mixture analysis also known as latent class analysis (LCA) describes how the probabilities of a set of observed categorical variables or indicators vary across groups of individuals, where group membership is not observed. For example, the observed categorical variables may

correspond to a set of dichotomous diagnostic (Muthén and Muthén, 2000). LCA refers to the unobserved groups of individuals as latent classes. LCA attempts to find the smallest number of latent classes that can describe the associations between the observed categorical variables. The analysis adds classes until the model fits the data well. The parameters of the model are the probabilities of being in each class and the probabilities of fulfilling each criterion given class membership. In addition, the latent class model provides estimates of class probabilities for each individual. These values are called posterior probabilities. LCA may relate the probability of class membership to a set of background variables.

The main objective of latent class analysis is: 1) to estimate the number and size of the latent classes in the mixture; 2) to estimate the response probabilities for each indicator given the latent class; 3) to examine the fit of the model to the data using various test statistics, and 4) to assign individuals to each latent class (Duncan *et al.*, 2002).

Muthén (2003) explored new ways of constructing more complex models that are better suited and more flexible to assessing change than conventional latent growth modeling. Muthén proposed an extension of current LGM methodology that includes relatively unexplored mixture models, such as growth mixture models, mixture structural equations, and models that combine latent class analysis and structural equation modeling (Muthén and Shedden, 1999, Duncan *et al.*, 2002; Muthén *et al.*, 2003). Muthén (2001) considers growth mixture modeling as a second generation of structural equation modeling.

The general framework proposed by Muthén and colleagues (Muthén and Shedden, 1999; Nagin, 1999; Muthén, 2001; Muthén, 2003; Muthén *et al.*, 2002; Muthén *et al.*, 2003) provides new opportunities for latent growth modeling. Growth mixture models (GMM) are applicable to longitudinal studies where individual growth trajectories are heterogeneous and belong to a finite number of unobserved groups. Muthén and colleagues (Muthén and Shedden, 1999; Muthén *et al.*, 2002; Muthén, 2003; Muthén *et al.*, 2003; Kreuter and Muthén, 2008) proposed the growth mixture modeling approach, combining categorical and continuous latent variables in the same model, in what they called a general latent growth mixture model framework (GGMM). The GGMM approach allows for unobserved heterogeneity in the sample, where different individuals can belong to different subpopulations.

Mixture modeling generally refers to modeling with categorical latent variables that represent mixtures of subpopulations in which population membership is not known a priori, but it is inferred from the data. In mixture modeling, unobserved heterogeneity in the development of a variable through time is captured by categorical and continuous latent variables. In particular, GMM relaxes the single-population assumption of conventional LGM method by allowing parameter differences across unobserved subpopulations. GMM allows within-class individual variation for the latent-growth factors that is captured by nonzero intercept and slope variances within latent classes. As in conventional LGM, the within-class individual variation is represented by random effects (i.e., continuous latent variables). Nevertheless, instead of considering individual variation around a mean growth curve, the growth-mixture model allows different classes of individuals to vary around different mean growth curves within each latent class (Wang and Bodner, 2007).

2.3.2 Model Specification

The GGMM modeling approach provides the joint estimation of: 1) a conventional finite mixture growth model where different growth trajectories can be captured by class-varying random coefficient means (the part that explains the latent classes); and 2) a logistic regression of outcome variables on the class trajectory (the conditional distribution in each latent class).

The model can be extended to estimate varying class membership probability as a function of a set of covariates (i.e., for each class the values of the latent growth parameters are allowed to be influenced by covariates) and to incorporate outcomes of the latent class variable (Duncan *et al.*, 2002).

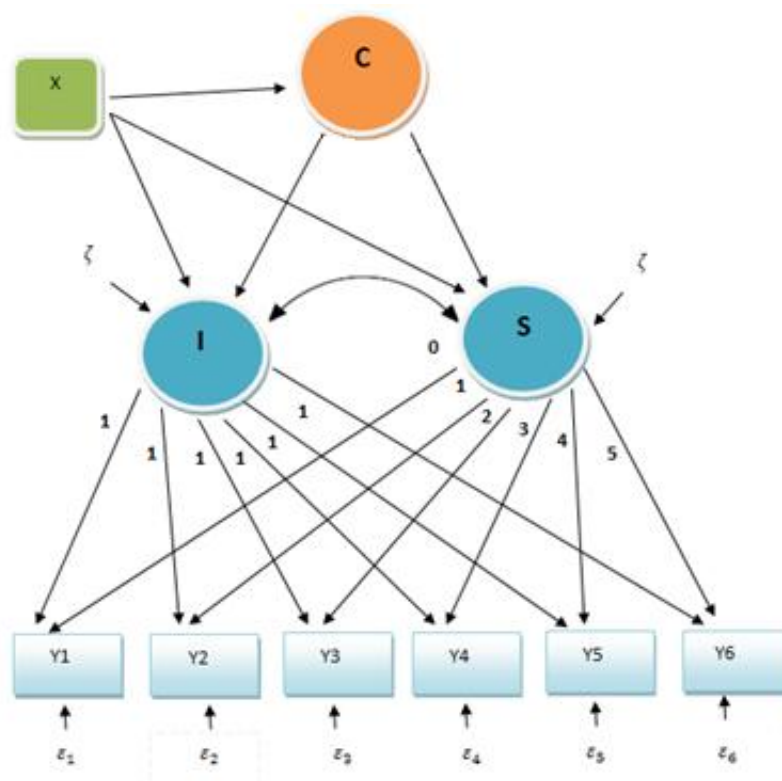


Figure 1 – Representation of the growth mixture model.

The full growth mixture model is depicted in Figure 1 (Muthén and Shedden, 1999; Muthén, 2001; Duncan *et al.*, 2002; Muthén, 2004b). This model contains a combination of a continuous latent growth variable, η_j ($j = \text{Intercept and Slope}$) and a latent categorical variable, C , with K classes, $C_i = (c_{i1} c_{i2}, \dots, c_{ik})$ where $c_{1k} = 1$ if individual i belongs to class k and zero otherwise. These latent variables are represented by circles in Figure 1.

In our case the continuous latent growth variable portion of the model represents conventional latent growth modeling with multiple indicators Y measured, let us say at six time points. The categorical latent variable is used to represent latent trajectory classes underlying the latent growth variable, η . Latent class and latent continuous variables can be predicted from a set of covariates (background variables), X , because the model allows the mixing proportions to depend on prior information and subject-specific variables.

In this model (Figure 1), the directional arrows from the latent trajectory classes to the growth factors indicate that the intercepts and slopes of the regressions of the growth factors on X vary across the classes of C .

Regarding the observed part of the model, multivariate normality is assumed for \mathbf{y}_i ¹ conditional on \mathbf{x}_i and class k ,

$$\begin{aligned}\mathbf{y}_i^{(k)} &= \mathbf{\Lambda} \boldsymbol{\eta}_i^{(k)} + \boldsymbol{\varepsilon}_i^{(k)} \\ \boldsymbol{\eta}_i^{(k)} &= \boldsymbol{\alpha}^{(k)} + \boldsymbol{\zeta}_i^{(k)},\end{aligned}$$

and by including covariates it turns out to be

$$\boldsymbol{\eta}_i^{(k)} = \boldsymbol{\alpha}^{(k)} + \boldsymbol{\Gamma}^{(k)} \mathbf{x}_i + \boldsymbol{\zeta}_i^{(k)}. \quad (15)$$

The model relates C to \mathbf{x}_i , by multinomial logistic regression for K classes (Muthén, 2004b):

$$P(C_{ik} = 1 | \mathbf{x}_i) = \frac{\exp(\delta_{k0} + \delta_{k1} x_{i1} + \dots + \delta_{kV} x_{iV})}{\sum_{k'=1}^K \exp(\delta_{k'0} + \delta_{k'1} x_{i1} + \dots + \delta_{k'V} x_{iV})}, \quad (16)$$

where V is the number of covariates. The multinomial logistic model is identified with null coefficients for class K , i.e., $\delta_{Kv} = 0, v = 0, \dots, V$.

The implied structure for a six-wave model is

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \\ y_{i6} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} \eta_{i1}^{(k)} \\ \eta_{i5}^{(k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1}^{(k)} \\ \varepsilon_{i2}^{(k)} \\ \varepsilon_{i3}^{(k)} \\ \varepsilon_{i4}^{(k)} \\ \varepsilon_{i5}^{(k)} \\ \varepsilon_{i6}^{(k)} \end{bmatrix}.$$

The mean vector of $\boldsymbol{\eta}^{(k)}$ (without covariates) is expressed as

$$E(\boldsymbol{\eta}^{(k)}) = \boldsymbol{\mu}_{\boldsymbol{\eta}}^{(k)} = \boldsymbol{\alpha}^{(k)}. \quad (17)$$

The covariances matrices of $\boldsymbol{\varepsilon}^{(k)} \sim N(\mathbf{0}, \boldsymbol{\Theta}^{(k)})$ and $\boldsymbol{\eta}^{(k)} \sim N(\mathbf{0}, \boldsymbol{\Psi}^{(k)})$ both assumed to be uncorrelated with other variables, and with each other, are

¹ Assuming that Y is an ordinal variable, y_i should be replaced by y_i^* .

$$\begin{aligned}\boldsymbol{\Theta}^{(k)} &= \text{COV}(\boldsymbol{\varepsilon}^{(k)}) = \text{diag}(\theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_T^{(k)}) \\ \boldsymbol{\Psi}^{(k)} &= \text{COV}(\boldsymbol{\eta}^{(k)}) = \begin{bmatrix} \psi_{00}^{(k)} & \psi_{01}^{(k)} \\ \psi_{10}^{(k)} & \psi_{11}^{(k)} \end{bmatrix}.\end{aligned}\quad (18)$$

Muthén and Shedden (1999) states that the model identification implies some parameter restrictions which have to be determined, as not all parameters can be estimated uniquely.

2.3.3 Model Estimation

Parameters are estimated by Maximum Likelihood (ML), using the Expectation-Maximization (EM) algorithm. The estimation of LGMM consists of two parts: the estimation of parameters related to the LGM and the estimation of class proportion (Yung, 1997; Muthén and Shedden, 1999; Dias, 2004). The log-likelihood function of observed data for the LGMM is

$$\log L = \sum_{i=1}^n \log L_i = \sum_{i=1}^n \log f(\mathbf{y}_i), \quad (19)$$

where the density function f is mixed from K density functions

$$f(\mathbf{y}_i) = \sum_{k=1}^K p^{(k)} f^{(k)}(\mathbf{y}_i), \quad (20)$$

and $p^{(k)}$ is the proportion of subpopulation k in the population. The density function in class k is

$$f^{(k)}(\mathbf{y}_i) \sim N_T(\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}), \quad (21)$$

where the observed data vectors \mathbf{y}_i are drawn from a multivariate normal distribution with mean vector $\boldsymbol{\mu}^{(k)}$ and covariance matrix $\boldsymbol{\Sigma}^{(k)}$ (in the case without covariates).²

Parameters are estimated using either the EM algorithm (Dempster *et al.*, 1977; Dias, 2004) or the Newton-Raphson algorithm. Each iteration of the EM algorithm involves two steps: the

² The default estimator for this type of analysis with ordinal variables is maximum likelihood with robust standard errors using a numerical integration algorithm. Note that numerical integration becomes increasingly more computationally demanding as the number of factors and the sample size increase (Muthén and Muthén, 1998-2007).

E–Step (Expectation Step) that estimates the conditional mean of the missing variable given the previous estimate of the model parameters and the observed data, and the M–Step (Maximization Step), that re-estimates the model parameters given the observed data and the soft clustering done by the E–Step. This algorithm is known for being a general method for ML estimation with incomplete data that reintroduces the additivity of the log-likelihood function using data augmentation (Dias, 2004).

2.4 *Model Assessment*

Model evaluation for latent growth mixture modeling (Duncan *et al.*, 2002) does not differ much from conventional SEM for homogeneous populations (Muthén *et al.*, 2002). For comparison of fit of models that have the same number of classes and are nested, the usual likelihood ratio chi-square difference test can be used. There exist a number of methods to evaluate the degree of fit of the hypothesized model to data and to assess whether the fit can be improved as a function of testing alternative models (Duncan *et al.*, 2006). Some of the indices of fit commonly used in latent growth modeling are the chi-square test statistic, the comparative fit index (CFI), and the root mean square error of approximation (RMSEA).

Testing and evaluating the overall fit is not possible, in the context of the mixture model, as it is in the framework of conventional structural equation models. Given the problems associated with the use of the likelihood ratio test for model selection, a variety of information-based criteria have been proposed to evaluate models with different numbers of latent classes. For this purpose the following criteria, provided by the current Mplus program, are used: Akaike’s information criterion (AIC), Bayesian information criterion (BIC), and adjusted BIC (aBIC). Moreover, the Lo-Mendell-Rubin likelihood ratio test is available as well (Lo *et al.*, 2001). Sample-size adjusted BIC (aBIC) has been shown to give superior performance in simulation studies for latent class analysis models (Muthén, 2003). The recommendation is to choose a model with the smallest AIC, BIC, or aBIC value (Li *et al.*, 2001).

2.4.1 Akaike information criterion (AIC)

Akaike information criterion, developed by Hirotugu Akaike under the name of "an information criterion" (AIC) in 1971 and published in Akaike (1974), is a measure of the goodness of fit of an estimated statistical model. This measure is intended for model comparison and not for the evaluation of an isolated model. AIC takes into account the statistical goodness-of-fit and the number of parameters that must be estimated to achieve that degree of fit. The model that produces the minimum AIC might be considered, in the absence of other substantive criterion, as the potentially more useful model (Duncan *et al.*, 2006).

The Akaike Information Criterion (AIC) is defined as

$$AIC = -2 \log L(\boldsymbol{\varphi}; \mathbf{y}) + 2d, \quad (22)$$

where d is the number of free parameters and $L(\boldsymbol{\varphi}; \mathbf{y})$ is the maximized value of the likelihood function for the estimated model (Muthén, 2004a). When comparing two competitive models, the best one is the one that has the lower AIC.

2.4.2 Bayesian Information Criterion (BIC)

As an alternative to AIC one can use BIC (Nagin, 1999). For a given model, BIC (Schwarz, 1978) is calculated as follows:

$$BIC = -2 \log L(\boldsymbol{\varphi}; \mathbf{y}) + d \log n, \quad (23)$$

where $L(\boldsymbol{\varphi}; \mathbf{y})$ is the value of the model's maximum likelihood, n is the sample size, and d is the number of parameters in the model (Nagin, 1999).

2.4.3 Adjusted Bayesian Information Criterion (*a*BIC)

The adjusted BIC (Sclove, 1987) replaces the sample size n in the BIC equation (24) with $(n + 2) / 24$, resulting in (Nylund *et al.*, 2007):

$$aBIC = -2 \log L(\boldsymbol{\varphi}; \mathbf{y}) + d \log[(n + 2) / 24] . \quad (24)$$

2.4.4 Lo-Mendell-Rubin likelihood ratio test – LMR LRT

Lo *et al.* (2001) proposed a likelihood ratio-based method for testing $K - 1$ classes versus K classes. The Lo–Mendell–Rubin likelihood ratio test (LMR LRT) *avoids* the classical problem of chi-square testing based on likelihood ratios. This concerns models that are nested but where the more restricted model is obtained from the less-restricted model by a parameter assuming a value on the border of the admissible parameter space, in the present case a latent class probability being zero. It is known that such likelihood ratios do not follow a chi-square distribution. LMR considers the same likelihood ratio but derives its correct distribution. A low p value indicates that the $(K-1)$ class model has to be rejected in favor of a model with at least K classes. The LMR LRT procedure was implemented in Mplus (Muthén and Muthén, 2001).

$$\begin{aligned} H_0 &: \text{number of classes is } K-1 \\ H_a &: \text{number of classes is } K. \end{aligned} \tag{25}$$

Lo *et al.* (2001) proposed that the above described Lo–Mendell–Rubin likelihood ratio test should adjust with the numbers of d_K and d_{K-1} of freely estimated parameters in K and $K-1$ classes, respectively, and sample size.

The adjusted test, called the LMR test, is then

$$\text{LMR} = \frac{\text{VLMR}}{1 + \{(d_K - d_{K-1}) \log n\}^{-1}}, \tag{26}$$

where

$$\text{VLMR} = 2 \sum_{i=1}^n \log \frac{f(X_i, \boldsymbol{\varphi}_K)}{g(X_i, \boldsymbol{\varphi}_{K-1})} \tag{27}$$

and, where $f(\cdot)$ and $g(\cdot)$ define the density function under K and K - component normal mixture density, respectively. The LMR test follows a weighted chi-squared distribution.

2.5 Synthetic Data Example

In this section the latent growth mixture model is applied to simulated data sets to demonstrate the framework delineated before. The observed data are composed by five measures, measured at five equally spaced time points (y_1 - y_5). This latent growth model includes two latent growth factors: intercept (η_I) and slope (η_S). x_1 and x_2 represent regression factors for intercept and slope, and w_1 is a covariate. The values of x_1 , x_2 , and w_1 were sampled from the standard normal: $N(0,1)$. This model was simulated in MATLAB for four data sets ($n=500$, $n=1000$, $n=5000$ and $n=10000$). The simulated data sets are estimated in Mplus (Muthén and Muthén, 1998-2007). This small simulation study illustrates the impact of sample size on parameter recovery. An LGMM is estimated for a two group solution.

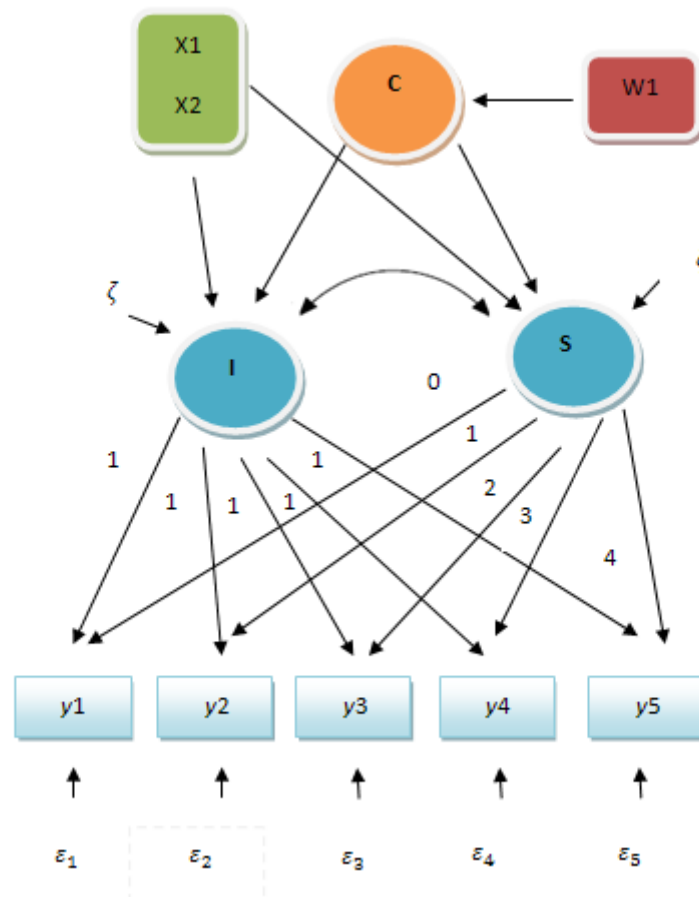


Figure 2 – Representation of the model being simulated.

The starting values of the parameter estimates were set at zero. The Mplus file used for the estimating of data set with $n = 500$ is given in Appendix A.1. We can see in Tables 1 and 2 the effect of sample size on parameter estimates.

Table 1 – Parameter estimates for latent class 1 (simulated data set).

Parameter	Estimates				True Values
	$n=500$	$n=1000$	$n=5000$	$n=10000$	
$\hat{\alpha}_{\eta_I}$	1.024	1.007	1.028	1.004	1.0
$\hat{\alpha}_{\eta_S}$	0.478	0.515	0.473	0.473	0.5
$\hat{\gamma}_{\eta_I x_1}$	1.974	2.014	2.000	2.018	2.0
$\hat{\gamma}_{\eta_I x_2}$	2.876	2.990	3.024	3.009	3.0
$\hat{\gamma}_{\eta_S x_1}$	0.270	0.270	0.296	0.297	0.3
$\hat{\gamma}_{\eta_S x_2}$	0.390	0.364	0.370	0.390	0.4
$\hat{\delta}_c w_1$	-0.308	-0.057	-0.001	-0.021	0.0
$\hat{\psi}_{\eta_I}$	0.791	0.878	0.985	0.961	1.0
$\hat{\psi}_{\eta_S}$	0.908	0.947	0.997	0.998	1.0
$\hat{\psi}_{\eta_I \eta_S}$	-0.242	-0.341	-0.484	-0.503	-0.5
$\hat{\theta}_1$	1.161	0.941	0.959	0.976	1.0
$\hat{\theta}_2$	1.123	0.992	0.976	0.984	1.0
$\hat{\theta}_3$	1.082	0.887	0.970	1.003	1.0
$\hat{\theta}_4$	0.707	1.138	1.083	0.942	1.0
$\hat{\theta}_5$	1.365	0.872	0.977	1.022	1.0

Table 2 – Parameter estimates for latent class 2 (simulated data set).

Parameter	Estimates				True Values
	$n=500$	$n=1000$	$n=5000$	$n=10000$	
$\hat{\alpha}_{\eta_I}$	-1.114	-0.948	-0.956	-0.996	-1.0
$\hat{\alpha}_{\eta_S}$	-1.476	-1.610	-1.573	-1.527	-1.5
$\hat{\gamma}_{\eta_I x_1}$	-2.224	-1.978	-2.011	-2.005	-2.0
$\hat{\gamma}_{\eta_I x_2}$	-2.974	-2.930	-2.957	-3.005	-3.0
$\hat{\gamma}_{\eta_S x_1}$	-0.282	-0.320	-0.284	-0.283	-0.3
$\hat{\gamma}_{\eta_S x_2}$	-0.453	-0.347	-0.433	-0.361	-0.4
$\hat{\delta}_c w_1$	-0.308	-0.057	-0.001	-0.021	0.0
$\hat{\psi}_{\eta_I}$	4.999	4.919	5.243	4.877	5.0
$\hat{\psi}_{\eta_S}$	5.648	4.863	4.968	5.166	5.0
$\hat{\psi}_{\eta_I \eta_S}$	0.072	0.719	0.512	0.537	0.5
$\hat{\theta}_1$	0.612	0.974	1.035	0.970	1.0
$\hat{\theta}_2$	1.097	1.012	1.004	1.005	1.0
$\hat{\theta}_3$	0.937	0.992	1.031	1.029	1.0
$\hat{\theta}_4$	1.218	1.235	1.001	0.915	1.0
$\hat{\theta}_5$	1.197	0.610	0.898	1.071	1.0

It was possible to retrieve the cluster structure (discrete latent variable) and the growth structure (continuous latent variables). As expected, increasing sample size improves the approximation to the true values.

3 Empirical Study

3.1 *Introduction*

The main goal of this empirical study is to analyze the shape of the learning process in pilot training and to identify the variables that can explain the pattern of growth. As not all the applicants adapt themselves to the demands of the aeronautic environment, we allow different patterns of learning, such as different levels of abilities (initial status of performance).

Psychological tests have been used in military pilot selection since the beginning of the 19th century (Martinussen, 2003). World War I, and especially World War II, were creative periods for research on both assessment program construction, personnel selection, and test and program validation. The search for better predictors of success in pilot training is probably a function of the high costs related to personnel training, human suffering from failure to complete, and the need for recruiting competent and well-suited people. Burke and Hunter (1995) divide the measures used in the pilot selection in four different categories: the cognitive measures, the personality tests, the psychomotor measures, and the measures of “job sample”. On the basis of 50 studies in the pilots selection field, the psychomotor measures were pointed out as the best predictors of the performance in flight. In a recent study, Kokorian (2008) looked at levels of prediction of simulator performance where simulators are used as one of the final criteria for acceptance into a training programme. The candidates had been previously assessed using the PILAPT system, and psychomotor ability and spatial aptitude were found to offer significant and practical predictions of performance at the simulator assessment stage.

The high importance of interactive /social skills is clearly reflected by the results of the job requirement studies of pilots. Most remarkable was that social/interactive capabilities seem to be as important for a successful pilot career as mental and psychomotor capabilities. Factors of interpersonal competence have been of growing importance over the last decades. The technical and procedural aspects have been more and more complemented by soft-factors like communication, teamwork, or situation awareness. This applies both to civil and military aviation, and certain aspects of personality nearly always have a decisive effect on achievement. The best predictors for initial training success were emotional measurement

scales (“Emotional Instability” as negative predictors”), and interpersonal scales (“Extraversion” and “Vitality” as positive predictors). Comparable results were also observed in a sample of Spanish flight students (Maschke, 2004). Stokes and Bohan (1995) stated that there is some evidence that personality variables, in particularly tension/anxiety, may be able to assist in discriminating between students who will pass or fail initial flight training.

3.2 *Hypotheses*

As seen before, previous studies concerning pilot performance in PoaF (Bártolo-Ribeiro *et al.*, 2004; Gomes *et al.*, 2008) identified key aptitudes as predictors of flight performance, such as psychomotor coordination (SMA) and spatial aptitude (INS2 and Hands). This research sets the following hypotheses:

Hypothesis 1: There is heterogeneity in the population, i.e., there exist different trajectories of growth and patterns of learning;

Hypothesis 2: Covariates from perceptive-cognitive and psychomotor dimension - INS2 (spatial aptitude), SMA (psychomotor coordination), and Hands (spatial aptitude), are predictors of latent growth curve, i.e., they can predict the intercept (initial status) and the slope (development over time):

H 2.1: Spatial aptitude (INS2) predicts the latent growth curve, i.e., predicts the intercept (initial status) and the slope (development over time);

H 2.2: Spatial aptitude (HANDS) predicts the latent growth curve, i.e., predicts the intercept (initial status) and the slope (development over time);

H 2.3: Psychomotor coordination (SMA) predicts the latent growth curve, i.e., predicts the intercept (initial status) and the slope (development over time);

Hypothesis 3: Covariate from personality and motivation dimension – *General Adaptability* predicts the slope (development over time) as a result of the adaptation to the demands of the aeronautic environment;

Hypothesis 4: Covariates predict cluster membership, in particular:

H 4.1: Spatial aptitude (INS2) predicts the cluster membership;

H 4.2: Spatial aptitude (HANDS) predicts the cluster membership;

H 4.3: Psychomotor coordination (SMA) predicts the cluster membership;

H 4.4: *General Adaptability* predicts the cluster membership.

3.3 *Data*

3.3.1 **Sample**

The selection process of pilot applicants to the Portuguese Air Force Academy is expensive and is composed by four evaluation stages (psychological, medical, physical and the final screening). These phases are sequential and eliminatory.

The sample consists of 297 applicants who had passed the psychological, medical and physical phases for the incorporation as Pilot Aviator (PILAV) in the Portuguese Air Force Academy (PoAFA), and flew the seven required flights (the first flight is experimental) of the Flight Screening (FS) from 2001 to 2008. The mean age is 18.4 years ($SD = 1.3$; $min = 17$; $max = 26$) and 98% were male applicants.

3.3.2 **Flight Screening (FS)**

The FS is the final stage of the selection process and it eliminates the applicants who cannot adapt themselves to the demands of the aeronautic environment. It takes place at the Air Force Academy for two weeks, and it encompasses two days of lectures with an eliminatory test and a flight performance assessment in seven flight missions. The first flight mission is for demonstration. The others six flight missions are evaluated by the flight instructor in a four point ordinal scale: Low (1), Satisfactory (2), Medium (3) and High (4).

Table 3 presents the number of applicants that reached the final stage (number of applicant per year), and also the final outcome (Pass/Fail). During this period, 80.5% of applicants successfully accomplish a positive classification in the Flight Screening stage.

Table 3 – Applicants per year and results of the FS.

		Pass	Fail	Total Year
	2001	27	12	39
	2002	20	8	28
	2003	39	10	49
Sample year	2004	31	6	37
	2005	27	5	32
	2006	29	9	38
	2007	26	5	31
	2008	39	4	43
	Total	239	58	297
	Total %	80.5%	19.5%	100%

The original measures correspond to the score (in a four point ordinal scale) of each flight mission (flights one to six). Figure 3 shows the results per flight (%), demonstrating that the score 2 (Low) is the most frequent one (see also Table 4).

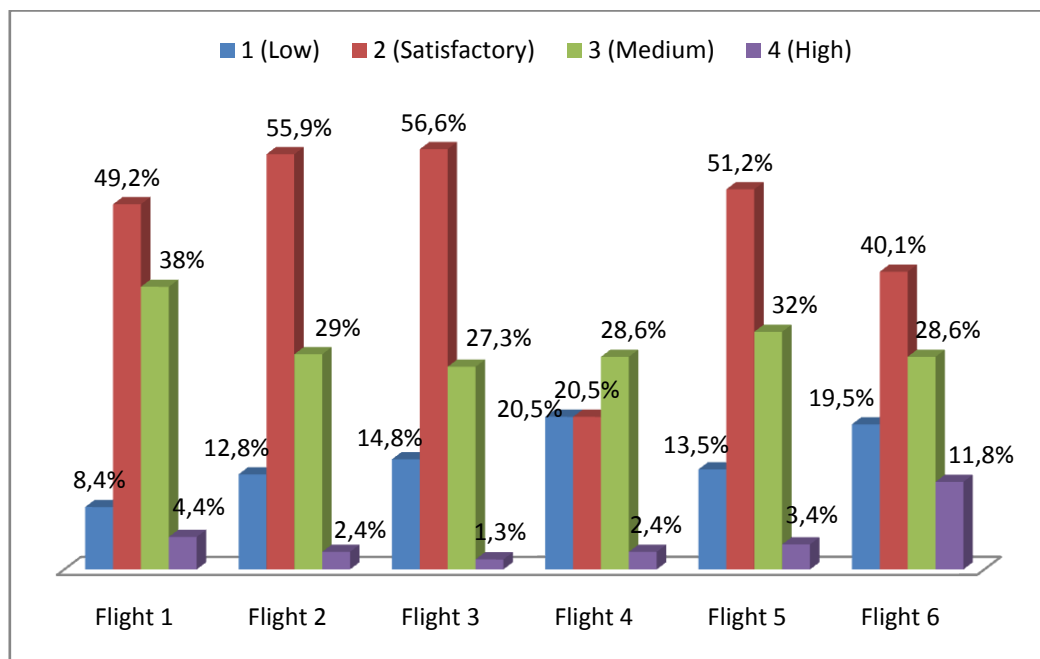


Figure 3 – Scores obtained in the selection flights.

Table 4 – Flight scores frequencies.

	Low (1)		Satisfactory (2)		Medium (3)		High (4)		Total	
	N	%	N	%	N	%	N	%	N	%
Flight 1	25	8.4%	146	49.2%	113	38%	13	4.4%	297	100%
Flight 2	38	12.8%	166	55.9%	86	29%	7	2.4%	297	100%
Flight 3	44	14.8%	168	56.6%	81	27.3%	4	1.3%	297	100%
Flight 4	61	20.5%	144	48.5%	85	28.6%	7	2.4%	297	100%
Flight 5	40	13.3%	152	51.2%	95	32%	10	3.4%	297	100%
Flight 6	58	19.5%	119	40.1%	85	28.6%	35	11.8%	297	100%

As we can see in Figure 3 and Table 4, category 4 (High) has few observations so the original scale was aggregated into three categories (see Figure 4 and Table 5): Low (1), Satisfactory (2), and Medium/High (3).

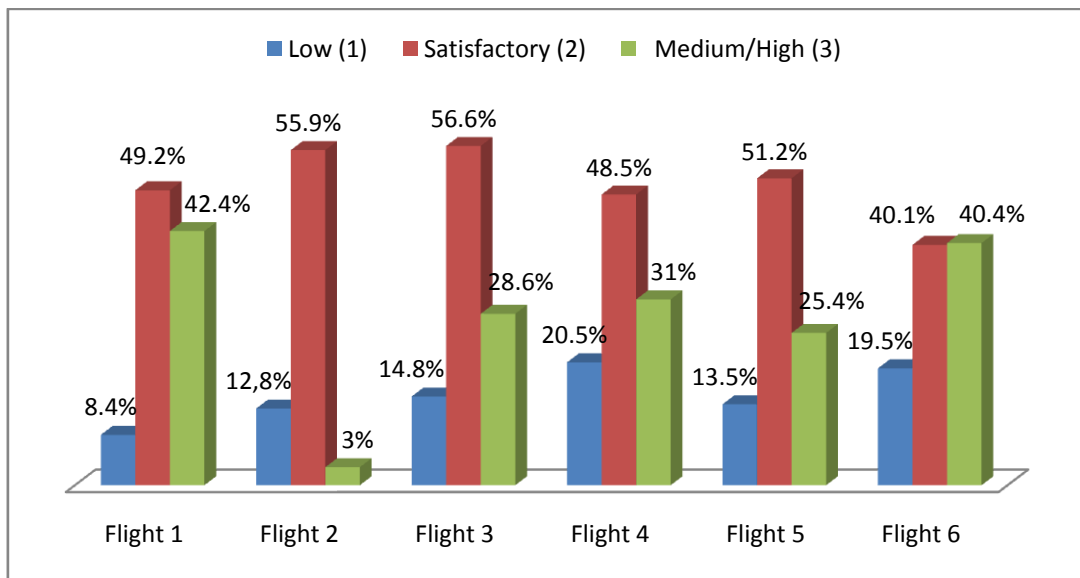


Figure 4 – Recoded scores obtained in the selection flights.

Table 5 – Recoded flight scores frequencies.

	Low (1)		Satisfactory (2)		Medium/High (3)		Total	
	N	%	N	%	N	%	N	%
Flight 1	25	8.4%	146	49.2%	126	4.4%	297	100%
Flight 2	38	12.8%	166	55.9%	93	31.3%	297	100%
Flight 3	44	14.8%	168	56.6%	85	28.6%	297	100%
Flight 4	61	20.5%	144	48.5%	92	31%	297	100%
Flight 5	40	13.5%	152	51.2%	105	35.4%	297	100%
Flight 6	58	19.5%	119	40.1%	120	40.4%	297	100%

We can see (Figure 4) that data proportions for *Low* grade (coded as 1) increased from time measurement 1 (flight 1) to time measurement 4 (flight 4). This can be due to the increasing operations complexity required from the pilot applicants during the flights. Or, in another way, it may suggest that if one has a good grade in a first flight, the tendency is to obtain worst results in the subsequent flights.

In what refers to category 2 – *Satisfactory* grade, the applicants are considered to have a reasonable performance at the flights. This is considered a positive evaluation, they still perform some mistakes in their flights. There exists an increasing tendency from flight 1 to flight 3, and after flight 5 the data proportions for this category decrease.

At last, we can see that the data proportion of pilot applicants who performed Medium/High grades decreases from flight 1 to flight 3, and increases from flight 3 onwards.

Looking at Figure 4, we can assume that category 1 (Low) and 3 (medium/high) have an increasing tendency throughout the six assessments in opposition to category 2 (satisfactory) that has a decreasing tendency.

Table 6 – Summary of categorical data proportions.

	Flight 1	Flight 2	Flight 3	Flight 4	Flight 5	Flight 6
Category 1	0.084	0.128	0.148	0.205	0.135	0.195
Category 2	0.492	0.559	0.566	0.485	0.512	0.401
Category 3	0.424	0.313	0.286	0.310	0.354	0.404

Table 6 presents the summary of categorical data proportions. It is possible to see that the category 2 is the most frequent one for all flights, with exception of flight 6.

3.3.3 Covariates

The psychological evaluation phase to select all the candidates who join the requisites is defined in the psychological profile for the PILAV (Pilot Aviator). The psychological evaluation is based on a traditional selection procedure that consists in the application of tests, group dynamics and psychological interview.

The psychological evaluation intends to evaluate three main psychological dimensions:

- a) Perceptive-cognitive dimension: evaluation of cognitive-perceptive aptitudes, through computerized tests standardized for the population in study;
- b) Psychomotor dimension: evaluation of psychomotor aptitudes through computerized tests;
- c) Personality and motivation dimension: evaluation based on simultaneously or separately use from the personality questionnaires, projective tests, group dynamics and individual interviews.

3.3.3.1 *Personality and Motivation Dimension*

The evaluation of this dimension is based on personality questionnaires, projective tests, group dynamics and psychological interviews. This dimension is operationalized in the classification of 12 factors (in a five-point scale) described in Table 7.

Table 7 – Personality and Motivation Dimensions.

Factor	Description
Contact and Presentation	<ul style="list-style-type: none"> - Courtesy, respect and education - Availability, attention and sense of opportunity - Poise and hygiene - Clothing and appropriate "accessories"
Communication	<ul style="list-style-type: none"> - Attitude and posture in favor of communication - Fluent and structured speech - Good elocution (diction, volume ...) - Conviction and persuasion
Interpersonal relationships and teamwork	<ul style="list-style-type: none"> - Empathy, assertive - Good listener, sensitive, flexible
Maturity	<ul style="list-style-type: none"> - Capacity for reflection and preparation before the situations - Adaptability and flexibility - Internal locus of control, based on a set of principles and values
Emotional stability	<ul style="list-style-type: none"> - Stability in the emotional-affective dimension - Self-control in situations of pressure, or stress evaluation
Activity level	<ul style="list-style-type: none"> - High general activity - Diverse interests and activities - Good occupation of leisure
Initiative	<ul style="list-style-type: none"> - Strong presence - Conviction and security in defense of their opinions - Initiative
Persistence and combativeness	<ul style="list-style-type: none"> - Persistence in pursuit of its objectives and activities - Proactive behavior - Pragmatic commitment and investment in the emotional tasks (ability to delay the reward)
Decision and practical intelligence	<ul style="list-style-type: none"> - Determined and decided - Pragmatic - Objectivity in the trial and mental agility - Good ability to react to pressure
Leadership	<ul style="list-style-type: none"> - Determination, initiative and persuasion - It is followed by others - Ability to coordinate and guide a group - Team spirit - Available and flexible to the other
Habits and attitudes	<ul style="list-style-type: none"> - Habits and positive attitudes towards: military status and situation of war, military commitments - Good attitude towards study - Good occupation of leisure - Regular physical activity
Motivation	<ul style="list-style-type: none"> - Information about the PoAfA - Enthusiasm and conviction.

Based on all these items the applicant obtains a global score – *General Adaptability* - that ranges from 1 to 20 (1-3: Unfavorable; 4-7: Strong reservations; 8-9: Acceptable with reservations; 10-11: Acceptable with minor reservations; 12-14: Acceptable; 15-17:

Favorable; and 18-20: Very positive). This score measures the personality and motivation dimension and aims to prognostic the adaptation of the individuals. Table 8 depicts the descriptive statistics for *General Adaptability*. The mean is 11.24 (SD = 2.38; min = 9; max = 20).

Table 8 – Descriptive statistics for General Adaptability.

	N	Minimum	Maximum	Mean	Std. Deviation
General Adaptability	297	9.00	20.00	11.24	2.38

3.3.3.2 *Perceptive-cognitive and Psychomotor Dimension*

The perceptive-cognitive and psychomotor dimensions were assessed at the Psychometric Laboratory of the Portuguese Air Force in a total of 13 tests. Variables included in this study were based on two rules: (a) showed no missing values; (b) proved to be predictors of flight performance in previous validity studies (Bártolo-Ribeiro *et al.*, 2004). Hence, the variables included in this study were INS2 (perceptive-cognitive dimension), SMA (psychomotor dimension), and Hands (perceptive-cognitive dimension). The descriptive statistics are depicted in Table 9.

- a) Hands: This test requires the candidate to use an audio search message to scan visual objects and declare how many of those objects meet that search message. This task requires the translation of verbal information into visual information, and while correlated with spatial tests, is essentially a basic working memory task testing how quickly someone can move from verbal to visual information and make an accurate decision (Figure 5).

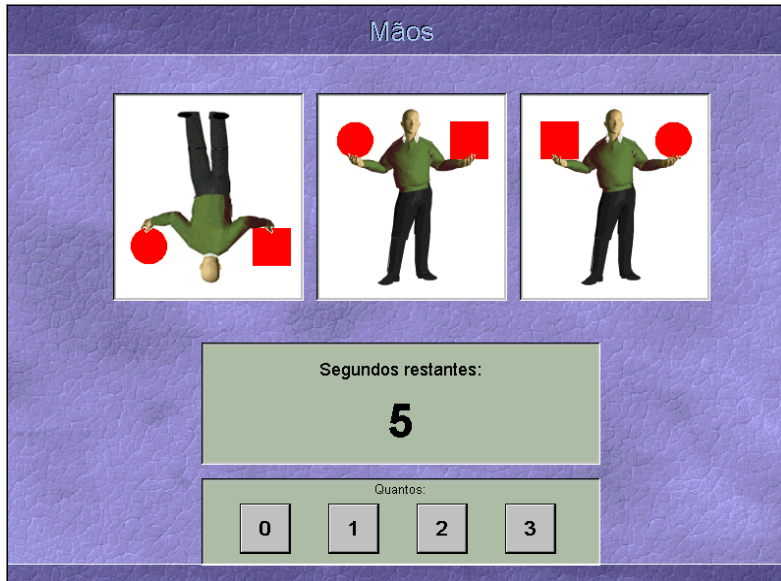


Figure 5 – Hands (spatial aptitude test).

- b) *Instruments Interpretation* (INSB1 and INSB2): Spatial aptitude, information processing; this is a test of the candidate’s spatial visualization ability using spatial, numerical, and verbal information. The test is based on the aviation board panel (Figure 6).



Figure 6 – Instruments interpretation.

Part 1 presents six aircraft instruments (altimeter, artificial horizon, airspeed, vertical speed, compass, and turn & bank) on the top half of the screen, whilst in the bottom half there are five verbal descriptions about the aircraft's orientation. The candidate must inspect the instrument readings and select the description that corresponds to the readings.

Part 2 presents line drawings of aircraft in five different orientations and two of the instruments used in Part 1 (artificial horizon and compass). The candidate must inspect the instrument readings and identify which of the five aircraft orientations accurately corresponds to the instrument readings.

- c) SMA – *Sensory Motor Apparatus*: multi-limbic motor coordination; SMA is a compensatory tracking test that measures eye-hand-foot coordination (psychomotor ability). The candidate uses a joystick and rudder pedals to move a pointer both horizontally and vertically on a visual display unit. Performance on the test is assessed using an error score (Figure 7).



Figure 7 – Sensory Motor Apparatus.

Table 9 depicts the descriptive statistics for the three variables (perceptive-cognitive and psychomotor dimension). For each test is presented the mean, standard deviation and the minimum and maximum value obtained at each test.

Table 9 – Descriptive statistics.

	N	Minimum	Maximum	Mean	Std. Deviation
Hands	297	4,00	133,00	57,02	27,50
SMA	297	27,00	255,00	133,08	42,60
INS2	297	17,00	59,00	45,80	9,10

3.4 *Empirical Results*

The analysis will be conducted in four main steps: a first part, where homogeneity in population is assumed; a second part, where population heterogeneity is checked, and the appropriate number of latent classes is selected; a third part, where the previous selected 2-class unconditional mixture model is analyzed; and at last step, where the conditional 2-class mixture model (including predictors) is analyzed.

3.4.1 **Single-class growth curve model – Assuming homogeneity**

This specific analysis assumes that all individuals belong to the same population. As recommended in latent growth mixture modeling (Duncan *et al.*, 2002), a step-by-step approach will be followed: first the unconditional analyses with no predictor variables for each class, and then the conditional analyses (taking into account the influence of predictor variables). Figure 8 represents the single-class mixture model. Input syntax for the single class mixture model is presented in Appendix B (Table B.1).

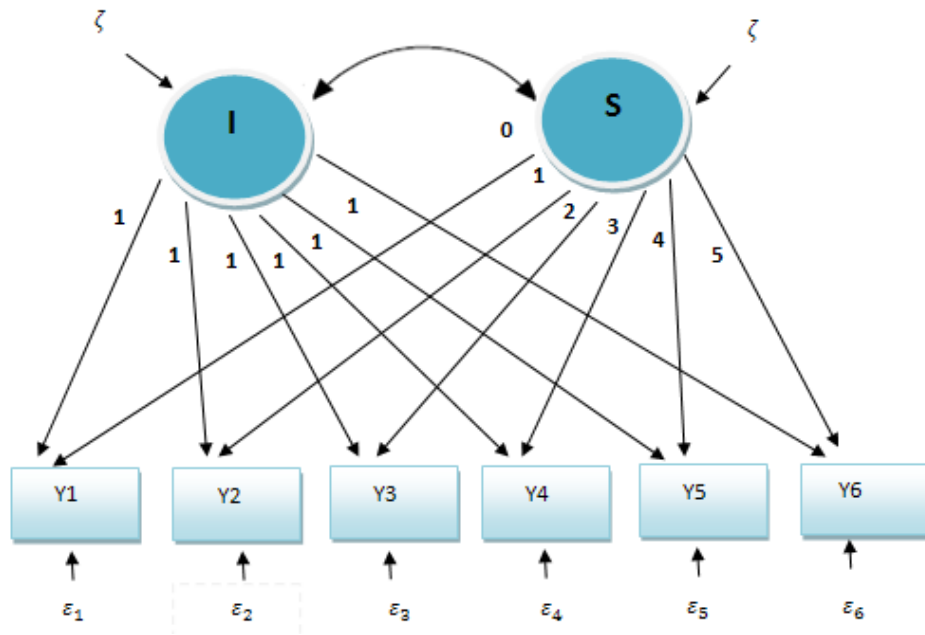


Figure 8 – Representation of the single class mixture model.

Model fitting procedures for the single-class mixture model resulted in a log likelihood H_0 value of -1583.419, a BIC of 3201.000, an aBIC of 3181.971, and an AIC value of 3178.837.

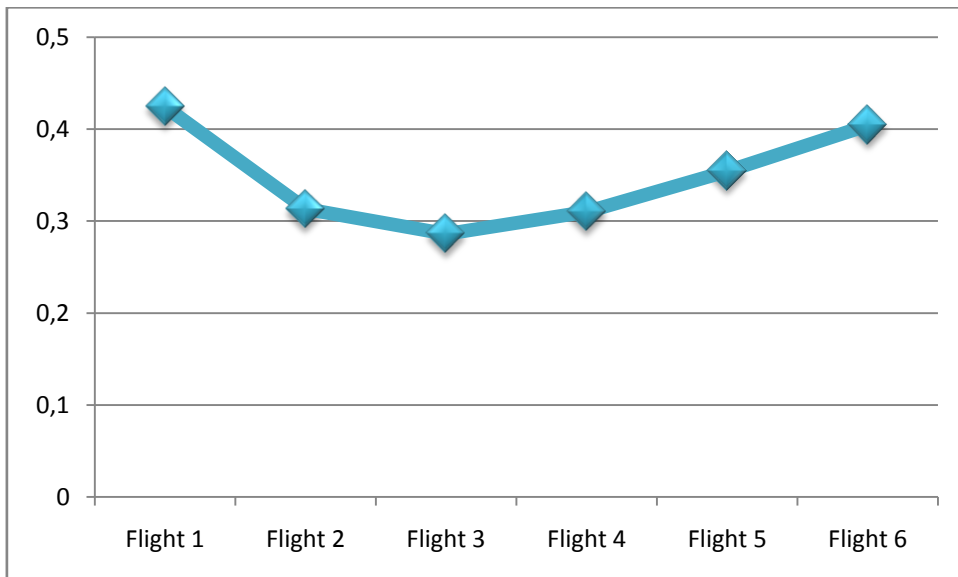


Figure 9 – Growth curve for a single latent class model.

Figure 9 presents the single class model trend, showing a decreasing tendency from flight 1 to flight 3 and an increasing tendency from flight 3 onwards.

Table 10 – Parameter estimates – single latent class.

Parameter	Estimates	S.E.	Estimates/S.E.
$\hat{\alpha}_{\eta_I}$	-	-	-
$\hat{\alpha}_S$	-0.059	0.039	1.510
$\hat{\psi}_{\eta_I}$	2.857	0.506	5.647
$\hat{\psi}_{\eta_S}$	0.163	0.045	3.614
$\hat{\psi}_{\eta_I\eta_S}$	-0.173	0.108	1.603
Thresholds			
$\tau_{1,1}$	-2.804	0.172	16.289
$\tau_{1,2}$	0.843	0.142	5.922

Table 10 reports the coefficients (estimates), their standard errors, and the z-test (Est./S.E.). If $|z|$ (|est./S.E.|) is greater than 1.96 the coefficient is statistically significant at the 0.05 level (for a two-tailed test), for a p-value of less than 0.01, $|z| > 2.75$ is necessary (Academic Technology Services, Statistical Consulting Group, 2006; Muthén and Satorra, 1995).

The estimates for the intercept mean is set at zero by default ($\hat{\alpha}_{\eta_I}$), and the slope mean was not significant at 0.05 level ($\hat{\alpha}_{\eta_S} = -0.059$, $|z| = 1.510$), indicating that no significant growth occurred over the six flights. The estimates for the intercept and slope variance were both significant at 0.01 level ($\hat{\psi}_{\eta_I} = 2.857$, $|z| = 5.647$, $\hat{\psi}_{\eta_S} = 0.163$, $|z| = 3.614$). This indicates substantial variation among individuals in initial status (first evaluated flight), and substantial variation among individuals referring to the rate change of flight performance. The covariance between intercept and slope factors was not significant at 0.05 level ($\hat{\psi}_{\eta_I\eta_S} = -0.173$, $|z| = 1.603$), indicating the inexistence of a strong relation between the latent factors.

3.4.2 Checking population heterogeneity

To assess whether there exists population heterogeneity, the first step is to test whether a model with two classes (heterogeneity) would perform better than a model with just one latent class (homogeneity).

The Vuong-Lo-Mendell-Rubin test tests the model with K classes (in our case 2) against a model with (K-1) classes (in our case, K - 1 = 1 class).

Table 11 – Vuong-Lo-Mendell-Rubin likelihood ratio test for 1 (H0) versus 2 classes

Vuong-Lo-Mendell-Rubin likelihood ratio	
H_0 Log-likelihood Value	-1583.417
2 Times the Log-likelihood Difference	9.181
Difference in the number of parameters	3
Mean	1.259
Standard Deviation	1.705
p -value	0.0026

The Vuong-Lo-Mendell-Rubin test and the Lo-Mendell-Rubin adjusted LRT test have a p -value of 0.0026 and 0.0035, respectively (Table 11 and Table 12).

There is a difference in the number of parameters known and the number of parameters unknown. The difference in the number of parameters is three because all parameters are held constant across latent classes with the exception of three parameters: $\hat{\alpha}_{\eta_I}$ (latent class 1), $\hat{\alpha}_{\eta_S}$ (latent class 1) and $\hat{\alpha}_{\eta_S}$ (latent class 2). Those tests suggest that a two-class solution is better than one latent class solution, indicating population heterogeneity ($p \leq 0.01$).

Table 12 – Lo-Mendell-Rubin Adjusted LRT test.

Value	P-value
8.673	0.0035

We conclude that there exists population heterogeneity, and therefore the single model is not appropriate in this case. We confirm our Hypothesis 1 that postulates population heterogeneity, and different trajectories growth.

3.4.3 Selecting the appropriate number of latent classes

After determining the existence of sample heterogeneity one needs to choose the appropriate number of latent classes. To select the appropriate number of latent classes, AIC and BIC are more appropriate than log-likelihood value, for models with different number of free

parameters. The sample size adjusted BIC is referred in literature as superior to BIC (Muthén, 2003; Duncan *et al.*, 2002). All model selection criteria are presented in Table 13.

Table 13 – Model selection criteria.

	1 class	2 classes	3 classes	4 classes
AIC	3178.837	3175.653	3174.148	3180.949
BIC	3201.000	3208.896	3218.473	3236.355
aBIC	3181.971	3180.354	3180.417	3188.785

Through the analyzes of Table 13 and concerning that aBIC has been shown to give superior performance in a simulation study (Muthén, 2003) we suggest that the best fitting model should be the 2-class model (aBIC). Moreover, AIC tends to overfit data (Dias, 2004).

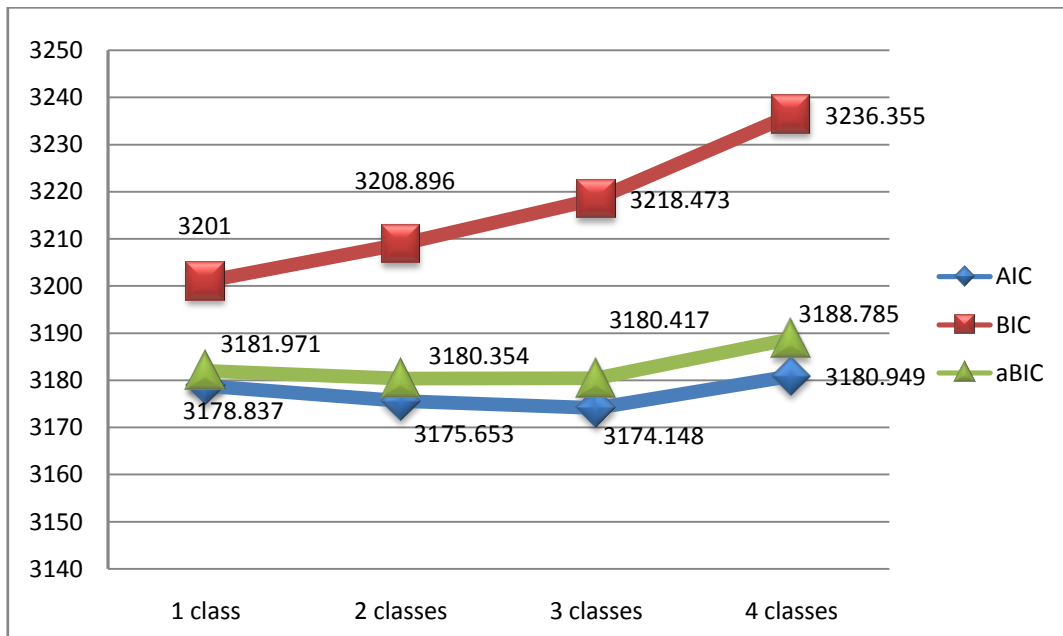


Figure 10 – Model assessment.

From Figure 10 we conclude that the 2-class model is the most parsimonious model as aBIC presents the minimum value for 2 latent classes.

Table 14 – Vuong-Lo-Mendell-Rubin likelihood ratio test for 2 (H0) versus 3 classes

H_0 Log-likelihood Value	-1578.826
2 Times the Log-likelihood Difference	7.504
Difference in the number of parameters	3
Mean	6.860
Standard Deviation	8.820
p -value	0.3025

The Vuong-Lo-Mendell-Rubin and Lo-Mendell-Rubin adjusted LRT tests of 2 latent classes against 3 the p -value is 0.3025 and 0.3187, respectively (Table 14 and Table 15). Those tests suggest that the 2-class solution is better than a 3-class solution ($p \geq 0.05$).

Table 15 – Lo-Mendell-Rubin Adjusted LRT test (2 versus 3 classes)

Value	P-value
7.089	0.3187

3.4.4 Two-class Mixture Model

In the parameterization of the latent growth model depicted here, the thresholds of the outcome variable at the six time points are held constant across latent classes as the default. The intercept of the growth factor is fixed at zero in the last class and is free to be estimated in the first class. Input statement for the two-class mixture model is presented in Appendix B (Table B.2).

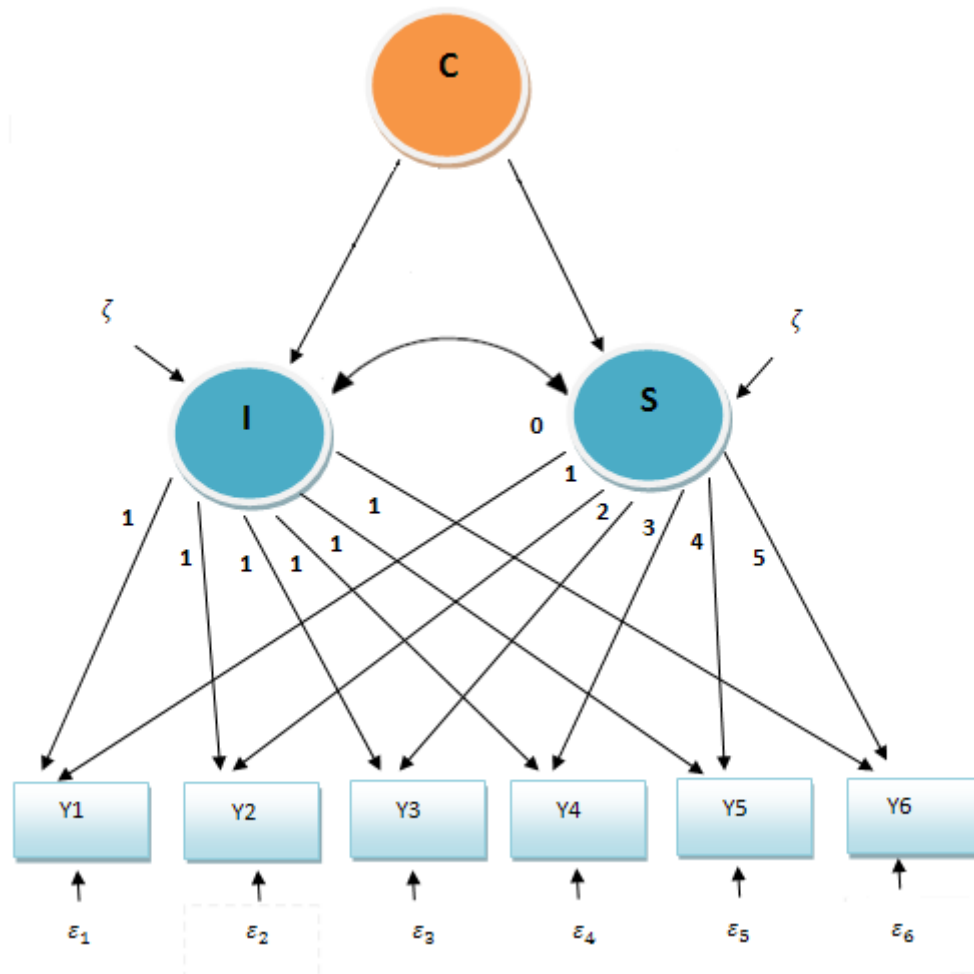


Figure 11 – Two-class mixture model.

Model fitting procedures for the two-class model with latent growth factors, resulted in a log-likelihood H_0 value of -1578.826, an aBIC of 3180.354, a BIC of 3208.896 and an AIC value of 3175.653.

Table 16 presents the final class proportions and classifications. The class proportions based on the estimated model and the class proportions based on the estimated posterior probabilities are very similar: the latent class 1 and latent 2 class contain 182 individuals (61%) and 115 individuals (39%), respectively. Using the classification of the individuals based on their most likely latent class membership there is no relevant difference, and there are 6 individuals that pass from class 2 to latent class 1.

Table 16 – Final latent class proportions and classification.

	Class proportions based on estimated model	Class proportions based on the estimated posterior probabilities	Classification based on their most likely latent class membership
Class 1	182 61%	182 61%	188 63%
Class 2	115 39%	115 39%	109 37%

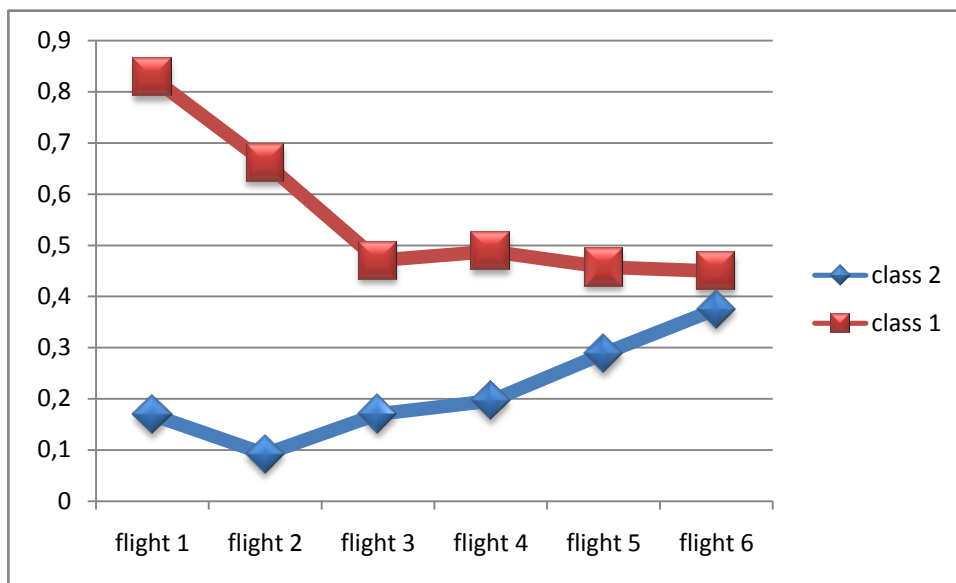


Figure 12 – Growth curve for each of the two latent class membership.

From Figure 12, we can conclude that latent class 2 shows an increasing growing tendency; while the latent class 1 presents a decreasing tendency in the growth curve (it decreases from flight 1 to flight 3 and stabilizes from flight 3 onwards). We can conclude that there exist two heterogeneous populations: latent class 1 presents a superior initial flight performance, and tends to stabilize, while latent class 2 depicted the worst initial performance with an increasing tendency.

Table 17 – Parameter estimates (unconditional model).

Parameter	Estimates	S.E	Estimates/S.E
Latent class 1			
$\hat{\alpha}_{\eta_I}$	-3.304	0.294	11.256
$\hat{\alpha}_{\eta_S}$	0.151	0.057	2.658
Latent class 2			
$\hat{\alpha}_{\eta_I}$	-	-	-
$\hat{\alpha}_{\eta_S}$	-0.382	0.127	3.011
Variances & covariances			
$\hat{\psi}_{\eta_I}$	0.408	0.191	2.141
$\hat{\psi}_{\eta_S}$	0.107	0.024	4.513
$\hat{\psi}_{\eta_I\eta_S}$	0.209	0.035	5.912
Thresholds			
τ_1	-4.824	0.337	14.297
τ_2	-1.165	0.311	3.749

For latent class 1 (Table 17) the estimate for the intercept mean was significant at 0.01 level ($\hat{\alpha}_{\eta_I} = 3.304$, $|z| = 11.256$), indicating significant performance in initial status (first evaluated flight), and the slope mean was also significant ($\hat{\alpha}_{\eta_S} = 0.151$, $|z| = 2.658$), indicating that significant growth occurred over the six flights.

For latent class 2, the estimates for the intercept mean was set at zero, and the estimate for the slope mean was significant at 0.01 level ($\hat{\alpha}_{\eta_S} = 0.382$, $|z| = 3.011$), indicating that significant growth occurred over the six flights.

For both classes, estimates for the intercept and slope variances were both significant at 0.01 level ($\hat{\psi}_{\eta_I} = 0.408$, $|z| = 2.141$, $\hat{\psi}_{\eta_S} = 0.107$, $|z| = 4.513$). This indicates substantial variation among individuals in initial status (first evaluated flight) and the presence of interindividual differences in the shape of growth of flight performance. The covariance between intercept and slope factors was significant at 0.01 level ($\hat{\psi}_{\eta_I\eta_S} = 0.209$, $|z| = 5.912$), indicating a significant positive relation between the latent factors, that is, the initial status of the pilot applicant is not independent of the growth occurred during the six assessments. As variances of the growth factors, were all significantly different from 0, we suggest the need to introduce covariates, in such way they could explain variability.

The thresholds shown at the bottom of the table indicate where the latent variable is cut to make the three categories that we observe in data. As we have seen before, variances, covariances and thresholds were held constant across classes as the default.

3.4.5 LGMM with Covariates

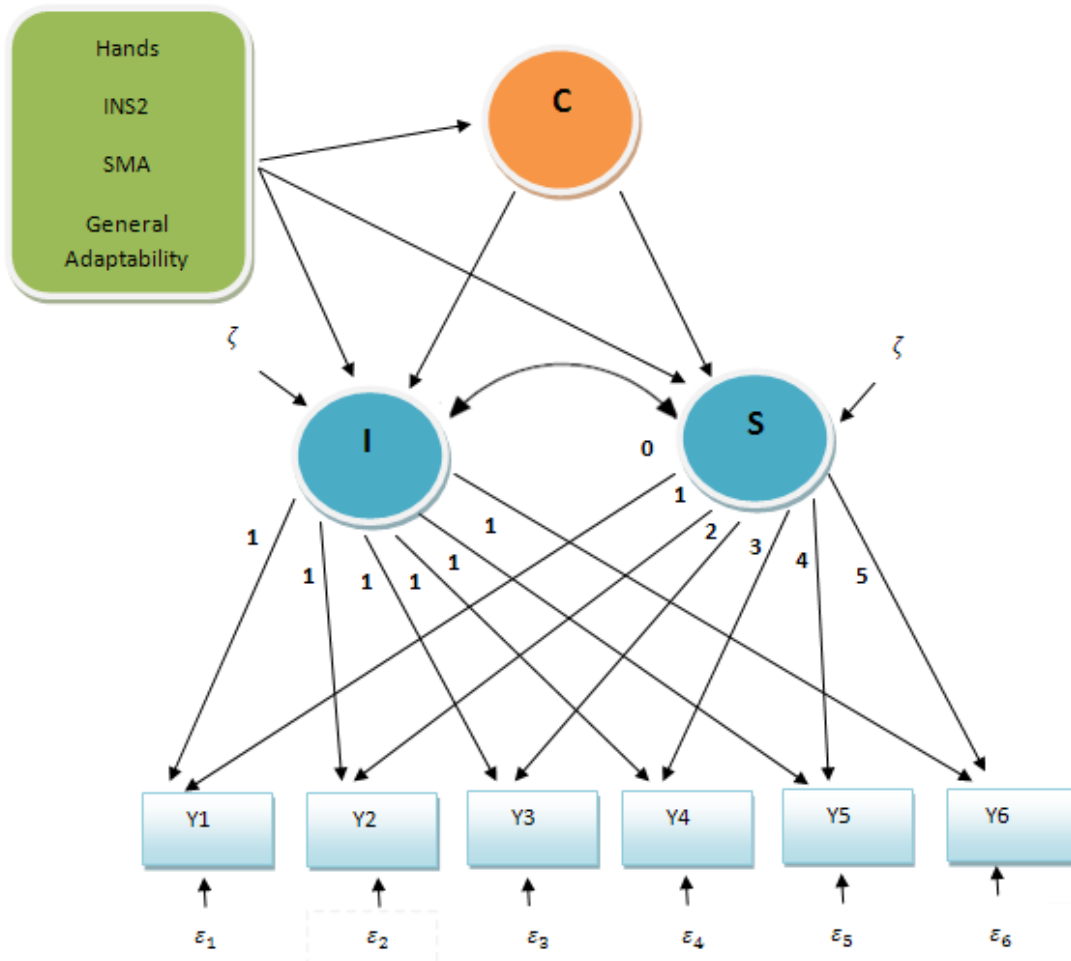


Figure 13 – Two-class mixture model with covariates.

We have seen before that the variances of the growth factors were all significantly different from 0, so we introduce here the covariates, and expect variability reduction, as a result of the covariates impact.

The model postulates the influence of the covariates Hands, SMA, INS2, and General Adaptability on the latent factors intercept and slope, and on C, the latent class variable. Input

statement for the two-class mixture model with covariates is presented in Appendix B (Table B.3).

Model fitting procedures for the two-class model with latent growth factors and covariates, resulted in a log-likelihood H_0 value of -1550.568, an aBIC of 3174.285, a BIC of 3266.254 and an AIC value of 3159.136.

Table 18 – Final latent class proportions and classification (conditional model).

	Class proportions based on estimated model	Class proportions based on the estimated posterior probabilities	Classification based on their most likely latent class membership
Class 1	197	197	213
	66%	66%	72%
Class 2	100	100	84
	34%	34%	28%

Table 18 presents the final class proportions and classifications. The class proportions based on the estimated model and the class proportions based on the estimated posterior probabilities are very similar: the latent class 1 and latent 2 class contain 197 individuals (66%) and 100 individuals (34%), respectively. Using the classification of the individuals based on their most likely latent class membership there are 16 individuals that pass to latent class 1.

Table 19 – Parameter estimates for latent class 1 (conditional model).

Parameter	Estimates	S.E.	Estimates/S.E.
$\hat{\alpha}_{\eta_1}$	12.084	4.811	2.512
$\hat{\alpha}_{\eta_5}$	0.277	1.175	0.235
$\hat{\gamma}_{\eta_1 Hands}$	0.007	0.007	0.883
$\hat{\gamma}_{\eta_1 INS2}$	0.019	0.030	0.620
$\hat{\gamma}_{\eta_1 SMA}$	-0.002	0.005	0.460
$\hat{\gamma}_{\eta_1 Adapt}$	0.407	0.165	2.461
$\hat{\gamma}_{\eta_5 Hands}$	-0.004	0.002	1.578
$\hat{\gamma}_{\eta_5 INS2}$	0.007	0.015	0.438
$\hat{\gamma}_{\eta_5 SMA}$	-0.001	0.004	0.333
$\hat{\gamma}_{\eta_5 Adapt}$	-0.015	0.038	-0.398

For latent class 1 (Table 19) the estimate for the intercept mean was significant at 0.01 level ($\hat{\alpha}_{\eta_I} = 12.084$, $|z| = 2.512$), indicating significant performance in initial status (first evaluated flight), but was not significant for the slope mean ($\hat{\alpha}_{\eta_S} = 0.277$, $|z| = 0.235$), indicating that no significant growth occurred over the six flights (conditional model).

The estimates are given in units of ordered logits, or ordered log odds. The regression coefficients for the covariates showed a significant effect of the covariate *General Adaptability* on the intercept ($\hat{\gamma}_{\eta_I Adapt} = 0.407$, $|z| = 2.641$), meaning a positive effect of the *General Adaptability* in the initial status of the individuals in latent class 1. In this case, we would say that for one unit increase in the *General Adaptability* we expect an increase of 0.407 in the log odds by moving from a given level of the initial performance of the pilot applicants.. The covariates SMA, Hands, and INS2 were not significant for the slope and intercept (Table 19).

Table 20 – Parameter estimates for latent class 2 (conditional model).

Parameter	Estimates	S.E	Estimates/S.E
$\hat{\alpha}_{\eta_I}$	-	-	-
$\hat{\alpha}_{\eta_S}$	1.957	2.171	0.902
$\hat{\gamma}_{\eta_I Hands}$	0.007	0.011	0.608
$\hat{\gamma}_{\eta_I INS2}$	0.060	0.037	1.626
$\hat{\gamma}_{\eta_I SMA}$	-0.016	0.008	2.006
$\hat{\gamma}_{\eta_I Adapt}$	-0.567	0.322	1.762
$\hat{\gamma}_{\eta_S Hands}$	0.006	0.006	0.897
$\hat{\gamma}_{\eta_S INS2}$	-0.013	0.037	0.356
$\hat{\gamma}_{\eta_S SMA}$	0.001	0.009	0.079
$\hat{\gamma}_{\eta_S Adapt}$	0.167	0.066	2.507

For latent class 2 (Table 19) the estimate for the intercept mean was set at zero (to keep the model identified), and the estimate for the slope mean was not significant at 0.05 level ($\hat{\alpha}_{\eta_S} = 1.957$, $|z| = 0.902$), indicating that no significant growth occurred over the six flights (conditional model).

The regression coefficients for the covariates showed a significant effect of the covariate SMA test on the intercept ($\hat{\gamma}_{\eta_I SMA} = -0.016$, $|z| = 2.006$), meaning a negative effect of the

SMA test on the initial status of the individuals in latent class 2. In this case, we can say that SMA test can explain the individual differences in the initial performance of the pilot applicants of latent class 2, and for a unit increase in SMA test, we would expect a 0.016 decrease in the expected value of the initial flight performance in the log odds scale, given that all of the other variables of the model are held constant. The covariate *General Adaptability* showed a significant effect on the slope ($\hat{\gamma}_{\eta_s Adapt} = 0.167$, $|z| = 2.507$). In this case, for one unit increase in *General Adaptability*, we would expect a 0.167 increase in the expected value of the growth performance (slope) in the log odds scale, given that all of the variables of the model are held constant. The covariates Hands and INS2 tests were not significant for the slope and intercept (Table 20).

Table 21 – Parameter estimates (conditional model).

Parameter	Estimates	S.E.	Estimates/S.E.
Variances & covariances			
$\hat{\psi}_{\eta_i}$	0.994	0.837	1.187
$\hat{\psi}_{\eta_s}$	0.076	0.018	4.258
$\hat{\psi}_{\eta_i \eta_s}$	0.075	0.182	0.409
Thresholds			
τ_1	-9.528	4.652	-2.048
τ_2	-5.895	4.645	-1.269

The estimates for the intercept and slope variance show only significant effect at 0.01 level for the slope variance ($\hat{\psi}_{\eta_i} = 0.994$, $|z| = 1.187$, $\hat{\psi}_{\eta_s} = 0.076$, $|z| = 4.258$). This indicates the inexistence of substantial variation among individuals in the initial status (first evaluated flight), but still denotes substantial variation in growth of flight performance (slope). The covariance between intercept and slope factors was not significant at 0.05 level ($\hat{\psi}_{\eta_i \eta_s} = 0.075$, $|z| = 0.409$), indicating that exists no significant relation between the latent factors, i.e., initial status of the pilot applicant is not related with the growth occurred during the six assessments (conditional model).

Table 22 – Parameter estimates for the effects of the covariates on C.

Parameter	Estimates	S.E.	Estimates/S.E.
$\hat{\delta}_C$	0.699	1.783	0.392
$\hat{\delta}_C Hands$	0.002	0.009	0.167
$\hat{\delta}_C Ins2$	0.032	0.000	0.000
$\hat{\delta}_C SMA$	-0.009	0.007	-1.233
$\hat{\delta}_C Adapt$	-0.029	0.189	0.152

The coefficients for the regression of the latent class (C) on the covariates were not significant for all variables (Table 22). In this case, we can say that Hands test (spatial aptitude), INS2 test (spatial aptitude), SMA test (psychomotor coordination) and Psychological Adaptation could not predict the latent class membership of the applicants. But in another way, we saw that SMA test explains the individual differences in the initial status (latent class 2), and that Psychological Adaptation explains the growth performance of the pilot applicants in latent class 2, and the initial status of pilot applicants in latent class 1.

Table 23 – Sample statistics weighted by estimated class probabilities.

	Hands	INS2	SMA	Adapt
class 1	58.92	46.97	126.83	11.32
class 2	53.26	43.43	145.42	11.11

Table 23 depicts sample statistics weighted by estimated class probabilities for class 1 and 2. We saw before that latent class 1 presents a superior initial flight performance, and tends to stabilize, while class 2 depicted the worst initial performance with an increasing tendency. Here we can conclude that class 1 also shows better results in the tests performed in the psychological phase (note that SMA test performance is assessed using an error score, so class 1 presents less errors) than class 2. Note that SMA test depicts a negative effect on the intercept of class 2, for one unit increase in the SMA test (error score) a decrease of 0.016 is expected in the value of the initial flight performance. The latent class 2 depicted the worst results in the SMA test. The class 1 presents higher aptitude proficiency, presenting also a superior flight performance, but they do not improve much from their first flight onwards. Otherwise, examinees in latent class 2 present some difficulties in the first flights, but they adapt themselves to the demands of the aeronautic environment, improving and developing their proficiency, and showing a positive effect of *General Adaptability* on the slope.

3.5 *Discussion and implications of the results*

Some assumptions were made based on previous studies concerning pilot performance (Bártolo-Ribeiro *et al.*; 2004, Gomes *et al.*, 2008; Martinussen, 2003; Maschke, 2004; Stokes and Bohan, 1995). We concluded that not all the applicants were able to adapt themselves to the demands of the aeronautic environment. Therefore, we found different patterns of learning and levels of ability (initial status of performance). In Section 3.2 four hypotheses were postulated: (1) there exists heterogeneity in the population, and different trajectories in growth; (2) Covariates from *perceptive-cognitive* and *psychomotor* dimension - INS2 test (spatial aptitude), SMA test (psychomotor coordination), and Hands test (spatial aptitude) are predictors of latent growth curve, i.e., they can predict the intercept (initial status) and the slope (development over time); (3) Covariate from personality and motivation dimension – *General Adaptability* predicts the slope (development over time) as a result of the adaptation to the demands of the aeronautic environment; and (4) SMA, Hands and INS2 tests and *General Adaptability* are predictors of cluster membership.

In what concerns the first hypothesis, there exists heterogeneity in the population and the best fitting model is a conditional 2-class mixture model (aBIC=3174.285), so we confirm hypothesis 1. Concerning the second hypothesis, only the SMA test (psychomotor coordination) showed significant effect on intercept (initial status), so the hypothesis 2 is partially confirmed. In what respects the third hypothesis, the covariate *General Adaptation* (personality and motivation dimension) predicts intercept (initial status) and slope (development over time). At last, for the fourth hypothesis the covariates did not showed significant effect on the discrete latent variable (C), and the Hands test (spatial aptitude), INS2 test (spatial aptitude), SMA test (psychomotor coordination) and *General Adaptation* could not predict the latent class membership of the applicants.

The results depicted the existence of two latent classes: latent class 1 (66%) and latent class 2 (34%). Latent class 1 presents a superior initial flight performance (comparing to latent class 2), decreases the performance from flight 1 to flight 3 and tends to stabilize onwards; while class 2 depicted the worst initial performance (comparing to latent class 1), slightly decreasing from flight 1 to flight 2, but recovering afterwards and presenting an increasing

tendency. The *General Adaptation* showed a significant effect on the intercept of latent class 1 (as a result, they showed a superior initial flight) but not on slope (we see that this class tends to stabilize and did not improve the performance). The SMA test presented a negative effect on the intercept of the latent class 2 (we can see that this class present a weaker initial flight performance and also depicted more error score in SMA test). Latent class 2 also showed a significant effect of *General Adaptation* on the slope (development over the six flights), improving and developing their proficiency. In such way, we saw that SMA test could explain the individual differences in the initial status (latent class 2), and that *General Adaptation* could explain the growth performance of the pilot applicants in latent class 2, and the initial status of pilot applicants in latent class 1.

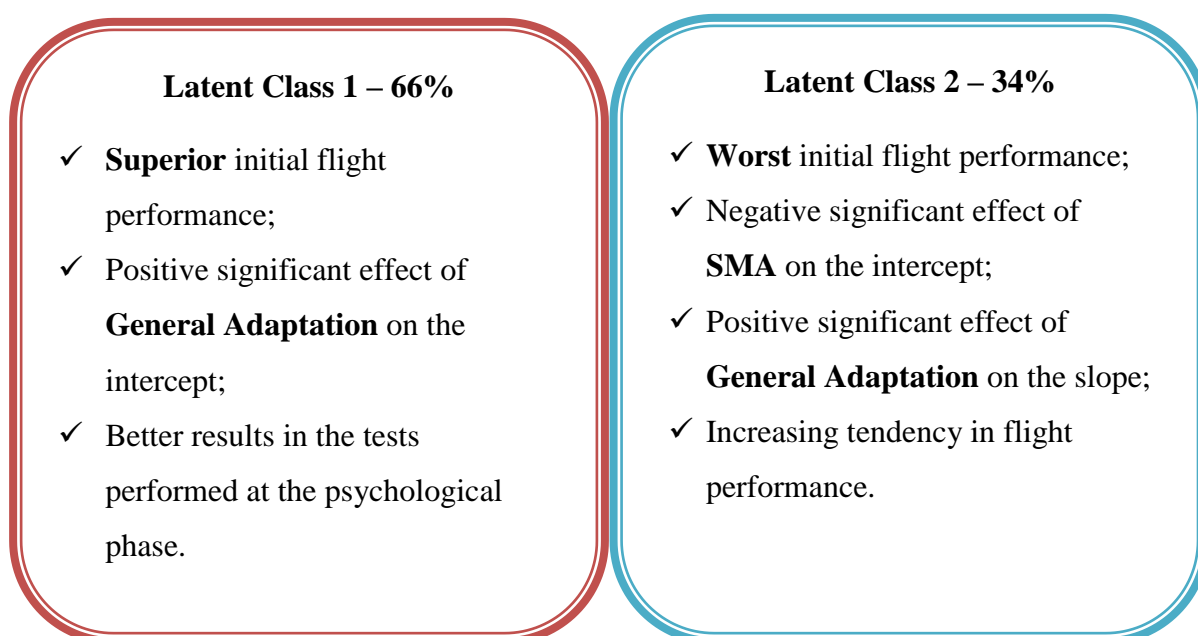


Figure 14 – Latent Class profiles.

In the unconditional model the variances of the growth factors were all significantly different from 0, and after introducing the covariates (conditional model) the variability reduction is confirmed, as a result of the covariates impact. In this case, the intercept variance and the covariance between the latent growth factors were no longer significantly different from 0.

4 Conclusions

4.1 *Empirical Findings*

The purpose of this thesis was to explore latent Growth Mixture Models (LGMM) in empirical research. In Chapter 2, with the simulation study, we conclude that the sample size has a great impact on the accuracy of the retrieving of the true parameters. The performance of LGMM was also examined in the aeronautic training field. This is a first attempt to model performance in flight using a longitudinal framework. In this field it was limited to a dichotomous split of training success into *Pass* versus *Fail*, regressed on different predictors (Roe, 2008), but without taking into account the learning process. The relevance of this type of research on pilot selection drifts essentially from the high cost associated to pilots training, and the inherent operations that apply for the use of aircraft (Martinussen, 2003).

This method (LGMM) helps answer two important questions in longitudinal data analysis that cannot be addressed by the conventional latent growth modeling (LGM). First, do multiple latent-growth trajectories exist and can the corresponding unobserved subpopulations be identified? Second, if so, can these unobserved subpopulations be predicted?

The findings of this study demonstrated the existence of two longitudinal change patterns in the pilot applicants' population.

We can conclude that SMA test proved to be a predictor of initial flight performance (*ab-initio abilities*) as it showed a significant effect on the initial status of the individuals (first evaluated flight). According to the theory (Bártolo-Ribeiro, 2004; Gomes *et al.*, 2008; Martinussen, 2003; Maschke, 2004; Stokes and Bohan, 1995), we confirm that motor coordination is predictor of success in flight. We also see that, there exists different patterns of learning among applicants, and that *General Adaptation* (personality and motivation dimension) can predict *learning capacities* and adaptation of the applicants. Given this, we can say that if an applicant possesses some factors such as maturity, emotional stability, initiative, combativeness and motivation, he can succeed in a stressful aeronautic environment, even if their *ab-initio abilities* were not brilliant.

As there are different patterns of learning, and different initial status among applicants, different programs of training can be suggested to adapt to each class needs.

We use the aeronautic training context to illustrate LGMM analysis, but LGMM can be applied easily to any research field dealing with longitudinal panel data with heterogeneous trajectories.

4.2 *Limitations and recommendations for further research*

Despite its advantages, the LGMM has several limitations that should be addressed. First, the computational load of LGMM estimation is extremely heavy, because (a) estimating the posterior distribution of the latent class variable requires a numerical integration technique that computes function and derivative values for the posterior distribution at every iteration of the estimation process, and (b) increasing the number of latent classes increases the number of free parameters to be estimated. As a result, some models are less stable or difficult to estimate, as evidenced by slow convergence, failed M-step iterations, a non-positive definite Fisher information matrix, or other computational problems.

Another issue in LGMM estimation that needs caution is that multiple local maxima may exist for the log-likelihood function (McLachlan and Peel, 2000). In other words, with different starting values, different maximizing results may be obtained for the log-likelihood function. Therefore, it is recommended that for a given model and data set, different sets of starting values are used to minimize the effect of potential local maxima.

It should be noted that, although several methods or indices are available to compare LGMM with different numbers of classes, the best way to guide LGMM selection is to test different models following theory-based hypotheses (i.e., a priori hypotheses; Muthén, 2004a). Model selection that is based solely on empirical results (i.e., in an exploratory manner) leads to selecting parsimonious models. However, parsimony is not simply a matter of statistical significance of increment in model fit in nested-model comparisons but also has theoretical and conceptual implications. Without theoretical backup, statistical parsimony will result in difficulty in terms of result interpretation. As such, we recommend using LGMM in a

confirmatory manner, via which theories of coexisting latent growth patterns can be evaluated by systematically testing different specific growth-mixture models representing competing ways of describing and explaining the data.

At last, it should be cautioned that similar to SEM, LGMM is a statistical procedure that usually is based on large samples and should not be used with small samples. In this study this is a limitation because the sample was composed by 297 individuals. Bentler and Chou (1987) recommended that the ratio of respondents to the parameters estimated should be at least 5:1 in SEM analyses. Although no such general guideline for sample sizes is available for LGMM because of the lack of research in this area, consideration of an appropriate sample size should be informed by the number of hypothesized latent groups and the differences in those growth parameters across groups (based on past research or theory). Generally, as the number of latent groups increases and as the differences in the growth parameters across groups decrease, larger sample sizes will be needed.

LGMM also can be extended to fit data simultaneously to multiple observed groups. For example, to examine whether the unobserved latent subpopulations are the same across gender, a multiple-group analysis can be conducted to estimate LGMM both for male and female participants simultaneously.

The greatest advantage of LGMM over LGM is that LGMM allows joint modeling of continuous and discrete latent variables to represent longitudinal data more realistically.

In summary, LGMM is a promising approach for modeling heterogeneous latent change processes across unobserved population subgroups. However, the search for these subgroups and the functional form of latent change should be guided by past research findings and relevant theories.

In our research study we include three covariates from perceptive-cognitive and psychomotor dimension (that measure spatial aptitude and motor coordination individually) and one covariate from the personality and motivation dimension (*General Adaptation*) that includes twelve factors (e.g., communication, leadership, emotional stability). In this area we suggest, the inclusion of individual personality variables (factors), that we can identify which variables have impact on the *learning capacities*. Finally, future research can take into account time-

varying covariates. For example, in our application temperature and time of the day of the flight may explain partially the performance of the applicant.

REFERENCES

- Academic Technology Services, Statistical Consulting Group (2006). Introduction to SAS. from <http://www.ats.ucla.edu/stat/sas/notes2/> (accessed on September 24, 2008).
- Akaike, H. (1974). A new look at the statistical model identification. *Transactions on Automatic Control* 19 (6), 716–723.
- Arbuckle, J. L. (2006). *Amos 7.0 User's Guide*. Chicago: SPSS.
- Aunola, K., Leskinen, E., Lerkkanen, M.K., and Nurmi, J.E. (2004). Developmental dynamics of math performance from preschool to grade 2. *Journal of Educational Psychology*, 96, 699–713.
- Bártolo-Ribeiro, R., Leal, R., Lafuente, D., Fonseca, M. J., Almeida, V., Piedade, P., Jamal, S., Antunes, A., and Valente, A. (2004). Validação de duas baterias de testes informatizados na selecção de pilotos militares. *Revista de Psicologia Militar*, 15, 193 – 208.
- Bauer, D. and Curran, P. (2003). Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. *Psychological Methods*, 8(3), 338–363.
- Bentler, P. M. (1995). *EQS Structural Equations Program Manual*. Encino, CA: Multivariate Software.
- Bentler, P. M. and Chou, C. H. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16, 78–117.
- Bollen, K. and Curran, P. (2006). *Latent Curve Models: A Structural Equation Perspective*. John Wiley & Sons, Inc.: New Jersey.
- Burke, E. F. & Hunter, D.R. (1995) *Handbook of Pilot Selection*. England: Avebury Aviation.
- Collins, L. (2005). Analysis of longitudinal data: the integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505 – 528.
- Curran, P. and Muthén, B. (1999). The application of latent curve analysis to testing developmental theories in intervention research. *American Journal of Community Psychology*, 27 (4), 567 – 595.
- Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1–38.
- Dias, J.G. (2004). *Finite Mixture Models. Review, Applications, and Computer-intensive Methods* (PhD Thesis), Groningen, the Netherlands: Research School Systems, Organization and Management, University of Groningen.

- Duncan, T. E., Susan, S. C., Strycker, L. A., and Okut, H. (2002). *Growth mixture modeling of adolescent alcohol use data: Chapter addendum to An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications*. Eugene, OR: Oregon Research Institute.
- Duncan, T., Duncan, S., and Stryckler, L. (2006). *An Introduction to Latent Variable Growth Curve Modeling – Concepts, Issues, and Applications*. Lawrence Erlbaum Associates: London.
- Gomes, A., Bártolo-Ribeiro, R. & Maroco, J. (2008). Utilização de Redes Neurais. Análise Discriminante e Regressão Logística como sistemas classificatórios na selecção de pilotos militares. *Revista de Psicologia Militar*, nº 17, 117 – 132.
- Hipp, J. R. and Bauer, D. J. (2006). Local solutions in the estimation of growth mixture models. *Psychological Methods*, 11(1), 36-53.
- Hoeksma, J. and Kelderman, H. (2006). On growth curves and mixture modeling. *Infant and Child Development*, 15, 627 – 634.
- Jöreskog, K. G., Sörbom, D., Du Toit, S., and Du Toit, M. (1999). *LISREL 8: New Statistical Features*. Chicago: Scientific Software International.
- Jung, T. and Wickrama, K.A.S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2, 302-317.
- Kreuter, F. and Muthén, B. (2008). Analyzing criminal trajectory profiles: Bridging multilevel and group-based approaches using growth mixture modeling. *Journal of Quantitative Criminology*, 24,1-31.
- Kokorian, A. (2008). Predicting simulator performance and the cost benefits of assessments for the selection of pilots. In T. Oliveira & P. Jorna (Eds.), *Proceedings of 28th Conference of European Association for Aviation Psychology (EAAP) – Enabling Operational Improvements*.
- Li, F., Duncan, T., Duncan, S., and Hops, H. (2001). Piecewise growth mixture modeling of adolescent alcohol use data. *Structural Equation Modeling*, 8(2), 175-204.
- Lo, Y., Mendell, N., and Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88 (3), 767-778.
- Lubke, G. and Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, 10 (1), 21-39.
- Martinussen, M. (2003). Psychological measures as predictors of military training performance. *The International Journal of Personality Assessment*, 80 (1),87-98.
- Maschke, P. (2004). Personality evaluation of applicants in aviation. In K. Goeters (Ed.), *Aviation Psychology: Practice and Research* (pp. 141–152). Hampshire, England: Ashgate Publishing Company.

- McArdle, J. and Epstein, D. (1987). Latent growth curves within developmental structural equation models. *Child Development*, 58, 110-133.
- McLachlan, G. J., and Peel, D. (2000). *Finite Mixture Models*. New York: John Wiley.
- McGrath J., Tschan F. (2004). *Temporal Matters in Social Psychology: Examining the Role of Time in the Lives of Groups and Individuals*. Washington, DC: APA.
- Mehta, P. D., Neale, M. C., & Flay, B. R. (2004). Squeezing interval change from ordinal panel data: Latent growth curves with ordinal outcomes. *Psychological Methods*, 9, 301-333.
- Meredith, W. and Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55,107–122.
- Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. and Sayer, A. (Eds.), *New Methods for the Analysis of Change* (pp. 291-322). Washington, D.C.: APA.
- Muthén, B. (2003). Statistical and substantive checking in growth mixture modeling. *Psychological Methods*, 8, 369-377.
- Muthén, B. (2004a). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.), *Handbook of Quantitative Methodology for the Social Sciences* (pp. 345–368). Newbury Park, CA: Sage Publications.
- Muthén, B. (2004b). *Mplus Technical Appendices*. Los Angeles, CA: Muthén & Muthén.
- Muthén, B. (2006). The potential of growth mixture modeling. Commentary. *Infant and Child Development*, 15, 623-625.
- Muthén, B. and Curran, P. (1997). General longitudinal modeling of individual differences in experimental design: A latent variable framework for analysis and power estimation. *Psychological Methods*, 2, 371-402.
- Muthén, B. and Khoo, S. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences*, 10(2), 73-101.
- Muthén, B. and Satorra, A. (1995). Technical aspects of Muthén's Liscomp approach to estimation of latent variable relations with a comprehensive measurement model. *Psychometrika*, 60(4), 489-503.
- Muthén, B. and Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55, 463-469.
- Muthén, B. and Muthén, L. (2000). Integrating person-centered and variable-centered analysis: growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and Experimental Research*, 24, 882-891.

- Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P., Kellam, S., Carlin, J., and Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. *Biostatistics*, 3, 459-475.
- Muthén, B., Khoo, S.T., Francis, D., and Boscardin, C. (2003). Analysis of reading skills development from Kindergarten through first grade: An application of growth mixture modeling to sequential processes. S.R. Reise and N. Duan (Eds). *Multilevel Modeling: Methodological Advances, Issues, and Applications*. Mahaw, NJ: Lawrence Erlbaum Associates, 71-89.
- Muthén, L.K. and Muthén, B.O. (1998-2007). *Mplus User's Guide*. Fifth Edition. Los Angeles, CA.
- Nagin, D. (1999) Analysing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, 4(2), 139 – 157.
- Nylund, K., Asparouhov, T., and Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: a monte carlo simulation study. *Structural Equation Modeling*, 14(4), 535 – 569.
- Raykov, T. (2005). Analysis of longitudinal studies with missing data using covariance structure modeling with full-information maximum likelihood. *Structural Equation Modeling*, 12(3), 493 – 505.
- Roe, R. (2008). Time in applied psychology. *European Psychologist*, 13(1), 37-52.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* 6(2), 461-464.
- Sclove, S. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52, 333 - 343.
- Stokes, A. F. & Bohan, M. (1995). Academic proficiency, anxiety, and information-processing variables as predictors of success in university flight training. *Proceedings of the 8th International Symposium on Aviation Psychology* (pp. 1107-1112). Columbus: The Ohio State University.
- Wang, M. and Bodner, T. E. (2007). Growth mixture modeling: identifying and predicting unobserved subpopulations with longitudinal data. *Organizational Research Methods*, 10(4), 635-656.
- Willett, J. B. and Sayer, A. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116, 363-381.
- Yung, Y. F. (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika*, 62(3), 297-330.

APPENDICES

Appendix A

Table A.1. Input statement for the simulated data

```

Title: Mixture modeling: 2 latent class [heterogeneity]
Data: File is "D:\mlgm_dataset1_c2_n500.dat";
Variable: Names are id y1 y2 y3 y4 y5 x1 x2 w1 gr;
Usevariables are y1 y2 y3 y4 y5 x1 x2 w1;
Classes = c(2);
Analysis: Type = mixture;
miteration = 300;
starts = 20 2;

```

```
MODEL:%overall%
```

```

int by y1-y5@1;
slp by y1@0 y2@1 y3@2 y4@3 y5@4;
[y1-y4@0];
y1 y2 y3 y4 y5;
int slp*0;
int with slp*0;
int on x1 x2;
slp on x1 x2;
C#1 on W1;
%c#1%
int*slp*;
y1 y2 y3 y4 y5;
int with slp;
%c#1%
[int*0 slp*0];
y1 y2 y3 y4 y5;
int*0 slp*0;
int with slp;
int on x1 x2;
slp on x1 x2;
%c#2%
[int*0 slp*0];
y1 y2 y3 y4 y5;
int*0 slp*0;
int with slp;

```

Appendix B

Table B.1. Input statement for the single class model

```
Data: File is " D:\geral_3cat_ap.dat"
VARIABLE:
  NAMES ARE ano ap idade hands ins1 ins2 sma v1-v6 u;
  USEVARIABLES ARE v1-v6;
  CATEGORICAL IS v1-v6;
  CLASSES = c(1);
ANALYSIS:
  TYPE IS MIXTURE;
  algorithm integration;
MODEL:%overall%
  i s | v1@0 v2@1 v3@2 v4@3 v5@4 v6@5;

OUTPUT: tech1 tech7;
PLOT:type is plot3;
series is v1(1) v2(2) v3(3) v4(4) v5(5) v6(6);
```

Table B.2. Input statement for the 2- class mixture model

```
Data: File is " D:\geral_3cat_ap.dat";
VARIABLE:
  NAMES ARE ano ap idade hands ins1 ins2 sma v1-v6 u;
  USEVARIABLES ARE v1-v6;
  CATEGORICAL IS v1-v6;
  CLASSES = c(2);
ANALYSIS:
  TYPE IS MIXTURE;
  algorithm integration;
MODEL:%overall%
  i s | v1@0 v2@1 v3@2 v4@3 v5@4 v6@5;

OUTPUT: tech1 tech7 tech11;
PLOT:type is plot3;
series is v1(1) v2(2) v3(3) v4(4) v5(5) v6(6);
```

Table B.3. Input statement for the 2- class mixture model with covariates

```
Data: File is "D:\geral_3cat_ap.dat";
VARIABLE:
  NAMES ARE ano ap idade hands ins1 ins2 sma v1-v6 u;
  USEVARIABLES ARE ap hands ins2 sma v1-v6;
  CATEGORICAL IS v1-v6;
  CLASSES = c(2);
ANALYSIS:
  TYPE IS MIXTURE;
  algorithm integration;
MODEL:%overall%
  i s | v1@0 v2@1 v3@2 v4@3 v5@4 v6@5;
  i s on ap hands ins2 sma;
  C#1 on ap hands ins2 sma;
  %C#1%
  i s on ap hands ins2 sma;
  %C#2%
  i s on ap hands ins2 sma;

OUTPUT: tech1 tech7;
PLOT:type is plot3;
series is v1(1) v2(2) v3(3) v4(4) v5(5) v6(6);
```
