Latent Transition Analysis
With Random Intercepts (RI-LTA)

Bengt Muthén and Tihomir Asparouhov *
Mplus

Version 4

Accepted for publication in Psychological Methods

September 18, 2020

*We are indebted to Michael Eid, David Kaplan and Stephanie Lanza for providing their data sets used in the paper. We thank Michael Eid, Ellen Hamaker and Mårten Schultzberg for helpful comments. Noah Hastings provided excellent support with tables and figures.

Abstract

This paper demonstrates that the regular LTA model is unnecessarily restrictive and that an alternative model is readily available that typically fits the data much better, leads to better estimates of the transition probabilities, and extracts new information from the data. By allowing random intercept variation in the model, between-subject variation is separated from the within-subject latent class transitions over time allowing a clearer interpretation of the data. Analysis of two examples from the literature demonstrates the advantages of random intercept LTA. Model variations include Mover-Stayer analysis, measurement invariance analysis, and analysis with covariates.

Key words: Hidden Markov, mixtures, transition probabilities, latent trait-state, measurement non-invariance.

Translational abstract

Modeling with latent classes over time is a common approach in Psychology when studying the development of for example mental states of happiness or depression over time. Latent transition analysis is a well-known approach for this purpose. A better statistical approach is presented here which represents the data better and more correctly assesses change and stability over time. Interpretations of psychological change processes are changed by this new methodology. Earlier LTA findings need to be revisited.

1 Introduction

Latent transition analysis (LTA) is frequently used in longitudinal studies to characterize changes over time in latent discrete states, also referred to as latent classes (see, e.g. Graham et al., 1991; Collins et al. 1992; Mooijaart, 1998; Reboussin et al. 1998; Langeheine & van de Pol, 2002; Kaplan, 2008; Lanza & Collins, 2008; and Collins & Lanza, 2010). The regular LTA model is, however, unnecessarily restrictive and an alternative model is readily available that typically fits the data much better, leads to better estimates of the transition probabilities, and extracts new information from the data.

The regular LTA is represented as a single-level, wide-format model. The alternative LTA model draws on the multilevel modeling idea of separating between-subject variation from within-subject variation. From a multilevel perspective, viewing time as the within level and subject as the between level, the latent class transitions are represented on the within level whereas the between level captures the variability across subjects. Essential parts of this multilevel idea, however, can be represented in a single-level model in line with the regular LTA model. Such an alternative single-level LTA model will be referred to as random intercept LTA (RI-LTA) because a key focus is allowing for variation across subjects represented by random intercepts.

The paper is structured as follows. Section 2 describes the regular single-level LTA model and gives a critique of it. Section 3 proposes the RI-LTA model. Section 4 places the proposed model in the context of other multilevel models with multiple indicators of latent variables. Section 5 provides a Monte Carlo study to investigate estimation of the new model. Section 6 shows applications of

Figure 1: LTA for 1 binary item at 5 time points (squares denote observed variables and circles latent variables).

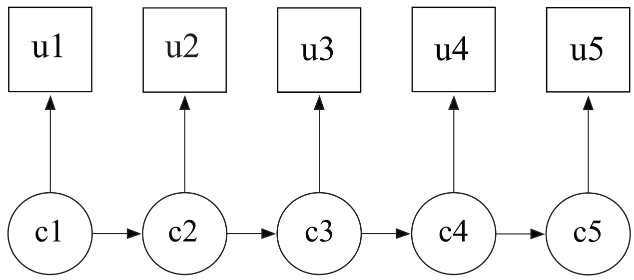
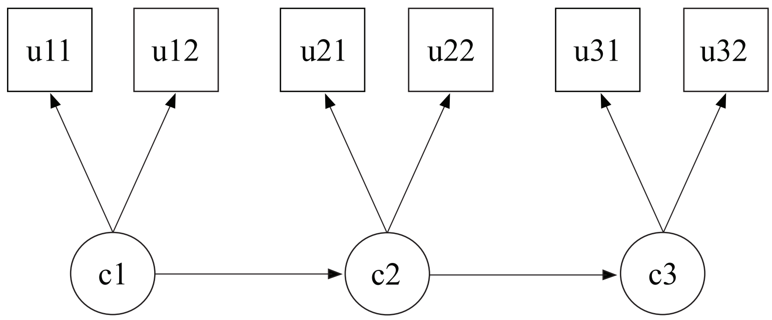


Figure 2: LTA for 2 binary items at 3 time points



RI-LTA to two data sets. Section 7 concludes with a discussion of computational aspects, other model variations, and the need for further research.

2 Regular LTA

Figure 1 and Figure 2 show model diagrams for two types of regular LTA models. In Figure 1 there is one binary indicator measured at five time points and in Figure 2 there are two binary indicators measured at three time points.

The regular LTA model has three parts. (1) The part for the latent class variable C_t at the first time point describes the initial status probabilities for the time 1 latent classes, $P(C_1)$. (2) The transition part describes the conditional probabilities of the latent class variable C_t at time t given the latent classes at time t-1, $P(C_2|C_1)$, $P(C_3|C_2)$, etc. Note that regular LTA allows only lag-1 relationships among the latent class variables, that is, C_t is influenced only by C_{t-1}, not C at any earlier time point. This is known as the Markov property. Stationarity, that is, invariance across time of the transition probabilities, is sometimes imposed. (3) The measurement part specifies the conditional probabilities $P(U_t|C_t)$ of the categorical latent class indicators U_t given the latent classes of C_t where the different latent class indicators U_t at time point t are independent conditioned on their respective latent class variable C_t. The latent class indicators U_t are typically assumed to be influenced only by C_t, the latent class variable at the same time point. Furthermore, measurement invariance for all latent class indicators is typically applied across all the time points. The model implies that the correlations across time for the latent class indicators are fully explained by the correlations among the latent class variables. Regular LTA is typically estimated using maximum-likelihood (ML) although Bayesian estimation can also be used.

Consider the parameters of the model represented in Figure 1. With 2 latent classes, this model has 5 parameters for the stationary version and 11 for the non-stationary version: 1 initial status parameter $P(C_1 = 1)$ with 2 transition parameters for the stationary model; $P(C_t = 1|C_{t-1} = 1)$, $P(C_t = 1|C_{t-1} = 2)$, and with 8 transition parameters for the non-stationary model, obtained as 2 times the 4 transitions; and 2 measurement parameters corresponding to the conditional probabilities $P(U_t = 1|C_t = 1)$ and $P(U_t = 1|C_t = 2)$. The 5 binary outcomes contribute $2^5 - 1 = 31$ pieces of information, that is, the unrestricted model for the 5 binary outcomes has $2^5 - 1 = 31$ parameters. With a large enough sample and a small enough total number of latent class indicators, it is possible to test fit between the observed and estimated frequency tables. This uses a likelihood-ratio or a Pearson chi-square test of the LTA model against the unrestricted model with degrees of freedom equal to the difference in the number of parameters for the unrestricted model and the LTA model. In other cases, model fit has to be assessed in more limited ways, e.g. via univariate and bivariate marginal frequency tables. The decision on the number of latent classes to use is typically based on BIC (Schwarz, 1978).

As an example, Table 1 gives the estimates for a life satisfaction example of Langeheine and van de Pol (2002) which corresponds to the Figure 1 model. Survey respondents were asked "How satisfied are you on the whole with your life" with answer categories unsatisfied and satisified. A two-class model was considered with classes labeled the same way as the answer categories.

The top part of the table shows the measurement parameters as the conditional probabilities of an unsatisfied/satisfied answer given membership in an unsatisfied/satisfied latent class. Each row shows the probabilities for the observed responses for the two latent classes. For each row, the large difference in these probabilities shows that the latent class indicators clearly discriminate between the two latent classes. The off-diagonal probabilities can be seen in the context of "measurement error" in that membership in a certain class does not necessitate an answer in the corresponding response category (Wiggins, 1973) but the probabilities are less than one. This discrepancy between latent and observed categories is a key feature of LTA and has given rise to the name hidden Markov modeling (see, e.g., MacDonald & Zucchini, 1997).

The bottom parts of the table show estimates for the latent classes. The latent class probabilities at the initial time point are estimated as 0.395 for the unsatisfied class and 0.605 for the satisfied class. The probability of staying in the same class between time 1 and time 2 is high, estimated as 1.000 and 0.874 for the unsatisfied and satisfied class, respectively. The latent class probabilities at the second time point are obtained as follows from the latent class probabilities at the first time point and the transition probabilities.

$$Unsatisfied: \; 0.395 \times 1.000 + 0.605 \times 0.126 = 0.471 \tag{1}$$

$$Satisfied: \; 0.605 \times 0.874 + 0.395 \times 0.000 = 0.529. \tag{2}$$

The transition probabilities for the other three transitions are of similar magnitude (although a test rejects invariance/stationarity).

2.1 A critique of the regular LTA model

The regular LTA model is analyzed in a single-level, wide format. It can, however, be viewed as a two-level model where time represents the within level (level 1) and subject represents the between level (level 2). In line with general two-level modeling, it is therefore important to separate between-level variation across subjects from within-level, across-time latent transitions. It is essential to remove between-subject differences that are stable over time from the within-subject process which is of primary interest. This general idea appears in several contexts with continuous observed and latent variables. For example, latent trait-state modeling (see, e.g., Kenny & Zautra, 1995; Cole et al., 2005; Eid

Table 1: LTA estimates for the Life satisfaction example

Measurement probabilities

Observed response	Latent class: Unsatisfied	Latent class: Satisfied
Unsatisfied	0.855	0.163
Satisfied	0.145	0.837

Time 1 latent class probabilities

Unsatisfied: 0.395 Satisfied: 0.605

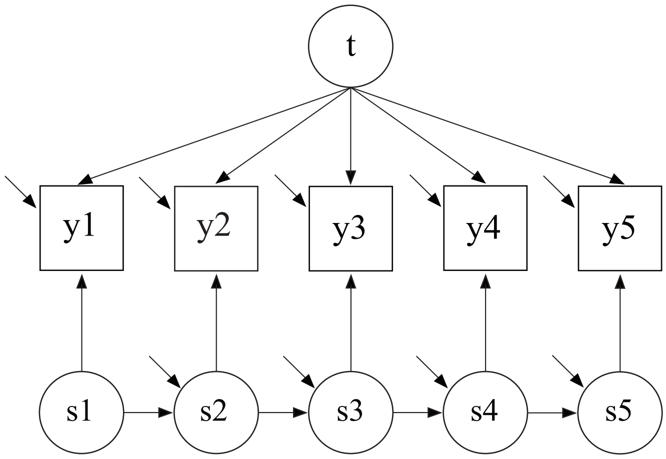
Transition probabilities for Time 1 (rows) to Time 2 (columns)

	Unsatisfied	Satisfied
Unsatisfied	1.000	0.000
Satisfied	0.126	0.874

et al., 2017) refers to the stable between-subject differences as a latent trait, a continuous latent variable. A related example is cross-lagged panel modeling (CLPM) where Hamaker et al. (2015) strongly advocates for separating out the stable between-subject differences referred to as random intercepts so that the cross-lagged relationships across time can be studied without interference of those between-subject differences. This is named the RI-CLPM approach and is the inspiration for the current paper. The idea of separating trait and states can be clearly seen in the Kenny-Zautra model shown in Figure 3. The latent trait is referred to as "T" and the latent states as "S" while the observed outcomes are denoted "Y" (other literature refers to this modeling as latent state-trait and defines states as the sum of the trait and the occasion-specific latent variables; see, e.g., Eid & Langeheine, 1999). Each observed outcome is the sum of trait, state, and a residual seen as measurement error. The key feature is that the latent trait influences the observed outcomes and not the latent states. In this way, the states are free of trait influence which means that the relationships between the states are not affected by stable differences between subjects.

The aim of the current paper is similar to the literature just cited, building on the idea of a stable trait in Kenny and Zautra (1995) and extracting between-subject variation in Hamaker et al. (2015). These two articles discuss continuous outcomes where you can split each outcome into a between and a within component of variation. This paper considers categorical variables where this split is more challenging. The split of the variation in the continuous-outcome case, however, is the same as using random intercept modeling and it is the random intercept idea that connects the continuous and categorical cases. The random intercept idea is common in the statistics and econometrics literature as

Figure 3: Kenny & Zautra (1995) latent trait-state model (t represents a latent trait variable, y1-y5 represent observed variables, s1-s5 represent continuous latent state variables).



a general way of representing unobserved heterogeneity (see, e.g., Fitzmaurice et al., 2011; Wooldridge, 2002). For categorical latent and observed variables, Eid and Langeheine (1999; 2003) consider latent trait-state modeling with a lag-1 structure for occasion-specific latent class variables which together with latent class variable traits contribute to the categorical outcomes. This is a type of latent transition model that uses a random intercept notion although not portrayed as such. Judging from the last two decades of applied LTA articles, however, the Eid-Langeheine model appears to have been overlooked and not adopted in latent transition analysis practice but will be one of the models studied here.

This paper focuses on the following two key aspects. First, it is of interest to study how much the latent transition probabilities are distorted in regular LTA when stable between-subject differences are ignored. Second, because LTA typically considers several indicators of the latent class variables, it is important to correctly assess the measurement quality of the indicators.

To summarize, because regular LTA does not separate out stable between-subject differences, it suffers from the risk of distorted estimates of the model's parameters, especially the transition probabilites. The alternative of random intercept LTA aims to avoid this distortion while staying in the single-level, wide analysis format.

3 Random intercept LTA (RI-LTA)

3.1 Continuous random intercept

Figure 4 shows two versions of continuous random intercept LTA (RI-LTA) for 2 binary latent class indicators measured at 3 time points. Here, $f1$, $f2$, and f are continuous latent variables (factors) where the loadings λ capture their different influence on the 2 latent class indicators. Each indicator's loading is held equal across time. In the top part of Figure 4, each latent class indicator has its own random intercept, $f1$, and $f2$, whereas in the bottom part of the figure, the indicators share the same random intercept factor f that has different effects on the two indicators. With many indicators, allowing each indicator to have its own random intercept makes the model unnecessarily complex and computationally cumbersome. This is discussed further in Section 4.

The focus of the paper is the proposed single-factor model in the bottom part of Figure 4. It is in the spirit of the Kenny-Zautra latent trait-state model for continuous observed and latent variables shown in Figure 3. Between-subject variation in the u outcomes is represented by a random interept factor and the

$c_1 - c_3$ model part represents the within-subject variation across time.

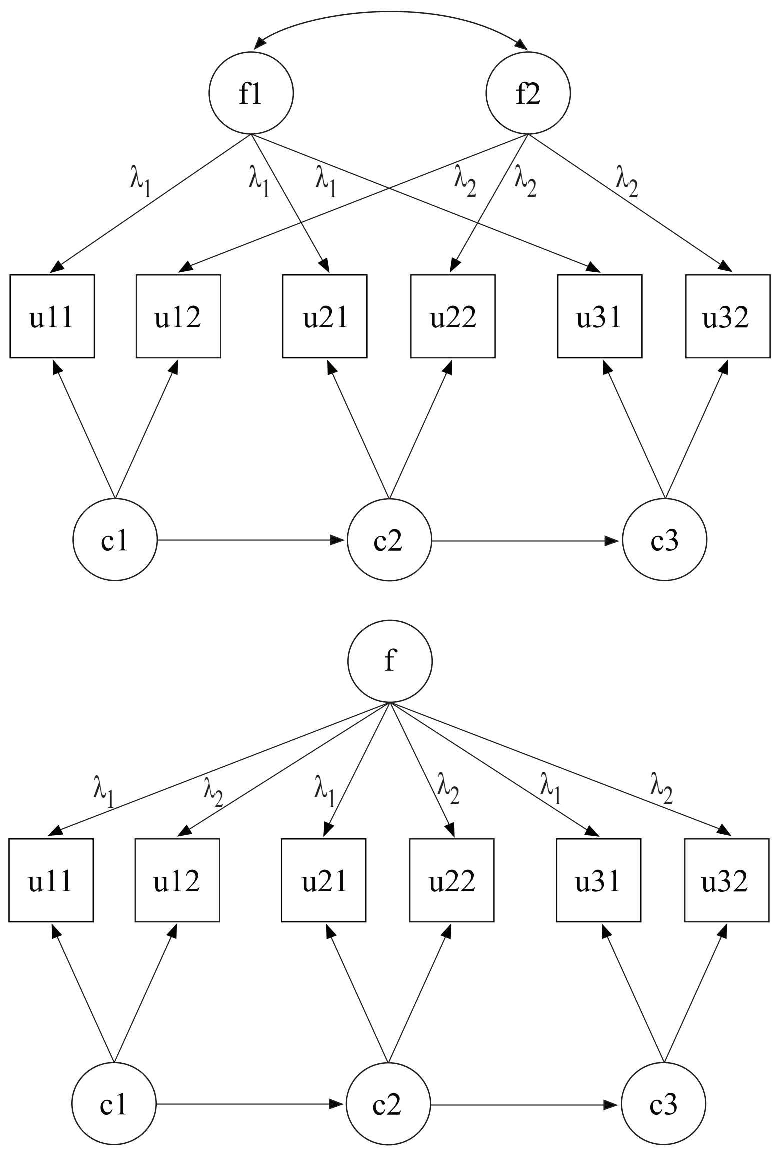
Formally presented, a single-factor continuous random intercept model version expressed in logit terms uses the following parameterization for a binary latent class indicator U_{rit} for indicator r, subject i, and time t measuring the latent class variable C_{it} for subject i and time t for class j,

$$logit[P(U_{rit} = 1|C_{it} = j, f_i)] = \alpha_{rj} + \lambda_r \, f_i, \tag{3}$$

where α_{rj} represent parameters that vary over latent class indicators and latent classes and f_i represent a subject-specific random intercept factor that does not vary across time point with factor loadings λ_r that vary across the latent class indicators. The factor is assumed to be distributed as $N(0, 1)$. Alternatively, each latent class indicator can have its own factor f_{ri} as in the top part of Figure 4 but this extension did not improve fit in the examples studied and the extension is not considered further here. The single random intercept factor version for R latent class indicators per time point adds only R parameters to regular LTA namely the factor loadings λ_r for each latent class indicator held equal across time. Note that in line with the concept of a random intercept, the factor loadings should not be different across time points because then the intercept factor does not reflect stable (time-invariant) individual differences (in contrast, latent trait-state modeling sometimes let loadings for traits be different across time). For simplicity, the factor loadings are also not allowed to change across the latent classes.

The factor f_i in equation (3) can be viewed as the trait of subject i, that is, a stable, time-invariant tendency where the λ factor loadings allow different effects of this trait on the different latent class indicators. For example, continuing the

Figure 4: RI-LTA for 2 binary latent class indicators at 3 time points with a random intercept: 2 continuous random intercept factors versus 1 (f denotes factors, u denotes observed binary variables, and c denotes latent class variables).



life satisfaction example and assuming a positive factor loading, a subject with a higher positive factor value has higher trait satisfaction and a subject with a higher negative factor value has lower trait satisfaction. If U is a binary latent class indicator where $U = 1$ represents satisfaction, a higher positive factor value means that the probability is higher of answering in the satisfied category and a higher negative value means that the probability is higher of answering in the unsatisfied category. In this way, large positive or negative factor values capture a tendency to not transition over time. This also implies that analysis using regular LTA of data generated by a RI-LTA model will tend to over-estimate the probabilities of staying in the same class.

Note also that the factor effect is specified to be the same for all latent classes which means that if there are two classes representing unsatisfied versus satisfied, the probability of answering in the satisfied category is increased for both classes. This results in the latent class indicator being somewhat less discriminating between the two classes in the sense that the $U = 1$ probabilities for the two classes are closer to each other. Typically, this effect is small and at the factor mean of zero, there is no such effect. Nevertheless, if the model fits better than regular LTA, this implies that regular LTA gives an inflated view of the class separation.

In regular LTA, the measurement part of the model considers the probability of a latent class indicator conditional on latent class, $P(U_{rt} = 1|C_t = j)$. For RI-LTA, this probability cannot be expressed in an explicit form but is obtained from equation (3) by integration over the factor. It is, however, possible to use the approximate logit to probit transformation obtained as

$$P(U = 1|C = j) = \Phi[\alpha_j / \sqrt{(3.2865 + \lambda^2 \, V(f))}], \tag{4}$$

where Φ is the standard normal distribution function and $3.2865 = \pi^2/3$, the variance of the logistic density[1]. Setting the metric of the factor as $V(f) = 1$, this shows that the larger the factor loading λ, the smaller the argument of Φ, that is, the closer the probability is to 0.5. This is another way to look at the lower latent class discrimination when the continuous random intercept factor is called for.

As equation (3) shows, different response probabilities are obtained for subjects with different f_i values. Because of this, the RI-LTA model allows for a certain form of measurement non-invariance across subjects (see also the discussion of related multilevel models in Section 4). In contrast, regular LTA implicitly imposes measurement invariance across subjects and this may be a too strict assumption.

3.2 Binary random intercept

Consider next the version of RI-LTA that has a binary random intercept represented by a latent class variable. This model also corresponds to the bottom part of Figure 4 but where f is a binary latent variable, here referred to as I. A simple model version expressed in logit terms uses the following parameterization for a binary latent class indicator U_{rt} for indicator r, time t, measuring the latent class variable C_t with latent class j and the single random intercept latent class

[1] Other transformations are also used in the literature, e.g. replacing the constant 3.2865 by 1.7^2. It is also possible to use a probit link in which case the Φ approximation using the constant 1 is exact.

variable I with latent class k ,

$$logit[P(U_{rt} = 1|C_t = j, I = k)] = \alpha_r + \beta_{rj} + \gamma_{rk}, \tag{5}$$

where $\beta_{r1} = 0$, $\gamma_{r1} = 0$ for identification purposes. Here, α_r is a parameter representing the effect of the latent class indicator r, β_{rj} is a parameter representing the effect of the latent class indicator r in combination with the latent class j of C, and γ_{rk} is a parameter representing the effect of the latent class indicator r in combination with the latent class k of I. An interaction term for the combination of j and k classes is omitted to keep the model parsimonious. As an example for 3 C classes and 2 I classes, the logits for a binary latent class indicator U_{rt} at time t are

$$logit[P(U_{rt} = 1|C_t = 1, I = 1)] = \alpha_r \tag{6}$$

$$logit[P(U_{rt} = 1|C_t = 2, I = 1)] = \alpha_r + \beta_{r2}, \tag{7}$$

$$logit[P(U_{rt} = 1|C_t = 3, I = 1)] = \alpha_r + \beta_{r3}, \tag{8}$$

$$logit[P(U_{rt} = 1|C_t = 1, I = 2)] = \alpha_r + \gamma_{r2}, \tag{9}$$

$$logit[P(U_{rt} = 1|C_t = 2, I = 2)] = \alpha_r + \beta_{r2} + \gamma_{r2}, \tag{10}$$

$$logit[P(U_{rt} = 1|C_t = 3, I = 2)] = \alpha_r + \beta_{r3} + \gamma_{r2}. \tag{11}$$

It is seen that the 6 logits are expressed in terms of 4 parameters. The parameters do not change over time. For the case of R latent class indicators per time point, J latent classes for C, and only 2 latent classes for I, this binary random intercept model has $R + R(J - 1) + R + 1$ parameters beyond those of the C part of the model: R α parameters, $R(J - 1)$ β parameters, R γ parameters, and 1 latent class parameter for I. The regular LTA model has $R\ J$ parameters beyond those of the C part of the model. This means that $R+1$ parameters are added to the regular LTA model when using 2 latent classes for I. This is irrespective of the number of response categories for U due to assuming a common shift for all response categories. Although not portrayed as a random intercept model, this is the parameterization used in the Eid and Langeheine (1999, 2003) studies of longitudinal mixture models.

3.3 Estimation and modeling considerations

Both the continuous and the binary random intercept models involve more heavy computations than regular LTA. Using maximum-likelihood estimation, the single continuous random intercept version leads to computations with one dimension of numerical integration. The binary random intercept version does not involve numerical integration but leads to one more latent class variable than regular LTA. Both the continuous and binary random intercept model versions of RI-LTA can be estimated using Mplus (Muthén & Muthén, 1998-2017). This draws on the general modeling framework described e.g. in Muthén and Asparouhov (2009). The analyses in this paper use maximum-likelihood estimation. Models cannot be compared using regular likelihood-ratio chi-square difference testing when they differ in the number of latent classes and/or when one model contains a continuous random intercept factor and the other does not. Also, due to having many cells in the frequency table for all the categorical outcomes, frequency table chi-square is not possible due to too many low frequency cells. The choice of model will instead be based on BIC (Schwarz, 1978) where smaller values are better,

$$BIC = -2\ loglikelihood + p\ ln\ N, \tag{12}$$

where p is the number of parameters, ln is the natural (e) log, and N is the sample size. BIC was found to perform well in Nylund et al. (2007) for related models with the large sample sizes usually encountered in latent transition analysis settings.

It should be noted that the regular LTA model is a special case of the RI-LTA model. In situations where there are no stable between-subject differences, the continuous random intercept model obtains zero factor loadings while the binary random intercept model does not find a latent intercept class.

It is clear from Figure 4 that the random intercept variable allows the indicators to correlate across time beyond what is captured by the latent class variables C_t being correlated across time in the latent transition part of the model. The indicator correlation across time is not a typical auto-regressive feature in that the correlation does not diminish with increasing time distance but is constant in line with representing a stable, time-constant, between-subject difference. Because it accounts for some of the correlation across time, it is clear that introducing this random intercept will affect the estimates of the latent transition probabilities, especially with respect to staying in the same latent class over time, that is, the diagonals of the transition probability matrices.

To some extent, random intercept modeling also relaxes the latent class assumption of conditional independence among the latent class indicators at a given time point. In this way, the continuous random intercept version is related to factor mixture modeling (see, e.g., Lubke & Muthén, 2005, Muthén & Asparouhov, 2006). The random intercept model does not, however, specify a factor for each time point but a factor that is in common for all time points. Using a factor mixture model for each time point as the measurement model may reduce the number of latent classes at each time point but is unlikely to reduce the number of latent classes in the analysis of all time points due to a one-factor construct being more restrictive than multiple latent classes in how across-time correlation is captured.

Unobserved heterogeneity in the form of between-subject variation in the latent class variable part of the model can be represented by adding a binary latent class variable where the two classes have different transition matrices. For each class, transitions can be viewed as a within-subject process using an RI-LTA model. The Mover-Stayer model (see, e.g. Langeheine & van der Pol, 2002) is an example of this where a latent class of Stayers is specified to stay in their time 1 latent class membership throughout all time points with probability 1. This attempt at capturing between-subject heterogeneity is in line with the random intercept theme of this paper, here applied to the latent class part of the model. The Mover-Stayer latent class variable can also be regressed on covariates. In the examples of Section 6, regular LTA with a Mover-Stayer addition is compared to the RI-LTA models with and without a Mover-Stayer addition.

Observed between-subject heterogeneity can be studied using groups and covariates and is discussed next.

3.4 Groups and covariates

3.4.1 Regular LTA

In regular LTA, it is possible to study group differences in the model parameters in line with Clogg and Goodman (1985) who presented an approach to a simultanous analysis of several groups. A strength of the multiple-group approach is its generality which allows any parameter to be equal or different across the groups. An example is the exploration of gender differences in the Lanza and Collins (2008) dating and sexual risk behavior study. The multiple-group approach can be used to test for measurement invariance across groups, that is, equality across time of latent class indicator probabilities conditional on latent class. An alternative approach is to let covariates representing subject characteristics such as gender, ethnicity, ses, and age influence the latent class indicators directly to thereby change their probabilities.

Covariates can also influence the latent class variables as well as their transition probabilities. This is carried out using the logit parameterizations shown in Table 2 for two latent class variables C_1 and C_2 where C_2 is regressed on C_1 and a covariate X. The regression is expressed as a multinomial logistic regression where

$$P(C_2 = c|C_1 = k, X = x) = e^{\alpha_c+\beta_{ck}+\gamma_{ck}x} / \sum_{j=1}^{J} e^{\alpha_j+\beta_{jk}+\gamma_{jk}x}, \tag{13}$$

with $\alpha_J = 0$, $\beta_{Jk} = 0$, $\beta_{c,J} = 0$, $\gamma_{Jk} = 0$. Here, α represents the intercepts for C_2, β represents the regression coefficients of C_2 regressed on C_1, and γ represents the regression coefficients of C_2 regressed on X. This translates the

Table 2: Logit parameterizations for C_2 regressed on C_1 and X: Interaction and main effect model versions

Interaction model

		C_2 1	2	3
C_1	1	$\alpha_1 + \beta_{11} + \gamma_{11}\mathrm{x}$	$\alpha_2 + \beta_{21} + \gamma_{21}\mathrm{x}$	0
	2	$\alpha_1 + \beta_{12} + \gamma_{12}\mathrm{x}$	$\alpha_2 + \beta_{22} + \gamma_{22}\mathrm{x}$	0
	3	$\alpha_1 + \gamma_{13}\mathrm{x}$	$\alpha_2 + \gamma_{23}\mathrm{x}$	0

Main effect model

		C_2 1	2	3
C_1	1	$\alpha_1 + \beta_{11} + \gamma_1\mathrm{x}$	$\alpha_2 + \beta_{21} + \gamma_2\mathrm{x}$	0
	2	$\alpha_1 + \beta_{12} + \gamma_1\mathrm{x}$	$\alpha_2 + \beta_{22} + \gamma_2\mathrm{x}$	0
	3	$\alpha_1 + \gamma_1\mathrm{x}$	$\alpha_2 + \gamma_2\mathrm{x}$	0

logit parameters into transition probabilities. Equation (13) model implies that the log odds comparing a certain C_2 category c to the last C_2 category J, is obtained as

$$log[P(C_2 = c|C_1 = k, X = x)/P(C_2 = J|C_1 = k, X = x)] = \alpha_c + \beta_{ck} + \gamma_{ck}\, x. \tag{14}$$

Exponentation gives the odds. The log odds and odds can also be computed with the diagonal of the transition table as the reference category showing the odds of transitioning relative to staying in the same class.

Table 2 shows two model variations. In the most general case shown at the top, an interaction is allowed between the X variable and the latent class variable C_1 so that the γ parameters vary across the different rows, that is the classes of C_1. Not allowing interactions but only main effects, the bottom part of the table shows that the γ parameters describing the influence of X are held equal across the C_1 classes. In this way, $\alpha_1 + \gamma_1 x$ and $\alpha_2 + \gamma_2 x$ can be seen as intercepts that are different for the C_2 classes whereas the regressions of C_2 on C_1 are not affected.

3.4.2 RI-LTA

With RI-LTA, the intent is to represent between-subject variation by random intercepts so that the relationships between the latent class variables are based on within-subject variation only. Because a random intercept of RI-LTA represents between-subject variation, it is therefore natural to let the random intercept have different means across groups in a multiple-group approach or be regressed on covariates in the covariate approach. The multiple-group approach, allowing for group specific transition probabilities in addition to group specific random intercept means, is suitable for the RI-LTA purpose because within each group, it can still be assumed that there is no between-subject variation in the relationships among the latent class variables. The covariate approach captures observed heterogeneity among subjects so that conditioning on the covariate values, the relationships among the latent class variables can be seen as within-subject relationships.

The next section places the proposed RI-LTA model in the context of other multiple indicator models with latent variables. This section is followed by a Monte Carlo simulation study of the proposed model. Readers more interested in applications can proceed directly to Section 6.

4 Related models: A multilevel perspective

Because LTA can be viewed as a model with variation across time and variation across subjects, it can be described as a two-level model. In this way, random intercept LTA can be related to other multiple-indicator latent variable models namely two-level factor analysis and two-level latent class analysis. This places the proposed random intercept LTA in a broader perspective.

A key modeling choice is if the random intercepts appear for the observed indicators of the latent variables or for the latent variables themselves. When the two-level modeling is applied to longitudinal data, this determines if the within-level relationships across time refer to within-level parts of the variables or to the combination of within- and between-level parts of the variables (the whole variables). The former approach is chosen in this paper as it gives a clearer representation of the data. The choice between the two modeling approaches is discussed below. The concept of measurement non-invariance expressed as random intercepts for the observed indicators is also emphasized.

4.1 Random intercepts in multilevel factor analysis

Consider a binary outcome U_{ij} for subject i in cluster j which is an indicator of a factor f_{ij} using e.g. logistic regression. A typical example is measurement of student performance in schools. Denoting the within- and between-level factors as $f_{W_{ij}}$ and f_{B_j}, the model can be expressed by the two equations

$$logit[P(U_{ij} = 1|f_{W_{ij}})] = \nu_j + \lambda_W f_{W_{ij}}, \tag{15}$$

$$\nu_j = \nu + \lambda_B f_{B_j} + \epsilon_{Bj}, \tag{16}$$

corresponding to the within- and between-level parts of a two-level model. This is in line with two-level regression where the intercept ν_j is random, varying across schools. The fact that the intercept ν_j is not the same for all schools can be seen as a type of measurement non-invariance (Jak et al. 2013, 2014; Muthén and Asparouhov, 2018). The model is shown in Panel (a) of Figure 5 for five factor indicators $u_1 - u_5$ where in line with Muthén and Muthén (1998-2017), the filled circles for the factor indicators on the within level show that their intercepts are random. On the between level, the random intercepts are shown as circles representing continuous latent variables (u in the figure corresponds to ν in (15) - (16)). The ϵ_B residuals on the between level are left out in the figure because they are often close to zero. Leaving them out typically has little consequence for the rest of the model. The extraction of between-level variation ensures that using

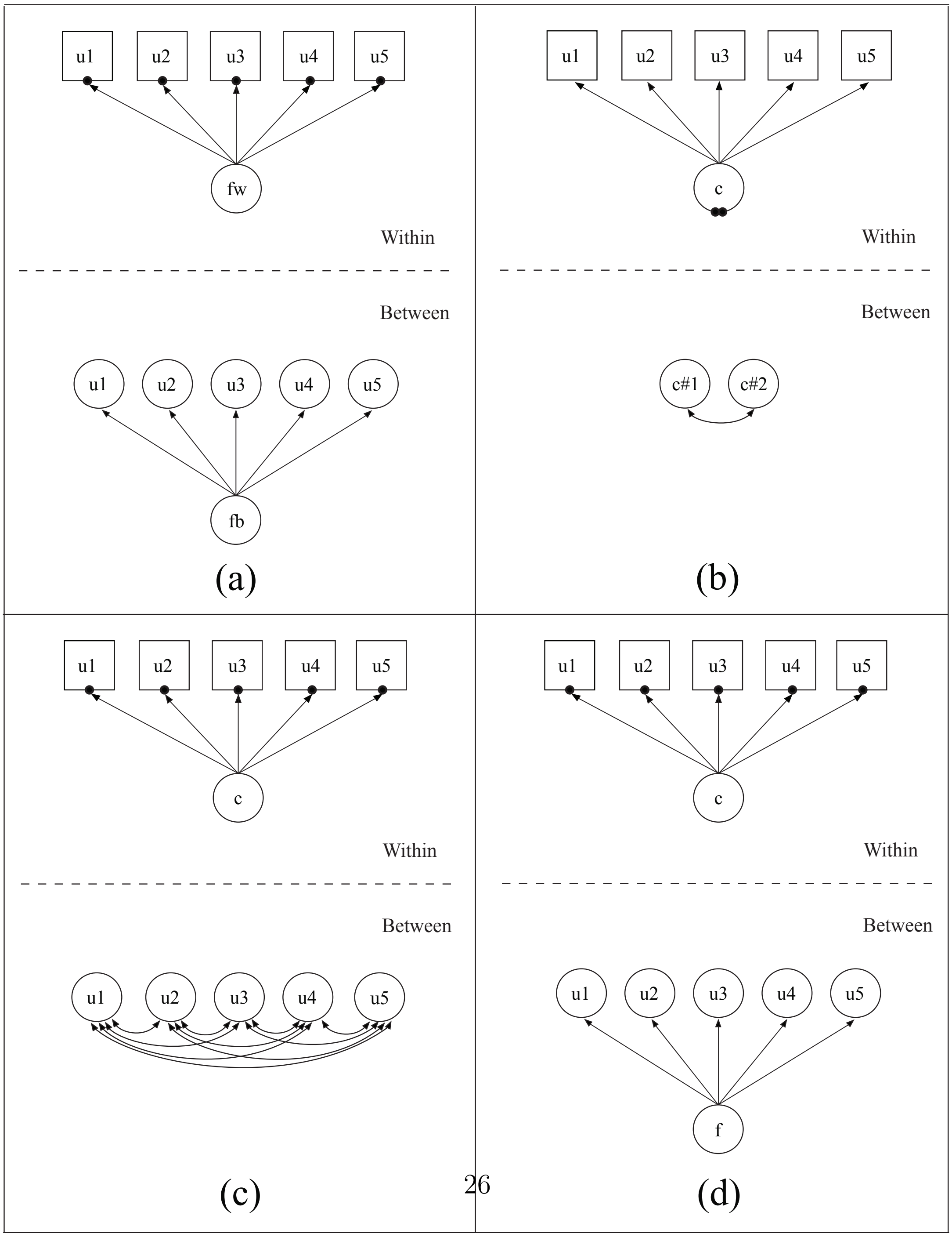
f_W as a predictor on the within level does not confound its effect by between-level variation.

Although the random intercept values are different for different clusters, the clusters are assumed to belong to the same population with the same mean and variance for the random intercepts. This view of measurement invariance/non-invariance is discussed in Asparouhov and Muthén (2016) and Muthén and Asparouhov (2018) and also relates to two-level modeling with random item parameters in Item Response Theory (see, e.g., de Jong, Steenkamp, and Fox 2007; de Jong & Steenkamp, 2010; Fox 2010).

4.2 Random intercepts in multilevel latent class analysis

Latent class analysis (LCA) has typically taken a different approach to multilevel modeling than factor analysis. As shown in Panel (b) of Figure 5, the variation across clusters is expressed via random intercepts/means for the classes of the latent class variable c instead of its indicators. This implies that cluster variation in the latent class indicators is sufficiently well accounted for by cluster variation in their underlying latent class variables. On the between level, the random intercepts/means $c\#1$ and $c\#2$ are continuous latent variables and are typically correlated as indicated by the double-headed arrow. The statistical underpinnings of multilevel latent class and latent transition analysis are discussed in e.g. Altman (2007), Asparouhov and Muthén (2008), Henry and Muthén (2010), and Vermunt (2003, 2008). The latent class variable c in Panel (b) contains both within-level and between-level variation. Using c as a predictor therefore confounds the two sources of variation. This modeling approach is therefore not suitable for latent

Figure 5: Panel (a) shows multilevel factor analysis and panels (b) - (d) show multilevel latent class analysis (squares represent observed u variables and circles represent latent variables; filled circles on the within level show that the variables have random intercepts; on the between level, the random intercept variables are latent variables; fw and fb represent factors on the two levels; $c\#$ variables represent the between-level random intercepts of the latent class variable c).



transition modeling with random intercepts.

The current paper draws on another multilevel LCA model that is in line with the multilevel factor analysis model presented earlier. The random intercepts will be specified for the latent class indicators instead of the latent classes as has been discussed in Asparouhov and Muthén (2008) and Henry and Muthén (2010). Consider a binary latent class indicator U_{ij} observed for student i in school j where the latent class variable C_{ij} represents different latent classes of students. Considering one of the five latent class indicators U, the random measurement intercept α_{cj} can be expressed via the logit of the conditional probability for U_{ij} given the latent class variable C_{ij} as

$$logitP(U_{ij} = 1|C_{ij} = c) = \alpha_{cj} = \alpha_c + \epsilon_j, \tag{17}$$

where the intercept α_c varies across the classes c and ϵ is a normally distributed random effect with mean zero and a variance that represents across-school variation.

This model is shown in Panel (c) of Figure 5. The filled circles at the bottom of the u boxes represent random measurement intercepts. On the between level, the random measurement intercept for each latent class indicator is shown as a circle u representing a continuous latent variable that varies across the between-level units, in this case schools. The random intercepts for the different items may correlate as indicated by the double-headed arrows. With a polytomous ordinal indicator, one can still specify a single random intercept shifting the probabilities of all response categories.

The model with random intercepts for the latent class indicators presents computational difficulties using maximum-likelihood estimation. With 5 indicators, it requires 5 dimensions of numerical integration corresponding to the 5 latent variables on the between level and this leads to very slow computations with low precision. A common solution to this problem is to place an intercept factor (a continuous latent variable) behind the set of latent variables as shown by the f intercept factor on the between level in Panel (d) of Figure 5. With zero residuals, this reduces the numerical integration to 1 dimension while allowing the random intercepts to correlate and estimating their factor loadings. A non-parametric version of this solution replaces the continuous intercept factors with a latent class variable to eliminate the numerical integration altogether and avoid a normality assumption for the factor. For example, a continuous factor can be seen as approximated by e.g. a 3-class latent class variable where the class proportions allow a non-symmetric distribution. In this paper, both the parametric approach using continuous factors and the non-parametric approach using latent classes are referred to as using random intercepts. The model in Panel (d) of Figure 5 is the latent class counterpart to the factor analysis model of Panel (a). The Panel (d) model is also the two-level representation of the model in the bottom part of Figure 4. In this context, the within level represents time and the between level represents subject.

Henry and Muthén (2010) provides an example of two-level LCA analyzing smoking behavior for 10,772 9th grade females in 206 rural communities across the United States. Six categorical latent class indicators measure three latent classes of student smoking behavior. Using random intercepts for the latent class indicators, they found significant variation across communities in the response probabilities for several of the indicators where the variation across communities was related to the proportion of youth living in poverty. For example, the indicator "Most friends are smokers" had a much larger probablity of being endorsed in communities with a large poverty proportion. In contrast, no significant differences across communities were found for the indicators "Parents would try to stop me from smoking" and "Smoking harms health". Using random intercepts/means for the latent classes, they also found differences across communities where communities in tobacco-growing states had a higher probability of being in the heavy smoking latent class.

As suggested by the Henry and Muthén (2010) smoking example, random intercept variation for the latent class indicators can be seen as a type of measurement non-invariance. In the LTA context, this non-invariance refers to different subjects having different response probabilities for a given latent class indicator.

5 Monte Carlo simulation

A small simulation study is carried out to assess the performance of RI-LTA at different sample sizes and number of time points. This is compared to the performance of regular LTA, both when data have been generated by a regular LTA model and when the data have been generated by an RI-LTA model.

The study uses 5 binary indicators to represent a situation with a moderate number of latent class indicators. Two latent classes are used where the indicators have the same logit values of 1 for class 1 and -1 for class 2. This translates to an indicator probability of 0.731 conditional on class 1 and an indicator probability of 0.269 conditional on class 2 when the regular LTA is considered. The large difference in probability means that the latent class indicators discriminate well between the classes. For RI-LTA, the indicators all have a factor loading of 2 which is of a magnitude seen in real data. The factor mean is zero and the factor variance is 1 (these are fixed parameters). By integrating over the factor, this corresponds to conditional probabilities of 0.644 for class 1 and 0.356 for class 2, that is, the indicators discriminate somewhat less well between the classes in the RI-LTA setup. The probabilities of class membership at the first time point are chosen as 0.5, 0.5. The transition probabilities are chosen as (rows represent starting class and columns ending class):

	Class 1	*Class* 2
Class 1	0.622	0.378
Class 2	0.500	0.500

This means that starting in class 1, it is more likely for a subject to stay in class 1 while starting in class 2, the subject is equally likely to stay as to change class. For simplicity in reporting, the study will focus on the 1, 1 diagonal element with population value 0.622 and the 2, 1 off-diagonal element with population value 0.500. These parameters are referred to as TRANS11 and TRANS21, respectively in the result tables. The Monte Carlo study uses sample sizes of 500, 1000, 2000, and 4000. Two and three time points are studied. With three time points the transition probabilities are the same for the last two time points as for the first two time points, reflecting a stationary LTA. Stationarity is not, however, imposed in the analysis. 500 replications are used. The simulations use Mplus where the Monte Carlo reporting gives the population value, the estimate mean over replications, the estimate standard deviation across replications (referred to as SD), the average standard error across replications (referred to as Ave SE), the mean square error, the 95% coverage, and the power to reject that the parameter is zero computed as the proportion of the replications where the confidence interval does not include zero. Bias in each estimated transition probability is reported as estimate minus population value. Key evaluation criteria are bias in the estimates; agreement between SD and Ave SE; and coverage.

5.1 Performance of regular LTA

A useful first step is to study the performance of regular LTA when data have been generated by a regular LTA so that the analysis model is correctly specified. The results are shown in the top part of Table 3. Only the case of N=500 is shown because the performance is good already at this sample size. For both T=2 and T=3, the bias is negligible, SD and Ave SE agree, and coverage is close to 0.95.

The bottom part of Table 3 shows the performance of regular LTA when data have been generated by the RI-LTA model with a continuous random intercept factor. It is seen that the performance is not acceptable because of the large bias which is 40% and 68% of the population probabilities, respectively. There is a strong over-estimation of the probabilities representing staying in the same class which is to be expected as discussed in Section 3.1. There is no improvement increasing from 2 to 3 time points. Increasing the sample size also does not help (results not shown).

Table 3: Analysis using regular LTA on data generated by regular LTA and by RI-LTA

Parameter	Est-pop bias	SD	Ave SE	MSE	Coverage	Power
Data generated by LTA						
			T=2, N=500			
TRANS11	0.003	0.052	0.052	0.003	0.940	1.000
TRANS21	0.001	0.054	0.054	0.003	0.948	1.000
			T=3, N=500			
TRANS11	0.003	0.051	0.050	0.003	0.948	1.000
TRANS21	0.001	0.049	0.051	0.002	0.962	1.000
Data generated by RI-LTA						
			T=2, N=500			
TRANS11	0.251	0.030	0.029	0.064	0.000	1.000
TRANS21	-0.328	0.030	0.032	0.108	0.000	1.000
			T=3, N=500			
TRANS11	0.261	0.028	0.028	0.069	0.000	1.000
TRANS21	-0.341	0.030	0.031	0.117	0.000	1.000

5.2 Performance of RI-LTA

Table 4 shows the results when data are generated by RI-LTA with a continuous random intercept factor and analyzed with this model. The top and bottom parts of the table shows the results for T=2 and T=3, respectively. For T=2, the parameter bias is large and the overall performance in terms of standard errors and coverage is not quite acceptable even for N=4000. The poor results for T=2, N=500 are in contrast with the Table 3 results for regular LTA when data are generated by a regular LTA. Note, however, that when data are generated by RI-LTA, the T=2, N=500 results in Table 4 for analysis using RI-LTA are considerably better in terms of both bias and coverage than those at the bottom of Table 3 using regular LTA.

For T=3, acceptable results are obtained already at N=500 and are very good for N=1000 and above. Further increasing the number of time points gives a small but practically negligible improvement of performance (not shown). A large sample size is most important.

6 Analysis of two examples

As a first step, analyses of two examples are described in terms of model fit, comparing regular LTA with RI-LTA using both a continuous random intercept and a binary random intercept. Next, estimates are presented and compared between regular LTA and RI-LTA.

The RI-LTA with a continuous random intercept factor uses the simple one-factor model version shown in equation (3) and in the bottom part of Figure 4. This is the same model as used in the Monte Carlo simulations. For the RI-LTA

Table 4: Analysis using RI-LTA on data generated by RI-LTA (population value for TRANS11=0.622, TRANS21=0.500)

Parameter	Est-pop bias	SD	Ave SE	MSE	Coverage	Power
T=2						
N=500						
TRANS11	0.121	0.335	0.358	0.127	0.576	0.514
TRANS21	-0.072	0.316	0.269	0.105	0.522	0.618
N=1000						
TRANS11	0.122	0.307	0.277	0.109	0.664	0.558
TRANS21	-0.058	0.275	0.275	0.079	0.642	0.652
N=2000						
TRANS11	0.090	0.251	0.243	0.071	0.764	0.656
TRANS21	-0.066	0.239	0.221	0.061	0.756	0.768
N=4000						
TRANS11	0.050	0.172	0.159	0.032	0.860	0.820
TRANS21	-0.044	0.182	0.174	0.035	0.850	0.910
T=3						
Parameter	Est-pop bias	SD	Ave SE	MSE	Coverage	Power
N=500						
TRANS11	0.021	0.141	0.147	0.020	0.924	0.914
TRANS21	-0.005	0.137	0.146	0.019	0.904	0.932
N=1000						
TRANS11	0.012	0.095	0.098	0.009	0.936	1.000
TRANS21	-0.002	0.101	0.101	0.010	0.920	0.992
N=2000						
TRANS11	0.006	0.068	0.068	0.005	0.926	1.000
TRANS21	-0.002	0.068	0.071	0.005	0.948	1.000
N=4000						
TRANS11	0.000	0.048	0.047	0.002	0.942	1.000
TRANS21	-0.003	0.048	0.050	0.002	0.946	1.000

with a binary random intercept, the parameterization of (6) - (11) is used. The Mover-Stayer models use one parameter more than the standard models with only movers due to adding a binary latent class variable of movers and stayers. All analyses are carried out using Mplus (Muthén & Muthén, 1998-2017) and scripts are available from the first author as well as at `http://www.statmodel.com/RI-LTA.shtml`

6.1 Analysis of the Mood data

The first example concerns ratings of mood. The data set is from a longitudinal study with N=494, 4 time points 3 weeks apart, and 2 binary latent class indicators measuring 2 latent classes at each time point (Eid & Langeheine, 2003). Participants rated their momentary sadness and unhappiness on a 5-point scale ranging from 1 (not at all) to 5 (very much). A dichotomized version of the two items was used in Eid and Langeheine (2003) as well as here (first category versus the other categories). A stationary model is chosen for this example because this is the model considered in Eid and Langeheine (2003).

6.1.1 Model fitting results for the Mood data

Table 5 compares the model fitting results of regular LTA with those of RI-LTA with a continuous and a binary random intercept. Models 1-3 in the top part of the table show standard analysis whereas models 4-6 in the bottom part show Mover-Stayer analysis. Model 2 is the same as model 2 in Table 1 of Eid and Langeheine (2003). The RI-LTA models are clearly better than the regular LTA both in terms of higher loglikelihood and lower BIC. BIC of the RI-LTA model 2

Table 5: Model fitting results for the Mood data

Model	# parameters	loglikelihood	BIC
	Standard		
1 Regular LTA	7	-2053	4150
2 RI-LTA, binary RI[1]	10	-2028	4118
3 RI-LTA, continuous RI	9	-2019	4093
	Mover-Stayer		
4 Regular LTA	8	-2037	4123
5 RI-LTA, binary RI[2]	11	-2017	4101
6 RI-LTA, continuous RI	10	-2017	4096

[1] Model 2 is model 2 in Table 1 of Eid and Langeheine (2003).
[2] Model 5 is model 5 in Table 1 of Eid and Langeheine (2003).

is better than that of the regular LTA model 1 and the RI-LTA model 3 with a continuous random intercept factor further improves on BIC. Model 3 in fact has a better loglikelihood value than model 2 despite having 1 parameter less.

Models 4-6 in the bottom part of Table 5 have the same BIC rank ordering as in the top part. For regular LTA, the Mover-Stayer version of model 4 is preferred over the regular LTA model 1 due to the better BIC. Similarly, using a binary random intercept, the Mover-Stayer RI-LTA model 5 is preferred over the RI-LTA model 2. Model 5 is model 5 of Table 1 of Eid and Langeheine (2003) and is the preferred model in that article. In contrast, using a continuous random intercept, the RI-LTA model 6 has a worse BIC than that of model 3 indicating no need for Mover-Stayer modeling.

6.1.2 Model estimates for the Mood data

It is interesting to compare the model estimates for some key models of Table 5, both in terms of measurement probabilities and transition probabilities. Table 6 shows estimates from regular LTA (model 1), regular LTA with a Mover-Stayer component (model 4), and the continuous version of RI-LTA without a Mover-Stayer component (model 3). Regular LTA gives quite different results than the better-fitting RI-LTA both in terms of measurement parameters and transition parameters.

The measurement parameter estimates of the two regular LTA models suggest that the two indicators Sad and Unhappy discriminate well between the two classes, labelled as Not sad/Happy and Sad/Unhappy. The probabilities are very low in the first class and very high in the second class. For the better-fitting RI-LTA model, the class differences in probabilities are smaller showing that regular LTA gives an inflated view of class separation as discussed in Section 3.1.

The factor loadings of the continuous version of RI-LTA are significant with estimates (SEs) 2.805 (0.251) and 6.775 (2.135). The larger loading for the unhappiness indicator compared to the sadness indicator suggests that unhappiness shows a higher degree of stability over time.

The estimated transition probabilities of regular LTA have higher diagonal values suggesting more stability in class membership over time than with RI-LTA. This is in line with the Monte Carlo simulation findings of over-estimated diagonal values with regular LTA shown in the bottom part of Table 3.

For RI-LTA, transitioning from the Sad/Unhappy class to the Not sad/Happy class is almost as likely as staying in the same class. The Mover-Stayer version

Table 6: Mood data estimates

Regular LTA		
Measurement probabilities		
	Classes	
Indicator	Not sad/Happy	Sad/Unhappy
Sad	0.089	0.902
Unhappy	0.038	0.857
Transition probabilities		
	Not sad/Happy	Sad/Unhappy
Not sad/Happy	0.803	0.197
Sad/Unhappy	0.248	0.752
Mover-Stayer LTA: Movers		
Measurement probabilities		
Iindicator	Not sad/Happy	Sad/Unhappy
Sad	0.089	0.898
Unhappy	0.031	0.860
Transition probabilities		
	Not sad/Happy	Sad/Unhappy
Not sad/Happy	0.664	0.336
Sad/Unhappy	0.446	0.554
RI-LTA, continuous RI		
Measurement probabilities		
Indicator	Not sad/Happy	Sad/Unhappy
Sad	0.286	0.748
Unhappy	0.164	0.804
Transition probabilities		
	Not sad/Happy	Sad/Unhappy
Not sad/Happy	0.691	0.309
Sad/Unhappy	0.486	0.514

of the regular LTA has similar transition probabilities but they are valid for only the movers, estimated as 61%; this model also has worse log likelihood and BIC.

Further insight into the RI-LTA model can be obtained by taking a closer look at the stayers in the Mood data. Of the N=494 subjects, 96 (19%) give the same Not sad/Happy answer to both latent class indicators at all four time points and 62 (13%) give the same Sad/Unhappy answer. In other words, about a third of the sample consists of stayers. The RI-LTA model captures the stayers by assigning large factor values to them. To see this, factor values can be estimated by the usual maximum a priori method. With the 0, 1 metric of the factor, the 96 consistently happy subjects have the by far lowest factor score estimate of -1.157 and the 62 consistently unhappy subjects have the by far highest factor score of 1.420. As discussed in Section 3.1, the large negative estimate implies a high probability of answering in the Not sad/Happy category ($U = 0$) at all time points and the large positive estimate implies a high probability of answering in the Sad/Unhappy category ($U = 1$) at all time points. Rather than the regular LTA categorization into movers and stayers, the random intercept factor of RI-LTA provides a continuum of more or less movement over time.

6.2 Analysis of the Dating data

The second example concerns dating and sexual risk behavior. The data set is from the National Longitudinal Survey of Youth (NLSY97) with N=2937, 3 time points one year apart, and 4 ordinal and binary items measuring 5 latent classes at each time point. An LTA analysis of these data appeared in an influential article by Lanza and Collins (2008), introducing SAS PROC LTA. The items are past-year number of dating partners (0, 1, 2 or more), past-year sex (no, yes), past-year number of sexual partners (0, 1, 2 or more), and exposed to STD in past year (no, yes). Covariates are gender and whether the respondent has used cigarettes, been drunk, or used marijuana in the past year. In the current analyses, the item Had sex in past year is dropped due to a no response necessitating a zero answer to the item Number of sexual partners, thereby avoiding an unnecessary violation of conditional independence. The regular LTA analyses still produce the same 5-class interpretation as in Lanza and Collins (2008). A stationary model is chosen for the Dating and sexual risk behavior example because unlike regular LTA as in Lanza and Collins (2008), stationarity cannot be rejected for the RI-LTA models.

6.2.1 Model fitting results for the Dating data

Table 7 compares the model fitting results of regular LTA with those of RI-LTA with a continuous and a binary random intercept. Models 1-3 in the top part of the table show standard analysis whereas models 4-6 in the bottom part show Mover-Stayer analysis. The RI-LTA models are clearly better than the regular LTA both in term of higher loglikelihood and lower BIC. For regular LTA, BIC points to a Mover-Stayer model whereas for the RI-LTA models it does not.

6.2.2 Model estimates for the Dating data

It is interesting to compare the model estimates for some key models of Table 7, both in terms of measurement probabilities and transition probabilities. Table 8 shows estimates from regular LTA (model 1) used in Lanza and Collins (2008) and the continuous version of RI-LTA without a Mover-Stayer component (model 3). Regular LTA gives quite different results than the better-fitting RI-LTA both

Table 7: Model fitting results for the Dating data

Model	# parameters	loglikelihood	BIC
	Standard		
1 Regular LTA	49	-16202	32796
2 RI-LTA, binary RI	53	-16056	32535
3 RI-LTA, continuous RI	52	-16043	32502
	Mover-Stayer		
4 Regular LTA	50	-16194	32787
5 RI-LTA, binary RI	54	-16053	32536
6 RI-LTA, continuous RI	53	-16041	32506

in terms of measurement parameters and transition parameters.

For the regular LTA, the measurement parameter estimates show a pattern of probabilities that is very similar to that of Lanza and Collins (2008) with latent class described as Nondaters, Daters, Monogamous, Multipartner safe, and Multipartner exposed where exposed refers to being exposed to STD in the past year. In parts, the RI-LTA model has a similar pattern of probabilities where the latent classes of Nondaters, Daters, and Multi-exposed can be seen. The Monogamous class is however somewhat different in that having 2 or more dating partners is a bit more likely than having 1 partner. Also, the Multi-safe class is not found but instead a class that can be described as Monogamous exposed.

The factor loadings for the continuous random intercept version of the RI-LTA model are significant with estimates (SEs) 1.574 (0.152), 4.194 (0.504), 1.507

Table 8: Dating data estimates

	Regular LTA				
	Nondaters	Daters	Monogamous	Multi-safe	Multi-exposed
Measurement Probabilities:					
# dating partners in past year					
0	0.789	0.008	0.096	0.031	0.025
1	0.166	0.214	0.641	0.026	0.044
>=2	0.045	0.778	0.262	0.943	0.932
# sex partners in past year					
0	0.975	0.948	0.000	0.104	0.000
1	0.014	0.048	0.961	0.258	0.118
>=2	0.011	0.004	0.039	0.638	0.882
Exposed to STD in past year					
No	1.000	1.000	0.385	1.000	0.187
Yes	0.000	0.000	0.615	0.000	0.813
Transition probabilities					
Non-daters	0.627	0.197	0.096	0.038	0.042
Daters	0.023	0.626	0.173	0.086	0.091
Monogamous	0.034	0.032	0.679	0.056	0.199
Multi-safe	0.036	0.000	0.177	0.584	0.203
Multi-exposed	0.021	0.033	0.201	0.055	0.690
	RI-LTA, continuous RI				
	Nondaters	Daters	Monogamous?	Mono-exposed	Multi-exposed
Measurement Probabilities:					
# dating partners in past year					
0	0.861	0.039	0.000	0.193	0.057
1	0.041	0.192	0.457	0.708	0.048
>=2	0.098	0.770	0.543	0.100	0.895
# sex partners in past year					
0	0.922	0.789	0.026	0.000	0.017
1	0.033	0.075	0.883	0.898	0.099
>=2	0.045	0.136	0.091	0.102	0.884
Exposed to STD in past year					
No	1.000	1.000	0.485	0.113	0.398
Yes	0.000	0.000	0.515	0.887	0.602
Transition probabilities					
Non-daters	0.583	0.264	0.029	0.058	0.067
Daters	0.014	0.660	0.183	0.013	0.130
Monogamous	0.008	0.067	0.405	0.200	0.320
Multi-safe	0.071	0.007	0.040	0.639	0.243
Multi-exposed	0.031	0.050	0.172	0.127	0.619

(0.246). The larger loading for the latent class indicator Number of sexual partners in the last year suggests that this indicator has a higher degree of stability over time. Viewed from the perspective of measurement non-invariance discussed in Section 3.1, it also indicates that this latent class indicator shows a larger amount of measurement non-invariance across subjects.

The difference in estimated transition probabilities do now show a clear pattern in the difference between regular LTA and RI-LTA. For the three latent classes with the same interpretation by the two models, the estimated transition probabilities show larger diagonal elements for regular LTA as compared to RI-LTA in two out of the three classes.

6.2.3 Measurement invariance testing for the Dating data

Lanza and Collins (2008) studied differences in LTA parameters for males and females. A first such analysis concerns measurement invariance across gender. This can be done using a multiple-group approach or equivalently by using a gender covariate that influences all latent class indicators directly. The latter approach is used here because it is more efficient computationally. To reduce the risk of distorting the measurement invariance testing, a reasonably flexible structural model for the latent class part is used here, namely, the main effect model described in the bottom part of Table 2.

Table 9 shows the model testing results for regular LTA and RI-LTA with a continuous random intercept factor. Measurement invariance can be checked using likelihood-ratio chi-square testing, in this case with 15 degrees of freedom corresponding to 3 latent class indicators and 5 latent classes for which there are gender differences. For these data, regular LTA and RI-LTA agree that invariance

Table 9: Measurement invariance testing for the Dating data: Males vs females

Model	#par's	LL	BIC	Test(df)	x^2
1 Regular LTA, invariance	61	-16123	32733		
2 Regular LTA, non-invariance	76	-16097	32800	1 vs **2**(15)	52
3 RI-LTA, continuous RI, invariance	64	-15977	32465		
4 RI-LTA, continuous RI, non-invariance	79	-15951	32532	3 vs **4**(15)	56

is rejected with $p < 0.005$. BIC, however, points to measurement invariance in both cases and for simplicity this model will used when adding other covariates.

6.2.4 Covariate influence for the Dating data

The Dating example has four binary covariates, gender and whether the respondent has used cigarettes, been drunk, or used marijuana in the past year. Table 10 shows the results of a second set of analyses that explores the influence of these covariates on the latent class variables and the transitions.

The regular LTA model 1 uses the main effect model shown at the bottom of Table 2. Model 2 uses the interaction effect model for regular LTA shown at the top of Table 2 but where the interaction is only with respect to gender and not the other three covariates. This interaction model was chosen because the possible gender effect on transitions was mentioned in Lanza and Collins (2008). Contrasting the models, both BIC and chi-square testing indicate that males and females do not have different transitions.

In the RI-LTA model 3, the covariates are allowed to influence the continuous

Table 10: Model testing using covariate analysis for the Dating and sexual risk behavior example: Regular LTA compared to RI-LTA with a continuous random intercept

Model	Covariate influence	# par's	LL	BIC	Test (df)	χ^2
1. Regular LTA	Main effects	81	-15630	31906		
2. Regular LTA	Main effects and gender interaction effects	97	-15621	32016	**1** vs 2 (16)	18
3. RI-LTA	Continuous RI	56	-15653	31753		
4. RI-LTA	Continuous RI and main effects	88	-15461	31624	3 vs **4** (32)	384
5. RI-LTA	Continuous RI, main effects, and gender interaction effects	104	-15454	31738	**4** vs 5 (16)	14

random intercept factor while in the RI-LTA model 4, the covariates also influence the latent class variables using the main effect parameterization shown at the bottom of Table 2. Comparing models 3 and 4 shows that the covariate influence on the latent class variables needs to be included in the model. Model 5 is the RI-LTA counterpart to the regular LTA model 2 which allows gender interaction effects on the transitions. This indicates that males and females do not have different transitions so in this case there is agreement with regular LTA. Comparing the best regular LTA model 1 and the best RI-LTA model 4, however, it is seen that both the loglikelihood and BIC are better for RI-LTA. In addition, model 1 and model 4 have different covariate effects. The effect of covariates on the latent class of multipartner-exposed is of special interest. In presenting these results, the log odds relates this class to the class of monogamous. For the regular LTA of model 1, significant and positive effects are seen for male and past-year marijuana usage at all time points, with an additional significant positive effect of past-year drunkeness for the first time point. Past-year cigarette use does not have a significant effect. For the RI-LTA model 4, only male has a significant effect and it is positive. The covariate effects on the continuous random intercept, however, are significant and positive for all the covariates. Positive effects increase the random intercept value which in turn increases the probability of the latent class indicators being in category 1 versus category 0 for binary indicators and increases the probabilities of the higher categories relative to the lower categories for the ordinal indicators. In other words, only male increases the latent class odds while all covariates increase the odds of answering in a more “extreme” category of the latent class indicators. The latter effect refers to a between-subject difference that is stable over time and is unrelated to latent class membership.

7 Discussion

This paper demonstrates the need for replacing regular LTA with random intercept LTA (RI-LTA). Most importantly, RI-LTA typically fits the data better as illustrated by the examples in this paper. This was also found to be the case using several other data sets, including the life satisfaction data of Langeheine and van de Pol (2002) mentioned in Section 2 and the reading data of Kaplan (2008). Apart from a better fit of the model to the data, RI-LTA gives a clearer interpretation. Regular LTA suffers from estimating transition probabilities that confound between- and within-subject influences. By allowing random intercept variation in the model, the between-subject variation is extracted from the latent class indicators so that latent class transitions over time refer to within-subject transitions. Because regular LTA does not include a random intercept, the probability of staying in the same class is typically over estimated. In addition, regular LTA overlooks information in the data which relates to measurement. Unlike regular LTA, RI-LTA allows for measurement non-invariance across subjects represented by the random intercepts. Regular LTA typically overstates the ability of latent class indicators to discriminate between latent classes. Regular LTA is also more likely to need an added Mover-Stayer component whereas the random intercept of RI-LTA captures tendencies to stay in the same latent class without such an added component. A limited simulation study indicates that for sample sizes of at least 500, RI-LTA performs well when there are three or more time points whereas with only two time points, a sample size of more than 4000 may be needed.

While the case of categorical latent class indicators has been discussed here, the same approach can also be applied to continuous, count, or nominal latent class indicators. Several additional aspects of modeling with random intercepts are of interest and are discussed below.

7.1 Computational aspects

The RI-LTA model requires a considerably longer computational time than regular LTA. The continuous random intercept version is the most time-consuming in that the maximum-likelihood estimation requires numerical integration but also because it needs more random starting values to replicate the best loglikelihood. While much faster than the continuous random intercept version, the binary random intercept version is also slower than regular LTA due to having one more latent class variable. Recent advances in CPU speed, multithreading, and algorithmic improvements, however, have made it practical to estimate RI-LTA models.

7.2 Other model variations

Several other variations of RI-LTA are possible in order to make the model more flexible. Following are five such variations that are possible in the latent variable framework of Mplus (Muthén & Muthén, 1998-2017). First, the typical assumption of a lag-1 relationship between the latent class variables C_t may be relaxed. Lag-2 effects were significant per likelihood-ratio chi-square testing in the examples using the three model types. Second, the assumption of uncorrelated latent class indicators across time conditional on the latent classes and the random intercept may be relaxed. Asparouhov and Muthén (2015) presented a method for this in a regular LTA setting, allowing correlated "residuals". Several instances of correlated residuals were found for these examples using both regular LTA and RI-LTA models. Third, with the use of a binary random intercept, RI-LTA can be generalized to more than two classes and more than one latent class variable. In the examples in this paper, however, there was no evidence that this was needed. Fourth, the model can be extended to include other model parts such as distal outcomes and multiple processes, the latter including the possibility to connect RI-LTA to the random intercept cross-lagged panel modeling of Hamaker et al. (2015). Fifth, a trend over time can be accomodated. In the continuous random intercept case, a slope can be added to the random intercept, e.g. by letting the slope influence the latent class indicators at each time point using the same loadings as for the random intercept and allowing a slope mean to influence the outcomes over time. Using a linear trend, this did not result in a better-fitting model for the examples of this paper.

7.3 Future research on RI-LTA

Despite the promising results obtained by replacing regular LTA with RI-LTA, further explorations and extensions of this new technique are warranted. It will be useful to have more extensive Monte Carlo simulation studies for different settings, studying the sample size requirements as a function of number of time points, number of latent class indicators, number of latent classes, covariates, etc. The susceptibility to model mis-specification should be studied. Class enumeration techniques need to be considered. It will be of interest to develop multi-step analyses for including covariates and distal outcomes in line with Asparouhov and

Muthén (2014) and Bakk and Kuha (2018). Multilevel versions of RI-LTA are needed when subjects are nested within schools, organizations, or communities.

References

[1] Altman, R, M. (2007). Mixed hidden Markov models: An extension of the hidden Markov model to the longitudinal data settting. Journal of the American Statistical Association, 102, 201-210

[2] Asparouhov, T. & Muthén, B. (2008). Multilevel mixture models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), Advances in latent variable mixture models, pp. 27-51. Charlotte, NC: Information Age Publishing, Inc.

[3] Asparouhov, T. & Muthén, B. (2014). Auxiliary variables in mixture modeling: Three-step approaches using Mplus. Structural Equation Modeling: A Multidisciplinary Journal, 21:3, 329-341.

[4] Asparouhov, T. & Muthén, B. (2015). Residual associations in latent class and latent transition analysis. Structural Equation Modeling: A Multidisciplinary Journal, 22:2, 169-177.

[5] Asparouhov, T. & Muthén, B. (2016). General random effect latent variable modeling: Random subjects, items, contexts, and parameters. in Advances in multilevel modeling for educational research: Addressing practical issues found in realworld applications, edited by J. R. Harring, L. M. Stapleton, and S. N. Beretvas. Charlotte, NC: Information Age Publishing, Inc.

[6] Asparouhov, T. & Muthén, B. (2019). Latent variable centering of predictors and mediators in multilevel and time-series models. Structural Equation Modeling: A Multidisciplinary Journal, 26, 119-142.

[7] Bakk, Z. & Kuha, J. (2018). Two-step estimation of models between latent classes and external variables. Psychometrika, 83, 871-892.

[8] Clogg, C.C. & Goodman, L.A. (1985). Simultaneous latent structural analysis in several groups. In Tuma, N.B. (ed.), Sociological Methodology, 1985 (pp. 81-110). San Francisco: Jossey-Bass Publishers.

[9] Cole, D.A., Martin. N.C., & Steiger, J.H. (2005). Empirical and conceptual problems with longitudinal trait-state models: Introducing a trait-state-occasion model. Psychological Methods, 10, 3-20.

[10] Collins, L. M., & Lanza, S. T. (2010). Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences. New York, NY: John Wiley & Sons, Inc.

[11] Collins, L.M. & Wugalter, S.E. (1992). Latent class models for stage-sequential clynamic latent variables. Multivariate Behavioral Research, 27, 13 1-1 57.

[12] de Jong, M. G. & J. B. E. M. Steenkamp. (2010). Finite Mixture Multilevel Multidimensional Ordinal IRT Models for Large-scale Cross-cultural Research. Psychometrika 75:3-32.

[13] de Jong, M. G., J. B. E. M. Steenkamp, & J. P. Fox. (2007). Relaxing Measurement Invariance in Cross-national Consumer Research Using a Hierarchical IRT Model. Journal of Consumer Research 34:260-78.

[14] Eid, M. & Langeheine, R. (1999). The measurement of consistency and occasion specificity with latent class models: A new model and its applicaton to the measurement of affect. Psychological Methods, 4, 100-116.

[15] Eid, M. & Langeheine, R. (2003). Separating stable from variable individuals in longitudinal studies by mixture distribution models. Measurement: Interdisciplinary Research and Perspectives, 1, 179-206.

[16] Eid, M., Holtmann, J., Santangelo, P. & Ebner-Priemer, U. (2017). On the definition of latent-state-trait models with autoregressive effects. European Journal of Psychological Assessment, 33, 285-295.

[17] Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. (2011). Applied Longitudinal Analysis, 2nd edition. Hoboken, NJ: Wiley.

[18] Fox, J. P. (2010). Bayesian Item Response Theory. New York: Springer.

[19] Graham, J.W., Collins, L.M., Wugalter, S.E., Chung, N.K., & Hansen, W.B. (1991). Modelling transitions in latent stage-sequential processes: a substance use prevention example. Journal of Consulting and Clinical Psychology, 59, 48-57.

[20] Hamaker, E. L., Kuiper, R. M. & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. Psychological Methods, 20, 102 - 116.

[21] Henry, K. L. & Muthén, B. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. Structural Equation Modeling, 17, 193-215.

[22] Jak, S., F. J. Oort, & C. V. Dolan. (2013). A Test for Cluster Bias: Detecting Violations of Measurement Invariance across Clusters in Multilevel Data. Structural Equation Modeling 20:265-82.

[23] Jak, S., F. J. Oort, & C. V. Dolan. (2014). Measurement Bias in Multilevel Data. Structural Equation Modeling: A Multidisciplinary Journal 21:31-39. doi:10.1080/10705511.2014.856694.

[24] Kaplan, D. (2008). An overview of Markov chain methods for the study of stage-sequential developmental processes. Developmental Psychology, 44, 457-467.

[25] Kenny, D. A. & Zautra, A. (1995). The trait-state-error model for multiwave data. Journal of Consulting and Clinical Psychology, 63, 52-59.

[26] Langeheine, R. & van de Pol, F. (2002). Latent Markov chains: In J. A. Hagenaars & A. L. McCutcheon (eds.), Applied latent class analysis (pp. 304-341). Cambridge, UK: Cambridge University Press.

[27] Lanza, S. T. & Collins, L. M. (2008). A new SAS procedure for latent transition analysis: Transitions in dating and sexual risk behavior. Developmental Psychology, 44, 446-456.

[28] Lubke, G.H. & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. Psychological Methods, 10, 21-39.

[29] MacDonald, I.L & Zucchini, W. (1997). Hidden Markov and Other Models for Discrete-valued Time Series. London: Chapman & Hall.

[30] Mooijaart, A. (1998). Log-linear and Markov modeling of categorical longitudinal data. In C. C. J. H. Bijleveld & T. van der Kamp (eds.), Longitudinal data analysis: Designs, models, and methods (pp. 318-370). Newbury Park, CA: Sage Publications.

[31] Muthén, B. & Asparouhov, T. (2006). Item response mixture modeling: Application to tobacco dependence criteria. Addictive Behaviors, 31, 1050-1066.

[32] Muthén, B. & Asparouhov, T. (2009). Growth mixture modeling: Analysis with non-Gaussian random effects. In Fitzmaurice, G., Davidian, M., Verbeke, G.& Molenberghs, G. (eds.), Longitudinal Data Analysis, pp. 143-165. Boca Raton: Chapman & Hall/CRC Press.

[33] Muthén, B. & Asparouhov, T. (2018). Recent methods for the study of measurement invariance with many groups: Alignment and random effects. Sociological Methods & Research, 47:4 637-664.

[34] Muthén, L. & Muthén, B. (1998-2017). Mplus User's Guide. Eighth edition. Los Angeles, CA: Muthén & Muthén.

[35] Nylund, K. L., Asparouhov, T. & 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, 535-569.

[36] Reboussin, B. A., Reboussin, D. M., Liang, K. Y. & Anthony, J. (1998). Latent transition modeling of progression of health-risk behavior. Multivariate Behavioral Research, 33, 457-478.

[37] Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6, 461-464.

[38] Vermunt, J.K. (2003). Multilevel latent class models. In R.M. Stoltzenberg (ed.), Sociological Methodology 2003 (pp. 213-239). Washington, D.C.: ASA.

[39] Vermunt, J.K. (2008). Latent class and finite mixture models for multilevel data sets. Statistical Methods in Medical Research, 17, 33-51.

[40] Wiggins, L .M. (1973). Panel Analysis. Amsterdam: Elsevier.

[41] Wooldridge, J. M. (2002). Econometrics analysis of cross sectional and panel data. Cambridge, MA: The MIT Press.