What Multilevel Modeling Can Teach Us About Single-Level Modeling: Latent Transition Analysis With Random Intercepts (RI-LTA)

Bengt Muthén and Tihomir Asparouhov *

bmuthen@statmodel.com

Submitted for publication

September 10, 2019

*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 four examples from the literature demonstrates the advantages of random intercept LTA. Model variations include Mover-Stayer analysis, multiple-group measurement invariance analysis, and analysis with covariates.

Key words: Hidden Markov, mixtures, transition probabilities, latent trait-state, two-level LCA, measurement non-invariance, Mover-Stayer.

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 four data sets from the LTA literature that will be used to demonstrate the advantage of RI-LTA over regular LTA. Section 3 describes the regular single-level LTA model and gives a critique of it. Section 4 discusses twolevel factor analysis and twolevel latent class analysis models which serve as background for the proposed RI-LTA in Section 5.

Section 6 shows applications of RI-LTA to the four data sets. Section 7 presents an extension to Mover-Stayer modeling. Section 8 discusses random intercept modeling extended to groups and covariates and shows applications. Section 9 concludes with a discussion of computational aspects, other model variations, and the need for further research.

2 Data sets

Following is a brief description of the four data sets from the LTA literature that will be used to demonstrate the advantage of RI-LTA over regular LTA. The data sets exemplify a variety of samples sizes (N), number of time points (T), number of outcomes (latent class indicators) per time point (R), and number of latent classes (J).

2.1 Life satisfaction (N = 5147, T = 5, R = 1, J = 2)

This data set is from the German Socio-Economic Panel with N=5147, 5 time points one year apart, and 1 binary latent class indicator measuring 2 latent classes at each time point (Langeheine & van de Pol, 2002). Survey respondents were asked "How satisfied are you on the whole with your life" with answer categories unsatisfied and satisfied.

2.2 Mood (N = 494, T = 4, R =2, J = 2)

This 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).

2.3 Reading proficiency (N=3574, T = 4, R = 5, J = 3)

This data set is from the Early Childhood Longitudinal Study with N=3574, 4 time points corresponding to Fall and Spring of Kindergarten and first grade, and 5 binary latent class indicators measuring 3 latent classes at each time point (Kaplan, 2008). The 5 indicators concern a stage-sequential process measured by the basic reading skills of letter recognition, beginning sounds, ending letter sounds, sight words, and words in context. A binary covariate indicates whether or not the child's household is above the poverty threshold.

2.4 Dating and sexual risk behavior (N = 2933, T = 3, R = 5, J = 5)

This data set is from the National Longitudinal Survey of Youth (NLSY97) with N=2937, 3 time points one year apart, and 5 ordinal and binary items measuring 5 latent classes at each time point (Lanza & Collins, 2008). 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.

Figure 1: LTA for 1 binary item at 5 time points

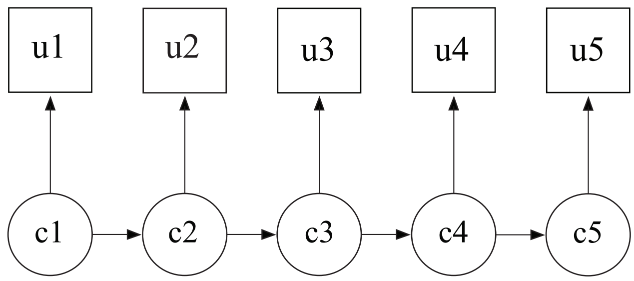
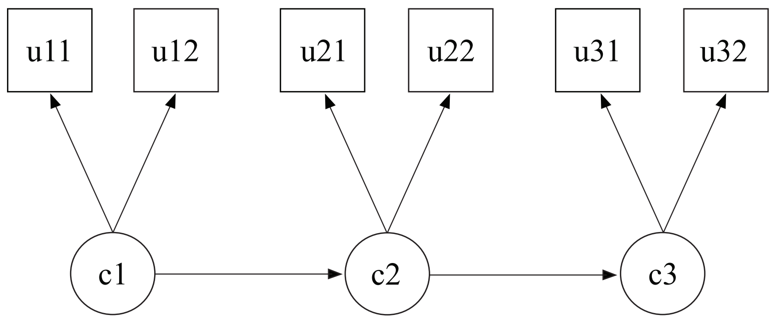


Figure 2: LTA for 2 binary items at 3 time points



3 Regular LTA

Figure 1 and Figure 2 show model diagrams for two types of regular LTA models. Figure 1 corresponds to the model for the Life satisfaction data with one binary indicator per time point and Figure 2 corresponds to the model for the Mood data with two binary indicators per time point.

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, $P(C_2|C_1)$, $P(C_3|C_2, C_1)$, 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; $P(C_t|C_{t-1} = 1)$, $P(C_t|C_{t-1} = 2)$, 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 the Life satisfaction example which corresponds to the Figure 1 model. 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). The bottom 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 difference in observed response probabilites for the two latent classes. For each row, the large difference in these probabilities

Table 1: Regular LTA estimates for the Life satisfaction example

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
Measurement probabilities		
Observed	Latent class	
response	Unsatisfied	Satisfied
Unsatisfied	0.855	0.163
Satisfied	0.145	0.837

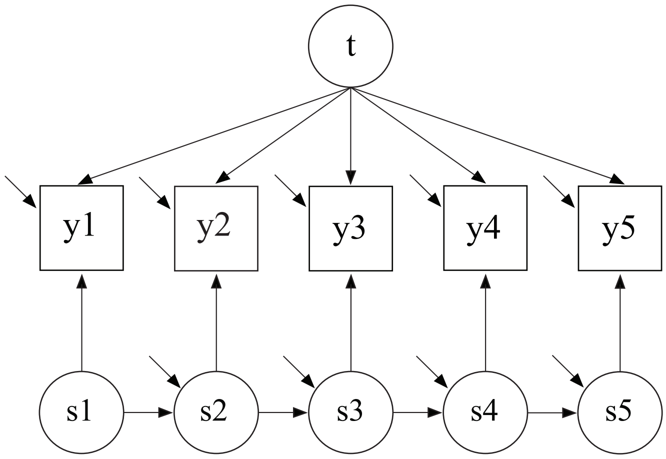
shows that the latent class indicators clearly distinguished between the two latent classes. The off-diagonal probabilities can be seen as "measurement error" in that membership in a certain class does not necessitate an answer in the corresponding response category (Wiggins, 1973). This discrepancy 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).

3.1 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 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

Figure 3: Latent trait-state model (Kenny & Zautra, 1995)



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 not possible. 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 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, because LTA typically considers several indicators of the latent class variables, measurement invariance/non-invariance across subjects needs to be considered. Allowing for a degree of measurement non-invariance should be of primary concern when studying structural relations, in this case relationships between the latent class variables. Second, it is of interest to study how much the latent transition probabilities are distorted in regular LTA when stable between-subject differences are ignored.

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.

3.2 A hypothetical example

A simple hypothetical example illustrates the effects of ignoring stable between-subject differences. The model that the data correspond to will be formally specified in Section 5 but can be conceptualized in line with Figure 3 although with T replaced by a binary random intercept variable I and the S's replaced by binary latent class variables C_t. In this example, there are 5 binary indicators of the latent class variable C_t at each time point, 2 time points, and 2 latent classes for C_t. The single binary latent class random intercept variable I influences all 5 indicators equally. The latent class variable probabilities and the transition probabilities are presented below. The random intercept variable I has a 73-27 split representing two types/classes of subjects. A continuous random intercept variable may be more realistic in many cases but as will be seen in the real-data analyses, the binary random intercept variable can serve as a reasonable approximation and makes the example simpler. For the 73% majority class, the probability of endorsing each item is 0.27 when in C_t class 1 and 0.73 when in C_t class 2. In this way, the latent class indicators clearly distinguish between the two C_t classes for the 73% majority class. For the 27% minority class, the probability of endorsing each item is 0.62 when in C_t class 1 and 0.92 when in C_t class 2. The latent class indicators distinguish between the two C_t classes also for the minority but at a higher level of endorsement probabilities. This suggests that subjects in the minority class interpret the latent class indicator questions differently than the majority class or have a different response style, perhaps related to their background characteristics. As will be shown in in Section 4.2, the measurement difference illustrated by the binary random intercept is in line with the multilevel latent class example in Henry and Muthén (2010) where different response behaviors are observed in different types of communities. The top part of Table 2 gives the population values of the probabilities of the latent class variable C_t and the transition probabilities. These probabilities are the same for the two types of subjects, that is, the structural part of the model is not affected by its measurement part.

A Monte Carlo study is carried out where data are generated according to this model for a sample of $N = 3000$. This sample size is in line with those of the Reading proficiency example and the Dating and sexual risk behavior example. Using 500 replications, the data are analyzed by maximum likelihood both using

Table 2: Latent class and transition probabilities for a hypothetical example using an RI-LTA model with a binary random intercept (standard errors are given in parentheses)

Population values for RI-LTA, binary RI model

	Class 1	Class 2
Class probabilities		
Time 1	0.500	0.500
Time 2	0.561	0.439
Transition probabilities		
Class 1	0.622	0.378
Class 2	0.500	0.500

Estimated regular LTA model

	Class 1	Class 2
Class probabilities		
Time 1	0.446 (.016)	0.554 (.016)
Time 2	0.497 (.016)	0.503 (.016)
Transition probabilities		
Class 1	0.670 (.021)	0.330 (.021)
Class 2	0.358 (.018)	0.642 (.018)

Estimated RI-LTA, binary RI model

	Class 1	Class 2
Class probabilities		
Time 1	0.498 (.035)	0.502 (.035)
Time 2	0.558 (.037)	0.442 (.037)
Transition probabilities		
Class 1	0.624 (.031)	0.376 (.031)
Class 2	0.497 (.068)	0.503 (.068)

the regular LTA which ignores the random intercept variable and using the correct random intercept LTA model. The focus of this example is how the transition probabilities are affected when using regular LTA on data generated by a random intercept LTA model.

The second part of Table 2 gives the estimated values using the regular LTA model. In parentheses are given the estimated standard errors obtained as the average standard error across the 500 replications. The table shows that regular LTA obtains biased parameter estimates. In particular, it obtains too high diagonal values for the transition probabilities, that is, it overstates the stability of class membership over time. The largest bias is for class 2 where the true model says that it is equally likely for a subject to transition to class 1 as it is to stay in class 2 while regular LTA says that it is distinctly less likely to transition than to stay. The time 1 latent class probabilities are somewhat biased as well (the time 2 latent class probability bias is a function also of the bias in the transition probabilities).

The bottom part of Table 2 shows the results for the RI-LTA which uses the correct model for the generated data. This shows that the population parameter values are well recovered. The standard errors are well estimated and the coverage values are also good (not shown). Comparing the standard errors to those of regular LTA shows that the standard errors are underestimated by regular LTA. For the class 2 transition probabilities, they are underestimated by a factor greater than 3. This underestimation of standard errors is in line with ignoring cluster effects in two-level data (see, e.g., Muthén & Satorra, 1995).

This hypothetical example illustrates the problem of regular LTA which ignores stable between-subject differences, in this case represented by a majority and a minority class of subjects. A researcher may argue that sometimes the interest is in the overall picture and not in the separate mixture components of different subject classes. However, the two classes have the same transition probabilities so that the mixture of the two obtained by regular LTA is not correct for either subject class but instead an uninterpretable blend is obtained.

The potential for bias in regular LTA is clearly seen in this hypothetical example. It remains to be seen, however, if this is a common phenomenon or not in real data. This will be studied in Section 6 where the four different data sets from the LTA literature are analyzed. Readers with a main interest in the applications may go straight to this section. In the next section, however, the multilevel background for the random intercept idea is discussed, followed by Section 5 which presents the details of the proposed random intercept LTA.

4 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 twolevel model. This idea will be approached in two steps, considering a twolevel factor analysis model and a twolevel latent class analysis model. This gives the background for the proposed random intercept LTA.

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. Decomposing the factor variation into within and between components as $f_{ij} = f_{W_{ij}} + f_{B_j}$, the model can be expressed by the two equations

$$logit P(U_{ij} = 1 | f_{W_{ij}}) = \nu_j + \lambda_W f_{W_{ij}}, \tag{3}$$

$$\nu_j = \nu + \lambda_B f_{B_j} + \epsilon_{Bj}, \tag{4}$$

corresponding to the within- and between-level parts of a multilevel model. This is in line with twolevel regression where the intercept ν_j is random, varying across schools. The fact that 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 Figure 4 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 continuous latent variables. The ϵ_B residual on the between level is left out in the figure because it is often close to zero. The extraction of between-level variation ensures that using the factor as a predictor on the within level does not confound its effect by between-level variation, that is, using f_W as the independent variable, not $f = f_W + f_B$.

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).

Figure 4: Multilevel factor analysis

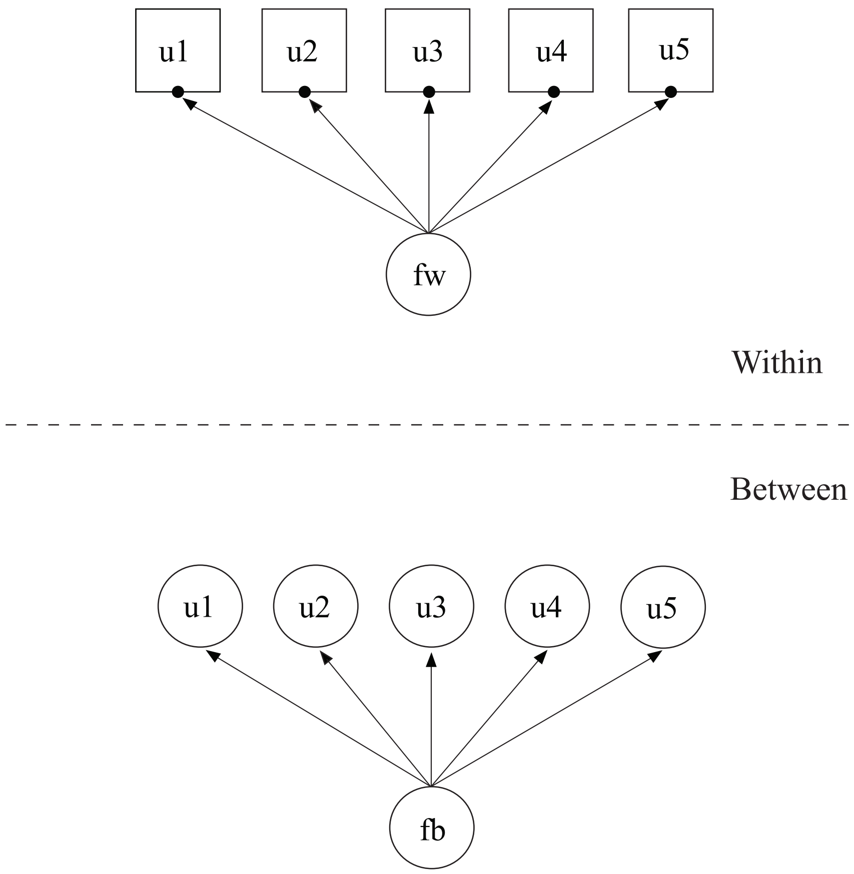
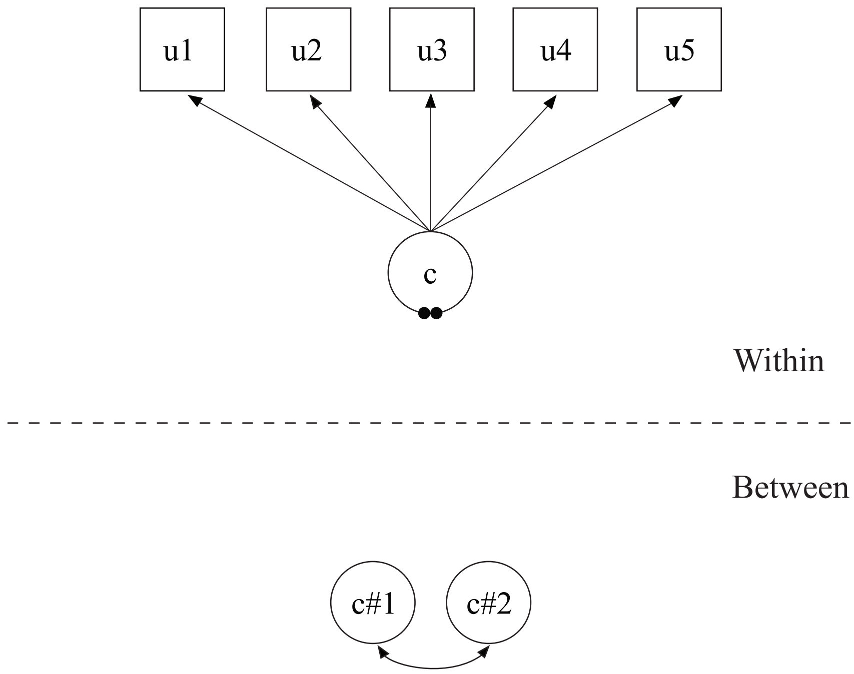


Figure 5: Twolevel LCA for 5 binary items and 3 latent classes with latent class random intercepts



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 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. 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 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

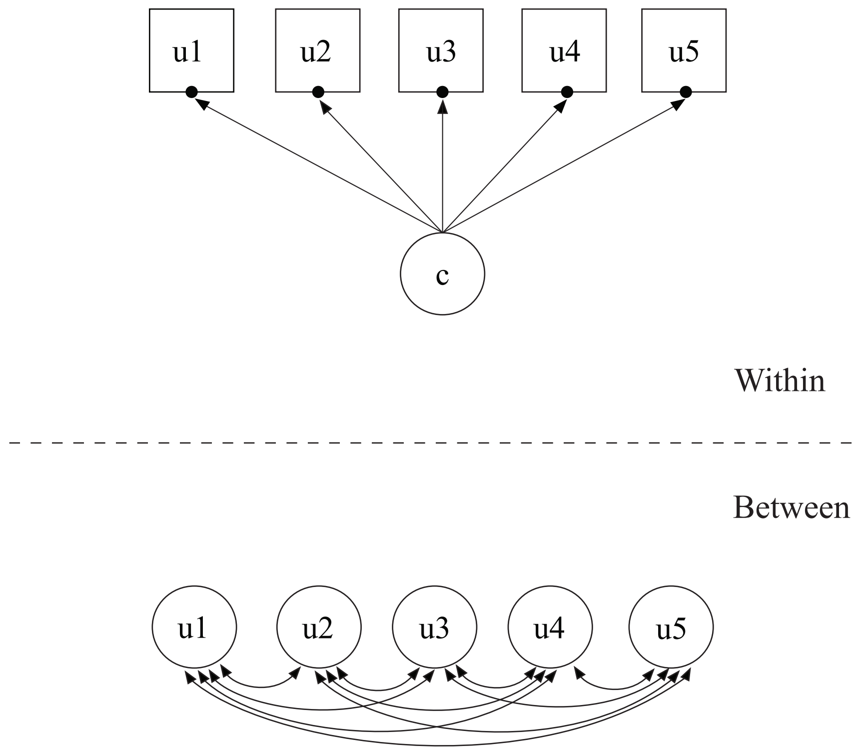
$$logit P(U_{ij} = 1 | C_{ij} = c) = \alpha_{cj} = \alpha_c + \epsilon_j, \tag{5}$$

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 Figure 6. 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

Figure 6: Twolevel LCA for 5 binary items with latent class indicator random intercepts

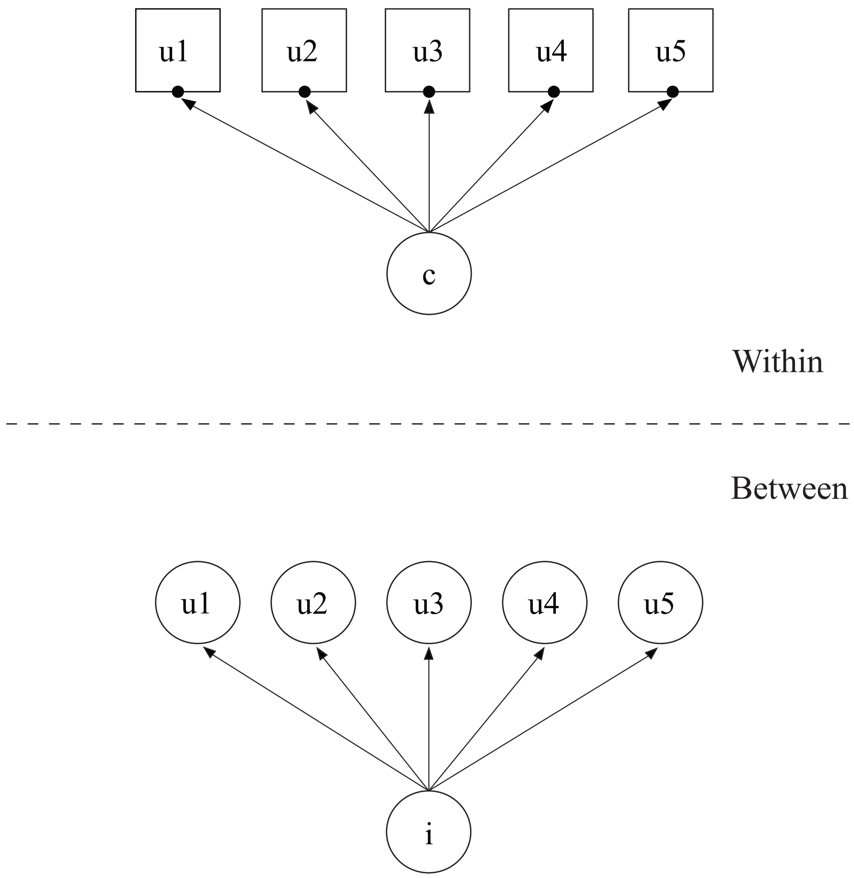


(a continuous latent variable) behind the set of latent variables as shown by the i intercept factor on the between level of Figure 7. With zero residuals, this reduces the numerical integration to 1 dimension while allowing the random intercepts to correlate and estimating their factor loadings. A non-parametric 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 will be referred to as using random intercepts.

It is possible to use random intercepts for both the latent class indicators and the latent classes, that is, a combination of Figure 5 and Figure 6. Such a twolevel LCA, however, requires large cluster sizes for the parameters to be well defined. In longitudinal settings where cluster size refers to the number of time points, Asparouhov and Muthén (2019) found that at least 10-20 time points were needed. In typical LTA applications, however, there are only 2-5 time points. Also, when a random intercept is specified to influence the latent class variables, the transitions refer to latent class variables that contain between-subject variation, thereby losing the between-within separation. For these reasons, the current paper proposes random intercepts for the latent class indicators only.

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

Figure 7: Twolevel LCA for 5 binary items with a factor for latent class indicator random intercepts



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.

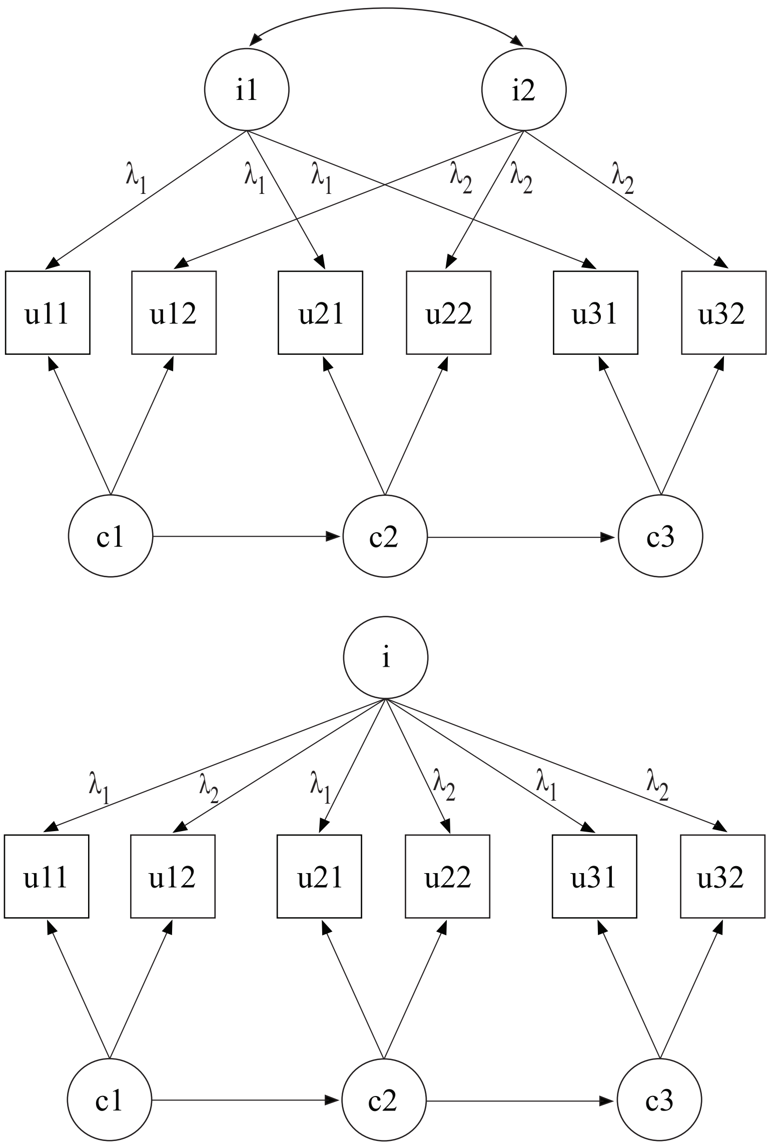
5 Random intercept LTA (RI-LTA)

Random intercept LTA (RI-LTA) borrows from the idea of indicator-level random intercepts shown in Figure 4, Figure 6, and Figure 7. Using a single-level wide analysis format, Figure 8 shows two versions of continuous random intercept RI-LTA for 2 binary latent class indicators measured at 3 time points as in the Mood example. The random intercepts $i1$, $i2$, i are continuous latent variables where the loadings λ capture their different influence on the 2 latent class indicators. Each indicator's loading is held equal across time. The two model versions correspond to the between-level part Figure 6 and Figure 7, respectively, in that each indicator has either its own random intercept or share the same random intercept factor that has different effects on the two latent class indicators. Note that the random intercept latent variables u on the between level in Figure 7 are not needed in the single-level setting because the i variable can point directly to the observed latent class indicators. In other words, the latent intercept variable i in the bottom part of Figure 8 can be seen as the counterpart to the i factor in Figure 7. It is interesting that the model in the bottom part of Figure 8 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 and the $c_1 - c_3$ model part represents the within-subject variation across time.

With a continuous intercept variable and the single random intercept factor version for R latent class indicators per time point, only R parameters are added namely the intercept factor loadings 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.

Consider next the version of RI-LTA that has a binary random intercept represented by a latent class variable. A simple model version expressed in logit terms uses the following parameterization for a binary latent class indicator U_t at time t, latent class j at time t for the latent class variable C, and latent class k

Figure 8: RI-LTA for 2 binary latent class indicators at 3 time points with a random intercept: 2 continuous random intercepts versus 1



for a single random intercept latent class variable I,

$$logit[P(U_t = 1|C_t = j, I = k)] = \alpha + \beta_j + \gamma_k, \tag{6}$$

where $\beta_1 = 0$, $\gamma_1 = 0$ for identification purposes. Here, α is a parameter specific to the latent class indicator, β_j is a parameter specific to the latent class indicator as well as the latent class of C, and γ_k is a parameter specific to the latent class indicator as well as the latent class 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_t at time t are

$$logit[P(U_t = 1|C_t = 1, I = 1)] = \alpha \tag{7}$$

$$logit[P(U_t = 1|C_t = 2, I = 1)] = \alpha + \beta_2, \tag{8}$$

$$logit[P(U_t = 1|C_t = 3, I = 1)] = \alpha + \beta_3, \tag{9}$$

$$logit[P(U_t = 1|C_t = 1, I = 2)] = \alpha + \gamma_2, \tag{10}$$

$$logit[P(U_t = 1|C_t = 2, I = 2)] = \alpha + \beta_2 + \gamma_2, \tag{11}$$

$$logit[P(U_t = 1|C_t = 3, I = 2)] = \alpha + \beta_3 + \gamma_2. \tag{12}$$

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 only 2 latent classes for I, J latent classes for C_t, and R latent class indicators per time point, 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 due to assuming a common shift for all response categories. This is the parameterization used in Eid and Langeheine (1999).

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).

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 8 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 a second-order latent class variable to represent different transition matrices. For each second-order latent class, transitions can be viewed as a within-subject process. The mover-stayer model (see, e.g. Langeheine & van der Pol, 2002) is an example of this with a particular structure for the transitions and it will be studied in Section 7. Observed between-subject variation can be studied using groups and covariates and will be discussed next.

5.1 Groups and covariates

5.1.1 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

Table 3: 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	$\alpha_1 + \beta_{11} + \gamma_{11}$x	$\alpha_2 + \beta_{21} + \gamma_{21}$x	0
	2	$\alpha_1 + \beta_{12} + \gamma_{12}$x	$\alpha_2 + \beta_{22} + \gamma_{22}$x	0
	3	$\alpha_1 + \gamma_{13}$x	$\alpha_2 + \gamma_{23}$x	0

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

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. An alternative approach is to let covariates representing subject characteristics such as gender, ethnicity, ses, and age influence the latent class variables as well as their transition probabilities. In this paper, analysis with covariates is carried out using the logit parameterizations shown in Table 3 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 $\alpha_J = 0$, $\beta_{Jk} = 0$, $\beta_{c,J} = 0$, $\gamma_{Jk} x = 0$. 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 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 3 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.

5.1.2 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.

6 Analyses of the four examples

As a first step, analysis of the four examples listed in Section 2 is described in terms of model fit, comparing regular LTA with RI-LTA using both a binary random intercept and a continuous random intercept. Next, latent class probability estimates are presented and compared between the models. Finally, the measurement estimates are discussed, showing the new information obtained by the RI-LTA approaches. All analyses are carried out using Mplus (Muthén & Muthén, 1998-2017) and scripts are available from the first author.

6.1 Model fit

Table 4 shows the model fitting results for the four examples. The choice of model will be based on BIC (smaller values are better). Because of differences in the number of latent classes and whether or not a continuous random intercept is present, models cannot be compared using likelihood-ratio chi-square. 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 exception being the Life satisfaction example which has only 32 cells and a large sample. For the RI-LTA with a binary random intercept, the parameterization of (6) - (12) is used. For the RI-LTA with a continuous random intercept, the simple model version shown in the bottom part of Figure 8 is used. A non-stationary model is chosen for the Life satisfaction and Reading proficiency examples because stationarity was rejected by a likelihoo-ratio chi-square difference test. A stationary model is chosen for the Mood example because this is the model considered in Eid and Langeheine (2003). A stationary model is also chosen for the Dating and sexual risk behavior example because unlike for regular LTA as in Lanza and Collins (2008), stationarity cannot be rejected for the RI-LTA models.

For all four examples, Table 4 shows that the RI-LTA with a continuous random intercept is preferrable based on BIC. The improvement in BIC is especially noteworthy for the Reading proficiency example (in Kaplan, 2008, a priori zero transition probabilities were specified for the lower triangle entries but are freely estimated here due to better BIC). For the Dating and sexual risk behavior example, the likelihood values are the same for the binary and continuous versions of the random intercept and the continuous case wins out in terms of

Table 4: Model fitting results

Model	# parameters	loglikelihood	BIC
Life Satisfaction (non-stationary). N=5147, T=5, R=1, J=5			
Regular LTA	11	-15326	30745
RI-LTA, binary RI	13	-15268	30646
RI-LTA, continuous RI	12	-15267	30637
Mood (stationary). N=494, T=4, R=2, J=2			
Regular LTA	7	-2053	4150
RI-LTA, binary RI	10	-2028	4118
RI-LTA, continuous RI	9	-2018	4093
Reading proficiency (non-stationary). N=3574, T=4, R=5, J=3			
Regular LTA	35	-21793	43873
RI-LTA, binary RI	41	-20916	42167
RI-LTA, continuous RI	40	-20329	40984
Dating and sexual risk behavior (stationary). N=2933, T=3, R=5, J=5			
Regular LTA	54	-16720	33871
RI-LTA, binary RI	59	-16580	33631
RI-LTA, continuous RI	58	-16580	33623

BIC merely due to using one parameter less. In the non-stationary version of this model, the continuous random intercept model encounters a problem of exploding loadings for the random intercept factor where the best loglikelihood value cannot be replicated. Discarding the solution with such inadmissible parameter values, the next best loglikelihood value is replicated with no loading problems. In this case, the continuous random intercept model has a somewhat better loglikelihood value and better BIC than the binary random intercept model.

It is instructive to consider in more detail the Life satisfaction example, comparing the regular LTA model and the RI-LTA with a continuous random intercept with respect to the fit to the 32 cells of the frequency table. For the regular LTA model, the model test of fit is 130.25 using the likelihood-ratio chi-square test for the frequency table and 131.03 using the Pearson chi-square. With 20 degrees of freedom, the regular LTA model is clearly rejected. Adding just one parameter, the RI-LTA obtains a dramatic improvement in chi-square fit with values of 13.62 and 13.64, respectively, with 19 degrees of freedom. Table 5 shows how this improvement in fit is obtained by listing the observed and estimated frequencies for each of the 32 response patterns for both models. The estimated frequencies track the observed ones much better for the RI-LTA. While the regular LTA has 14 instances of standardized residual z-tests greater than 2, RI-LTA has none.

The Reading proficiency and Dating and sexual risk behavior examples which use three and five latent classes, respectively, raise the question if fewer latent classes can be used when analyzed by RI-LTA instead of regular LTA. As judged by BIC, this was not the case, however.

Table 5: Model fit for the Life Satisfaction example

Response #	Response pattern	Observed frequency	Regular LTA estimate	Regular LTA z-score	RI-LTA, cont's RI estimate	RI-LTA, cont's RI z-score
1	00000	891	786.41	4.05	877.71	0.49
2	00001	176	222.59	-3.19	177.52	-0.12
3	00010	119	149.75	-2.55	123.54	-0.41
4	00011	106	120.41	-1.33	104.40	0.16
5	00100	111	139.99	-2.48	111.77	-0.07
6	00101	60	53.67	0.87	60.55	-0.07
7	00110	52	40.97	1.73	55.99	-0.54
8	00111	92	97.37	-0.55	97.86	-0.60
9	01000	120	151.88	-2.63	135.21	-1.33
10	01001	64	49.54	2.06	60.08	0.51
11	01010	51	33.59	2.59	45.33	0.85
12	01011	187	209.53	-1.59	179.77	0.55
13	01100	54	41.62	1.93	58.38	-0.58
14	01101	50	48.08	0.28	53.79	-0.52
15	01110	49	44.92	0.61	52.40	-0.47
16	01111	176	202.62	-1.91	164.81	0.89
17	10000	237	290.08	-3.21	242.85	-0.38
18	10001	107	98.42	1.99	110.90	-0.37
19	10010	68	61.67	0.81	71.80	-0.45
20	10011	107	78.55	3.23	110.67	-0.35
21	10100	80	65.70	1.78	72.45	0.89
22	10101	75	56.17	2.53	67.77	0.88
23	10110	51	50.80	0.03	59.83	-1.15
24	10111	200	213.19	-0.92	187.57	0.92
25	11000	136	142.83	-0.58	127.40	0.77
26	11001	95	75.14	2.31	98.93	-0.40
27	11010	64	62.57	0.18	68.96	-0.60
28	11011	187	209.53	-1.59	179.77	0.55
29	11100	99	102.73	-0.37	103.01	-0.40
30	11101	165	209.71	-3.15	162.02	0.24
31	11110	172	203.64	-2.26	154.23	1.45
32	11111	1066	992.23	2.61	1084.81	-0.64

6.2 Latent class estimates

Table 6 - Table 9 show the estimated latent class and transition probabilities for each of the four examples and each of the three models. For the Life satisfaction example of Table 6, the high degree of stability seen in the regular LTA solution is reduced in the two better-fitting random intercept models with the continuous random intercept model producing much higher probabilities for transitioning between the Unsatisfied and Satisfied latent classes (0.305 and 0.268 versus 0.126 and 0.000). Note also that the probability in the Satisfied class is higher at both time points for the continuous random intercept model. A similar picture emerges in Table 7 for the Mood example.

In the Table 8 Reading proficiency example, the 3 classes Low, Medium and High correspond to low alphabet knowledge, early word reading, and early reading comprehension (Kaplan, 2008; p. 464). In this example, the continuous random intercept model was strongly preferred and shows several transition probabilities that are different from the regular LTA. Here, the diagonal elements of the transition tables are not uniformly higher for regular LTA compared to the continuous random intercept model. Transitioning from Fall to Spring in Kindergarten, regular LTA underestimates the probability of moving from the Low to the Medium class and underestimates the probability of staying in the Medium class. A similar picture emerges for transitions between Spring Kindergarten and Fall of 1st grade. For the transition table for Fall 1st grade and Spring 1st grade, the regular LTA underestimates the probability of transitioning from the Low to the High class. Note also the distinct difference in latent class probabilities. Compared to the regular LTA, the random intercept model has

Table 6: Latent class and transition probabilities for Life satisfaction example (transitions from first to second wave)

Regular LTA		
	Unsatisfied	Satisfied
Class probabilities		
Time 1	0.395	0.605
Time 2	0.471	0.529
Transition probabilities		
Unsatisfied	1.000	0.000
Satisfied	0.126	0.874
RI-LTA, binary RI		
	Unsatisfied	Satisfied
Class probabilities		
Time 1	0.362	0.638
Time 2	0.474	0.526
Transition probabilities		
Unsatisfied	0.953	0.047
Satisfied	0.203	0.797
RI-LTA, continuous RI		
	Unsatisfied	Satisfied
Class probabilities		
Time 1	0.294	0.706
Time 2	0.430	0.570
Transition probabilities		
Unsatisfied	0.732	0.268
Satisfied	0.305	0.695

Table 7: Latent class and transition probabilities for Mood example

Regular LTA		
	Sad/Unhappy	Not sad/Happy
Class probabilities		
Time 1	0.492	0.508
Time 2	0.470	0.530
Time 3	0.458	0.542
Time 4	0.451	0.549
Transition probabilities		
Sad/Unhappy	0.752	0.248
Not sad/Happy	0.197	0.803

RI-LTA, binary RI		
	Sad/Unhappy	Not sad/Happy
Class probabilities		
Time 1	0.444	0.556
Time 2	0.419	0.581
Time 3	0.410	0.590
Time 4	0.407	0.593
Transition probabilities		
Sad/Unhappy	0.626	0.374
Not sad/Happy	0.254	0.746

RI-LTA, continuous RI		
	Sad/Unhappy	Not sad/Happy
Class probabilities		
Time 1	0.426	0.574
Time 2	0.384	0.616
Time 3	0.375	0.625
Time 4	0.373	0.627
Transition probabilities		
Sad/Unhappy	0.507	0.493
Not sad/Happy	0.293	0.707

Table 8: Latent class and transition probabilities for Reading proficiency example

	Regular LTA								
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Class Probabilities									
Time 1	0.694	0.284	0.023						
Time 2	0.235	0.635	0.130						
Time 3	0.142	0.627	0.232						
Time 4	0.041	0.154	0.805						
Transition Probabilities	Fall K -– Spring K			Spring K -– Fall 1st			Fall 1st -– Spring 1st		
Low	**0.338**	0.649	0.012	**0.596**	0.401	0.002	**0.263**	0.505	0.232
Medium	0.001	**0.652**	0.348	0.002	**0.837**	0.161	0.005	**0.132**	0.863
High	0.000	0.000	**1.000**	0.002	0.003	**0.994**	0.001	0.000	**0.999**

	RI-LTA, binary RI								
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Class Probabilities									
Time 1	0.824	0.162	0.014						
Time 2	0.208	0.692	0.100						
Time 3	0.122	0.723	0.155						
Time 4	0.041	0.065	0.894						
Transition Probabilities	Fall K -– Spring K			Spring K -– Fall 1st			Fall 1st -– Spring 1st		
Low	**0.253**	0.738	0.010	**0.584**	0.404	0.012	**0.308**	0.464	0.228
Medium	0.000	**0.520**	0.480	0.000	**0.923**	0.077	0.004	**0.012**	0.984
High	0.000	0.000	**1.000**	0.004	0.000	**0.996**	0.003	0.000	**0.997**

	RI-LTA, continuous RI								
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Class Probabilities									
Time 1	0.948	0.049	0.003						
Time 2	0.161	0.818	0.022						
Time 3	0.040	0.880	0.080						
Time 4	0.010	0.017	0.973						
Transition Probabilities	Fall K -– Spring K			Spring K -– Fall 1st			Fall 1st -– Spring 1st		
Low	**0.170**	0.820	0.010	**0.240**	0.742	0.018	**0.154**	0.000	0.845
Medium	0.000	**0.819**	0.181	0.001	**0.931**	0.068	0.004	**0.019**	0.977
High	0.000	0.000	**1.000**	0.022	0.000	**0.978**	0.009	0.000	**0.991**

a higher probability of being in the Low class in Fall of Kindergarten, a higher probability of being in the medium class in Spring of Kindergarten and Fall of 1st grade, and a higher probability of being in the high class in Spring of 1st grade.

In the Table 9 Dating and sexual risk behavior example, the major differences in the transition probabilities appear for the last two latent classes. Both refer to having multiple sexual partners, differing in whether they were exposed to sexually transmitted diseases (Multi-exposed) or not (Multi-safe). Regular LTA overestimates the probability of staying in the Multi-safe latent class and underestimates the probability of staying in the Multi-exposed class. The time 3 probability of the Multi-exposed class is higher for the random intercept models than for regular LTA.

6.3 Measurement Estimates

The measurement model determines the interpretation of the latent classes. The measurement model of regular LTA is the conditional distribution $[U_t|C_t]$ for the outcome U_t at time t conditioned on the latent class variable C_t at the same time point. With RI-LTA, the distribution of U_t depends not only on C_t but also on the random intercept I. To compare to regular LTA, the distribution $[U_t|C_t]$ is obtained when integrating (for a continuous random intercept) or summing (for a latent class random intercept) over the random intercept I in the distribution $[U_t|C_t, I]$. An advantage of RI-LTA is that it also makes it possible to study how the subject variation captured by the random intercept creates different measurement models. When I is a latent class variable, $[U_t|C_t]$ can be studied for each latent class of I to understand the differences in measurement models for the

Table 9: Latent class and transition probabilities for Dating and sexual risk behavior example

	Regular LTA				
	Nondaters	Daters	Monogamous	Multi-safe	Multi-exposed
Class probabilities					
Time 1	0.179	0.295	0.115	0.236	0.174
Time 2	0.134	0.229	0.223	0.209	0.206
Time 3	0.108	0.187	0.280	0.183	0.242
Transition probabilities					
Nondaters	**0.618**	0.177	0.108	0.078	0.019
Daters	0.023	**0.553**	0.171	0.194	0.059
Monogamous	0.040	0.049	**0.665**	0.054	0.192
Multi-safe	0.040	0.104	0.170	**0.557**	0.128
Multi-exposed	0.013	0.020	0.207	0.000	**0.760**
	RI-LTA, binary RI				
	Nondaters	Daters	Monogamous	Multi-safe	Multi-exposed
Class probabilities					
Time 1	0.205	0.237	0.076	0.216	0.236
Time 2	0.156	0.206	0.173	0.191	0.275
Time 3	0.127	0.166	0.214	0.177	0.316
Transition probabilities					
Nondaters	**0.592**	0.181	0.092	0.129	0.006
Daters	0.054	**0.515**	0.103	0.241	0.088
Monogamous	0.049	0.028	**0.542**	0.167	0.214
Multi-safe	0.060	0.112	0.236	**0.408**	0.184
Multi-exposed	0.014	0.021	0.145	0.000	**0.820**
	RI-LTA, continuous RI				
	Nondaters	Daters	Monogamous	Multi-safe	Multi-exposed
Class probabilities					
Time 1	0.189	0.283	0.090	0.199	0.239
Time 2	0.143	0.220	0.204	0.152	0.281
Time 3	0.116	0.177	0.260	0.123	0.325
Transition probabilities					
Nondaters	**0.620**	0.166	0.087	0.117	0.011
Daters	0.026	**0.548**	0.138	0.197	0.091
Monogamous	0.040	0.035	**0.642**	0.040	0.243
Multi-safe	0.058	0.126	0.247	**0.357**	0.212
Multi-exposed	0.016	0.023	0.172	0.000	**0.789**

different random intercept classes. When I is a continuous latent variable, the variation in the measurement model can be studied for example at one standard deviation away from the mean of zero.

6.3.1 The Reading proficiency example

The Reading proficiency example provides a good illustration of how to interpret the measurement model using RI-LTA. It is of special interest because both RI models fit considerably better than the regular LTA model as seen in Table 4 but the meaning of the latent classes of C is not fundamentally changed. Instead, a richer understanding is provided of how the responses to certain latent class indicators vary across subjects.

It is interesting to compare regular LTA with RI-LTA using a continuous random intercept. The random intercept factor loadings for the latent class indicators are positive, significant and of similar magnitude. The top part of Table 10 shows the estimated conditional probabilities $P(U_t = 1|C_t)$ using regular LTA on the left and using continuous random intercept RI-LTA on the right. The RI-LTA estimates are obtained by integrating out the random intercept. It is seen that these two sets of estimates give the same interpretation. In the low class of low alphabet knowledge (Class 1), the only moderately high probability of mastery is seen for letter recognition. Compared to the low class, the medium class (Class 2) of early word reading adds high probabilities of mastery also for beginning sounds and ending letter sounds. Compared to the medium class, the high class (Class 3) of early reading comprehension adds a high probability of mastery also for sight words and a moderately high probability for words in context (WIC).

The middle part of Table 10 shows the measurement model at one standard

Table 10: Estimated measurement model for Reading proficiency example (Class 1 = low alphabet knowledge, Class 2 = early word reading, Class 3 = early reading comprehension)

	Regular LTA			RI-LTA, continuous RI		
	Classes			Classes		
	1	2	3	1	2	3
Letrec	0.505	0.994	1.000	0.627	0.939	0.979
Begin	0.066	0.917	0.984	0.303	0.806	0.941
Ending	0.013	0.661	0.972	0.167	0.630	0.904
Sight	0.000	0.051	0.985	0.020	0.208	0.808
WIC	0.000	0.001	0.509	0.005	0.058	0.460
	RI-LTA, continuous RI					
	-1 SD			+1 SD		
Letrec	**0.097**	0.931	0.990	0.992	1.000	1.000
Begin	0.012	0.515	0.915	0.744	0.996	1.000
Ending	0.004	**0.176**	0.817	0.391	0.974	0.999
Sight	0.000	**0.001**	**0.464**	0.008	0.592	0.999
WIC	0.000	0.000	**0.015**	0.001	0.059	0.965
	RI-LTA, continuous RI, poverty covariate					
	Poverty (19%)			Non-poverty (81%)		
Letrec	**0.220**	0.973	0.996	0.861	0.999	1.000
Begin	0.025	0.692	0.960	0.227	0.962	0.996
Ending	0.008	**0.302**	0.903	0.071	0.807	0.989
Sight	0.000	0.002	0.706	0.000	0.061	0.984
WIC	0.000	0.000	**0.042**	0.000	0.003	0.544

deviation away from the mean of the random intercept. The bold entries of the -1 SD case indicate instances of large differences in probabilities compared to the probabilities in the top part. For each of the three classes, the probabilities are distinctly lower for the more advanced topics/indicators of the class. A similar picture is seen in the bottom part of the table which is based on a regression of the random intercept variable on the binary covariate poverty (modeling with covariates is discussed in Section 8). The random intercept is strongly related to the poverty covariate and the low and high random intercept values are obtained as the means of the random intercept for poverty=1 and 0, respectively. It is seen that the measurement model for poverty=1 has the same pattern of probabilities as the -1 SD case. In this way, the random intercept variable can be thought of as an achievement or reading preparedness dimension on which subjects vary.

6.3.2 The Dating and sexual risk behavior example

The Dating and sexual risk behavior example illustrates how the measurement model varies over the two latent classes of a binary random intercept. The estimated random intercept latent class percentages are 68 and 32. The corresponding two sets of measurement estimates are shown in Table 11. The interpretation of the measurement model is similar to that given in Lanza and Collins (2008) where the 5 latent classes are described as Nondaters (I), Daters (II), Monogamous (III), Multipartner safe (IV), and Multipartner exposed (V). Some details of the interpretations of the 5 classes are, however, challenged which is noteworthy given the Table 4 finding that this RI-LTA has a considerably better BIC than the regular LTA.

For the majority class (68 %), Multipartner safe (IV) is no longer clearly

Table 11: Estimated measurement model for the Dating and sexual risk behavior example using RI-LTA with a binary random intercept

	Random intercept class 1 (68%)					Random intercept class 2 (32%)				
	I	II	III	IV	V	I	II	III	IV	V
# dating partners in past year										
0	0.900	0.000	0.126	0.137	0.016	0.267	0.000	0.006	0.006	0.001
1	0.095	0.193	0.874	0.023	0.053	0.617	0.010	0.994	0.001	0.002
>=2	0.005	0.807	0.000	0.840	0.931	0.115	0.990	0.000	0.992	0.997
Had sex in past year										
No	0.979	0.999	0.000	0.000	0.000	0.986	0.000	0.000	0.000	0.000
Yes	0.021	0.001	1.000	1.000	1.000	0.014	1.000	1.000	1.000	1.000
# sex partners in past year										
0	1.000	1.000	0.000	0.033	0.000	1.000	1.000	0.000	0.029	0.000
1	0.000	0.000	0.952	0.580	0.118	0.000	0.000	0.946	0.550	0.104
>=2	0.000	0.000	0.048	0.388	0.882	0.000	0.000	0.054	0.421	0.896
Exposed to STD in past year										
No	1.000	1.000	0.221	0.509	0.185	1.000	1.000	0.806	0.938	0.768
Yes	0.000	0.000	0.779	0.491	0.815	0.000	0.000	0.194	0.062	0.232

defined by the two latent class indicators Number of partners in past year (0, 1, 2 or more) and Exposed to STD in past year (No, Yes). For the partners indicator, 1 instead of 2 or more partners is more likely and for the STD exposure indicator, the probability of STD exposure is as likely as not exposed. For the minority class (32%), the difference relative to Lanza-Collins is bigger and is again with respect to the Number of partners and Exposed to STD indicators. The Nondaters class is not found but is instead similar to the Daters class but with 1 partner instead of 2 or more. The Monogamous class is similar to Lanza-Collins but with low probability of STD exposure. The two Multipartner classes are not as clearly distinguishable.

7 Mover-Stayer modeling

In regular LTA, researchers sometimes explore the need for more than one set of latent transition probabilities to represent the data well. A common example is the Mover-Stayer model 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 can be viewed as an attempt to capture between-subject heterogeneity and is therefore 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. It is of interest to compare regular LTA with two sets of transition probabilities to the RI-LTA models for the four examples. This can be done using BIC.

For the Life satisfaction example, the observed response patterns of Unsatisfied at all time points and Satisfied at all time points are observed for 891 and 1066 subjects, respectively. This is 38% of the total sample of 5147 respondents. It is therefore natural to explore a Mover-Stayer model. As reported in Table 4, the regular LTA model has 11 parameters, loglikelihood -15326, and BIC 30745. Adding the Stayer class gives 13 parameters, loglikelihood -15268, and BIC 30646. The two added parameters are the probability of the Stayer class and the regression of the latent class variable at the first time point on the Mover-Stayer class variable. The likelihood-ratio and Pearson chi-square test values for the frequency table are 14.30 and 14.53, respectively with 18 degrees of freedom. The Mover-Stayer feature is clearly a big improvement to the regular LTA. 52% are estimated as belonging to the Stayer class, that is, a higher number than the 38% stayers observed in the data. It is interesting, however, that the BIC value and the chi-square fit values are not better than for the two RI-LTA models which do not have an additional Mover-Stayer feature. As reported in Table 4, the binary version of the RI-LTA has the same BIC of 30646 with almost the same chi-square test values of 14.18 for both tests with 18 degrees of freedom. The continuous version of RI-LTA has a somewhat better BIC of 30637 and a slightly lower chi-square test value of 13.63 for both tests with 19 degrees of freedom. The choice among these different models is not as clear cut in this example due to the limited information available with a single binary latent class indicator at each time point.

For the Mood example, the regular LTA model with stationarity has 7 parameters, loglikelihood value -2053, and BIC value 4150 as reported in Table 4. Adding the Mover-Stayer feature to the model gives 9 parameters, loglikelihood -2036, and BIC 4127. The better BIC says that there is a need for Mover and Stayer classes when using the regular LTA model. The class percentages are estimated as 64% for Movers and 36% for Stayers. However, the BIC of 4127 is worse than for both RI-LTA models which have BICs of 4118 and 4093 for the binary and continuous random intercept version, respectively. For the binary RI-LTA model, the addition of the Mover-Stayer component has 12 parameters, loglikelihood -2016, and BIC 4105. This BIC is better than without Mover-Stayer where BIC was 4118. The continuous RI-LTA model, however, gives a different picture. Adding the Mover-Stayer component results in 11 parameters, loglikelihood -2016, and BIC 4100. This BIC is worse than without Mover-Stayer where BIC was 4093. In conclusion, using the best-fitting RI-LTA model, there is no need for a Mover-Stayer component in the model.

For the Reading proficiency example, the regular LTA model has 35 parameters, loglikelihood value -21793, and BIC value 43873 as reported in Table 4. Adding the Mover-Stayer component to the model gives 38 parameters, loglikelihood -21725, and BIC 43761. The improved BIC indicates that there is a need for adding the Mover-Stayer feature. The class percentages are estimated as 88% for Movers and 12% for Stayers. However, the BIC of 43761 is worse than for both RI-LTA models which without a Mover-Stayer component have BICs of 42167 and 40984 for the binary and continuous random intercept version, respectively.

For the Dating and sexual behavior example, the regular LTA model has 54 parameters, loglikelihood value -16720, and BIC value 33871 as reported in Table 4. Adding the Mover-Stayer component to the model gives 59 parameters, loglikelihood -16702, and BIC 33876. Because this BIC value is worse than for the regular LTA, a Mover-Stayer component is not needed. As reported in Table 4 the two RI-LTA models have better BICs of 33631, and 33623, respectively.

In conclusion, none of the four examples show a need for a Mover-Stayer model when the best RI-LTA model is used. In contrast, a need for a Mover-Stayer model is indicated for three of the four examples when using regular LTA.

8 Groups and Covariates

Two of the four examples have information on covariates. The Reading proficiency example has a binary covariate indicating whether or not the child's household is above the poverty threshold. The Dating and sexual risk behavior example has four binary covariates, gender and whether the respondent has used cigarettes, been drunk, or used marijuana in the past year.

A first set of analyses uses multiple-group analysis to explore measurement invariance across the poverty groups of the Reading proficiency example. 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 3. For RI-LTA, results are presented for only the continuous random intercept model to save space.

Table 12 shows the model fitting results when testing measurement invariance for the Reading proficiency example. Testing measurement invariance across the poverty groups by comparing models 1 and 2, regular LTA rejects invariance with a likelihood-ratio chi-square value of 312 for 15 degrees of freedom. BIC also favors non-invariance. RI-LTA also rejects measurement invariance but the chi-square value is considerably smaller and BIC favors the invariance model. This illustrates that regular LTA and RI-LTA can lead to different decisions on measurement invariance. RI-LTA has a considerably better BIC value than either of the regular LTA models.

Table 12: Measurement invariance testing using multiple-group analysis with poverty groups for the Reading proficiency example: Regular LTA compared to RI-LTA with a continuous random intercept

Model	Measurement	# par's	LL	BIC	Test (df)	χ^2
1. Regular LTA	Invariance	43	-21584	43519		
2. Regular LTA	Non-invariance	58	-21428	43330	1 vs **2** (15)	312
3. RI-LTA	Invariance	49	-20104	40608		
4. RI-LTA	Non-invariance	64	–20088	40700	3 vs **4** (15)	32

Table 13: 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

Table 13 shows the results of a second set of analyses that explores the influence of covariates in the Dating and sexual risk behavior example. [1] The regular LTA model 1 uses the main effect model shown at the bottom of Table 3. Model 2 uses the interaction effect model for regular LTA shown at the top of Table 3 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 indicates that males and femailes do not have different transitions.

In the RI-LTA model 3, the covariates are allowed to influence the continuous random intercept while in the RI-LTA model 4 the covariates can also influences the latent class variables using the main effect parameterization shown at the bottom of Table 3. 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 femailes 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

[1] In these analyses, the item Had sex in past year was 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 analyses still produce the same 5-class interpretation.

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.

9 Discussion

This paper demonstrates the need for replacing regular LTA with random intercept LTA. Regular LTA suffers from estimating transition probabilities that confound between- and within-subject influences. In addition, it overlooks information in the data which relates to measurement. Regular LTA typically assumes measurement invariance across time but what has been less clear is that it also implicitly assumes measurement invariance across subjects. Analysis of the four examples shows that this is not a realistic assumption. 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. This gives a clearer interpretation as well as a better fit of the model to the data. 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.

9.1 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.

9.2 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 four examples using the three model types with the exception of the continuous random intercept model for the Life satisfaction example. 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 four 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 resulted in a better-fitting model for the Reading example but not for the other examples.

9.3 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 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. (2018). Recent methods for the study of measurement invariance with many groups: Alignment and random effects. Sociological Methods & Research, 47:4 637-664.

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

[34] Muthén, B. & Satorra, A. (1995). Complex sample data in structural equation modeling. Sociological Methodology, 25, 267-316.

[35] 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.

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

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

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

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

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