Exploring Dynamics in Mood Regulation—Mixture Latent Markov Modeling of Ambulatory Assessment Data

CLAUDIA CRAYEN, MS, MICHAEL EID, PHD, TANJA LISCHETZKE, PHD, DELPHINE S. COURVOISIER, PHD, AND JEROEN K. VERMUNT, PHD

Objective: To illustrate how fluctuation patterns in ambulatory assessment data with features such as few categorical items, measurement error, and heterogeneity in the change pattern can adequately be analyzed with mixture latent Markov models. The identification of fluctuation patterns can be of great value to psychosomatic research concerned with dysfunctional behavior or cognitions, such as addictive behavior or noncompliance. In our application, unobserved subgroups of individuals who differ with regard to their mood regulation processes, such as mood maintenance and mood repair, are identified. **Methods:** In an ambulatory assessment study, mood ratings were collected 56 times during 1 week from 164 students. The pleasant-unpleasant mood dimension was assessed by the two ordered categorical items unwell-well and bad-good. Mixture latent Markov models with different number of states, classes, and degrees of invariance were tested, and the best model according to information criteria was interpreted. **Results:** Two latent classes that differed in their mood regulation pattern during the day were identified. Mean classification probabilities were high (>0.88) for this model. The larger class showed a tendency to stay in and return to a moderately pleasant mood state, whereas the smaller class was more likely to move to a very pleasant mood state and to stay there with a higher probability. **Conclusions:** Mixture latent Markov models are suitable to obtain information about interindividual differences in stability and change in ambulatory assessment data. Identified mood regulation patterns can serve as reference for typical mood fluctuation in healthy young adults. **Key words:** ambulatory assessment, experience sampling method, mood regulation, latent class analysis, hierarchical latent Markov model, mixture distribution.

AA = ambulatory assessment; **MLM** = mixture latent Markov; **MMQ** = Multidimensional Mood Questionnaire; **BIC** = Bayesian Information Criterion; **AIC3** = modified Akaike Information Criterion.

INTRODUCTION

Affective states (e.g., pleasant-unpleasant mood, calm-tense mood), body states (e.g., blood pressure, sleep quality), cognitions (e.g., appraisals, self-esteem), and behaviors (e.g., treatment compliance, drinking behavior) typically fluctuate over time. Importantly, individuals may differ in the specific pattern of fluctuations they show (e.g., slow versus fast transitions between states), and these individual differences are of key interest to psychosomatic research. For instance, in health psychology, one might be interested in the patterns of instability of a specific type of health behavior, or in psychiatry, one might be interested in specific patterns of self-destructive behavior over time.

A way to explore these patterns and to gain insight into their circumstances is the repeated measurement of individuals' affective states, body states, cognitions, and behaviors via ambulatory assessment (AA). Ambulatory assessment studies provide intensive longitudinal data (e.g., several measurements a day across a period of 2 weeks) that permit researchers to analyze individual differences in patterns of change and stability (1). There are several statistical approaches appropriate for analyzing intensive longitudinal data. A summary of their main strengths and limitations can be found in Table 1 (2–11). Of particular importance to the selection of a statistical approach is the type of variable that is to be analyzed. In AA studies, states and behaviors of interest are often categorical in nature (e.g., compliant versus noncompliant behavior) or are assessed by only few items with a categorical response format (e.g., very bad mood, rather bad mood, rather good mood, very good mood). The key aim of this article was to explain and illustrate one particular approach to the analysis of intensive longitudinal data that is appropriate for categorical observed and categorical latent variables and that is able to separate variability due to occasion-specific influences from variability due to measurement error: mixture latent Markov (MLM) models (12). Originally, latent Markov models were developed for the analysis of panel data (13). Until recently, the application of MLM models required very large sample sizes and was restricted to few measurement occasions. Methodological developments by Vermunt and colleagues (12) make it now possible to apply these models to the analysis of interindividual differences in intraindividual fluctuations in intensive longitudinal studies. Previous applications based on the new approach include models with 23 measurement occasions, but models for much longer time series can be dealt with. Dias and colleagues (14) applied these methods to financial time series consisting of almost 2000 time points (days). In this article, we will show how MLM models can be applied to AA data with many measurement occasions of many individuals. We will use the models to test hypotheses about the existence of subgroups differing in their pattern of mood fluctuations over time, which can be conceptualized as indicating different mood regulation competencies. These differences can have important consequences for subjective well-being and psychological health. Many forms of psychopathology (e.g.,

From the Department of Education and Psychology (C.C., M.E., T.L.), Freie Universität Berlin, Berlin, Germany; Division of Clinical Epidemiology (D.S.C.), University of Geneva, Geneva, Switzerland; and Department of Methodology and Statistics (J.K.V.), Tilburg University, Tilburg, the Netherlands.

Address correspondence and reprint requests to Claudia Crayen, MS, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. E-mail: claudia.crayen@fu-berlin.de

This work was supported by the German Research Foundation (Grant LI-182771-1).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.psychosomaticmedicine.org).

Received for publication December 28, 2010; revision received December 28, 2011.

DOI: 10.1097/PSY.0b013e31825474cb

0033-3174/12/7404–0366

TABLE 1. Methodological Approaches to the Analysis of Ambulatory Assessment Data

Type of Data Analysis	Advantages	Limitations	Application
Time series analysis, frequency domain analysis (2,3)	– Large number of occasions	– Few individuals, no interindividual differences	– Effects of daily stress on well-being (4)
	– Single-case estimates	– Measurement error not considered	– Mood change frequency (5)
	– Weekly or daily cycles	– Continuous outcomes	
Dynamic factor analysis (6)	– Structure of observed variables can be tested	– Few individuals, no interindividual differences	– Daily emotions after a romantic breakup (7)
	– Complex relations between variables	– Continuous outcomes	
Multilevel analysis (8,9)	– Many individuals and occasions	– Measurement error not considered	– Intraindividual variability in positive and negative affect (10)
	– Many different types of change processes	– Measurement invariance assumed	
	– Intraindividual and interindividual differences		
Structural equation modeling (11)	– Latent variables free of measurement error	– Usually restricted to continuous outcomes	– Mobile phone assessment of mood in daily life (11)
	– Unbiased estimates of stability and variability	– No qualitative differences in change	

depression, phobias) may arise from, and be maintained by, unsuccessfully implemented mood regulation (15).

The remainder of this section is organized as follows. First, we consider mood and mood regulation processes and review empirical findings on individual differences in mood regulation competencies. Second, we present the properties of the MLM model and demonstrate how the model can be applied to assess mood regulation patterns by AA data.

Mood and Mood Regulation

Mood states are diffuse and unfocused affective states which shape the background of our moment-to-moment experience (16,17). Structural models of mood assume that mood states can be described by a few dimensions (e.g., Schimmack and Grob (18)). The three-dimensional model of mood (19,20), for instance, includes wakefulness-tiredness, relaxation-tension, and pleasant-unpleasant mood as basic dimensions. In our application, we will focus on the pleasant-unpleasant dimension of mood.

Mood has both a stable and a variable aspect; that is, individuals have a characteristic (habitual) level of mood (also called set point), and their momentary mood fluctuates around this set point (15). Research has demonstrated that individuals differ both in their set point of mood and in their pattern of mood fluctuations (e.g., Eid and Diener (21) and Eid et al. (22)). This pattern of fluctuation is partly due to mood regulation behavior (15,17). Research has demonstrated that individuals differ considerably in their ability to effectively improve a negative mood or maintain a positive mood (23–25).

To date, most research on mood regulation competencies has attempted to measure stable individual differences in negative mood repair and positive mood maintenance by self-report questionnaires (26–28). As an alternative, AA allows us to measure mood regulation competencies indirectly by drawing on information of individuals' mood course over a longer period. Mixture latent Markov models are well suited to investigate interindividual differences in the intraindividual course of mood as an indirect measure of mood regulation because these models take into account that the fluctuation process might differ between individuals and might depend on the specific state that is maintained or modified. For example, the information that over time, individuals have a high probability of changing their mood state has to be judged differently if this refers to a negative mood state or to a positive mood state.

Aim of the Study

The aim of the present work was to show how MLM models could be used to assess interindividual differences in mood regulation. We expected to find several latent classes of individuals who differ in their fluctuation pattern. According to current theories of mood regulation, we expected classes that differ in their ability to maintain their positive mood and to repair their negative mood. Classes with high mood maintenance should show a high probability to stay in a pleasant mood state. Classes with high mood repair should show a high probability to leave an unpleasant mood state.

The Mixture Latent Markov Model

In Markov models, stability and change are represented by transition probabilities, which describe the probabilities of staying in the same category over time or moving to another response category. The transition probabilities are estimated based on a set of measurement occasions. Applied to mood measurement, the model estimates the overall probability of being in a certain mood state given the state at the previous time point.

The MLM model is an extension of the simple Markov model and the latent Markov model. The simple Markov model describes relations among categories at different points in time in a so-called Markov chain. Only two kinds of parameters are needed to describe this process: the initial probabilities Prob(Response_0), which contain information about the size of each category at the very first time point; and the transition probabilities Prob(Response_t | Response_{t-1}), which indicate the probability of a category given a certain previous category. In the Markov model with time-homogeneous transition probabilities, it is assumed that the transition probabilities are constant across different time points (time-homogeneous or stationary). To illustrate this point with a simple example, the probability of switching from a positive to a negative mood state between Time Point 1 and Time Point 2 (e.g., from morning to noon) is the same as between Time Point 2 and Time Point 3 (e.g., from noon to afternoon). The general structure of a simple Markov model with time-homogeneous transition probabilities for two observed categories is depicted in Figure 1. Here, the coefficient τ_{11} is the probability to stay in the first category, τ_{22} is the probability to stay in the second category, τ_{12} is the probability to move from the first to the second category, and τ_{21} is the probability to move from the second to the first category. In other words, the transition probabilities between the same category at different time points (τ_{11} and τ_{22}) describe stability. The ones between different categories (τ_{12} and τ_{21}) contain information about change.

The assumption of time-homogeneous transition probabilities can be tested by comparing a model with time-homogeneous transition probabilities with a model assuming time-heterogeneous transition probabilities. One should keep in mind that time-homogeneous transitions are only sensible if the time points are equidistant. Assuming the same influence on a current state by a state 1 hour ago or a state 5 hours ago is very restrictive. We will return to this point in the discussion. Another notable assumption in this model is the fact that a first-order Markov process is assumed. This means that the probability of a state someone is in on a certain occasion depends only on the previous occasion (and not, for example, on the occasion before that).

A disadvantage of simple Markov models is that it is unclear whether change is due to measurement error or to true change

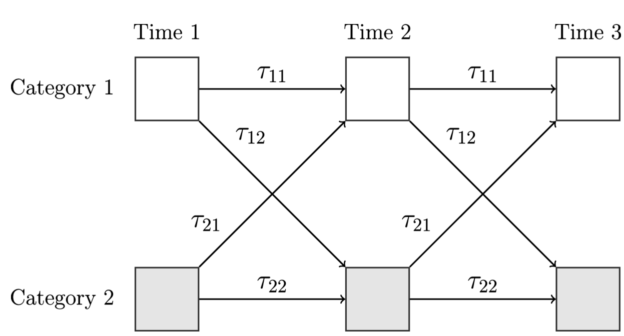

Figure 1. Simple Markov chain for a manifest response variable with two categories and three occasions of measurement. τ_{ij} is the transition probability from category i to j. Depicted is a time-homogeneous Markov process, because τ_{ij} are independent of the specific time point.

processes. Because most measures in psychological and clinical research are afflicted with measurement error, some of the observed change between categories may be attributable to measurement error instead of reflecting true change. To separate measurement error from true change, latent Markov models have been developed.

Latent Markov models are multiple indicator extensions of simple Markov models. At each measurement occasion, at least two indicators are linked to a "true" latent state variable by state-specific response probabilities. In the latent Markov model, the Markov process takes place on the level of error-free latent categories (categorical latent state variables). Because the observed (manifest) indicators have to be linked to the latent state, the latent Markov model contains an additional type of parameter: the conditional response probabilities Prob(Response_t | State_t). They describe how likely an observed category is, given a certain latent state at the same point in time. Even if, for example, someone is in a pleasant mood state, he or she might not respond with the according category *I feel well*. This would be reflected in a response probability Prob (Response_t = well | State_t = pleasant) lower than 1. Whenever an observed category is linked to a corresponding latent category indicating the same state, deviations of the response probabilities from 1 indicate the influence of measurement error, and the response probabilities indicate reliability. There are as many response probabilities as combinations of observed categories and latent states. Whether these response probabilities are constant over time or not is a question of whether it is the same construct that is measured over time (29,30). In AA studies, time lags are usually too short to assume that the meaning of the latent states or the properties of the measurement instrument changes over time. Nevertheless, the assumption of measurement invariance has to be tested.

The latent Markov model assumes that all individuals show the same fluctuation pattern. If distinct subgroups are observed, for example, via a patient/control group variable, a multigroup model with a chain for each group could be considered and differences in parameters between the chains could be formally tested. However, the groups that differ in their fluctuation pattern are often not identifiable by means of observed variables. To identify unobserved subgroups in latent Markov processes, MLM models have been defined. In MLM models, each subpopulation (latent chain or latent class) of individuals is characterized by a latent Markov model (31). The aim of the analysis is to detect the number of latent chains (classes) that differ in their parameters (initial state probabilities, response probabilities, and state transition probabilities (32,33)). As additional parameters, the MLM model contains the probabilities of belonging to a particular latent class Prob(Class), which is also referred to as the size of the class. All other parameters are conditional on the latent class membership in the MLM model.

The MLM model can be extended by including covariates (12). These covariates can be either time constant, such as measures of stable personality traits or sex, or time varying, such as situational factors collected via AA (e.g., events, physiological measures). Covariates can be used to predict the different types of parameters in the model, for example, the transition probabilities. Individuals might be more likely to move to a more pleasant mood state in social situations compared with nonsocial situations. The effect of covariates could also differ between latent classes of individuals.

An attribute of AA data that requires special attention is the nesting of measurement occasions in days and the dependency between days. Not only is the time lag between the last signal at night and the first signal in the morning much longer than the time lags within the day, but the processes that operate at night might also be different. These transitions on different levels can be accounted for by treating the measurement occasions as nested within days. Such a hierarchical model was suggested by Rijmen and colleagues (34) in their application of MLM models to AA data. They reported results for a study with 32 female patients and 63 signals during the course of a week, assessing emotional states. The structure of such a hierarchical MLM model is depicted in Figure 2. In our illustrated example, there are two subgroups or latent day classes and three latent states for each measurement occasion. Two Markov processes are operating, one on each level: On the lower within-day level, transitions between latent states occur only during the day but not across days. On the upper between-day level of latent day classes, transitions between latent classes occur only across days but not during the day. The latent day classes for each day are obtained by separating individuals who differ in their pattern of fluctuations between latent states over the day.

In a simplified example, there may be two classes on each day: one with individuals who had a bad day (high initial probability and high stability of unpleasant mood state) and one with individuals who had a good day (high initial probability and high stability of pleasant mood state). Because of the transition between days on the upper level of the latent classes, a person can have a good day after a bad day and vice versa, independent of the person's last state on the previous night.

To sum up, the hierarchical MLM model contains the following parameters (disregarding covariates):

1. Prob(Class_0) is the initial class probability, that is, the probability of belonging to a particular class at $d = 0$ (e.g., the first day of assessment; circled "A" in Fig. 2).
2. Prob($\text{Class}_d \mid \text{Class}_{d-1}$) is the latent transition probability of being in a certain latent class given the latent class on the previous day (circled "B" in Fig. 2).
3. Prob($\text{State}_{d0} \mid \text{Class}_d$) is the initial state probability on the beginning of each day that depends on the latent class of the same day (circled "C" in Fig. 2).
4. Prob($\text{State}_{dt} \mid \text{State}_{dt-1}, \text{Class}_d$) is the latent transition probability on the latent state level, that is, the probability of being in a certain state given the previous state and the class on that day (circled "D" in Fig. 2).
5. Prob($\text{Response}_{jdt} \mid \text{State}_{dt}, \text{Class}_d$) is the response probability for the observed categories of indicator j, given the latent state on that particular time point and the latent class on that day (not depicted in Fig. 2).

APPLICATION

The data analyzed here are a subset of data from a larger study on mood regulation processes that combined a laboratory session with a 14-day AA period. In the laboratory session, various personality variables were assessed via self-report. During the AA period, the focal construct of momentary mood was measured, as well as a number of additional variables. In this application, we will use data of the first week of the AA period only.

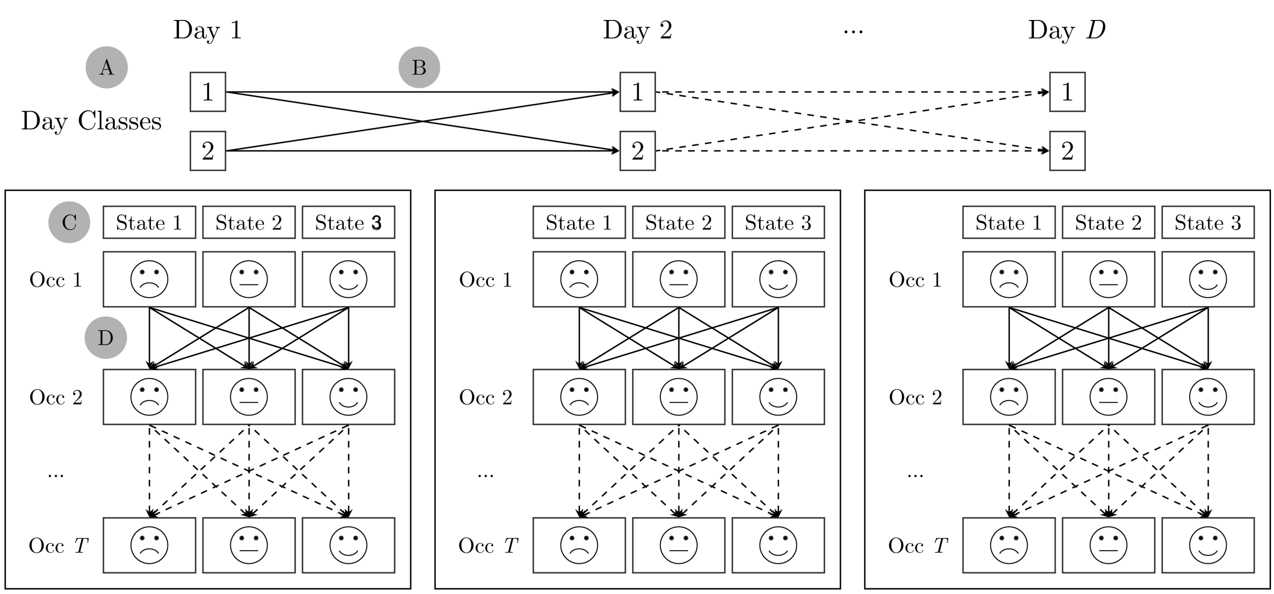

Figure 2. A hierarchical mixture latent Markov model with three latent states and two latent day classes. The measurement part of the model has been omitted. Occ = occasion within a day; D = number of days; T = number of occasions. Parameters denoted by gray-shaded letters are referred to in the text: A = initial day class probabilities; B = day class transition probabilities; C = initial state probabilities; D = state transition probabilities.

Participants

A total of 165 participants were recruited from the Freie Universität Berlin via a notice posted on campus. Criteria for inclusion were student status in a subject other than psychology and German as a native tongue. Data from one participant were excluded from the analyses because this person's average mood level across the AA period was exceptionally low (−5 SD). The final sample consisted of 164 students (88 women) with an M (SD) age of 23.7 (3.31) years (min = 18 years, max = 35 years). Students received 80 EUR in exchange for their participation, and an additional 20 EUR if at least 80% of the field signals were answered.

Procedure

Initial laboratory sessions were done in groups of one to six. Participants gave informed consent and completed a computer-based questionnaire assessing several personality dimensions. After the computer-based part, participants were given detailed instructions in the use of the handheld device and the ambulatory questionnaire. The AA period started for all participants on the Wednesday after the laboratory session, which was either the next or the next but 1 day. Data collection took place between late October 2009 and early May 2010, covering mostly the winter half year.

During the first week of the AA period, momentary mood was assessed eight times per day using signal-contingent time sampling. Participants were requested to respond on handheld devices (HP iPAQ rx 1950 Pocket PCs) when signaled by an alarm (software: Izybuilder, IzyData Ltd., Fribourg, Switzerland). The signal sounded pseudo-randomly within a 13-hour period during the day. Participants were able to choose the period according to their waking hours. The delay between adjacent signals could vary between 60 and 180 minutes[1] (M [SD] = 100.24 [20.36] minutes, min = 62 minutes, max = 173 minutes). Responses had to be made within a 30-minute time window after the signal on the touch screen of the device using a stylus. If participants failed to respond within the 30-minute time window, the session was counted as missing. On average, the 164 participants responded to 51 (of 56) signals (M [SD] = 51.07 [6.05] signals, min = 19 signals, max = 56 signals). In total, there were 8374 nonmissing measurement occasions in the present analysis.

Measures

Momentary Mood

At each measurement occasion during the AA period, participants rated their momentary mood on an adapted short version of the Multidimensional Mood Questionnaire (MMQ) (20,35). Instead of the original monopolar mood items, a shorter bipolar version was used to fit the need for brief scales in an AA study (36). Several studies (20,35) have shown that the items belonging to the same scale of the MMQ but different poles are strongly negatively correlated when momentary mood is assessed, resulting in a common factor. Hence, building bipolar items on the basis of monopolar items of the same scale of the MMQ is acceptable. Four items assessed pleasant-unpleasant mood (happy-unhappy, content-discontent, good-bad, and well-unwell). Participants rated how they momentarily feel on a 4-point bipolar intensity scales (e.g., *very unhappy*, *rather unhappy*, *rather happy*, *very happy*). For the current analysis, we focused on the items well-unwell and good-bad to keep the model simple. Preliminary analysis of the response category frequencies showed that the lowest category (i.e., *very bad* and *very unwell*) was only chosen in approximately 1% of all occasions. We therefore decided to collapse the two lower categories together into one *unwell* and *bad* category, respectively. The following analyses are based on the recoded items with three categories.

Trait Mood Regulation

In the laboratory session, participants completed an 11-item scale measuring perceived effectiveness in mood regulation (37). Six items assessed negative mood repair (e.g., "It is easy for me to improve my bad mood") and five items assessed positive mood maintenance (e.g., "It is easy for me to maintain my good mood for a long time"). The items were answered on 4-point frequency scales (ranging from *almost never* to *almost always*).

Data Analysis

Software

To estimate MLM models with many occasions and feasible sample sizes, the special forward-backward EM algorithm as described by Vermunt and colleagues (12) has to be integrated in the software. For all analyses, the Latent GOLD 4.5 software package (38) was used. Syntax for Latent GOLD and the corresponding code for the R system (39,40) to run WinBUGS (41) is provided in the online appendix (Supplemental Digital Content 1, http://links.lww.com/PSYMED/A42). Rijmen and colleagues (34) provided the according functions for the Matlab environment. Functions that are able to estimate similar models in the R system may be found in the depmixS4 package (42).

Determining the Number of Latent States and Classes

Following a bottom-up strategy, we started building models for each of the 7 days separately to see which combination of number of latent states and number of latent classes would fit the observed data best. We expected the number of latent states to mirror the three observed states (i.e., item categories). Alternative models with two and four latent states were tested in addition. The number of latent states can deviate from the number of observed categories if the observed variables vary in their difficulties (43). The number of latent classes in the models we tested ranged from 1 to 4. The best fitting model was selected according to information criteria (44,45). Other fit statistics that rely on the χ^2 distribution are not applicable here because of sparse contingency tables (46). This problem occurs

[1]Two signals that violated this rule owing to device malfunctioning were excluded from the analysis.

TABLE 2. Fit Measures for the Estimated Models

Model	LL	BIC	AIC3	n_{par}
A. Unrestricted baseline model	−10,080	20,930	20,613	151
B. Response probabilities restricted	−10,097	20,902	20,610	139
C. State transition probabilities restricted	−10,151	20,643	20,502	67
D. Initial state probabilities restricted	−10,163	20,546	20,456	43
E. Class transition probabilities restricted	−10,166	20,500	20,430	33
F. No class transition allowed	−10,188	20,535	20,470	31
G. Model E + covariates	−10,152	20,483	20,410	35

LL = log-likelihood; BIC = Bayesian Information Criterion based on log-likelihood; n_{par} = number of parameters.

because, with many categories and time points, many combinations of the categories of the observed variables that would theoretically be possible do not appear in the data when the sample size is not extremely large. The distribution of fit statistics is no longer known and cannot be used for calculating valid *p* values. Information criteria depend on the fit of the model as well as its complexity. According to information criteria, the best-fitting model is the simplest model showing an adequate fit. This model can be found by comparing the information criteria of different models and selecting the model with the smallest value of the information criterion considered. There are many different information criteria that differ in how model complexity is penalized. For latent class models, the Bayesian Information Criterion (BIC) (47) has been shown to perform well (48,49). For the special case of MLM models, there is some evidence (50) suggesting the use of the modified Akaike Information Criterion (AIC3) (51).

Testing Invariance

Next, we combined the single-day models into a model of the first week. Days were linked by a transition between latent classes at the beginning of the day.[2] We proceeded in several steps to test parameter invariance across days, starting from a baseline model, in which all parameters (class- and state-dependent response probabilities, state transition probabilities, initial state probabilities at the beginning of the day, and transition probabilities between classes) were allowed to differ between days. Subsequently, we imposed equality constraints across days on the response probabilities (Model B), the latent state transition probabilities (Model C), the initial state probabilities (Model D), and the day class transition probabilities (Model E). Finally, we analyzed a model without transitions between day classes (probabilities to stay in the same class constrained to 1; Model F).

RESULTS

The results for the single-day models showed that, in general, a model with three latent states and two latent classes can be adopted for each day. Transition probabilities were assumed to be time-invariant during the course of a day, and response probabilities were allowed to differ between latent classes. The single-day models were combined into a single 7-day model. The BIC, the AIC3, and the number of parameters for each 7-day model tested are reported in Table 2. Because both information criteria were in agreement in our application, from here on, we will only refer to the AIC3. In the baseline model (Model A), all parameters (class- and state-dependent response probabilities, state transition probabilities, initial state probabilities at the beginning of the day, and transition probabilities between classes) were allowed to differ between days. It should be noted that this baseline model was not completely unrestrictive. Some equality assumptions had to be made to secure model identification. Model identification refers to the situation where it is possible to obtain unique estimates for all free parameters in the model (see Langeheine and Van de Pol (52), where identification of latent Markov models is discussed). One restriction in the baseline model concerned the response probabilities. They were restricted to be equal across measurement occasions within the same day and the same class in all models.

Next, we tested equality constraints. In Model B, we tested whether the measurement part of the model, the link between latent states and observed response categories, remained stable across days. We restricted the response probabilities to be equal across days (but not classes). Model B had a lower AIC3 than Model A did, implying that equal response probabilities could be assumed. The mood states did not change their meaning across days, but they were slightly different for the two classes. Next, the same procedure was applied to test homogeneity across days concerning the state transition probabilities. Model C contained these restrictions and yielded a lower AIC3 than Model B. In the following step, initial state probabilities were

TABLE 3. Estimated Conditional Response Probabilities in Model E

	Item "Well"			Item "Good"		
	Unwell	Rather Well	Very Well	Bad	Rather Good	Very Good
Class 1						
State 1	0.90	0.10	0.00	0.93	0.07	0.00
State 2	0.05	0.94	0.02	0.02	0.96	0.01
State 3	0.00	0.39	0.61	0.00	0.48	0.52
Class 2						
State 1	0.72	0.25	0.03	0.64	0.34	0.02
State 2	0.01	0.93	0.05	0.00	0.91	0.09
State 3	0.00	0.16	0.84	0.00	0.11	0.89

Probabilities may not add up to one due to rounding error.

[2]We tested whether an additional link allowing the last mood state of a day to influence the first state of the following day would improve the model, but it did not.

restricted to be equal across all days. The obtained Model D showed an even lower AIC3 than Model C did. The class transitions between days were set equal in Model E. Again, this restriction led to a lower AIC3. Because the stability of the classes was very high (the probability of staying in the same class across days was larger than 0.9), we tested whether it was necessary to even let people change classes between days or whether these classes could be seen as trait classes rather than day-specific classes linked by a Markov process. Compared with Model E with equal class transitions, Model F with no class transition allowed (the probabilities to stay in the same class were restricted to 1) had to be rejected based on a higher AIC3. We therefore kept Model E as our final model, which is described in more detail in the next paragraph.

In Model E, the mean classification probabilities were high for the latent states (0.94, 0.92, and 0.88) and the latent classes (0.96 and 0.93), showing that this model yielded a reliable classification of individuals. Mean classification probabilities were calculated in the following way. First, for each individual, the classification probabilities to belong to the different classes and states were calculated based on his or her response vector. Then, an individual was assigned to the latent state on each occasion and the latent class on each day for which his or her classification probability is maximum. Then, the mean of the assignment probabilities of all individuals belonging to the same class and state were calculated. The closer the mean probabilities are to 1, the better is the classification of individuals. From the perspective of psychological assessment, the classification of individuals to different types of mood regulation is a very important task. The high mean classification probabilities show that this assessment could be reliably done based on the model selected.

Two latent classes that differed with respect to their within-day mood fluctuation patterns were identified. These two classes characterize the pattern of mood change within a single day. The transition probabilities did not differ between days but individuals were allowed to change classes between days. We named the larger of the two classes as Class 1 and the smaller one as Class 2. The size of Class 1 was 0.68 on the first day of our AA week and remained very stable across days (the probability to stay in this class between 2 days was 0.98). Accordingly, the smaller Class 2 had a size of 0.32 and was a little less stable (the probability to stay in this class between 2 days was 0.90). To determine the character of a class, it was crucial to first characterize the latent states in the class by looking at the response probabilities for the two different items *well* and *good*. For Model E, these can be found in Table 3.

Latent mood State 1 of Class 1 was characterized by a high probability of choosing the observed categories *unwell* (0.90) and *bad* (0.93), respectively. This state could be interpreted as "unpleasant mood." On the other hand, latent mood State 2 in Class 1 was associated with a very high probability of choosing the observed categories *rather well* (0.94) and *rather good* (0.96). We labeled this state "rather pleasant mood." For the last latent mood State 3 in Class 1, the probabilities of choosing the observed categories *very well* (0.61) and *very good* (0.52) were not as high. There was still a considerable probability of choosing the middle categories *rather well* (0.39) and *rather good* (0.48). This means that individuals in this state were in a mood that is somewhat between a rather pleasant and a very pleasant mood. Because the probabilities were highest for the last item categories, we labeled this state "very pleasant mood."

Class 2 differed from Class 1 in the response probabilities given the latent states. In latent mood State 1, individuals in Class 2 had a lower probability of responding with the lowest observed categories *unwell* (0.72) and *bad* (0.64) than individuals in Class 1. Instead, there was a tendency toward the middle categories *rather well* (0.25) and *rather good* (0.34).

Class 1

Initial	*State − 1*	*State* 1	2	3
.07	1	.53	.44	.03
.75	2	.09	.82	.10
.18	3	.05	.43	.52

Class 2

Initial	*State − 1*	*State* 1	2	3
.14	1	.33	.43	.24
.47	2	.14	.57	.29
.39	3	.05	.21	.74

Figure 3. Estimated initial state probabilities and state transition probabilities in Model E. The upper part of the figure gives the estimates in Class 1, and the lower part of the figure gives the estimates for Class 2. The first column provides the initial probabilities for each state. For example, a member of Class 1 has a 75% chance of starting in State 2. The numbers and shaded circular areas in the grid represent the transition probabilities from the former state (State − 1, rows) to the current state (State, columns). For example, the probability for members of Class 1 to stay in State 2 is 0.82. Probabilities may not add up to 1 due to rounding error.

Compared with the "unpleasant mood" state in Class 1, this latent State 1 in Class 2 was between an unpleasant and a rather pleasant mood. Because the probabilities were highest for the first item categories, we named this state "unpleasant mood." Latent State 2 was very similar to the "rather pleasant mood" state in Class 1 and characterized by a very high probability of choosing the middle categories *rather well* (0.93) and *rather good* (0.91). In latent State 3, individuals in Class 2 had a clearly higher probability of choosing the highest observed categories *very well* (0.84) and *very good* (0.89) than individuals in Class 1. Keeping this difference in mind, this state in Class 2 was also labeled "very pleasant mood." In Class 2, the latent mood states seemed to reflect a higher basic mood level.

The initial state probabilities and state transition probabilities are depicted in Figure 3. In Class 1, the initial probability for the rather pleasant mood state was by far the highest (0.75). Looking at the transition probabilities, this mood state was very stable (0.82). The probabilities of changing one's rather pleasant mood state for the better (0.10) or the worse (0.09) were equally low. In this class, there was a high probability to start in a rather pleasant mood in the morning and to stay in this rather pleasant mood over the day. By comparison, the unpleasant and very pleasant mood states showed lower stabilities (0.53 and 0.52), and there was a tendency toward returning to the rather pleasant mood state. In sum, the rather pleasant mood state prevailed in Class 1.

We found a different pattern looking at Class 2. Here, people started off in a very pleasant mood (0.39) almost as often as in a rather pleasant mood (0.47). This class was more likely to remain in a very pleasant mood state (0.74) than to decline from there. In mood regulation terms, this number quantifies the extent of positive mood maintenance. Also, compared with Class 1, Class 2 had a higher probability of entering the very pleasant mood state (0.24 and 0.29) coming from one of the other two mood states that were not as stable. Negative mood repair (the transition probabilities of moving to a better mood state if in an unpleasant mood state) added up to 0.67 compared with 0.47 in Class 1. Overall, this class had a higher mean mood level and showed a pattern of regulation toward an elevated mood state. This pattern may very well be representative for people who are exceptionally skilled in regulating their mood. They are highly able to repair their negative mood and to maintain their positive mood.

To further test this interpretation of mood regulation patterns, we included measures from the laboratory session as time-constant covariates into the model. Specifically, we regressed the class proportions on the first day on self-reported mood repair and mood maintenance by means of a binary logistic regression model (Model G in Table 2). Class 1 served as the reference category. The positive and significant regression weights for both predictors, mood repair (b = 1.02; standard error [SE] = 0.43, p = .017) and mood maintenance (b = 1.36; SE = 0.44, p = .002), reflect that individuals with high self-reported mood regulation competencies were much more likely to start in the class with the very positive regulation pattern than in the moderately positive class.

DISCUSSION

In our empirical application, the MLM model allowed a reliable classification of individuals to different classes of mood fluctuation patterns. Moreover, it also allowed a reliable assignment to latent mood states within a day. We will review the substantive findings first and conclude with prospects of the modeling approach.

Individual Differences in Mood Regulation

The results revealed that there were only two classes (or patterns) of mood fluctuations in our sample. We found a class with pronounced abilities to repair negative and maintain positive mood and a class that was very stable in a moderately positive mood state. Individuals in this class were somewhat able to repair their negative mood and to maintain their very positive mood, and there seemed to be a high ability to maintain a moderately positive mood. These classes appeared to be distinct both in their habitual mood level (or set point) and their fluctuation pattern. The smaller class with the higher habitual mood level exhibited a higher rate of overall fluctuation. This fluctuation was mainly a result of the improvement of mood states. Whether these mood fluctuations were due to mood regulation behavior or other influences (rhythmic processes due to biologic and social factors, activities and situations, positive and negative daily events) cannot be answered by the present analysis. We are able to see that people were successful in regulating their mood, but we do not know how they achieved this. Investigating this question would require incorporating information about the situation or on individuals' regulation behavior into the analysis. Ways in which the presented model can be extended to include this type of information are discussed below. An indication that mood regulation competencies do play a role in these different patterns comes from the self-reported trait measures of mood regulation. Higher reported competency is linked to the class with the higher habitual mood level. With this interpretation, one should keep in mind that individuals were allowed to change classes between days. Even if the assignment to the classes was very stable across days, there remains a day-to-day variation that cannot be accounted for by trait measures. It would be interesting to determine conditions of this day-to-day variability in future research.

It is also of interest that there was no class of individuals with a very high probability of staying in an unpleasant mood state (i.e., with low mood repair competence). This might be because we have analyzed a sample from a nonclinical population. In a clinical population, one might expect different classes, for example, a class with high probability to stay in a negative mood (depression) or a class characterized by unusual high variability of mood (borderline personality disorder). Depending on the symptoms considered, many other classes are conceivable. The classes we found may serve future clinical studies as an example of standard regulation patterns.

In addition, the results provide some insight into the relationship between the two mood regulation abilities under consideration: positive mood maintenance and negative mood repair. The fact that we did not find a class in which only one ability was present but not the other shows that they do not seem to occur independently from each other. One reason may lie in the concurrent acquisition of these competencies or in regulation strategies that can be valuable for positive mood maintenance as well as negative mood repair (e.g., social sharing). Mixture latent Markov models allow analyzing mood regulation in an indirect way on the basis of repeated measurements. This has many advantages over the more traditional way of assessing mood regulation competencies by self-report questionnaires. The results might be less distorted by memory effects and are representative for individuals' daily life. Nevertheless, our results also demonstrate the convergent validity of our assessment method with traditional questionnaires because the mood regulation questionnaires predicted class membership.

The MLM Model

There are at least three ways in which the MLM model applied here can be extended. First, the model can be modified to allow for between-day differences in the number and structure of the latent classes. In the application presented, we found homogeneous structures across the different days. However, this might not be the case for other constructs.

Second, time-varying covariates could be included. Mood fluctuations depend not only on mood regulation but also on situational influences. Time-varying covariates characterize the situations in which individuals are and could be used to investigate whether the latent classes differ in the way they react to these situational influences. There might be classes of individuals with high resilience that do not react strongly to negative events but also classes that might be very reactive to situational influences (53). Mixture latent Markov models could be applied to measure resilience in an indirect way by separating latent subgroups that differ in the way situational factors influence behavior and feelings. If covariates are included, parameters are accordingly conditioned on the set of covariates. Hence, the meaning of the parameters can change when covariates are included. In which way the meaning of the parameter changes depends on the specific model considered. If covariates are observed variables, this does usually not affect the identifiability of the model parameters. However, if the covariates are latent variables, measurement models have to be specified for them as well.

Finally, we have assumed in our application that the transition probabilities between two states are the same for all individuals within a class. However, individuals might differ in the time lag between two occasions of measurement. If the time lag varies between individuals, the assumption of homogeneous transition probabilities might be inappropriate. The probability to stay in the same state might be higher for shorter time lags than for longer time lags. Although there were individual differences in time lags in our study, the intraindividual distributions of the time lags were homogeneous. Therefore, the model seems to be appropriate in our application. If there are large interindividual differences in time lags, the model has to consider these differences. A way to adequately include interindividually varying time lags in MLM models is suggested by Vermunt (54). Another option in this case would be to employ continuous time Markov models (55) that do not assume equal or at least similarly spaced measurements within and between individuals, as opposed to discrete time Markov models (like the MLM model we applied here).

From a more general point of view, the application illustrates some properties of the model that are attractive for AA studies. In contrast to more traditionally used models for analyzing AA data such as classic multilevel analyses, the model has three major advantages. First, it separates change due to measurement error from true change. Second, it allows single categories (states) to differ in the process of change. The degree of state-specific stability and change can easily be modeled with MLM models, whereas it is much harder to model state-specific change processes with other statistical approaches. Third, the model permits population heterogeneity with respect to the change process. In contrast to multiple group analysis, the subpopulations do not have to be known but are rather a result of the analyses.

Recommendations

When is it appropriate and beneficial to employ MLM models? First of all, the data at hand should show characteristics that are suitable for Markov processes: Is it sensible to assume qualitatively distinct states (categories) for each measurement occasion? Can the switching process among these categories considered to be autoregressive? If this is the case, one can start to think in greater detail about the model. If the occasions exhibit a nested structure that is likely to affect the change process, it has to be accounted for. In our application, measurement occasions were nested in days, and we found the hierarchical approach to provide a suitable solution. If the measures used are likely to contain measurement error, researchers should include at least two indicators of the same construct so that a measurement model for the latent state can be included. With this information, a latent Markov model that adequately reflects the basic structure of the data can be constructed.

The next step involves determining the number of latent states and latent classes, often the core question in such an analysis. The number of latent states is often expected to reflect the number of observed categories. As mentioned before, this number may increase if the observed variables vary in their difficulties, thereby capturing in-between states. Determining the number of latent classes may well be the hardest part, because as of yet, statistical fit criteria that ought to guide this decision are not well scrutinized. Theoretical considerations should be involved: How many classes can be expected? Do the classes of a particular solution make sense? Can the classes be separated well, as indicated by high mean classification probabilities? In the context of AA studies, one might even expect differences in the number or profile of the latent classes between, for example, weekdays and the weekend. In this case, one has to exercise reasonable care in arranging the data according to weekdays for all individuals in the study.

Once one has decided on the number of latent states and latent classes, one can go about to test specific assumptions, such as a homogeneous change process, by comparing models with different restrictions. In an additional step, one could include time-constant and/or time-varying covariates to gain insight into factors influencing the identified fluctuation patterns.

The authors thank Martin Schultze for his help in preparing the figures.

REFERENCES

1. Walls TA, Höppner BB, Goodwin MS. Statistical issues in intensive longitudinal data analysis. In: Stone A, Shiffman S, Atienza A, Nebelling L, editors. The Science of Real-time Data Capture. New York, NY: Oxford University Press; 2007:338–60.
2. Molenaar PCM, Sinclair KO, Rovine MJ, Ram N, Corneal SE. Analyzing developmental processes on an individual level using non-stationary time series modeling. Dev Psychol 2009;45:260–71.
3. Shumway RH, Stoffer DS. Time Series Analysis and Its Applications. With R Examples. New York, NY: Springer; 2011.
4. Reicherts M, Pihet S. Job newcomers coping with stressful situations: a micro-analysis of adequate coping and well-being. Swiss J Psychol 2000;59:303–16.
5. Larsen RJ. The stability of mood variability: a spectral analytic approach to daily mood assessments. J Pers Soc Psychol 1987;52:1195–204.
6. Nesselroade JR, Molenaar PCM. Applying dynamic factor analysis in behavioral and social science research. In: Kaplan D, editor. The Sage Handbook of Quantitative Methodology for the Social Sciences. Thousand Oaks, CA: Sage; 2004:335–44.
7. Sbarra DA, Ferrer E. The structure and process of emotional experience following nonmarital relationship dissolution: dynamic factor analyses of love, anger, and sadness. Emotion 2006;6:224–38.
8. Singer J, Willett J. Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence. Oxford, NY: Oxford University Press; 2003.
9. Walls TA, Jung H, Schwartz J. Multilevel models and intensive longitudinal data. In: Walls TA, Schafer JS, editors. Models for Intensive Longitudinal Data. New York, NY: Oxford University Press; 2006:3–37.
10. Röcke C, Li S-C, Smith J. Intraindividual variability in positive and negative affect over 45 days: do older adults fluctuate less than young adults? Psychol Aging 2009;24:863–78.
11. Eid M, Courvoisier D, Lischetzke T. Structural equation modeling. In: Mehl MR, Conner TS, editors. Handbook of Research Methods for Studying Daily Life. New York, NY: Guilford; 2012:384–406.
12. Vermunt JK, Tran B, Magidson J. Latent class models in longitudinal research. In: Menard S, editor. Handbook of Longitudinal Research: Design, Measurement and Analysis. Burlington, MA: Elsevier; 2008: 373–85.
13. Wiggins LM. Panel Analysis. Amsterdam, the Netherlands: Elsevier; 1973.
14. Dias JG, Vermunt JK, Ramos S. Mixture hidden Markov models in finance research. In: Fink A, Lausen B, Seidel W, Ultsch A, editors. Advances in Data Analysis, Data Handling and Business Intelligence. Berlin, Germany: Springer; 2010:451–9.
15. Parkinson B, Totterdell P, Briner RB, Reynolds S. Changing Moods. New York, NY: Addison, Wesley, Longman; 1996.
16. Frijda NH. Varieties of affect: emotions and episodes, moods, sentiments. In: Ekman P, Davidson R, editors. The Nature of Emotion: Fundamental Questions. New York, NY: Oxford University Press; 1994:59–67.
17. Morris WN. Mood: The Frame of Mind. New York, NY: Springer; 1989.
18. Schimmack U, Grob A. Dimensional models of core affect: a quantitative comparison by means of structural equation modeling. Eur J Pers 2000; 14:325–45.
19. Matthews G, Jones DM, Chamberlain AG. Refining the measurement of mood: the UWIST Mood Adjective Checklist. Br J Psychol 1990; 81:17–42.
20. Steyer R, Schwenkmezger P, Notz P, Eid M. Testtheoretische analysen des mehrdimensionalen befindlichkeitsfragebogens (MDBF). Diagnostica 1994;40:320–8.
21. Eid M, Diener E. Intraindividual variability in affect: reliability, validity, and personality correlates. J Pers Soc Psychol 1999;76:662–76.
22. Eid M, Notz P, Steyer R, Schwenkmezger P. Validating scales for the assessment of mood level and variability by latent state-trait analyses. Pers Individ Dif 1994;16:63–76.
23. Josephson BR, Singer JA, Salovey P. Mood regulation and memory: repairing sad moods with happy memories. Cognit Emot 1996;10: 437–44.
24. Larsen RJ. Toward a science of mood regulation. Psychol Inq 2000; 11:129–41.
25. Showers CJ, Kling KC. Organization of self-knowledge: implications for recovery from sad mood. J Pers Soc Psychol 1996;70:578–90.
26. Bryant FB. A four-factor model of perceived control: avoiding, coping, obtaining and savoring. J Pers 1989;57:773–97.
27. Lischetzke T, Eid M. Why are extraverts happier than introverts? Exploring the role of mood regulation processes. J Pers 2006;74:1127–62.
28. Salovey P, Mayer JD, Goldman SL, Turvey C, Palfai TP. Emotional attention, clarity, and repair: exploring emotional intelligence using the Trait Meta-Mood Scale. In: Pennebaker JW, editor. Emotion, Disclosure, and Health. Washington, DC: American Psychological Association; 1995:125–54.
29. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 2002;9:233–55.
30. Meredith W. Measurement invariance, factor analysis and factorial invariance. Psychometrika 1993;58:525–43.
31. Van de Pol F, Langeheine R. Mixed Markov latent class models. Sociol Methodol Oxford: Blackwell;1990:213–47.
32. Eid M. Longitudinal confirmatory factor analysis for polytomous item responses: model definition and model selection on the basis of stochastic measurement theory. Methods Psychol Res 1996;1:65–85.
33. Eid M. Latent-class models for analyzing variability and change. In: Ong AD, van Dulmen MHM, editors. Oxford Handbook of Methods in Positive Psychology. New York, NY: Oxford University Press; 2007:591–607.
34. Rijmen F, Vansteelandt K, de Boeck P. Latent class models for diary method data: parameter estimation by local computations. Psychometrika 2008;73:167–82.
35. Steyer R, Schwenkmezger P, Notz P, Eid M. MDBF – Mehrdimensionaler Befindlichkeitsfragebogen. Göttingen, Germany: Hogrefe; 1997.
36. Piasecki TM, Hufford MR, Solhan M, Trull TJ. Assessing clients in their natural environments with electronic diaries: rationale, benefits, limitations, and barriers. Psychol Assess 2007;19:25–43.
37. Lischetzke T, Eid M. Is attention to feelings beneficial or detrimental to affective well-being? Mood regulation as a moderator variable. Emotion 2003;3:361–77.
38. Vermunt JK, Magidson J. LG-Syntax User's Guide: Manual for Latent GOLD 4.5 Syntax Module. Belmont, MA: Statistical Innovations Inc.; 2008.
39. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna; 2010. Available at: http://www.R-project.org/.
40. Sturtz S, Ligges U, Gelman A. R2WinBUGS: a package for running WinBUGS from R. J Stat Softw 2005;12:1–16. Accessed May 3, 2011.
41. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput 2000;10:325–37.
42. Visser I, Speekenbrink M. depmixS4: an R package for hidden Markov models. J Stat Software 2010;36:1–21.
43. Rost J. Mixed and latent Markov models as item response models. MPR Online 2002;7:53–72.
44. Eid M, Langeheine R, Diener E. Comparing typological structures across cultures by latent class analysis: a primer. J Cross Cult Psychol 2003; 34:195–210.
45. Read TRC, Cressie NAC. Goodness-of-fit Statistics for Discrete Multivariate Data. New York, NY: Springer; 1988.
46. Agresti A, Yang M. An empirical investigation of some effects of sparseness in contingency tables. Comput Stat Data Anal 1986;5:9–21.
47. Schwarz G. Estimating the dimension of a model. Ann Stat 1978;6:461–4.
48. Nylund KL, Asparouhov T, Muthen B. Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. Struct Equ Modeling 2007;14:535–69.

49. Li F, Cohen AS, Kim S-H, Cho S-J. Model selection methods for mixture dichotomous IRT models. Appl Psychol Meas 2009;33:353–73.
50. Dias JG. Model selection criteria for model-based clustering of categorical time series data: a Monte-Carlo study. In: Decker R, Lenz HJ, editors. Advances in Data Analysis. Proc. 30th Annual conference of the Gesellschaft für Klassifikation. Berlin, Germany: Springer; 2007:23–30.
51. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions. Psychometrika 1987; 52:345–70.
52. Langeheine R, Van de Pol F. Latent Markov chains. In: Hagenaars JA, McCutcheon AL, editors. Applied Latent Class Analysis. Cambridge, MA: Cambridge University Press; 2002:304–41.
53. Courvoisier D, Eid M, Nussbeck FW. Mixture distribution state-trait models. Psychol Methods 2007;12:80–104.
54. Vermunt JK. Longitudinal research using mixture models. In: van Montfort K, Oud JHL, Satorra A, editors. Longitudinal Research With Latent Variables. Berlin, Germany: Springer; 2010:119–52.
55. Böckenholt U. A latent Markov model for the analysis of longitudinal data collected in continuous time: states, durations, and transitions. Psychol Methods 2005;10:65–83.