MULTIVARIATE BEHAVIORAL RESEARCH, *42*(4), 729–756

Bayesian Estimation of Categorical Dynamic Factor Models

Zhiyong Zhang and John R. Nesselroade
University of Virginia

Dynamic factor models have been used to analyze continuous time series behavioral data. We extend 2 main dynamic factor model variations—the direct autoregressive factor score (DAFS) model and the white noise factor score (WNFS) model—to categorical DAFS and WNFS models in the framework of the underlying variable method and illustrate them with a categorical time series data set from an emotion study. To estimate the categorical dynamic factor models, a Bayesian method via Gibbs sampling is used. The results show that today's affect directly influences tomorrow's affect. The results are then validated by means of simulation studies. Differences between continuous and categorical dynamic factor models are examined.

Analyzing change through modeling systematic fluctuation and patterns of intraindividual variability has become a familiar way to study many psychological processes, such as the decade-to-decade development of cognitive abilities, year-to-year change in adolescent substance abuse, and week-to-week fluctuations in mood (e.g., Baltes, Reese, & Nesselroade, 1977; Boker, 2002; Jones & Nesselroade, 1990). Given the goal of understanding patterns of change, research design and data analysis should attend to the following four considerations: First, because of the complexity of change processes, a multivariate approach to measurement is often necessary to capture an adequately detailed view of change (Baltes & Nesselroade, 1973; Jones & Nesselroade, 1990). Second, repeated measurements of the same individual are necessary to capture intraindividual

Correspondence concerning this article should be addressed to Zhiyong Zhang, Department of Psychology, P.O. Box 400400, University of Virginia, Charlottesville, VA 22903-4400. E-mail: zz5m@virginia.edu

change (Nesselroade & Ghisletta, 2000). Third, investigation of change at the level of the single subject is a promising way to inform the process of generalization. Fourth, appropriate statistical models need to be used (developed, if need be) to analyze the data. These considerations rather inexorably lead to multivariate time series data collection (e.g., Ferrer & Nesselroade, 2003; Molenaar, 1994; also see Jones & Nesselroade, 1990, for a review) and advanced statistical analytic techniques (e.g., Cattell, Cattell, & Rhymer, 1947; McArdle, 1982; Molenaar, 1985).

Compared with the wide availability of multivariate time series data from behavior research, an arsenal of powerful statistical models appears to be less developed. Although a variety of time series models has been well defined in economic research (e.g., Hamilton, 1994), appropriate models for behavioral data are still relatively scarce because of the complexity of behavioral phenomena and the involvement of latent structures in psychological theorizing (e.g., Jones & Nesselroade, 1990). In fact, despite some efforts to employ standard multivariate time series models (e.g., Schmitz, 1990), only two sets of models have been widely accepted: P-technique factor models and their extensions—dynamic factor models (DFMs). The P-technique factor model was originally developed to identify individual traits and has been applied in numerous studies (Cattell et al., 1947; see Jones & Nesselroade, 1990, and Luborsky & Mintz, 1972, for reviews). The major criticism of P-technique was that it only reflected concurrent relationships among variables and ignored any lagged relationships (Anderson, 1963; Cattell, 1963; Holtzman, 1963). The dynamic factor models were subsequently developed to incorporate both latent variables and lagged structures in analyzing multivariate time series data (e.g., Brillinger, 1975, 1981; McArdle, 1982; Molenaar, 1985).

Primarily, two kinds of DFMs have been discussed in the literature. One was called the direct autoregressive factor score model (DAFS) by Nesselroade, McArdle, Aggen, & Meyers (2001) and the process factor model by Browne and Nesselroade (2005; see also Engle & Watson, 1981; McArdle, 1982; Molenaar, 1985). The other was called the white noise factor score model (WNFS) by Nesselroade et al. (2001) and the shock factor model by Browne and Nesselroade (2005; see also Brillinger, 1975, 1981; Geweke & Singleton, 1981; Molenaar, 1985; Priestley, Rao, & Tong, 1973). Different procedures have been proposed for estimating the parameters of DFMs. Engle and Watson (1981) showed that what was later called the DAFS model was a special case of the state space model and can be estimated by maximum likelihood estimation (MLE) methods using a Kalman filter algorithm. Geweke and Singleton (1981) proposed an MLE for what was later called the WNFS model in the frequency domain. Molenaar (1985) employed MLE in the structural equation modeling (SEM) framework and the time domain based on the estimation of a block-Toeplitz matrix for both DAFS and WNFS models (see also Browne & Nesselroade, 2005; Molenaar &

Nesselroade, 1998; Nesselroade et al., 2001; Wood & Brown, 1994). Recently, Markov Chain Monte Carlo methods have become more widely accepted for estimating both DAFS and WNFS models (e.g., Justiniano, 2004; Kim & Nelson, 1999, 2001; West, 2000; Zhang, Hamaker, & Nesselroade, 2008). More recently, Browne and Zhang (2007) proposed a least squares estimation method and provided a computer program (DyFA) to implement it (Browne & Zhang, 2005).

The increasing interest in studying intraindividual behavioral variation by collecting multivariate response data across multiple occasions provides a promising opportunity to utilize dynamic factor models more widely (e.g., Ferrer & Nesselroade, 2003; Jones & Nesselroade, 1990; Luborsky & Mintz, 1972; Nesselroade et al., 2001). However, the models and the estimation methods mentioned previously have rested on the assumption that the observed variables are continuous and normally distributed. Considering that a lot of observed data in social and behavioral science are based on self-report measurements (e.g., Ferrer & Nesselroade, 2003; Lebo & Nesselroade, 1978) and are likely to be categorical at best and that treating categorical variables as continuous variables may render misleading results (e.g., Olsson, 1979; Song & Lee, 2002), it is desirable to develop and evaluate models appropriate to the measurement properties of such data. Using a rating scale data set from previous dynamic factor analysis (Nesselroade et al., 2001; Nesselroade & Molenaar, 2003), the authors aim to construct and evaluate what would seem to be suitable dynamic factor models for observed categorical time series data and to develop the appropriate estimation methods to support their use.

Although there is no significant amount of literature focused on the dynamic factor analysis of categorical time series data, the factor analysis of cross-sectional categorical data by latent variable models has been well explored (e.g., Jöreskog, 1994; Jöreskog & Moustaki, 2001; Lee, Poon, & Bentler, 1990, 1995; Lee & Song, 2003; Moustaki, 2000; Moustaki & Knott, 2000; Muthén, 1984; Shi & Lee, 1998). The so-called underlying variable (UV) method, for example, typically assumes that there is a normally distributed continuous variable underlying each observed categorical variable. Both maximum likelihood estimation methods (MLE; e.g., Jöreskog & Moustaki, 2001; Lee et al., 1990, 1995) and Bayesian estimation methods (e.g., Lee & Song, 2003; Shi & Lee, 1998) have been used to estimate the factor models with categorical data. Furthermore, dynamic cumulative models (Fahrmeir & Tutz, 1994) and a similar non-Gaussian model (Durbin & Koopman, 2001) have been used to analyze categorical time series data. These models can be viewed as particular versions of the categorical DAFS model discussed in this article.

Here we explicitly derive categorical DFMs from the widely used continuous DFMs and present a Bayesian approach for parameter estimation. First, the pertinent literature on the UV method is reviewed. The basics of the continuous data DFMs—both DAFS and WNFS models—are summarized. Then the UV method is applied to specify the DFMs for analyzing categorical time series data. Corresponding to continuous variable DAFS and WNFS, categorical DAFS and categorical WNFS models are presented. A Bayesian approach to the estimation of the models using Gibbs sampling is then presented. Data originally published by Lebo and Nesselroade (1978)[1] and Nesselroade et al. (2001) will be fitted using the proposed models. Finally, simulation studies will be carried out based on the results from the data analysis to evaluate the models validities further.

BACKGROUND AND HISTORICAL PERSPECTIVE

We first review some essential concepts on which our new developments rest. For convenience, throughout the article, p denotes the number of observed manifest variables; q denotes the number of the latent factors; N denotes the number of the observations; T denotes the largest number of occasions; L denotes the number of lags; and m_i denotes the number of categories for the ith observed variable unless defined specifically.

The Underlying Variable Approach to Factor Analysis

In the underlying response variable method (e.g., Jöreskog & Moustaki, 2001; Muthén, 1984; Shi & Lee, 1998), every observed ordinal variable is assumed to be generated by an underlying, unobserved continuous variable. The factor model is the same as the classical factor analysis model for the underlying variables,

$$\mathbf{z}_n = \mathbf{\Lambda f}_n + \mathbf{u}_n, \quad n = 1, \ldots, N,$$

where $\mathbf{z}_n$ is a $p \times 1$ underlying score vector; $\mathbf{\Lambda}$ is a $p \times q$ factor loading matrix; $\mathbf{f}_n$ is a $q \times 1$ factor score vector; and $\mathbf{u}_n$ is a $p \times 1$ uniqueness factor vector. The relationship between observed data $\mathbf{Y} = (y_{in})_{p\times N}$ and underlying data $\mathbf{Z} = (z_{in})_{p\times N} = (\mathbf{z}_1, \ldots, \mathbf{z}_N)$ is given by

$$y_{in} = k \Leftrightarrow \tau_{i,k-1} < z_{in} \leq \tau_{i,k},$$

$$i = 1, \ldots, p, \quad k = 1, \ldots, m_i, \quad n = 1, \ldots, N,$$

[1]We are grateful to Dr. Michael A. Lebo for permission to use these data.

where $-\infty \leq \tau_{i,0} < \tau_{i,1} < \cdots < \tau_{i,m_i-1} < \tau_{i,m_i} \leq +\infty$ are thresholds. Usually, the underlying variable is assumed to be normally distributed although it can have any continuous distribution.

The Dynamic Factor Models

Common factor analysis of multivariate time series data (called P-technique factor analysis) has evolved over the past 60 years. Cattell et al. (1947) fitted a factor model to the covariation of multiple variables measured across time on only one individual. An obvious drawback of P-technique is that it cannot capture the order relationships of a process (Anderson, 1963; Holtzman, 1963). It represents only the concurrent relationships (i.e., between factors and observed variables) of a process. Dynamic factor models (Brillinger, 1975, 1981; Browne & Nesselroade, 2005; Engle & Watson, 1981; Geweke & Singleton, 1981; McArdle, 1982; Molenaar, 1985; Nesselroade et al., 2001; Priestley et al., 1973), by contrast, can effectively represent a process by incorporating lagged relationships among the variables, both manifest and latent, involved.

The direct autoregressive factor score model (DAFS). A version of what was later called the DAFS model was proposed by Engle and Watson (1981) and then proposed as an SEM with psychological variables by McArdle (1982). The DAFS model specification is written as

$$\mathbf{y}_t = \mathbf{\Lambda f}_t + \mathbf{u}_t, \tag{1}$$

$$\mathbf{f}_t = \sum_{l=1}^{L} \mathbf{B}_l \mathbf{f}_{t-1} + \mathbf{v}_t, \tag{2}$$

where $\mathbf{y}_t$ is a p-variate observed time series measured at time $t(t = 1, \ldots, T)$, $\mathbf{\Lambda}$ is a $p \times q$ matrix of factor loadings, $\mathbf{f}_t$ is a q-variate factor vector at time $t(t = 1, \ldots, T)$, $\mathbf{f}_{t-l}$, $(l = 1, \ldots, L)$ is a q-variate factor vector l occasions prior to occasion t, $\mathbf{B}_l$ is the autoregressive and cross-regressive coefficient matrix on the factors l occasions prior to the current time, and $\mathbf{u}_t$ and $\mathbf{v}_t$ are uniqueness following a multivariate normal distribution with mean 0 and covariance matrix $\mathbf{Q}$ and $\mathbf{D}$.

The white noise factor score model (WNFS). What we are referring to as the WNFS model was first explicitly represented by Geweke and Singleton (1981) and further developed and rendered practical by Molenaar (1985), although particular versions of this model can be traced back to Priestley et al.

(1973) and Brillinger (1975, 1981). This model assumes that the factors influence both current observed variable scores and future observed scores directly. The model specification is written as

$$\mathbf{y}_t = \sum_{l=0}^{L} \mathbf{\Lambda}_l \mathbf{f}_{t-l} + \mathbf{e}_t, \tag{3}$$

where $\mathbf{y}_t$ is a p-variate observed time series measured at time $t(t = 1, \ldots, T)$, $\mathbf{\Lambda}_l$ is a $p \times q$ matrix of factor loadings at lag l, $\mathbf{f}_t$ is a q-variate factor vector at time $t(t = 1, \ldots, T)$, $\mathbf{f}_{t-l}(l = 0, \ldots, L)$ is a q-variate factor vector l occasions prior to occasion t, and $\mathbf{u}_t$ is the unique factor vector with mean 0 and covariance matrix $\mathbf{Q}$.

CATEGORICAL DYNAMIC FACTOR MODELS

The extension of DFMs from continuous data to categorical data is straightforward given an understanding of the underlying variable method and the specifications of dynamic factor models. The basic idea is that the relations between observed categorical time series variables and the underlying continuous variables are first established and then the DFMs are applied to the underlying continuous variables. Corresponding to their continuous dynamic factor models, the categorical direct autoregressive factor score (CDAFS) model and the categorical white noise factor score (CWNFS) model are presented here.

In accord with the UV model described earlier, we first construct the relationship between observed categorical data $\mathbf{Y} = (y_{it})_{p\times T}$ and underlying latent continuous data $\mathbf{Z} = (z_{it})_{p\times T} = (\mathbf{z}_1, \ldots, \mathbf{z}_T)$ via the thresholds $-\infty \le \tau_{i,0} < \tau_{i,1} < \cdots < \tau_{i,m_i-1} < \tau_{i,m_i} \le +\infty$,

$$\begin{gathered} y_{it} = k \Leftrightarrow \tau_{i,k-1} < z_{it} \le \tau_{i,k}, \\ i = 1, \ldots, p, \quad k = 1, \ldots, m_i, \quad t = 1, \ldots, T. \end{gathered} \tag{4}$$

Then we can construct a DAFS model for the underling variables according to Eqs. (1) and (2).

$$\mathbf{z}_t = \mathbf{\Lambda} \mathbf{f}_t + \mathbf{u}_t, \tag{5}$$

$$\mathbf{f}_t = \sum_{l=1}^{L} \mathbf{B}_l \mathbf{f}_{t-l} + \mathbf{v}_t. \tag{6}$$

We call the model in Eqs. (4)–(6) the CDAFS model. Particular versions of this model can be found in Fahrmeir and Tutz (1994) and Durbin and Koopman (2001).

Similar to the CDAFS model, the CWNFS model can be expressed using Eqs. (7) and (8),

$$y_{it} = k \Leftrightarrow \tau_{i,k-1} < z_{it} \leq \tau_{i,k}, \quad i = 1, \ldots, p, \quad k = 1, \ldots, m_i, \quad t = 1, \ldots, T, \tag{7}$$

$$\mathbf{z}_t = \sum_{l=0}^{L} \boldsymbol{\Lambda}_l \mathbf{f}_{t-1} + \mathbf{u}_t. \tag{8}$$

As in the static factor analysis with categorical data (e.g., Lee et al., 1990, 1995), we consider the identification problem first. Using conditional distributions facilitates the discussion of this problem. Conditionally on the factor scores $\mathbf{f}_t$, $\mathbf{z}_t$ is multivariate normally distributed with mean $\boldsymbol{\Lambda}\mathbf{f}_t$ and covariance matrix $\mathbf{Q}$ for the CDAFS model and with mean $\sum_{l=0}^{L} \boldsymbol{\Lambda}_l \mathbf{f}_{t-1}$ and covariance matrix $\mathbf{Q}$ for the CWNFS model. Then the cell probability for the CDAFS and CWNFS models at time t can be expressed as

$$\Pr(y_{it} = k) = \Pr(\tau_{i,k-1} < z_{it} \leq \tau_{i,k})$$
$$= \begin{cases} \boldsymbol{\Phi}\left(\dfrac{\boldsymbol{\tau}_{i,k} - \Lambda_i \mathbf{f}_t}{q_i}\right) - \boldsymbol{\Phi}\left(\dfrac{\boldsymbol{\tau}_{i,k-1} - \Lambda_i \mathbf{f}_t}{q_i}\right) & \text{for the CDAFS model and} \\ \boldsymbol{\Phi}\left(\dfrac{\boldsymbol{\tau}_{i,k} - \sum_{l=0}^{L} \Lambda_{li} \mathbf{f}_{t-l}}{q_i}\right) - \boldsymbol{\Phi}\left(\dfrac{\boldsymbol{\tau}_{i,k-1} - \sum_{l=0}^{L} \Lambda_{li} \mathbf{f}_{t-l}}{q_i}\right) & \text{for the CWNFS model} \end{cases}, \tag{9}$$

where Λ_i is the ith row of the factor loading matrix, Λ_{li} is the ith row of the lag l factor loading matrix, and q_i is the standard deviation of the ith underlying continuous variable and is also the square root of the ith diagonal element of $\mathbf{Q}$, conditionally on the factor scores.

The models in Eq. (9) will not be identified unless additional constraints on $\tau_{i,k}$ or q_i are introduced. Two kinds of constraints can be used to identify the models. One involves fixing $q_i \equiv 1$ for all i and another is to fix the threshold to be a pre-assigned constant number. As pointed out by Lee et al. (1990), the first option imposes complex nonlinear constraint on the covariance structure. Thus, for our analysis, we adopt the second way. For a complete discussion of the constraint problem in the factor analysis of categorical data, see Lee et al. (1990) and Shi and Lee (1998).

Different methods can be used to determine the threshold values (e.g., Jöreskog, 1994; Lee et al., 1990; Olsson, 1979; Shi & Lee, 1998). In our example to follow, we adopted the method used in Olsson (1979) to obtain the thresholds. In this case,

$$\tau_{i,k} = \mathbf{\Phi}^{-1}(p_{i,1} + p_{i,2} + \cdots + p_{i,k}), \quad k = 1, 2, \ldots, m_i - 1, \tag{10}$$

where $p_{i,s}$ is the corresponding percentage of responses in category s for the ith observed variable. It is shown that this two-step method obtained similar estimation of thresholds compared with the one-step method (e.g., Jöreskog & Moustaki, 2001). Using these thresholds obtained here as the pre-defined values will speed the mixing of the posterior in the Gibbs sampling discussed later (Albert & Chib, 1993).

As in any factor analysis, we also need to put constraints on either factor loadings or factor variances to identify the dynamic factor models. For the CDAFS model, we fix the variances of the residual errors for the factors in Eq. (6) to be 1 (Molenaar, 1985). For the CWNFS model, we fix the covariance matrix for the factors in order to identify all the factor loadings (Molenaar, 1985). If we aim to estimate the covariances among factors, as in the empirical analysis given herein, we can fix the variances of the factors to be 1 and some of the factor loadings to be 0 as in the usual factor analysis (e.g., Jöreskog, 1969). Then the unknown parameters are $\Theta = (\mathbf{\Lambda}, \mathbf{B}_l, \mathbf{Q}, \mathbf{D}^-)$ for the CDAFS model and $\Theta = (\mathbf{\Lambda}_l, \mathbf{Q}, \mathbf{D}^-)$ for the CWNFS model with $\mathbf{D}^-$ representing the off diagonal elements of the covariance matrix of $\mathbf{v}_t$ for the CDAFS model and the off diagonal elements of the covariance matrix of factors for the CWNFS model. Note that some of the elements in $\mathbf{\Lambda}$ or $\mathbf{\Lambda}_l$ must be 0. In the following section, a Bayesian estimation method is outlined to estimate the CDAFS and CWNFS models.

BAYESIAN ESTIMATION OF THE MODELS

Bayesian estimation of statistical models has been previously applied to categorical time series. For example, Fahrmeir and Tutz (1994) used a Gibbs sampling procedure to estimate dynamic cumulative models. Durbin and Koopman

(2001) also discussed an importance sampling procedure for the state space models involving categorical data. Here we outline the procedure for obtaining the parameter estimates for the categorical dynamic factor models.

Let $p(\Theta)$ be the prior distribution of all the parameters and $\mathbf{Y}$ represent the observed categorical data. Standard Bayesian approach requires the evaluation of the posterior distribution of Θ given $\mathbf{Y}$: $p(\Theta|\mathbf{Y}) \propto p(\Theta)p(\mathbf{Y}|\Theta)$. For our model, it is extremely difficult to evaluate this posterior distribution directly because of the involvement of the latent variables and categorical measurements. Therefore, data augmentation and Gibbs sampling strategies are used (Albert & Chib, 1993; Geman & Geman, 1984; Tanner & Wong, 1987).

Let $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_T)$ and $\mathbf{F} = (\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T)$ represent the underlying continuous data and latent factor scores, respectively. Based on the data augmentation ideas given in Tanner & Wong (1987), the observed data, $\mathbf{Y}$, are augmented with $(\mathbf{Z}, \mathbf{F})$. As pointed out by Tanner & Wong, the posterior distribution $p(\Theta|\mathbf{Y}, \mathbf{Z}, \mathbf{F})$ is easier to handle than $p(\Theta|\mathbf{Y})$. After augmenting the data, the Gibbs sampling procedure (Gelfand & Smith, 1990; Geman & Geman, 1984) is then used to generate observations from the conditional distributions using the following algorithm.

At the $(i+1)$th iteration with current values $(\Theta^{(i)}, \mathbf{Z}^{(i)}, \mathbf{F}^{(i)})$, generate[2]

$$\Theta^{(i+1)} \text{ from } p(\Theta|\mathbf{Y}, \mathbf{Z}^{(i)}, \mathbf{F}^{(i)})$$

$$\mathbf{Z}^{(i+1)} \text{ from } p(\mathbf{Z}|\mathbf{Y}, \Theta^{(i+1)}, \mathbf{F}^{(i)})$$

$$\mathbf{F}^{(i+1)} \text{ from } p(\mathbf{F}|\mathbf{Y}, \Theta^{(i+1)}, \mathbf{Z}^{(i+1)}).$$

This iteration can be repeated I times. Geman & Geman (1984) showed that for sufficiently large I, $(\Theta^{(I)}, \mathbf{Z}^{(I)}, \mathbf{F}^{(I)})$ can be viewed as a simulated observation from the posterior distribution $p(\Theta, \mathbf{Z}, \mathbf{F}|\mathbf{Y})$ under mild conditions. There are different ways to determine I. In practice, the "eyeball" method, which monitors the convergence by visually inspecting the plots of the generated sequences, is commonly used. Here, in addition to the "eyeball" method, the convergence was also evaluated by Gewekes convergence diagnostic (Geweke, 1992; see also Cowels & Carlin, 1996) in CODA (Plummer, Best, Cowels, & Vines, 2005). If Gewekes diagnostic is less than 1.96, the sequence is considered to have converged.

The simulated observations after I are then recorded for further analysis. For convenience, these observations are denoted as $(\Theta^{(m)}, \mathbf{Z}^{(m)}, \mathbf{F}^{(m)})$, $m = 1, 2, 3, \ldots$. Sometimes there are highly positive autocorrelations between the successive observations. To reduce the autocorrelations, we can select the observations with fixed interval a indexed 1, $1+a$, $1+2a$, $1+3a, \ldots$ on which

[2]The full conditional distributions can be obtained by request or viewed at http://dfa.psychstat.org

to perform further analysis. After generating the observations, the Bayesian estimates are calculated by

$$\hat{\Theta} = \hat{E}(\Theta|\mathbf{Y}) = \frac{1}{N}\sum_{m=0}^{N-1}\Theta^{1+ma},$$

with variance (or covariance matrix)

$$\hat{V}(\Theta|\mathbf{Y}) = \frac{1}{N-1}\sum_{m=0}^{N-1}(\Theta^{1+ma} - \hat{\Theta})(\Theta^{1+ma} - \hat{\Theta})^t.$$

For the CDAFS and CWNFS models, the Bayesian method using Gibbs sampling is a very convenient tool for obtaining the parameter estimates. For example, to estimate the models, we need to specify the initial factor scores. Specifically, for the one-lag models, the initial scores $\mathbf{f}_0$ are required. For the two-lag models, the initial scores $\mathbf{f}_0$ and $\mathbf{f}_{-1}$ are required. We have found that it is not very easy to handle this problem in MLE. However, with the Bayesian method, we can estimate these initial scores as unknown parameters directly (Zhang, Hamaker, et al., 2007).

For the analysis of empirical data, three indexes for model selection can be used to compare and select models—DIC, EBIC, and EAIC, defined as follows. Deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & Linde, 2002) is a widely used criterion for model selection in the Bayesian framework. DIC is defined as a Bayesian measure of model fit with a penalty for model complexity p_D,

$$\text{DIC} = \overline{D(\Theta)} + p_D = D(\overline{\Theta}) + 2p_D,$$

where $\overline{D(\Theta)}$ is the posterior mean of -2(LogLikehood function) and $D(\overline{\Theta})$ is -2(LogLikehood function) calculated at the posterior mean of Θ.

Extensions of the Bayesian information criterion (BIC; Raftery, 1993; Schwarz, 1978) and Akaike's information criterion (AIC; Akaike, 1973) are also used to compare model fits. In the Bayesian context, one kind of extension of the AIC and BIC is the EAIC and EBIC (e.g., Spiegelhalter et al., 2002), which are calculated by

$$\text{EBIC} = \overline{D(\Theta)} + 2p$$

and

$$\text{EAIC} = \overline{D(\Theta)} + \ln(T)p,$$

where T is the length of the time series and p is the number of parameters to be estimated. For all three fit indexes, smaller value means better model fit.

AN EMPIRICAL STUDY OF AFFECTIVE VARIABILITY

Dynamic factor models have been successfully applied to time series of affect (e.g., Ferrer & Nesselroade, 2003; Nesselroade et al., 2001; Shifren, Hooker, Wood, & Nesselroade, 1997). Affect data seem especially well suited for DFMs for several reasons. First, the individual is likely to experience many substantial changes in his or her mood over even relatively short time periods (Ferrer & Nesselroade, 2003). Second, mood variation tends to be multiple dimensional (Shifren et al., 1997). Third, it is very reasonable to hypothesize that current affect (e.g., today) will influence future effect (e.g., tomorrow), and this lagged effect has been found in previous research (Ferrer & Nesselroade, 2003; Nesselroade et al., 2001; Shifren et al., 1997). However, most, if not all, sets of emotion data are measured categorically instead of continuously. Thus, there should be good fit between such data and categorical DFMs.

To illustrate the application of the categorical DFMs and the Bayesian estimation method, the affect data (Lebo & Nesselroade, 1978) that were analyzed by Nesselroade et al. (2001) were used. The data are for a single participant who reported her emotional status daily on six 5-point Likert scales for 103 successive days—*active* (A), *lively* (L), *peppy* (P), *sluggish* (S), *tired* (T), and *weary* (W). The data for these six variables are plotted in Figure 1. The thresholds calculated by Eq. (10) are summarized in Table 1.

Using the threshold values in Table 1 as the pre-assigned parameters, seven more or less likely models listed in Table 2 are fitted and compared. The first four models, whose path diagrams are given in Figure 2, are CDAFS models and the last three models, whose path diagrams are given in Figure 3, are CWNFS mod-

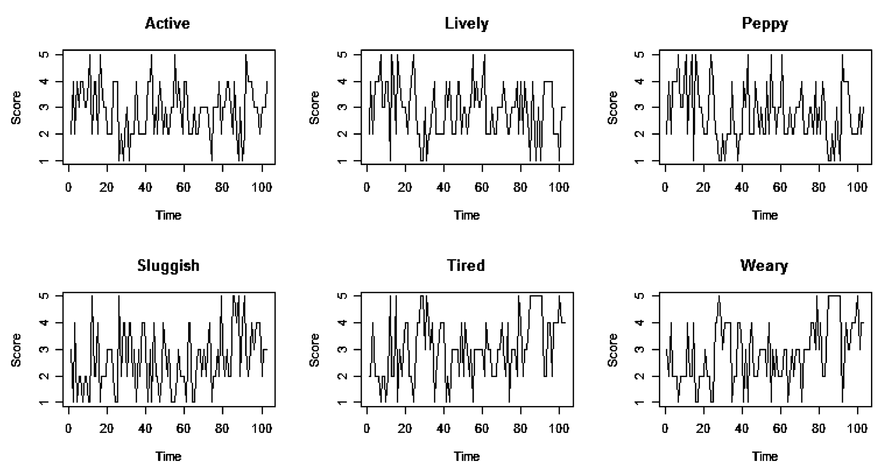

FIGURE 1 Time series plot of the observed data.

TABLE 1
Thresholds for the Six Observed Categorical Variables

	$\tau_{,1}$	$\tau_{,2}$	$\tau_{,3}$	$\tau_{,4}$
Active	−1.5696	−1.4209	0.5216	1.6591
Lively	−1.4209	−0.2584	0.4666	1.5096
Peppy	−1.3571	0.0122	0.6072	1.3571
Sluggish	−0.8985	0.0365	0.6669	1.4911
Tired	−1.3571	−0.3347	0.3091	1.0988
Weary	−1.1926	−0.2333	0.4666	1.2983

els. Model M1 is a one-factor (U) dynamic model with all six observed variables loading on one factor. The factor score has a one-lag autoregression structure. Models M2–M4 are two-factor models with the first three variables loading on the first factor and the last three variables loading on the second factor. The first factor is labeled energy (E) and the second one is labeled fatigue (F) following Lebo & Nesselroade (1978) and Nesselroade et al. (2001). Both M2 and M3 have a one-lag structure, whereas M2 has only an autoregressive structure and M3 also has the cross-regressive structure. M4 is a two-lag model with only an autoregressive structure. M5 is a one-factor CWNFS model with one lag. M6 and M7 are two-factor models with one lag and two lags, respectively.

Parameters of the seven models were estimated using the Bayesian method via Gibbs sampling as outlined earlier. The Gibbs sampling procedures were implemented in WinBUGS (Spiegelhalter, Thomas, Best, & Lunn, 2003) with

TABLE 2
Model Fit Statistics

Models	*Number of Parameters*	*DIC*	*EAIC*	*EBIC*
CDAFS models				
M1: 1 factor, 1 lag	13	1302	1328	1362
M2: 2 factors, 1 lag, no cross	15	1171	1201	1240
M3: 2 factors, 1 lag, cross	17	1175	1209	1254
M4: 2 factors, 2 lags, no cross	19	1172	1210	1260
CWNFS models				
M5: 1 factor, 1 lag	18	1302	1338	1385
M6: 2 factors, 1 lag	19	1180	1218	1268
M7: 2 factors, 2 lags	25	1180	1230	1296

Note. The number of parameters does not include the number of thresholds.

DIC: Deviance information criterion; EAIC: Extended Akaike's information criterion; EBIC: Extended Bayesian information criterion; CDAFS: Categorical direct autoregressive factor score; CWNFS: White noise factor score.

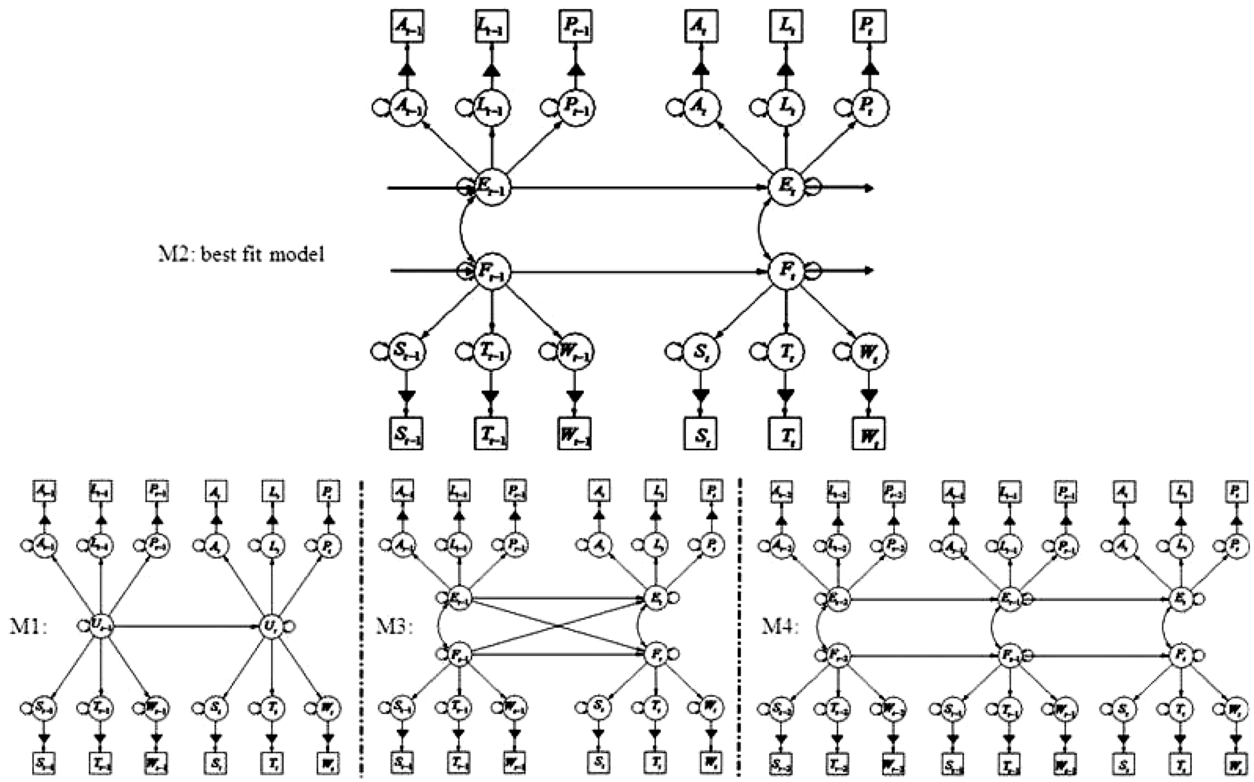

FIGURE 2 Four possible categorical direct autoregressive factor score (CDAFS) models. The squares with the first letter of the six observed variables represent the observed categorical data. The cycles with the same letters represent the latent continuous variables underlying the observed variables.

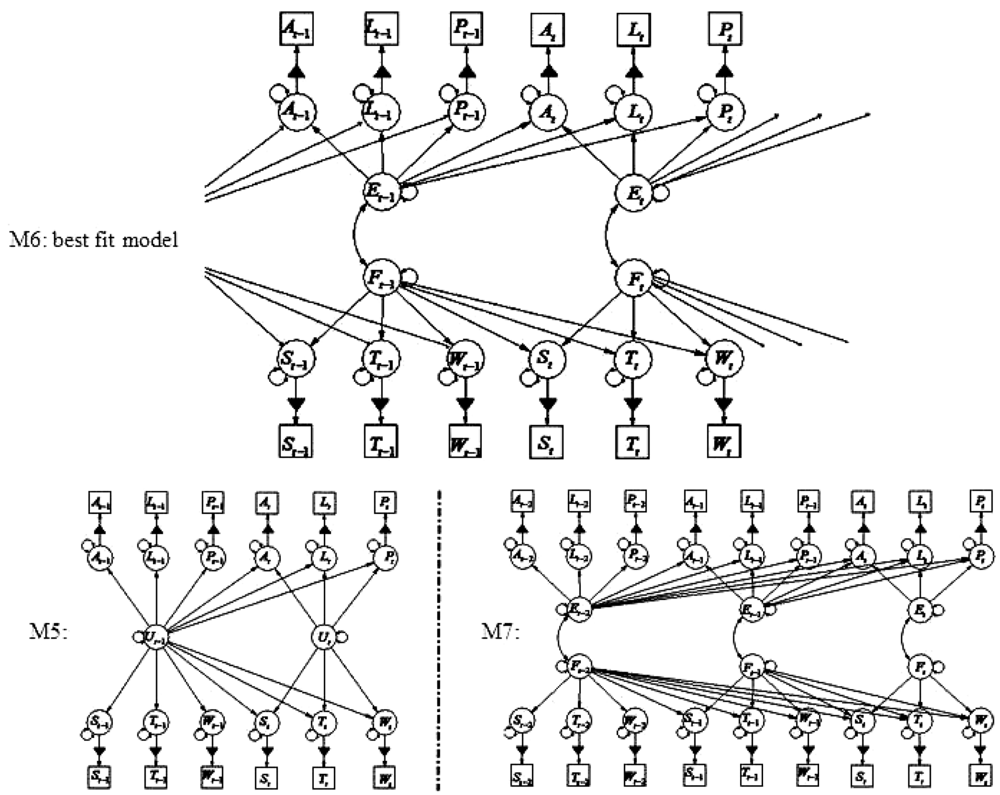

FIGURE 3 Three possible categorical white noise factor score (CWNFS) models. The squares with the first letter of the six observed variables represent the observed categorical data. The cycles with the same letters represent the latent continuous variables underlying the observed variables.

the convergence evaluated using CODA (Plummer et al., 2005) in R (R Development Core Team, 2005). In the current case, all the parameters are given non-informative priors. The prior distributions for $\mathbf{\Lambda}$ and $\mathbf{B}$ are specified as normal $N(0, 1.0\text{E-}6)$, which can be viewed as non-informative priors (Congdon, 2003). The prior distributions for $\mathbf{Q}$ were specified as the inverse Gamma distribution $IG(.0001, .0001)$, which is also a non-informative prior (Congdon, 2003). For the elements in $\mathbf{D}^{-}$, the uniform distribution $U(-1, 1)$ was used (Chib & Greenberg, 1998). See Appendix for WinBUGS codes. A formal discussion on the selection of priors can be found in Kass and Wasserman (1996).

The model fit statistics are summarized in Table 2. Based on the fit statistics, for the CDAFS models, the two-factor model with one lag and no cross-regression is the best fitting model. For the CWNFS models, the two-factor model with one lag is the best fitting model. However, the goodness of fit for M2, M3, and M4 is very close in value. The goodness of fit for M6 and M7 is also very close. Thus, we present the results for these five models for the purpose of demonstration. The model estimates for the CDAFS models are summarized in Table 3 and the model estimates for the CWNFS models are given in Table 4. We focus on the results from M2 and M6.

For both the CDAFS model and CWNFS model, a two-factor model with only one lag appears to represent the data best. The correlation between the energy and fatigue factors is very high and negative ($-.87$ and $-.85$ for CDAFS and CWNFS, respectively). For the CDAFS model, the autoregressive coefficients for energy and fatigue factors are .38 and .47, respectively, indicating that the fatigue factor tends to dissipate more slowly than the energy factor. The corresponding relationships are indicated in the CWNFS model by the fact that the factor loadings for the Lag 1 fatigue factor are larger than those for the Lag 1 energy factor, relative to the factor loadings at Lag 0.

EVALUATION OF CATEGORICAL DFMS AND COMPARISON WITH CONTINUOUS DFMS

To evaluate further the performance of the models and the estimation method in the empirical study, we simulated the data based on the best fitting models. The data were simulated from Model M2 for the CDAFS model and from Model M6 for the CWNFS model. The true parameter values are based on the estimates from Lebo data but with simplification. For the CDAFS model, $\mathbf{B}_1 = (b_{ij})_{2\times2} = \begin{pmatrix} .4 & 0 \\ 0 & .5 \end{pmatrix}$, $\mathbf{D} = (d_{ij})_{2\times2} = \begin{pmatrix} 1 & -.85 \\ -.85 & 1 \end{pmatrix}$, $\mathbf{\Lambda} = (\lambda_{ij})_{6\times2} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & .9 & .9 & .9 \end{pmatrix}^t$, and $\mathbf{Q} = (q_{ij})_{6\times6} = diag\begin{pmatrix} .5 & .1 & .05 & .3 & .2 & .2 \end{pmatrix}$.

TABLE 3
Model Estimates for Categorical Direct Autoregressive Factor Score (CDAFS) Models

3a. CDAFS Model With 2 Factors, 1 Lag, and No Cross-Regression Coefficients (M2)

	Factor Loadings		
Variable	*Energy*	*Fatigue*	*Uniqueness Variance*
Active	1.19(.12)		.45(.10)
Lively	0.95(.09)		.13(.04)
Peppy	1.01(.09)		.07(.04)
Sluggish		0.83(.10)	.32(.07)
Tired		0.92(.10)	.17(.05)
Weary		0.91(.10)	.17(.05)

	Factor Score Autoregression	
	Energy (t − 1)	*Fatigue (t − 1)*
Energy (t)	0.38(.10)	
Fatigue (t)		0.47(.09)

Note. The number in parenthesis is the standard error. The correlation between factor energy and fatigue is −.85.

3b. CDAFS Model With 2 Factors, 1 Lag, and Cross-Regression (M3)

	Factor Loadings		
Variable	*Energy*	*Fatigue*	*Uniqueness Variance*
Active	1.18(.12)		.44(.10)
Lively	0.94(.09)		.13(.04)
Peppy	1.01(.09)		.07(.04)
Sluggish		0.79(.10)	.32(.07)
Tired		0.88(.10)	.18(.05)
Weary		0.88(.09)	.17(.05)

	Factor Score Autoregression	
	Energy (t − 1)	*Fatigue (t − 1)*
Energy (t)	0.22(.23)	−0.004(.21)
Fatigue (t)	−0.46(.24)	0.73(.21)

Note. The number in parenthesis is the standard error. The correlation between factor energy and fatigue is −.86.

TABLE 3
(*Continued*)

3c. CDAFS Model With 2 Factors, 2 Lags, and Without Cross-Regression (M4)

	Factor Loadings		
Variable	*Energy*	*Fatigue*	*Uniqueness Variance*
Active	1.17(.12)		.45(.10)
Lively	0.93(.09)		.13(.04)
Peppy	1.00(.09)		.06(.04)
Sluggish		0.82(.10)	.32(.07)
Tired		0.91(.10)	.17(.05)
Weary		0.89(.10)	.17(.05)

	Factor Score Autoregression			
	Energy (t − 1)	*Fatigue (t − 1)*	*Energy (t − 2)*	*Fatigue (t − 2)*
Energy (t)	0.27(.10)	—	0.17(.10)	—
Fatigue (t)	—	0.35(.10)	—	0.19(.09)

Note. The number in parenthesis is the standard error. The correlation between factor energy and fatigue is −.86.

For the CWNFS model, $\mathbf{\Lambda}_0 = (\lambda_{0ij})_{6\times 2} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}^t$, $\mathbf{\Lambda}_1 = (\lambda_{1ij})_{6\times 2} = \begin{pmatrix} .2 & .2 & .2 & 0 & 0 & 0 \\ 0 & 0 & 0 & .2 & .2 & .2 \end{pmatrix}^t$, $\mathbf{D} = (d_{ij})_{2\times 2} = \begin{pmatrix} 1 & -.85 \\ 0.85 & 1 \end{pmatrix}$, and $\mathbf{Q} = (q_{ij})_{6\times 6} = diag\begin{pmatrix} .5 & .1 & .05 & .3 & .2 & .2 \end{pmatrix}$. For both models, the thresholds are set at $\tau_{i,} = \begin{pmatrix} -1.5 & -.5 & .5 & 1.5 \end{pmatrix}$. The simplification of the parameters will not change the features of the simulation studies. The continuous data were first generated from the DAFS and WNFS models and then divided into categorical data according to the thresholds. For each model, 100 replications with the length of $T = 100$ were generated.

For each replication, the data were analyzed using the following strategy. The generated data were first fitted by the true model, M2 for the CDAFS model and M6 for the CWNFS model. The parameter estimates along with the standard errors were obtained. Then for each parameter in one replication, we can obtain its estimate P_i and standard error $SE_i (i = 1, \ldots, 100)$. The mean parameter estimate across 100 replications is calculated by $ME = \overline{P} = \sum_{i=1}^{100} P_i/100$, the standard deviation for this parameter estimate is calculated by

TABLE 4
Model Estimates for Categorical White Noise Factor Score (CWNFS) Models

4a. CWNFS Model With 2 Factors and 1 Lag (M6)

	Factor Loadings				
	Factor (Lag 0)		*Factor (Lag 1)*		
Variable	*Energy*	*Fatigue*	*Energy*	*Fatigue*	*Uniqueness Variance*
Active	1.21(.13)		0.23(.14)		.46(.10)
Lively	0.98(.09)		0.17(.10)		.11(.04)
Peppy	1.01(.10)		0.27(.11)		.08(.04)
Sluggish		0.91(.11)		0.19(.12)	.30(.07)
Tired		0.96(.10)		0.31(.10)	.16(.05)
Weary		0.93(.10)		0.32(.10)	.20(.06)

Note. The number in parenthesis is the standard error. The correlation between factor energy and fatigue is −.85.

4b. CWNFS Model with 2 Factors and 2 Lags (M7)

	Factor Loadings						
	Factor (Lag 0)		*Factor (Lag 1)*		*Factor (Lag 2)*		
Variable	*Energy*	*Fatigue*	*Energy*	*Fatigue*	*Energy*	*Fatigue*	*Uniqueness Variance*
Active	1.24(.14)		0.21(.16)		0.32(.17)		.64(.11)
Lively	1.00(.10)		0.15(.12)		0.22(.12)		.15(.04)
Peppy	1.06(.10)		0.25(.12)		0.21(.13)		.07(.04)
Sluggish		0.96(.12)		0.15(.13)		0.10(.13)	.40(.07)
Tired		0.98(.11)		0.26(.13)		0.25(.14)	.33(.06)
Weary		0.94(.11)		0.29(.12)		0.26(.13)	.31(.06)

Note. The number in parenthesis is the standard error. The correlation between factor energy and fatigue is −.86.

$SD = \sum_{i=1}^{100}(P_i - \overline{P})^2/99$, and the mean of the standard errors for this parameter is calculated by $MSE = \sum_{i=1}^{100} SE_i/100$. Furthermore, the generated data were analyzed using the corresponding continuous DAFS and WNFS models. Finally, the data from the CDAFS model were analyzed using the alternative model M3 and the data from the CWNFS model were analyzed using the alternative model M7. The model fit indexes, DIC, EBIC, and EAIC, were compared with those fit indexes from M2 and M6 to determine how accurately the true models can be selected.

TABLE 5
Simulation Results for the Categorical Direct Autoregressive Factor Score (CDAFS) Models

	TRUE	*ME*	*SD*	*MSE*	*AS-CON*[a]
b_{11}	0.4	0.38	0.084	0.089	0.98(.01)
b_{22}	0.5	0.47	0.081	0.086	0.97(.02)
r_{12}	−0.85	−0.83	0.047	0.045	−0.57(.19)
λ_{11}	1	1.04	0.126	0.124	1.08(.10)
λ_{21}	1	1.03	0.096	0.097	1.09(.09)
λ_{31}	1	1.03	0.098	0.095	1.09(.10)
λ_{41}	0.9	0.92	0.089	0.104	0.95(.08)
λ_{52}	0.9	0.92	0.087	0.096	0.94(.08)
λ_{62}	0.9	0.92	0.084	0.098	0.94(.08)
q_{11}	0.5	0.54	0.121	0.118	0.51(.09)
q_{22}	0.1	0.10	0.038	0.045	0.18(.04)
q_{33}	0.05	0.07	0.032	0.039	0.14(.04)
q_{44}	0.3	0.34	0.071	0.085	0.38(.06)
q_{55}	0.2	0.21	0.063	0.065	0.28(.05)
q_{66}	0.2	0.23	0.072	0.067	0.29(.06)

	Model Selection[b]		
	DIC	*EBIC*	*EAIC*
M2	69	98	91
M3	31	2	9

Note. The number in parenthesis is the standard error.

[a]The simulated categorical data were analyzed as continuous data.

[b]The number means how many times the model (M2 or M3) was chosen based on the fit statistics.

TRUE: True parameter value; ME: Mean parameter estimate; SD: Standard deviation; MSE: Mean standard error; DIC: Deviance information criterion; EAIC: Extended Akaike's information criterion; EBIC: Extended Bayesian information criterion.

The results for the CDAFS model are presented in Table 5. For every parameter, the one standard error confidence interval contains the true parameter value. Furthermore, the differences between the true parameter values and the mean estimates are quite small. The SEs for the autoregressive parameters and the correlation between factors are underestimated, and the SEs for the factor loadings and the uniqueness variances are almost all overestimated, although the differences are very small. Overall, the Bayesian estimation method works well for this model. When the data were analyzed as continuous data, the estimates for the autoregressive parameters are highly overestimated and the estimate for the factor correlation is highly underestimated. Finally, when comparing the true model (M2) and the alternative model (M3), only 69 out 100 replications prefer M2 based on DIC. EBIC is the most accurate among the three indexes, as it correctly chose the true model 98 times out of 100 replications.

The results for the CWNFS model are presented in Table 6. The most noticeable result is that all the parameters are consistently overestimated. However, one SE confidence interval contains the true value for each parameter. For the SE, there is no consistent difference. Furthermore, when analyzing the data as continuous data, the correlation between factors is also underestimated. Finally, when comparing the true model (M6) and the alternative model (M7), only 48 out of 100 replications yielded the true model based on DIC. EBIC and EAIC correctly distinguished the true model 59 times out of 100 replications from the alternative model.

CONCLUSION AND DISCUSSION

We have illustrated the modeling of categorical time series data using dynamic factor models based on constructing relationships between the observed categorical variables and the underlying continuous variables. The corresponding CDAFS and CWNFS models can be constructed on the underlying continuous variables and estimated using Bayesian methods via Gibbs sampling. The application of the models and estimation method were demonstrated on the affective data. Complementary simulation studies demonstrated the effective performance of the models and the estimation method.

The merits of dynamic factor models have been discussed in previous articles (e.g., Molenaar, 1985; Nesselroade et al., 2001) and will not be repeated here. Because the same data have been analyzed in Nesselroade et al., assuming that the data are continuous, we focus on the comparisons of the results from our current analysis and Nesselroade et al. First, the estimation method used in Nesselroade et al. is the asymptotic MLE, which violates the independence of observations assumption although it can obtain consistent parameter estimates (Zhang, Hamaker, & Nesselroade, 2008). Our Bayesian estimation method works directly on the raw data and actually considers the dependence of the data. Second, our estimate of the factor inter-correlation is consistently larger than that from Nesselroade et al., which is shown to be attributable to the difference between the categorical model and continuous model in the simulation studies. Thus, modeling the categorical data set using categorical DFMs presented in this study seems more appropriate than modeling them with continuous DFMs.

We constructed the categorical DFMs within an underlying variable framework. To identify the model, we put constraints on the thresholds. In this appli-

TABLE 6
Simulation Results for the CWNFS Models

	TRUE	*ME*	*SD*	*MSE*	*AS-CON*[a]
λ_{011}	1	1.07	0.146	0.134	2.30(.22)
λ_{021}	1	1.07	0.125	0.105	2.35(.18)
λ_{031}	1	1.06	0.107	0.103	2.33(.18)
λ_{042}	1	1.07	0.154	0.121	2.33(.19)
λ_{052}	1	1.06	0.122	0.113	2.33(.19)
λ_{062}	1	1.05	0.105	0.111	2.33(.20)
λ_{111}	.2	0.25	0.109	0.115	1.35(.29)
λ_{121}	.2	0.26	0.093	0.098	1.33(.28)
λ_{131}	.2	0.26	0.084	0.097	1.36(.29)
λ_{142}	.2	0.26	0.100	0.109	1.48(.37)
λ_{152}	.2	0.27	0.085	0.104	1.49(.38)
λ_{162}	.2	0.26	0.089	0.103	1.48(.37)
q_{11}	.5	0.54	0.108	0.119	.52(.08)
q_{22}	.1	0.12	0.046	0.049	.17(.05)
q_{33}	.1	0.12	0.046	0.049	.19(.05)
q_{44}	.3	0.33	0.088	0.087	.37(.08)
q_{55}	.2	0.22	0.069	0.069	.28(.07)
q_{66}	.2	0.22	0.075	0.068	.27(.08)
r_{12}	−.85	−0.85	0.044	0.040	−.65(.07)

	Model Selection[b]		
	DIC	*EBIC*	*EAIC*
M6	48	59	59
M7	52	41	41

Note. The number in parenthesis is the standard error.

[a]The simulated categorical data were analyzed as continuous data.

[b]The number means how many times the model (M6 or M7) was chosen based on the fit statistics.

TRUE: True parameter value; ME: Mean parameter estimate; SD: Standard deviation; MSE: Mean standard error; DIC: Deviance information criterion; EAIC: Extended Akaike's information criterion; EBIC: Extended Bayesian information criterion.

cation of the Bayesian estimation procedure, we first calculated the thresholds and then fixed all of them, although we only needed to fix some of them to identify the model (e.g., Shi & Lee, 1998). Both from previous studies (e.g., Olsson, 1979; Shi & Lee, 1998) and our experience, the estimated thresholds were very close to the calculated thresholds from Eq (10). However, fixing all the thresholds shortened computation time.

Prior elicitation perhaps plays a crucial role in Bayesian inference (e.g., Gelman, 2006; Gelman, Carlin, Stern, & Rubin, 2003), but it is yet under researched. Prior elicitation is usually dependent on the models. For our categorical DFMs, only the non-informative priors were used. Based on the data analysis and simulation study, those priors seemed appropriate and useful. However, the non-informative priors may result in slow mixing of the posteriors if more complex models are adopted and missing data are present. In this case, different priors can be compared to choose the proper priors (e.g., Gelman, 2006; Kass & Wasserman, 1996) or informative priors can be constructed based on available information (e.g., Ibrahim & Chen, 2000; Zhang, Hamagami, Wang, Grimm, & Nesselroade, 2007).

As pointed out earlier, continuous dynamic factor models have been estimated as an SEM model through both MLE and least squares estimation methods. Similarly, categorical DFMs can be viewed as SEM models with categorical variables. To estimate the models, we can first calculate the lagged polychoric correlation matrix and then estimate the parameters using SEM software. However, the limitations in estimating continuous DFMs (e.g., Browne & Nesselroade, 2005; Zhang, Hamaker, et al., 2008) still exist and may even become worse with categorical models. The proposed Bayesian estimation method, instead, does not have these limitations.

The dynamic DFMs can also be constructed in the framework of latent trait models (e.g., Moustaki, 2000; Moustaki & Knott, 2000). Dunson (2003) proposed a dynamic latent trait model that can analyze mixture data of continuous, ordinal, and count variables. However, the dynamic latent trait model in Dunson is not a strict dynamic model because the data used are actually longitudinal data with only several waves of observation instead of time series data (at least more than 50 occasions of observations). Thus, Dunson's model requires the number of participants to be large enough to perform the analysis. Our model, however, requires one participant with adequately large repeated observations.

Finally, as was pointed out by Nesselroade et al. (2001) regarding the WNFS and DAFS models, neither the CDAFS model nor the CWNFS model specification can be viewed as a "winner" in the modeling process. The choice of the application of the two models should be mainly based on substantive considerations. From our empirical experience, the modeling of the CWNFS model is quite a bit easier in terms of computation. Furthermore, the lag relation between the latent factors and the observed variable is totally explicit for the CWNFS model. However, the CDAFS model can reveal more information about the dynamic processes of the latent factors by modeling the direct autoregressive and cross-regressive structure of factors. This is why Browne and Nesselroade (2005) labeled it "the process model."

Some unsolved problems remain but they are beyond the scope of this article. First, although the estimation method is good enough for us to make valid inference given that the length of the time series is only 100 and the bias of the estimates is relatively small, the inconsistencies of SEs for the CDAFS model and the bias for the CWNFS model are troubling and should be further investigated. Second, in comparing models, the claim for all three fit indexes is that smaller values mean better fit. However, there is no criterion to decide how large an improvement in a given fit index makes the model selection determinant. For the CWNFS model, especially, we found that none of the three criteria worked very well. Thus, the alternative model fit statistics need to be considered. Last but not least, although measurement error (usually refers to misspecification) has seemed to be a neglected research area in the study of categorical data (Gustafson, 2004; Liu & Agresti, 2005), it is a topic that deserves serious attention if categorical DFMs are to gain more favor with researchers.

ACKNOWLEDGMENTS

We are grateful to John J. McArdle and Peter C. M. Molenaar for their helpful suggestions and comments.

REFERENCES

Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association, 88,* 669–679.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Anderson, T. W. (1963). The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika, 28,* 1–25.

Baltes, P. B., & Nesselroade, J. R. (1973). The developmental analysis of individual differences on multiple measures. In J. R. Nesselroade & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological issues* (pp. 219–251). New York: Academic Press.

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods.* Monterey, CA: Brooks/Cole.

Boker, S. M. (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. *Multivariate Behavioral Research, 37*(3), 405–422.

Brillinger, D. R. (1975). *Time series: Data analysis and theory.* New York: Holt, Rinehart, & Winston.

Brillinger, D. R. (1981). *Time series: Data analysis and theory* (expanded edition). San Francisco: Holden-Day, Inc.

Browne, M. W., & Nesselroade, J. R. (2005). Representing psychological processes with dynamic factor models: Some promising uses and extensions of ARMA time series models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Advances in psychometrics: A Festschrift for Roderick P. McDonald* (pp. 415–452). Mahwah, NJ: Lawrence Erlbaum Associates.

Browne, M. W., & Zhang, G. (2005). DyFA: Dynamic factor analysis of lagged correlation matrices, Version 2.00 [Computer software and manual]. Retrieved from http://quantrm2.psy.ohio-state.edu/browne/

Browne, M. W., & Zhang, G. (2007). Developments in the factor analysis of individual time series. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical developments and future directions* (pp. 249–264). Mahwah, NJ: Lawrence Erlbaum Associates.

Cattell, R. B. (1963). The structuring of change by P- and incremental-R technique. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 167–198). Madison: University of Wisconsin Press.

Cattell, R. B., Cattell, A. K. S., & Rhymer, R. M. (1947). P-technique demonstrated in determining psychophysical source traits in a normal individual. *Psychometrika, 12*(4), 267–288.

Chib, S., & Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika, 85*(2), 347–361.

Congdon, P. (2003). *Applied Bayesian modeling.* New York: Wiley.

Cowels, M. K., & Carlin, B. P. (1996). Markov Chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association, 91,* 883–904.

Dunson, D. (2003). Dynamic latent trait models for multidimensional longitudinal data. *Journal of the American Statistical Association, 98*(463), 555–563.

Durbin, L., & Koopman, S. J. (2001). *Time series analysis by state space methods.* Oxford, UK: Oxford University Press.

Engle, R., & Watson, M. (1981). A one-factor multivariate time series model of metropolitan wage rates. *Journal of American Statistical Association, 76,* 774–781.

Fahrmeir, L., & Tutz, G. (1994). *Multivariate statistical modelling based on generalized linear models.* New York: Springer-Verlag.

Ferrer, E., & Nesselroade, J. R. (2003). Modeling affective processes in dyadic relations via dynamic factor analysis. *Emotion, 3,* 344–360.

Gelfand, A., & Smith, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85,* 398–409.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis, 1*(3), 515–534.

Gelman, A., Carlin, J. B., Stern, J. S., & Rubin, D. B. (2003). *Bayesian data analysis* (2nd ed.). New York: Chapman & Hall/CRC.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6,* 721–741.

Geweke, J. F. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. O. Berger, J. M. Bernardo, A. P. David, & A. F. M. Smith (Eds.), *Bayesian statistics 4* (pp. 169–194). Oxford, UK: Oxford University Press.

Geweke, J. F., & Singleton, K. J. (1981). Maximum likelihood "confirmatory" factor analysis of economic time series. *International Economic Review, 22,* 37–54.

Gustafson, P. (2004). *Measurement error and misclassification in statistics and epidemiology: Impacts and Bayesian adjustments.* New York: Chapman & Hall/CRC.

Hamilton, J. D. (1994). *Time series analysis.* Princeton, NJ: Princeton University Press.

Holtzman, W. C. (1963). Statistical models for the study of change in the single case. In C. W. Harris (Ed.), *Problems in measuring change* (pp. 19–211). Madison: University of Wisconsin Press.

Ibrahim, J. G., & Chen, M-H. (2000). Power prior distributions for regression models. *Statistical Science, 15*(1), 46–60.

Jones, C. J., & Nesselroade, J. R. (1990). Multivariate, replicated, single-subject, repeated measures designs and P-technique factor analysis: A review of intraindividual change studies. *Experimental Aging Research, 16*(4), 171–183.

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analsyis. *Psychometrika, 34,* 183–202.

Jöreskog, K. G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix. *Psychometrika, 59,* 381–389.

Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research, 36*(3), 347–387.

Justiniano, A. (2004). *Estimation and model selection in dynamic factor analysis.* Unpublished doctoral dissertation, Princeton University.

Kass R. E., & Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association, 91*(45), 1343–1370.

Kim, C. J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications.* Cambridge, MA: MIT Press.

Kim, C. J., & Nelson, C. R. (2001). A Bayesian approach to testing for Markov-Switching in univariate and dynamic factor models. *International Economic Review, 42*(4), 989–1013.

Lebo, M. A., & Nesselroade, J. R. (1978). Intraindividual differences dimensions of mood change during pregnancy identified by five p-technique factor analyses. *Journal of Research in Personality, 12,* 205–224.

Lee, S.-Y., Poon, W.-Y., & Bentler, P. M. (1990). Full maximum likelihood analysis of structural equation models with polytomous variables. *Statistics and Probability Letters, 9,* 91–97.

Lee, S.-Y., Poon, W.-Y., & Bentler, P. M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology, 48,* 359–370.

Lee, S.-Y., & Song, X.-Y. (2003). Bayesian analysis of structural equation models with dichotomous variables. *Statistics in Medicine, 22,* 3073–3088.

Liu, I., & Agresti, A. (2005). The analysis of ordered categorical data: An overview and a survey of recent developments. *Sociedad de Estadstíca e Investigación Operativa Test, 14*(1), 1–73.

Luborsky, L., & Mintz, J. (1972). The contribution of P-technique to personality, psychotherapy, and psychosomatic research. In R. M. Dreger (Ed.), *Multivariate personality research: Contribution to the understanding of personality in honor of Raymond B. Cattell.* Baton Rouge, LA: Claitors Publishing Division.

McArdle, J. J. (1982). *Structural equation modeling of an individual system: Preliminary results from "a case study in episodic alcoholism."* Unpublished manuscript, Department of Psychology, University of Denver.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. *Psychometrika, 50,* 181–202.

Molenaar, P. C. M. (1994). Dynamic latent variable models in developmental psychology. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 155–180). Newbury Park, CA: Sage.

Molenaar, P. C. M., & Nesselroade, J. R. (1998). A comparison of pseudo-maximum likelihood and asymptotically distribution-free dynamic factor analysis parameter estimation in fitting covariance-structure models to Block-Toeplitz matrices representing single-subject multivariate time-series. *Multivariate Behavioral Research, 33*(3), 313–342.

Moustaki, I. (2000). A latent variable model for ordinal variables. *Applied Psychological Measurement, 24*(3), 211–223.

Moustaki, I., & Knott, M. (2000). Generalized latent trait models. *Psychometrika, 65*(3), 391–411.

Muthén, B. (1984). A general structual equation model with dichotomous, ordered categorical and continuous latent variables indicators. *Psychometrika, 49,* 115–132.

Nesselroade, J. R., & Ghisletta, P. (2000). Beyond static concepts in modeling behavior. In L. R. Bergman & R. B. Cairns (Eds.), *Developmental science and the holistic approach* (pp. 121–135). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Nesselroade, J. R., McArdle, J. J., Aggen, S. H., & Meyers, J. (2001). Dynamic factor analysis models for multivariate time series analysis. In D. S. Moskowitz & S. L. Hershberger (Eds.), *Modeling individual variability with repeated measures data: Advances & techniques* (pp. 235–265). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Nesselroade, J. R., & Molenaar, P. C. M. (2003). Applying dynamic factor analysis in behavioral and social science research. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 622–639). London: Sage Publications.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika, 44*(4), 443–460.

Plummer, M., Best, N., Cowles, K., & Vines, K. (2005). Output analysis and diagnostics for Markov Chain Monte Carlo simulations (Version 0.9-2). Link: http://cran.r-project.org/src/contrib/Descriptions/coda.html

Priestley, M. B., Rao, T. S., & Tong, H. (1973). Identification of the structure of multivariate stochastic systems. In P. R. Krishnaiah (Ed.), *Multivariate analysis-III: Proceedings of the Third International Symposium on Multivariate Analysis held at Wright State University, Dayton, Ohio, June 19–24, 1972* (pp. 351–368). New York and London: Academic Press.

R Development Core Team. (2005). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org

Raftery, A. E. (1993). Bayesian model selection in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 163–180). Newbury Park, CA: Sage.

Schmitz, B. (1990) Univariate and multivariate time-series models: The analysis of intraindividual variability and intraindividual relationships. In A. v. Eye (Ed.), *Statistical methods in longitudinal research* (pp. 351–386). New York: Academic Press.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464.

Shi, J.-Q., & Lee, S.-Y. (1998). Bayesian sampling-based approach for factor analysis models with continuous and polytomous data. *British Journal of Mathematical and Statistical Psychology, 51,* 233–252.

Shifren, K., Hooker, K., Wood, P., & Nesselroade, J. R. (1997). Structure and variation of mood in individuals with Parkinsons disease: A dynamic factor analysis. *Psychology and Aging, 12,* 328–339.

Song, X.-Y., & Lee, S.-Y. (2002). Bayesian estimation and model selection of multivariate linear model with polytomous variables. *Multivariate Behavioral Research, 37*(4), 453–477.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. v. d. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583–639.

Spiegelhalter, D. J., Thomas, A., Best, N., & Lunn, D. (2003). WinBUGS manual version 1.4. MRC Biostatistics Unit, Institute of Public Health, Robinson Way, Cambridge CB2 2SR, UK, http://www.mrc-bsu.cam.ac.uk/bugs

Tanner, M. A., & Wong, W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association, 82,* 528–550.

West, M. (2000). Bayesian dynamic factor models and portfolio allocation. *Business and Economic Statistics, 18*(3), 338–357.

Wood, P., & Brown, D. (1994). The study of intraindividual differences by means of dynamic factor models: Rationale, implementation, and interpretation. *Psychological Bulletin, 116*(1), 166–186.

Zhang, Z., Hamagami, F., Wang, L., Grimm, K. J., & Nesselroade, J. R. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development, 31*(4), 374–383.

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (2008). Comparisons of four methods for estimating a dynamic factor model. *Structural equation modeling, 15*(1).

APPENDIX

(A) WinBUGS Codes for the CDAFS Models

```
model{
     for (t in 1:T){
              for (p in 1:P){
                          y[t,p]~dcat(pp[t,p,1:C])
                   for (c in 1:C-1){
                          qqphi[t,p,c]<-(thresh[p,c]
                             -mu[t,p])/sqrt(q[p])
                          qq[t,p,c]<-phi(qqphi[t,p,c])
                          }
                pp[t,p,1]<-qq[t,p,1]
                for ( c in 2:C-1){ pp[t,p,c]<-qq[t,p,c]
                                      -qq[t,p,c-1]}
                       pp[t,p,C]<-1-qq[t,p,C-1]}
              for (p in 1:3){mu[t,p]<-lam[p]*f[t,1]}
               for (p in 4:6){ mu[t,p]<-lam[p]*f[t,2]}}
     for (t in 2:T){
              f[t,1:2]~dmnorm(muf[t,1:2],tauv[1:2,1:2])
               muf[t,1]<-b[1]*f[t-1,1]+b[2]*f[t-1,2]
               muf[t,2]<-b[3]*f[t-1,1]+b[4]*f[t-1,2] }
     f[1,1:2]~dmnorm(f0[1:2], tauv[1:2,1:2])
     f0[1]~dnorm(0, 1.0E-6); f0[2] ~dnorm(0, 1.0E-6)
     tauv[1:2,1:2]<-inverse(R[1:2,1:2])
     R[1,1]<-1; R[2,2]<-1; R[1,2]<-R[2,1]; R[2,1]<-r21
     r21~dunif(-1,1)
     for (p in 1:6){
             lam[p]~dnorm(0,.001)
             q[p]~ dgamma(0.001, .001) }
     for (i in 1:4){ b[i]~dnorm(0,.001)}
     for (i in 1:4){Par[i]<-b[i]}
     Par[5]<-r21
     for (i in 1:6){
       Par[5+i]<-lam[i]
       Par[11+i]<-q[i]}
}
```

(The starting values and data are omitted from here for the sake of space.)

(B) WinBUGS Codes for the CWNFS Models

```
model{
     for (t in 1:T){
              for (p in 1:P){
                          y[t,p]~dcat(pp[t,p,1:C])
                   for (c in 1:C-1){
                           qqphi[t,p,c]<-(thresh[p,c]
                              -mu[t,p])/sqrt(q[p])
                           qq[t,p,c]<-phi(qqphi[t,p,c]) }
                   pp[t,p,1]<-qq[t,p,1]
                   for ( c in 2:C-1){
                          pp[t,p,c]<-qq[t,p,c]-qq[t,p,c-1]}
                   pp[t,p,C]<-1-qq[t,p,C-1] }}
      for (t in 2:T){
              for (p in 1:3){
                      mu[t,p]<-lam0[p]*f[t,1]+lam1[p]
                                 *f[t-1,1] }
               for (p in 4:6){
                      mu[t,p]<-lam0[p]*f[t,2]+lam1[p]
                                 *f[t-1,2]}
              }
               for (p in 1:3){
                      mu[1,p]<-lam0[p]*f[1,1]+lam1[p]*f0[1]}
               for (p in 4:6){
                      mu[1,p]<-lam0[p]*f[1,2]+lam1[p]*f0[2]}
      for (t in 1:T){
              f[t,1:2]~dmnorm(muf[1:2],tauv[1:2,1:2])}
      muf[1]~dnorm(0,1.0E-6); muf[2] ~dnorm(0,1.0E-6)
      f0[1:2]~dmnorm(muf[1:2],tauv[1:2,1:2])
      tauv[1:2,1:2]<-inverse(R[1:2,1:2])
      R[1,1]<-1; R[2,2]<-1; R[1,2]<-R[2,1]; R[2,1]<-r21
      r21~dunif(-1,1)
      for (p in 1:6){
             lam0[p]~dnorm(0,.001)
             lam1[p]~dnorm(0,.001)
             q[p]~ dgamma(0.001, .001) }
      for (i in 1:P){
          Par[i]<-lam0[i]
          Par[6+i]<-lam1[i]
          Par[12+i]<-q[i] }
      Par[19]<-r21
}
```

(The starting values and data are omitted from here for the sake of space.)