Emotion
2018, Vol. 18, No. 5, 739–754

© 2017 American Psychological Association
1528-3542/18/$12.00 http://dx.doi.org/10.1037/emo0000384

A Multivariate Statistical Model for Emotion Dynamics

Tanja Krone and Casper J. Albers
University of Groningen

Peter Kuppens
KU Leuven-University of Leuven

Marieke E. Timmerman
University of Groningen

In emotion dynamic research, one distinguishes various elementary emotion dynamic features, which are studied using intensive longitudinal data. Typically, each emotion dynamic feature is quantified separately, which hampers the study of relationships between various features. Further, the length of the observed time series in emotion research is limited and often suffers from a high percentage of missing values. In this article, we propose a vector autoregressive Bayesian dynamic model that is useful for emotion dynamic research. The model encompasses 6 elementary properties of emotions and can be applied with relatively short time series, including missing data. The individual elementary properties covered are within-person variability, innovation variability, inertia, granularity, cross-lag regression, and average intensity. The model can be applied to both univariate and multivariate time series, allowing one to model the relationships between emotions. One may include external variables and non-Gaussian observed data. We illustrate the usefulness of the model on data involving 50 participants self-reporting on their experience of 3 emotions across the period of 1 week using experience sampling.

Keywords: Bayesian, vector autoregressive, experience sampling, longitudinal data

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Emotions are an important part of our daily lives. The importance of emotions for our health and well-being is recognized more and more (Grühn, Lumley, Diehl, & Labouvie-Vief, 2013; Lewis, Haviland-Jones, & Feldman Barret, 2008, Chapter 29; Tugade, Fredrickson, & Feldman Barrett, 2004). Unlike, perhaps, personality and values, emotions fluctuate across time, changing both within and between days under the influence of external events and internal regulation. As such, a central function of emotions is to alert us to important events and changes, and to motivate us to deal with them (Frijda, 2007; Kuppens & Verduyn, 2015; Larsen, 2000; Scherer, 2009). Understanding the dynamics of emotions is therefore important, not in the least because it provides a window on how emotions may become dysregulated, which is considered a central feature of several mental disorders (Houben, Van Den Noortgate, & Kuppens, 2015; Wichers, Wigman, & Myin-Germeys, 2015).

To study emotion dynamics, intensive longitudinal data are used, sampled sufficiently frequently to characterize the dynamics of interest (Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015). Technological advantages facilitate the collection of such data, both in an experimental setting in a lab and in daily life. In lab studies, for instance, video-mediated recall and physiological recording can provide information on the dynamics of emotional episodes. In daily life, the widespread availability of mobile devices, first palmtops and now smartphones, enables researchers to collect multiple measurements per day in so-called ecological momentary assessment (EMA; also known as *experience sampling*) studies (Bolger & Laurenceau, 2013; Bos, Schoevers, & aan het Rot, 2015; Larson & Csikszentmihalyi, 1983; Shiffman, Stone, & Hufford, 2008).

When intensive longitudinal data are gathered with a structure as complex as found in emotion data, the choice of a proper analysis is of paramount importance. This choice is far from straightforward given the diversity of techniques available (Hamaker et al., 2015). The analysis typically focuses on identifying particular elementary features of emotion dynamics, with the aim to reveal distinct information on affective functioning and regulation (Kuppens & Verduyn, 2015). For instance, one may be interested in the level of variability emotions display within an individual or in how different emotions covary across time. However, choosing a proper analysis is hampered by the fact that these elementary features can often be quantified in different ways. Further, the quantifications of these elementary features are typically considered separately. This implies that relationships between these features remain hidden.

To provide a good picture of emotion dynamics, we propose using a single model in which most parameters have a clear interpretation in terms of a number of key features that are considered central to emotion dynamics. To this end, we propose using Bayesian dynamic modeling (West & Harrison, 1997). The

This article was published Online First December 21, 2017.

Tanja Krone and Casper J. Albers, Heymans institute for Psychology, Department of Psychometrics and Statistics, University of Groningen; Peter Kuppens, Faculty of Psychology and Educational Sciences, KU Leuven-University of Leuven; Marieke E. Timmerman, Heymans institute for Psychology, Department of Psychometrics and Statistics, University of Groningen.

This work was funded by the Netherlands Organization for Scientific Research (NWO), Grant 406-11-018.

Correspondence concerning this article should be addressed to Tanja Krone, who is now at the Department RAPID, Utrechtseweg 48, 3704 HE Zeist, the Netherlands. E-mail: tanja.krone@gmail.com

Bayesian dynamic model (BDM) we propose offers a representation of multivariate time series and may be applied to multiple individuals. Furthermore, the BDM as proposed in this article can be conveniently interpreted in terms of six important emotion dynamic features. This offers insight into the dynamics of single emotions as well as the dynamics between multiple emotions within an individual. By applying the model to data from multiple individuals, one can achieve insight into interindividual differences in emotion dynamics. Using Bayesian estimation offers flexibility with regard to the distributions used in specifying the model.

Emotion Dynamic Features

The patterns and regularities of an individual's experienced emotions across time can be captured by various elementary properties. We denote these elementary properties as emotion dynamic features (EDFs). There is a vast range of EDFs discussed in the literature (Brose, de Roover, Ceulemans, & Kuppens, 2015; Carstensen, Pasupathi, Mayr, & Nesselroade, 2000; Grühn et al., 2013; Houben et al., 2015; Kuppens & Verduyn, 2015). The taxonomy as discussed by Kuppens and Verduyn (2015) organizes these EDFs into four categories: emotional variability, emotional inertia, emotional cross-lag, and emotional granularity. If we complement these four with the average emotional intensity, we provide a fairly complete picture of an individual's experience of emotions across time. Our aim is to propose a way to succinctly capture the EDFs from the five categories in a single model. In the following section, we discuss each category and how it is captured in our model.

Emotional Variability

Emotional variability reflects to what extent the intensity of an emotion as experienced by an individual varies across time. Emotional variability has been found to increase with increasing stress levels (Scott, Sliwinski, Mogle, & Almeida, 2014) and decrease with increasing age (Brose et al., 2015; Carstensen et al., 2000; Scott et al., 2014). High emotional variability has been linked with lower emotional well-being and higher prevalence and severity of mood disorders (Houben et al., 2015). To quantify emotional variability, the within-person variance or standard deviation is typically used (Carstensen et al., 2000; Grühn et al., 2013; Kuppens & Verduyn, 2015; Röcke, Li, & Smith, 2009; Scott et al., 2014).

The within-person variance can be seen as a global summary of the degree of emotional variability. This variability can be decomposed into various elements (Jahng, Wood, & Trull, 2008). In this context, it is useful to distinguish the predictable part from the random part. The predictable part can be interpreted in terms of the emotional inertia and cross-lag, as discussed in the next few paragraphs. The random part covers the instantaneous change, and consists of the innovation variance and the white noise variance.

The innovation variance and the white noise variance express the sizes of the instantaneous changes at each measurement point. The key difference between the two is that the innovation variance captures the part of the change that is carried through to the next measurement point, whereas the white noise variance captures the part of the change that is not carried through to the next measurement point. Thus, a person showing a high white noise variance and low innovation variance is characterized by large variability at successive measurement points. Vice versa, a low noise variance with high innovation variance also indicates a large variability across the whole time span observed, but at a much slower rate, yielding much fewer oscillations in scores at successive measurement points.

As the global summary measure of the within-person variability, we use the within-person variance. Though our model includes both the innovation variance and white noise variance, we only interpret the innovation variance as a measure of innovation variability. We leave aside the white noise variance in our interpretation, because change because of white noise is typically attributed to measurement error.

Emotional Inertia

Emotional inertia refers to the tendency of an emotion to carry over from one moment to the next, reflecting resistance to change (Cook et al., 1995; Kuppens, Allen, & Sheeber, 2010; Suls, Green, & Hillis, 1998). High inertia has been linked to impaired emotion regulation (Gross, 2015; Koval et al., 2015; Kuppens, Allen, et al., 2010; Suls et al., 1998), inflexibility in adapting emotions (Kashdan & Rottenberg, 2010), and rumination (Koval, Kuppens, Allen, & Sheeber, 2012). Emotional inertia is generally quantified as the autoregression between successive measurements of an emotion. Note that although the concept of inertia is linked to the autocorrelation, it is generally quantified as a (autoregressive) regression variable. Therefore, autoregression is the adequate name for this variable, although it has also been referred to as *autocorrelation* in the literature (e.g., Kuppens, Allen, et al., 2010; Kuppens & Verduyn, 2015).

Emotional Cross-Lag

Emotions can be regulated through feedback loops: The increase of one emotion may infer an increase or decrease in another emotion (Gross, 2015; Kuppens & Verduyn, 2015; Pe & Kuppens, 2012). Although few studies have yet been conducted on emotional cross-lag, it is an important part of emotion regulation (Gross, 2015; Kuppens & Verduyn, 2015). For example, it has been found to be increased in major depression patients in terms of higher levels of overall emotion network density (Pe et al., 2015). Emotional cross-lag is quantified via the cross-lag regression—the lagged regression between two emotions (Pe & Kuppens, 2012). Analogously to the term autoregression, it is sometimes called the *cross-lag correlation* (e.g., Kuppens & Verduyn, 2015), whereas the *cross-lag regression* is generally used to quantify the emotional cross-lag. When the cross-lag regression is positive, this is called *augmentation*: The experience of one emotion increases the strength of another emotion on a later time point. A negative cross-lag regression is called *blunting*: The experience of one emotion decreases the strength of another emotion on a later time point (Pe & Kuppens, 2012).

Emotional Granularity

Emotional granularity refers to the ability of differentiating between different emotions and identifying emotions with speci-

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

ficity and precision. This is also known as *emotional differentiation* and is often measured in terms of emotional covariation (Barrett & Gross, 2001; Barrett, Gross, Christensen, & Benvenuto, 2001; Feldman, 1995; Kuppens & Verduyn, 2015). Higher emotional granularity is linked to increased emotion regulation (Barrett et al., 2001) and different, more effective coping mechanisms (Tugade et al., 2004). Furthermore, higher emotional granularity is associated with lower levels of neuroticism (Carstensen et al., 2000), and lower incidence of social anxiety disorder (Kashdan & Farmer, 2014) and depression (Erbas, Ceulemans, Lee Pe, Koval, & Kuppens, 2014).

Emotional granularity has been quantified in different ways. For example, the differentiation index equals the number of components found by a principal component analysis (PCA) on the covariances between emotions of a single individual (Brose et al., 2015; Grühn et al., 2013). Related to this is the concept of the unshared variance: the percentage of variance unexplained by the first component of such a PCA (Grühn et al., 2013). In practice, the choice between the differentiation index and the unshared variance is a pragmatic one. For a large number of emotions, the differentiation index appears to be more informative, and for a small number of emotions, the unshared variance is more informative. For these measures, higher scores indicate a higher granularity.

Other quantifications can be calculated directly from the observed data: the covariance between two emotions within a person (Erbas et al., 2014; Grühn et al., 2013), the correlation between two emotions (Barrett et al., 2001), and the intraclass correlation between all emotions (Erbas et al., 2014; Tugade et al., 2004). A higher covariance, correlation, and intraclass correlation indicate a lower level of differentiation between emotions, and thus a lower granularity. The correlation has the advantage of being standardized, allowing for a direct comparison between pairs of emotions both within and between individuals. However, a low within-person variance reduces the size of the absolute correlation, which renders interpretation difficult. This issue is not encountered when using the covariance (Scott et al., 2014). As all named quantifications measure the covariation between emotions, we do not need to include them all. In our model, the granularity will be quantified via the correlation.

Average Emotional Intensity

The EDFs discussed thus far capture the dynamics of emotions over time. In addition, how strong an emotion is felt on average may also differ, both between emotions within an individual and between individuals, and can provide important information on people's emotional lives. The average emotional intensity for positive emotions is positively related to emotion regulation (Barrett et al., 2001), as well as with extraversion, agreeableness, and conscientiousness, but negatively related to neuroticism (Carstensen et al., 2000). The average emotional intensity of negative emotions is higher for individuals with social anxiety disorder (Kashdan & Farmer, 2014) as well as for individuals with depression, high scores on neuroticism, and low scores on self-esteem (Erbas et al., 2014). To assess the average emotional intensity across time, we take into account the average intensity (Barrett et al., 2001; Carstensen et al., 2000), quantified as the mean score over time (Erbas et al., 2014; Kashdan & Farmer, 2014).

This Article

Each of these features provides unique information on how emotions (co)vary, carry over from one moment to the next, or mutually influence each other, and together they provide insight into many crucial aspects of emotional functioning and flexibility. As such, we propose a BDM that captures within-person variability, innovation variability, inertia, cross-lag, granularity, and average intensity for multiple emotions and individuals in a single model. First, we introduce the model and its possibilities. Then, we present an empirical application of the model. We conclude with a discussion on the model, its advantages and disadvantages, and recommendations for future research.

Model

To combine the aforementioned concepts for multiple variables, we use a Bayesian interpretation of a state space model, called the *Bayesian dynamic model* (BDM; West & Harrison, 1997). We use the BDM to estimate a vector autoregressive model, which can be rewritten into a state space model (Durbin & Koopman, 2012; Harvey, 1990).

The BDM has two equations: the observation equation and the system equation. For univariate data, the inclusion of so-called white noise in the observation equation improves estimation of the autoregression (Schuurman, Houtveen, & Hamaker, 2015). Following this, we include white noise in our multivariate BDM as well.

Single Individual

For each individual, a separate model can be formulated. We model $y_{i,t,n}$, the score on emotion i ($i = 1, 2, \ldots, I$), at time point t($t = 1, 2, \ldots, T_n$), for individual n ($n = 1, 2, \ldots, N$). The first equation, the observation equation, links the observed score $Y_{i,t,n}$ to the latent variable $\boldsymbol{\theta}_{i,t,n}$. The observation equation for the score vector $\mathbf{y}_{t,n} = [y_{1,t,n}, y_{2,t,n}, \ldots, y_{I,t,n},]$ is as follows:

$$\mathbf{y}_{t,n} = \boldsymbol{\mu}_n + \boldsymbol{\theta}_{t,n} + \boldsymbol{\varepsilon}_{t,n}, \qquad \boldsymbol{\varepsilon}_{t,n} \sim N(0, \mathbf{H}_n), \tag{1}$$

where $\boldsymbol{\mu}_n$ ($I \times 1$) denotes the mean vector of the I emotions, $\boldsymbol{\theta}_{t,n}$ ($I \times 1$) denotes the latent variable vector, $\boldsymbol{\varepsilon}_{t,n}$ ($I \times 1$) denotes the white noise vector, and $\mathbf{H}_n$ denotes the covariance matrix of $\boldsymbol{\varepsilon}_{t,n}$. As $\boldsymbol{\varepsilon}_{i,t,n}$ is assumed to be independent across emotions, $\mathbf{H}_n$ is a $I \times I$ diagonal matrix with $\sigma^2_{\varepsilon_{i,n}}$ as diagonal elements.

The system equation models the autoregression and cross-lag regression and the innovation over time of the latent variable $\boldsymbol{\theta}_{t,n}$:

$$\boldsymbol{\theta}_{t,n} = \boldsymbol{\Phi}_n \times \boldsymbol{\theta}_{t-1,n} + \boldsymbol{\eta}_{t,n}, \qquad \boldsymbol{\eta}_{t,n} \sim N(0, \mathbf{Q}_n), \tag{2}$$

where $\boldsymbol{\Phi}_n$ ($I \times I$) is the autoregression and cross-lag regression matrix, $\boldsymbol{\eta}_{t,n}$ ($I \times 1$) is the innovation vector, and $\mathbf{Q}_n$ ($I \times I$) is the covariance matrix of the innovation. All error terms, $\boldsymbol{\eta}_{t,n}$ and $\boldsymbol{\varepsilon}_{t,n}$, are assumed to be mutually independent. The model includes one lag, meaning that the autoregression is in respect to the previous time point only, and this model is denoted as the VAR(1)-BDM. A graphical representation of the model is shown in Figure 1.

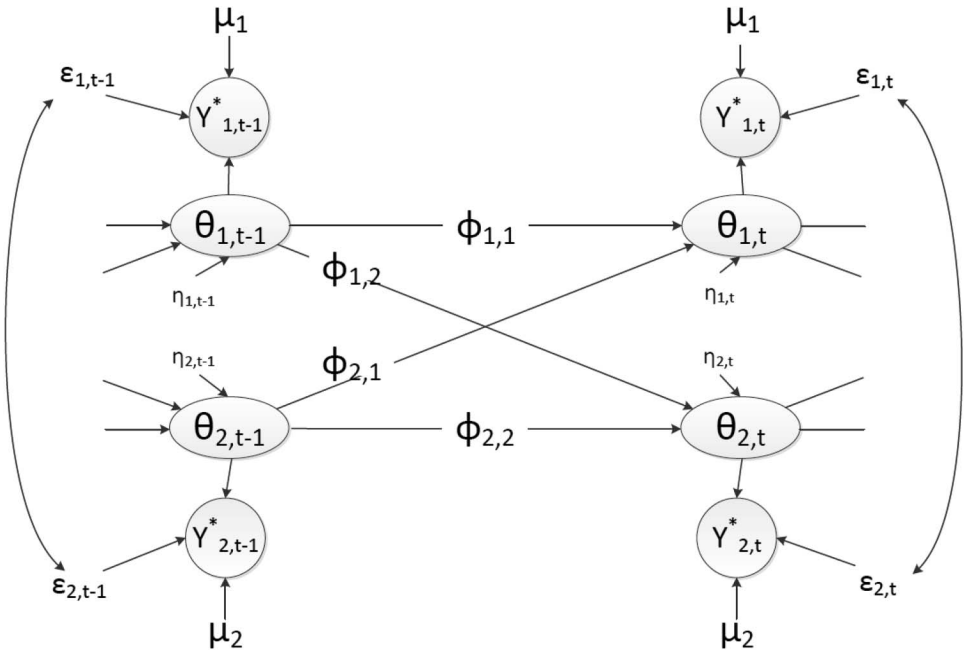

Figure 1. Schematic representation of the model as expressed in Equations 1 and 2 for a single individual and $i = 2$ emotions. As the focus lies on a single individual, all variables would have "n" in their subscript. For clarity, this is omitted. We use standard notation for referring to elements of vectors and matrices. For instance, $\theta_{1,\,t-1}$ represents the first element of the vector $\boldsymbol{\theta}_{t-1}$.

We remark that Σ_n ($I \times I$) is the model implied variance–covariance matrix of the observed scores for individual n, which is computed through

$$\mathrm{vec}\left(\sum_{\mathrm{n}}\right) = \left(\boldsymbol{I} - \boldsymbol{\Phi}_{\mathbf{n}}^{'} \otimes \boldsymbol{\Phi}_{\mathrm{n}}^{'}\right)^{-1} \mathrm{vec}(\mathbf{Q}_{\mathrm{n}} + \mathbf{H}_{\mathrm{n}}), \tag{3}$$

where $\mathrm{vec}(\boldsymbol{\Sigma}_{\mathrm{n}})$ is the vectorized version of $\boldsymbol{\Sigma}_{\mathrm{n}}$ and $\otimes$ denotes the Kronecker product. This is an adaptation of the Lyapunov equation used for the traditional vector autoregressive (VAR) model with only one error term, as discussed in Hamilton (1994, p. 265).

The VAR(1)-BDM includes dynamic parameters that are set to be constant across time. This implicit assumption seems to be most likely met when the dynamics of the individual under study do not change drastically (e.g., as a result of an intervention or life event) and the time interval between observations is roughly the same. A study using simulated data showed that the VAR(1)-model, and with that, the VAR(1)-BDM, is very robust against violations of the assumption of equal interval width (Albers, 2017). If time intervals differ widely (e.g., ranging from 1 hr up to 1 week), one may use an adapted model that explicitly allows for nonequidistant time points (see Kuppens, Oravecz, & Tuerlinckx, 2010; Oravecz, Tuerlinckx, & Vandekerckhove, 2011).

VAR(1)-BDM modeling can also be done in the presence of incidental missing measurements. To this, a link function is used that links the observed $y_{\mathrm{i,t,n}}$ to a latent $y^{*}_{\mathrm{i,t,n}}$. In the VAR(1)-BDM model, $y_{\mathrm{i,t,n}} = y^{*}_{\mathrm{i,t,n}}$, because $y_{\mathrm{i,t,n}}$ and $y^{*}_{\mathrm{i,t,n}}$ are assumed to be equally distributed. At time points with missing data, the latent variable $y^{*}_{\mathrm{i,t,n}}$ is not linked to the observed variable $y_{\mathrm{i,t,n}}$. The price to be paid for the missingness is that the uncertainty increases with more missing data points in a row, because the estimation cannot be checked against the observed data anymore. Further, the generalizability of the model to the "population of time points" is assured only if the missing data are missing completely at random or missing at random (see, e.g., Schafer & Graham, 2002).

The parameters of this statistical model are immediate translations of the EDFs, providing a direct link with emotion theory and application to empirical data. The linkage between each of the discussed EDFs and the model is summarized in Table 1. The within-person variability for individual n and emotion i is expressed via $\Sigma_{\mathrm{ii,n}}$ and the innovation variability via $\mathbf{Q}_{\mathrm{ii,n}}$. The autoregression for individual n and emotion i is $\boldsymbol{\Phi}_{\mathrm{ii,n}}$, and the cross-lag regression is $\boldsymbol{\Phi}_{\mathrm{ij,n}}$ for $i \neq j$. The correlation, used for the granularity, is obtained via $\boldsymbol{\Sigma}_{\mathrm{ij,n}}$ for $i \neq j$. The average intensity for individual n on emotion i is the mean $\mu_{\mathrm{i,n}}$. Hence, the model enables the simultaneous study of all discussed emotion dynamics.

Table 1
Quantification of Emotion Dynamics Features for Emotion I, in Relation to Emotion J Where Applicable, in the Notation of Equations 1, 2, and 3

Concept	Quantification	Parameter
Within-person variability	Variance of $Y_{\mathrm{i,n}}$	$\sum_{\mathrm{ii,n}}$
Innovation variability	Innovation of $Y_{\mathrm{i,n}}$	$\mathbf{Q}_{\mathrm{ii,n}}$
Inertia	Autoregression	$\Phi_{\mathrm{ii,n}}$
Emotional cross-lag	Cross-lag regression	$\Phi_{\mathrm{ij,n}}$
Granularity	Covariance of $y_{\mathrm{i,n}}$	$\sum_{\mathrm{ij,n}}$
	Correlation of $y_{\mathrm{ij,n}}$	$\mathrm{Cor}(\mathbf{y}_{\mathrm{ij,n}})$
Intensity	Mean estimated score	$\mu_{\mathrm{i,n}}$

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Multiple Individuals

The model can be applied without any difficulty for multivariate data collected among multiple individuals. This can be done by estimating the model of each individual separately. This implies that no assumptions are made with regard to the sampling of the individuals.

As an alternative, one may assume that the individuals are drawn at random from a certain population. As such, the parameters of the individuals can be assumed to be drawn randomly from the population distribution of the parameter concerned. These assumptions may be expressed in the model via a Level 2 model, for example, by assuming, as is standard in multilevel modeling, that each $\mathbf{\Phi}_{ij,n}$ is drawn from a normal distribution: $\mathbf{\Phi}_{ij,n} \sim N(\mathbf{\Phi}_{ij}, \sigma_{\Phi ij,n})$; Lodewyckx, Tuerlinckx, Kuppens, Allen, & Sheeber, 2011). The potential advantage is that the parameters can be estimated with more precision, provided that the distributional assumptions on the individual parameters are met. In practice, these assumptions may be too strict, yielding the first approach more attractive.

Both approaches have a heavy computational burden, with the second approach being even worse than the first, as modeling all individuals jointly takes more time per iteration and more iterations to reach convergence. This is why we decided upon the first approach in this article.

When the dynamics of a large number of individuals are studied, it typically will be of interest to examine relations between the different individual parameters. As illustrated in our empirical example, this can be done, for instance, via a cluster analysis on the individual parameters.

Possible Extensions of the VAR(1)-BDM Model

The VAR(1)-BDM model as defined in Equations 1 and 2 offers a rather flexible model for stationary individual time series with about normally distributed fluctuations that are constant dynamics across time. If the nature of the data requires a more flexible approach, the model can be extended in various ways.

Time-Varying Parameters

In the VAR(1)-BDM model, parameters are assumed to be equal over time, implying that emotion dynamics are assumed to be constant across the time span measured. In case this assumption would be too rigid, alternative models are available. For example, in regime switching or threshold models, the autoregression may change as the state changes (De Haan-Rietdijk, Gottman, Bergeman, & Hamaker, 2016; Hamaker & Grasman, 2012). Such an extension would be useful for modeling significant personal changes, for instance, as the consequence of clinical intervention. In time-varying autoregressive models, more gradual changes in parameters across time can occur (Bringmann et al., 2017), reflecting that the dynamics of psychological processes are not stationary in the long run.

Nonnormal Distributions

The VAR(1)-BDM model can be extended to nonnormal distributions in two ways. First, when the observations are assumed to be non-Gaussian realizations of an underlying Gaussian process, a link function can be used to transform the latent Gaussian scores into estimated observed non-Gaussian scores. Examples are a probit link for ordinal data (Chaubert, Mortier, & Saint André, 2008), or the log link for count data (assuming a Poisson distribution; Krone, Albers, & Timmerman, 2016a; Terui, Ban, & Maki, 2010). However, this adds more complexity to the model, requiring a larger sample size to estimate the model parameters with reasonable precision. Second, when the underlying process is assumed to be non-Gaussian, the distributions used in the model, for example, the white noise and innovation distributions, can be adjusted accordingly (Durbin & Koopman, 2012; West & Harrison, 1997, Chapters 13–14).

More Than a Single Lag

The VAR(1) model includes an autoregression on the previous time point only. The model can be extended with autoregressive effects from earlier time points as well, yielding a VAR(p) model, with p the number of previous time points regressed upon. To detect such effects, one would need very intensively sampled data. Therefore we deem this extension to be of possible use to model for instance psychophysiological measures, but not for emotion ratings, as this implies a very heavy burden on respondents.

External Variables

Research questions often probe how dynamic features of emotions are related to, or a function of, other variables, such as experimental manipulation or individual differences reflecting personality or well-being. Because of the flexible nature of the model, external variables can be dealt with in two ways. First, the external variable can be included in the model as an active covariate. This can be done as a direct effect, for example, letting $\boldsymbol{y}_{t,n}$ being dependent on $\boldsymbol{\mu}_n$, $\boldsymbol{\theta}_{t,n}$, and a covariate, and as a moderator effect, for example, letting elements of $\mathbf{\Phi}_n$ be dependent on the level of a covariate. Second, inactive covariates can be implemented post hoc, by examining the relation between any model parameter and an external variable after the model estimation. This can be done, for instance, by using partial correlations or linear regression, thereby accounting for confounding variables.

Modeling Empirical Time Series

Model Estimation

The VAR(1)-BDM is a Bayesian model that can be estimated using Bayesian Markov chain Monte Carlo (MCMC). To this end, we use Hamiltonian Monte Carlo (HMC), a generalization of the Metropolis-Hastings algorithm (Hastings, 1970; Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) that allows for an efficient estimation of the parameters (Gelman et al., 2013). This is incorporated in the software RStan (R Core Team, 2015; Stan Development Team, 2014). Example R code and Stan code for the VAR(1)-BDM can be found in Appendixes A and B, respectively.

In Bayesian modeling, prior distributions, quantifying the a priori degree of belief in parameter values, have to be specified. In empirical situations in which there is relevant context information (such as results of a pilot study), this information can be incorporated by specifying informative priors. Alternatively, one can aim

for weak-informative priors, to reduce the influence of the choice of priors on the estimates.

Estimating the model to empirical data may yield convergence problems. In general, these problems may be because of the identifiability of the model and/or a lack of data. The VAR(1)-BDM as expressed in Equations 1 and 2 is identified. Therefore, when estimation issues arise with this model, this is because of a lack of data. It is impossible to offer a general guideline on the minimal number of measurements required, because this would depend on the specific values of the population parameters and the required precision.

To check the convergence, two methods may be used: assessment of the potential scale reduction factor, $\widehat{R}$, and visual inspection of the trace plots. The $\widehat{R}$ shows the ratio of how much the estimation may change when the number of iterations is doubled, with an (ideal) value of 1 indicating that no change is expected (Gelman & Rubin, 1992; Stan Development Team, 2014). Trace plots show the MCMC estimates for each parameter at each iteration. If a parameter reaches convergence, the estimates over iterations are highly similar across chains. As a result, the trace plot will look like a fat caterpillar, in which all chains completely overlap, except at the fringe of the caterpillar.

Model Selection

In modeling empirical emotion dynamics, one might ask whether to use the VAR(1)-BDM, or a constrained version thereof, or even a more extended model. As in any statistical model specification (see Snijders & Bosker, 1999, p. 91), the steering wheels for model selection are substantive and statistical considerations.

With respect to the substantive considerations, the core advantage of the VAR(1)-BDM to characterize changes in emotions over time is that its parameters can be interpreted directly in terms of emotion dynamics features. Subject-matter-related considerations may indicate the need for an extended model, as described above. For example, when the outcome variables pertain to counts, the assumption of a normal distribution would be too far off, and a Poisson distribution would be a better choice.

With respect to the statistical considerations, it is useful to distinguish tests for individual parameters from measures of model fit. A test for an individual parameter can be used to assess the evidence for a particular value of a specific parameter. In this way, one could assess whether all parameters of the VAR(1)-BDM would be actually needed to describe the time series of a single subject, or whether one could do with a constrained version of the VAR(1)-BDM, for example, by fixing an innovation term for a specific variable at 0 or by setting the cross-lag between two specific emotions at 0.

Measures of model fit indicate the fit in a global way instead of focusing on a single parameter. As an absolute fit measure, we consider the predictive value of the model for the next measurement via the root mean squared error for emotion i of individual *n* ($RMSE_{i,n}$), as

$$RMSE_{i,n} = \sqrt{\sum_{t=2}^{T_n} (y_{i,t,n} - \hat{y}_{i,(t-1),n})^2}, \quad (4)$$

with $\hat{y}_{i,(t-1),n}$ the model predicted score at time point *t*, on the basis of the observed score at time point t–1. An RMSE value of 0 indicates perfect prediction, and larger values imply a lower predictive value.

In case substantive considerations would yield various competing models (e.g., VAR[1]-BDM vs. VAR[2]-BDM), information criteria such as the Bayesian information criterion (BIC; Schwarz, 1978) or the Watanabe-Akaike information criterion (WAIC; Watanabe, 2010) are of use. Such criteria add a penalty for model complexity to the model fit (as expressed via minus two times the log likelihood of the model). A smaller BIC/WAIC value points at better model; thus, the model with the lowest BIC/WAIC is favored. Which criterion to select is an ongoing debate between statisticians (Gelman, Hwang, & Vehtari, 2014), but in practice, the criteria usually point into the same direction.

Empirical Example

In this article, we reanalyze data described in Brans, Koval, Verduyn, Lim, and Kuppens (2013) and Erbas et al. (2014). As part of a larger study, the emotions of 50 individuals were self-reported using experience sampling. The data collection process consisted of three parts. In a first lab session, the participants signed an informed consent form and were handed out a palmtop, which they would use to record their emotions, along with instructions for its use. Second, during at least 7 days (extending to a maximum of 10 days for some participants), participants carried the palmtops and self-reported their emotions using the Experience Sampling Program (ESP; Barrett & Barrett, 2000). The waking hours of the individuals on each day were divided into 10 intervals. During each interval, at a random time point, the ESP would ask them to rate their emotions in terms of how angry, depressed, and stressed, for example, they felt at that moment. This yielded observations at roughly similar time intervals. Each emotion was rated on a 6-point Likert scale, ranging from 0 to 5, with higher values indicating a stronger feeling of that emotion. Finally, in a second lab session, the participants returned the palmtops and were each given AU$ 40 ($31) for their participation.

Sample Data

We selected the data on the three negative emotions "angry," "depressed," and "stressed" for reanalysis. The number of time points observed of the 50 individuals ranged from 20 to 90, with incidental missing data. To illustrate the sample data, the observed scores of three individuals are depicted in Figure 2. As shown, the first individual shows missing data all through the sample, the second shows large patches of complete and missing data, and the third shows little missing data.

Model Specification

To model the trivariate (i.e., of angry, depressed, stressed) time series of the 50 individuals, we applied the VAR(1)-BDM for each of the 50 individuals separately. This implies that there is no assumption with regard to the sampling of the individuals. In view of the limited length of the observed time series, we refrained from considering more complicated models (e.g., VAR[2]-BDM, or explicitly modeling the discrete nature of the dependent variable), as this would involve even more parameters to estimate yielding unstable results.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

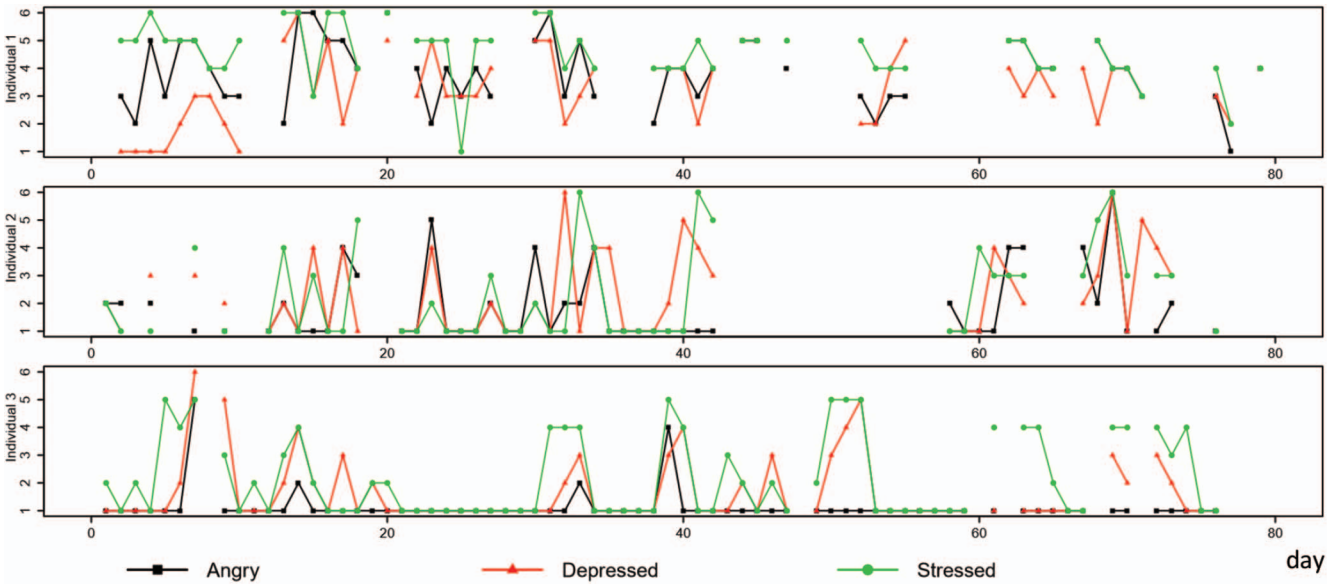

Figure 2. Observed time series for three of the individuals. See the online article for the color version of this figure.

The priors for the parameters were specified as follows. For the elements of $\mathbf{\Phi}_n$ we used a symmetrized reference prior (Berger & Yang, 1994); for the scale parameter of $\mathbf{H}_n$, a half-Cauchy(0, 2.5) prior; for the correlation matrix of $\mathbf{H}_n$, a Lewandowski-Kurowicka-Joe correlation prior (Lewandowski, Kurowicka, & Joe, 2009); for the diagonal elements of $\mathbf{Q}_n$, a $\Gamma(3,3)$; and for $\mu_{i,n}$, a normal prior with a mean of 0 and variance of 4.

Results

We start by presenting results concerning the estimation (i.e., convergence and computation time) and the absolute fit of the individual time series by as expressed via the RMSE. Then, we summarize the parameter estimates related to the emotion dynamic features across the 50 individuals.

Estimation: Convergence and Computation Time

In our analysis, we used four MCMC-chains of 30,000 iterations each. To check whether convergence was reached, we used $\widehat{R}$ and the trace plots. All elements of the model parameters ($\mathbf{\Phi}_n$, $\mathbf{Q}_n$, and $\boldsymbol{\mu}_n$) reached convergence for all individuals and emotions, with an $\widehat{R}$ below 1.02 for all estimated model parameters. One of the trace plots, representable for all relevant trace plots, is given in Figure 3, which, as can be seen, gives the expected fat caterpillar.

The total computation time (computer with 24 Intel Xeon 2.5 GHz cores) was 20.7 days, with a mean (*SD*) computation time per individual of 9.93 (4.70) hr. Note that parallel computing can be used in order to get the waiting time considerably smaller than the computing time.

The RMSE value had a mean (*SD*) across the 50 individuals of .20 (.21) for angry, .44 (.47) for depressed, and .50 (.45) for stressed. Given that each of the emotions are rated on a 6-point Likert scale, this indicates a reasonable absolute predictive fit of the models.

Summary of Emotion Dynamic Features

The VAR(1)-BDM of the emotions angry, depressed, and stressed of each individual involved 21 parameters of interest. For each of the 21 parameters of the 50 individuals, we computed the mean posterior parameter estimate, briefly denoted as parameter

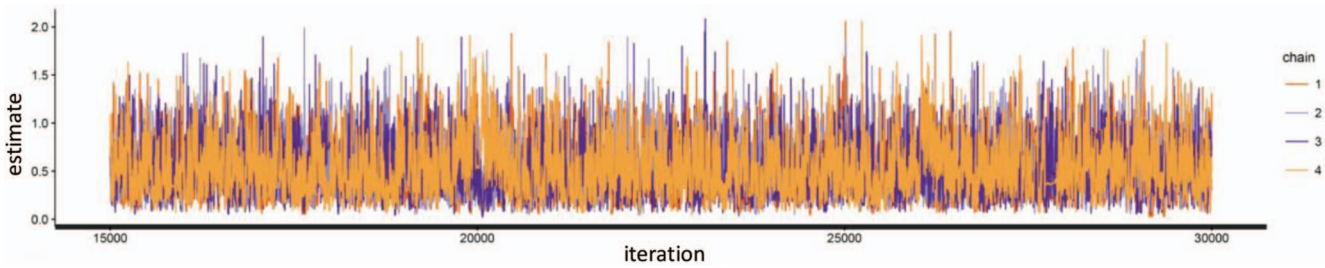

Figure 3. Trace plot of $Q_{11.1}$ as estimated in the empirical data, using six chains and 10.000 iterations (5.000 warm-up). See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

estimate in what follows. In Table 2 (Columns 3, 4, and 5), the means, standard deviations, and 95% confidence intervals of all parameter estimates across all individuals are presented. As can be seen, the standard deviations are rather large for all parameters, implying that there is a large variability in individual dynamics.

To achieve an insightful summary of the individual similarities and differences of emotion dynamic features, we performed a *K*-means cluster analysis (MacQueen, 1967) on the parameter estimates (in R, using 100 random starts), after scaling to ensure an equal weight of the variables to the cluster solution. As variables in the *K*-means analysis, we included all parameters except for the within-person variability ($\boldsymbol{\Sigma}_{ij,n}$) because of some outlying values (e.g., seven values above 6). These outliers are likely because of instable model implied estimates. Based on a scree plot of the total within-cluster sum of squares versus the number of clusters, we selected five clusters. The cluster size ranged from nine to 12 individuals; for each cluster, the mean parameter estimate and the within cluster standard deviation of each of the 21 parameters related to the emotion dynamic features are presented in Table 2 (Columns 5 to 9); we ordered Clusters 1 to 5 according to their average intensities (i.e., μ_{in}) for angry, depressed and stressed, with Clusters 1 and 2 showing low, Clusters 3 and 4 showing medium, and Cluster 5 showing high levels of intensity.

To depict the nature of the time series of the VAR(1)-BDM for each of the five clusters, we simulated time series, using the mean parameter estimates of that cluster a parameter value of the simulated model. The simulated time series for 70 time points for each of the variables and clusters are depicted in Figure 4. The appearance of the time series differs between the series, but the nature of the differences in dynamics is, of course, expressed in more detail via the differences in the model parameters. Close inspection of Figure 4 reveals that the intensity increases from Cluster 1 to 5 (by definition, as we ordered them as such), and more importantly, that the short-term and long-term dynamics characterizing the clusters differ.

We discuss the emotion dynamic features in the same order as in the introduction, that is, the parameters related to the within-person variability, innovation variability, inertia, cross-lag, granularity, and intensity.

Emotional variability is quantified using the EDFs within-person variability ($\boldsymbol{\Sigma}_{ii,n}$) and innovation variability ($\mathbf{Q}_{ii,n}$). As shown in Table 2, the mean within-person variability and innovation variability differs substantially across emotions and clusters. The within-person variability ($\boldsymbol{\Sigma}_{ii,n}$), indicating the overall variability across all measurements, of angry is smallest, whereas depressed and stressed show roughly equal variability across clusters. In contrast, the computed innovation variability, expressing the amount of change that carries over to the next measurement, is about similar for angry and depressed in all clusters, and in Cluster 3, substantially larger than in the other clusters.

To see to what extent within-person variability and innovation variability would be related across all individuals, we computed

Table 2
Mean Parameter Estimates per EDF, Across All Individuals (N = 50) and for Each of the Five Clusters, Resulting from the K-Means Analysis

Cluster:	50		1	2	3	4	5
Number of individuals:	9	12	9	10	10		
	M (*SD*)	95% CI	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Within-person variability $\Sigma_{ii,n}$							
A	1.34 (2.58)	[.61, 2.07]	.34 (.43)	2.32 (5.05)	1.36 (1.15)	.91 (.57)	1.47 (.92)
D	2.05 (3.25)	[1.12, 2.97]	1.02 (.86)	3.41 (6.34)	1.81 (.89)	2.06 (1.57)	1.53 (.71)
S	2.56 (4.33)	[1.33, 3.79]	1.48 (1.36)	4.49 (8.44)	2.61 (.86)	2.23 (2.14)	1.51 (.85)
Innovation variability $Q_{ii,n}$							
A	.37 (.27)	[.29, .45]	.21 (.26)	.21 (.13)	.51 (.28)	.36 (.23)	.58 (.24)
D	.37 (.23)	[.31, .44]	.27 (.18)	.19 (.12)	.53 (.23)	.38 (.17)	.52 (.23)
S	.42 (.22)	[.35, .48]	.42 (.21)	.25 (.13)	.70 (.19)	.34 (.13)	.43 (.19)
Inertia $\Phi_{ii,n}$							
A	.16 (.30)	[.07, .24]	.05 (.13)	.27 (.32)	.22 (.26)	−.08 (.26)	.29 (.32)
D	.07 (.29)	[−.02, .15]	.15 (.25)	.17 (.19)	−.33 (.24)	.12 (.26)	.17 (.23)
S	.25 (.37)	[.14, .36]	.15 (.22)	.27 (.37)	.51 (.31)	.50 (.21)	−.18 (.29)
Emotional cross-lag $\Phi_{ij,n}$							
A_{t-1} on D_t	.01 (.24)	[−.06, .08]	.02 (.11)	.06 (.24)	−.21 (.25)	.10 (.08)	.06 (.31)
A_{t-1} on S_t	.10 (.22)	[.04, .17]	.03 (.09)	.08 (.27)	.17 (.21)	.25 (.17)	−.01 (.23)
D_{t-1} on A_t	.04 (.36)	[−.06, .14]	−.05 (.38)	.27 (.29)	.26 (.26)	−.30 (.24)	.00 (.30)
D_{t-1} on S_t	.19 (.34)	[.09, .28]	.13 (.19)	.16 (.40)	.32 (.25)	.49 (.21)	−.15 (.23)
S_{t-1} on A_t	.12 (.32)	[.03, .21]	.11 (.30)	.29 (.27)	.33 (.25)	−.23 (.23)	.07 (.26)
S_{t-1} on D_t	.08 (.33)	[−.01, .18]	.18 (.26)	.12 (.28)	−.34 (.34)	.19 (.14)	.21 (.33)
Granularity Cor($y_{ij,n}$)							
A,D	.24 (.24)	[.17, .31]	.02 (.10)	.45 (.23)	.28 (.26)	.25 (.18)	.14 (.19)
A,S	.31 (.23)	[.25, .38]	−.03 (.07)	.47 (.19)	.35 (.14)	.38 (.13)	.34 (.21)
D,S	.52 (.24)	[.45, .59]	.37 (.28)	.70 (.20)	.46 (.10)	.58 (.16)	.44 (.27)
Intensity $\mu_{i,n}$							
A	1.47 (.65)	[1.29, 1.66]	1.12 (.13)	1.09 (.47)	1.45 (.60)	1.64 (.53)	2.09 (.82)
D	2.06 (.83)	[1.82, 2.30]	1.33 (.23)	1.51 (.62)	2.24 (.69)	2.27 (.71)	3.01 (.58)
S	2.09 (.85)	[1.85, 2.33]	1.45 (.26)	1.45 (.67)	2.29 (.41)	2.39 (.82)	2.95 (.76)

Note. EDF = emotion dynamic feature; CI = confidence interval; A = angry; D = depressed; S = stressed.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

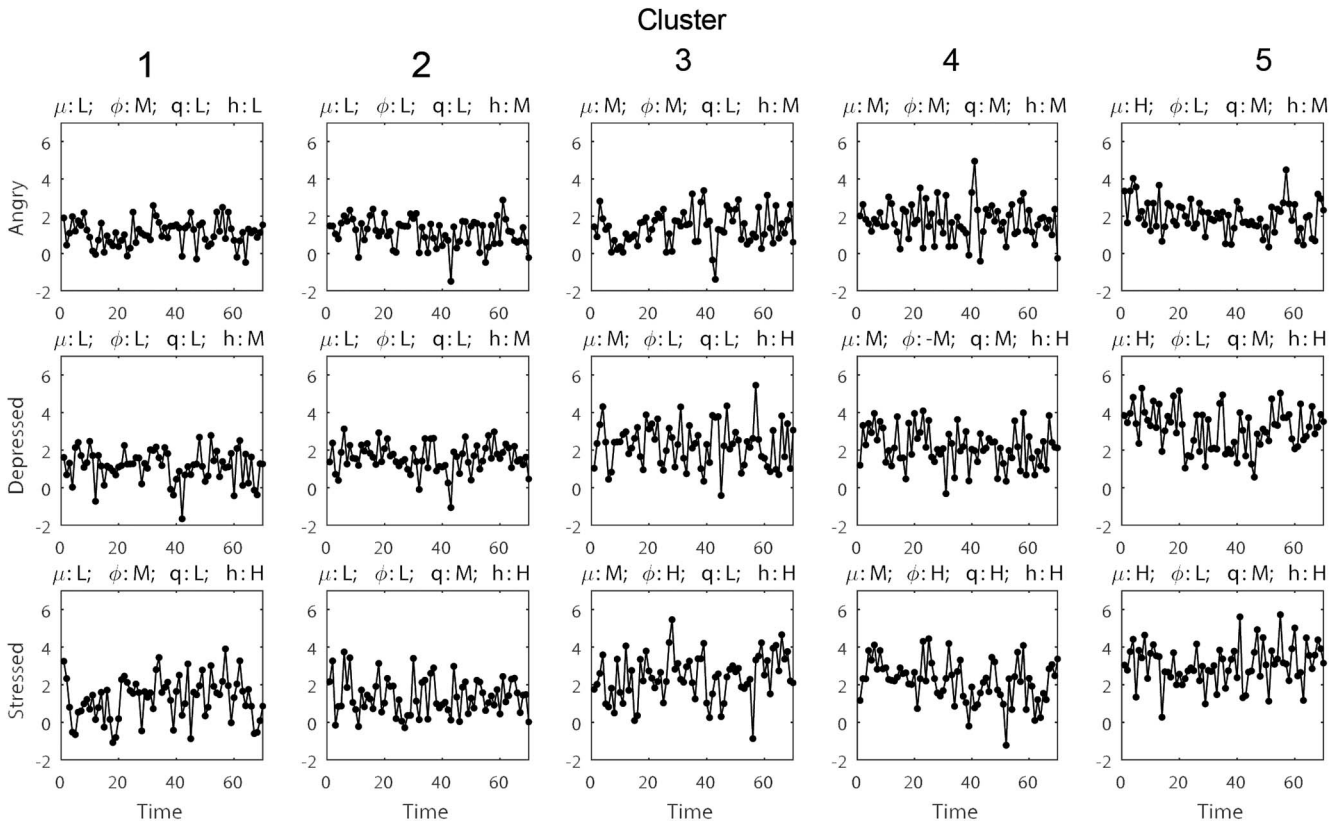

Figure 4. Typical time series for the emotions angry, depressed, and stressed (in rows), for each of the five clusters (in columns) of the K-means analysis, using the parameter means per cluster as parameters of the VAR(1)-BDM. Size of parameter (relative to other variables and clusters) is classified as low (L), medium (M), or high (H). μ = intensity; φ = autoregression; q = innovation variance; h = white noise variance.

Spearman's rho between the within-person variability ($\mathbf{\Sigma}_{ii,n}$) and innovation variability ($\mathbf{Q}_{ii,n}$), per emotion. The values were 0.80 for angry, 0.43 for depressed, and 0.40 for stressed. This suggests that the individual differences in dynamics for these three emotions are, to some extent, time-scale-invariant, that is, individuals with a high variability over the whole time period also show high variability between two time points. This has been found elsewhere as well (Kuppens, Oravecz, et al., 2010). This effect is much larger for angry than for depressed and stressed. This indicates that short-term anger fluctuations are more informative for its long-term fluctuations compared with those of depression and stress feelings, perhaps because the latter typically change more slowly (or last longer; Verduyn & Lavrijsen, 2015).

Emotional inertia is quantified using the autoregression. As shown in Table 2, the mean autoregression differs substantially across emotions and clusters. Generally, the mean values are positive, indicating that a relative high-scored emotion (relative in deviation from the overall mean of the time series for that individual) is followed by a relatively high-scored emotion on the next measurement. The highest mean values are found for stressed for Clusters 3 and 4 (about .50), suggesting a relative large stability of reported stress levels over time. Taken together, these results provide evidence that emotions are self-related over time. This is reflected in the literature. Indeed, most previous research has found that emotional states tend to be mildly or strongly predictive over time in daily life (e.g., Koval et al., 2012; Kuppens, Allen, et al., 2010; Suls et al., 1998). In exceptional cases (e.g., depressed in Cluster 3), the autoregressive parameter is negative, indicating individuals whose emotions seem to contrast themselves from one moment to the next. An interesting avenue for future research would consist of understanding what contributes to such dynamic patterns in these individuals.

Emotional Cross-Lag

The impact of one emotion on another is measured through the cross-lag regressions. As can be seen in Table 2, the mean cross-lag regressions differ across emotions and clusters, including their directions (i.e., positive and negative). This suggests large individual differences in degree of augmenting and blunting effects of emotions. The substantial individual differences that are found in the extent to which different emotions augment or blunt each other across time are consistent with previous research (Pe & Kuppens, 2012). The effects seem largest in Clusters 3 and 4, which have a medium intensity in emotion ratings.

A useful tool in interpreting the autoregression matrices—both the emotion inertia and the emotion cross-lag—is network visualization (cf. Bringmann et al., 2016, for an example from emotion

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

psychology). Figure 5 displays the mean autoregressive relations across all 50 individual models: Of the three emotions, stress clearly has the strongest inertia. The strongest relation between emotions is a positive relation between stress and depressed: Higher levels of stress precede higher levels of depression, on average. Lagged relations between angry and depressed are virtually absent.

Where Figure 5 displays the network for the average of the 50 participants, Figure 6 provides the network graph for each of the five clusters. This visualization demonstrates where the differences between the clusters are most prominent. Interestingly, even though the overall average network (see Figure 5) only has positive connections, four of five clusters display at least one negative relation, and thus instances in which one emotion blunts the subsequent experience of another emotion. The first cluster can be classified as one in which the "stressed → depressed" relation is reciprocated by a "depressed → stressed" relation. The second cluster mainly has stronger relations than the average individual. Cluster 3 can be typified as having negative outgoing relations from depression to all three emotions. Cluster 4, on the other hand, has negative outgoing relations from angry. Cluster 5 has strong autoregressions but relatively small cross-regressions. Note that the first two clusters are closer to the average cluster than the other three: The value of the distance metric $\sum_{i,j}|\mathbf{\Phi}_{c;i,j} - \overline{\mathbf{\Phi}}_{i,j}|$, with $\mathbf{\Phi}_c$ denoting the autoregression matrix for cluster c, and $\mathbf{\Phi}$ the autoregression matrix of the average, is 0.62 for Cluster 1, 0.78 for Cluster 2, 1.38 for Cluster 5, 1.87 for Cluster 4, and 1.99 for Cluster 3.

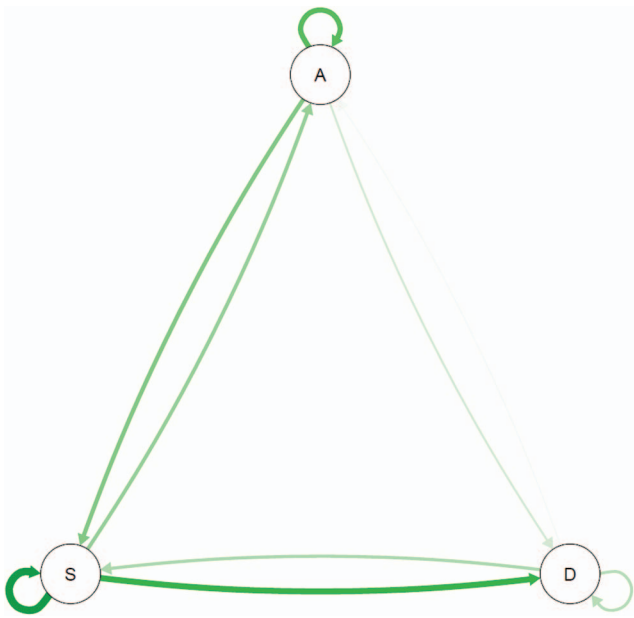

Figure 5. Network visualization of the auto- and cross-regressive effects for the average of the 50 individuals. The thickness of the arrows is proportional to the size of the effect. The thickness of the curves is proportional to the value. Visualization constructed using the R-package qgraph (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012). A = angry; D = depressed; S = stressed. See the online article for the color version of this figure.

Emotional Granularity

Granularity expresses the covariation between emotions and is quantified via the bivariate correlation between the different couples of emotions per individual. Here, higher values indicate a lower granularity. The correlations between the paired emotions are generally positive (see Table 2). Except for Cluster 1, the correlations are moderate (M = .25 for angry, depressed in Cluster 4) to strong (M = .70 for depressed, stressed in Cluster 2). This strongly resonates with previous research showing generally positive relations between like-valenced emotional states across time within individuals (e.g., Brose et al., 2015; Carstensen et al., 2000; Vansteelandt, Van Mechelen, & Nezlek, 2005). The variation in correlations across individuals indicate individual differences in the level of emotion differentiation or granularity, thought to be indicative of differences in emotion regulation and functioning (e.g., Barrett et al., 2001; Erbas et al., 2014).

Emotional intensity, quantified as $\mu_{i,n}$, appears to differ substantially across clusters (and thus individuals), but their order seems rather similar, with angry having the lowest intensity, and depressed and stressed having about equal intensity. The intensity (see Table 2) is lowest for Clusters 1 and 2, medium for Clusters 3 and 4, and highest for Cluster 5.

In earlier studies, a higher average intensity over all emotions seemed related to a lower average correlation over all emotion pairs (Carstensen et al., 2000; Erbas et al., 2014; Kashdan & Farmer, 2014). This finding was not replicated in our sample, with Spearman's rho values between the average correlation over all emotion pairs and the intensity of .02 for angry, .06 for depressed, and .05 for stressed.

Discussion

In this article, we proposed to use a BDM to analyze intensive longitudinal data to capture the patterns and regularities of an individual's experience of emotions across time. To accommodate for missing data, a link function was introduced, linking the observed to the latent variable only when the data are indeed observed. With this VAR-BDM, we analyzed a data set consisting of three emotions for 50 individuals. Using cluster analysis, we subsequently constructed five clusters of typical emotion dynamics. Our data set consisted of self-report data, but the VAR-BDM model is also applicable to behavioral and physiological measures of emotions over time as well as to data sets containing a mixture of such variables.

The study of emotion dynamics and the relation between EDFs is an important topic in psychological research. Most of the earlier studies regarding emotion dynamics have computed summary statistics for several EDFs (e.g., Barrett et al., 2001; Carstensen et al., 2000; Erbas et al., 2014; Scott et al., 2014). A complete model that encompasses all EDFs capturing the essential dynamics of multivariate data and that can be applied to EMA data with its inevitable limitations was lacking so far. Models that do capture multiple EDFs at once often require rather large sample sizes of at least 100 time points without missing values (e.g., De Haan-Rietdijk et al., 2016; Hamaker & Grasman, 2012). Another option, which does deal with the small sample size, is the use of a multilevel model. However, these models generally require a large number of individuals measured and are based on distributional assumptions on the individual parameters (e.g., Bringmann et al., 2013, 2016;

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

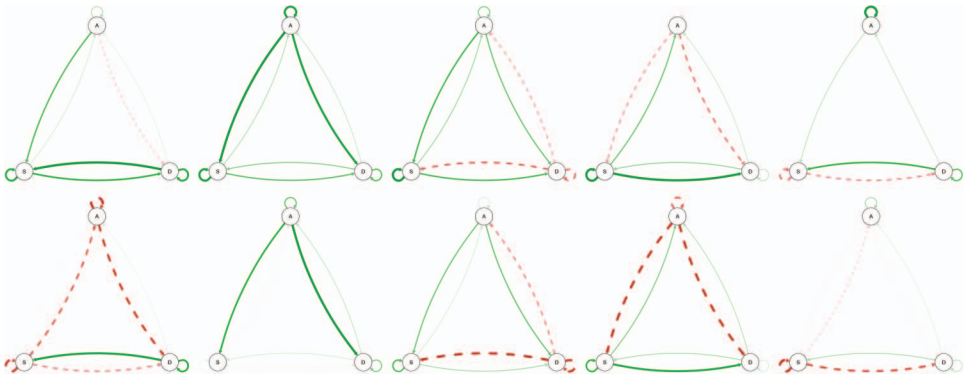

Figure 6. Network visualization (constructed using the qgraph package) for each of the five clusters. Top row: visualization of the $\mathbf{\Phi}$-matrix per cluster; bottom row: visualization of the deviation of $\mathbf{\Phi}_c$ (c = 1, . . . , 5) from the overall average of $\mathbf{\Phi}$. Green solid arrows indicate positive relations, and red dashed arrows indicate negative ones. Thickness of arrows is proportional to size of effects. A = angry; D = depressed; S = stressed. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Kuppens, Allen, et al., 2010; Lodewyckx et al., 2011). The multilevel VAR model of Schuurman, Ferrer, de Boer-Sonnenschein, and Hamaker (2016) lacked an important element, namely, white noise. Further, Schuurman et al. did not link the model elements to the EDFs as we did.

The results found in our empirical study largely concur with the results in previous literature. With our model, we strive to aid in the research on emotion dynamics, allowing for testing the proposed theories on empirical data. The theories on emotion dynamics are growing in number, but also in complexity (Gross, 2015; Kuppens & Verduyn, 2015). Because of the expansiveness of our model, it is possible to identify relations between emotions within individuals and between individuals, but also among EDFs, such as the relation between inertia and variability, or between EDFs and external variables, such as the relation between the average inertia and the level of mood disorder or other forms of psychopathology. Note that to quantify such relationships, a data set with a large number of individuals is needed. This allows for a more direct way of validating theories, stating relations between EDFs.

Further advantages of the model lie in its flexibility. The model, as shown in Equations 1 and 2, can be used with non-Gaussian data through the implementation of the link function or through adjustment of the distributions used. The link function also allows for data with missing values, without the need to adapt the model. Furthermore, the model can be used to implement external variables, both as active and inactive covariates. The final advantage lies in the wide range of data for which the model is applicable: one individual for whom one emotion is measured is enough for estimation. However, more individuals and emotions can be added.

Although the expansiveness of the proposed model is an advantage, it also introduces issues with estimation. Although the proposed VAR-BDM (Equations 1 and 2) can be estimated with relatively little data points, the estimated credible intervals, and thus the uncertainty, is large for data sets with short time series. The estimation of the cross-lag regression carries another issue. In empirical data, it may occur that the cross-lag regressions of the different variables are substantial. This introduces multicollinearity, which complicates estimation.

Further studies should focus on the estimation properties of the model. The conditions under which the model reaches convergence while estimating are unknown, as are the conditions needed to estimate the parameters with reasonable precision. Factors that may be of influence on the estimation and convergence of the model are the number of time points, the number of emotions, the values of $\mathbf{\Phi}_n$, and the number and pattern of missing data in the data set. These same factors may influence the precision with which the parameters may be estimated.

The cluster analysis clearly shows the multidimensional nature of emotion dynamics. By focusing on a single (set of) variable(s) only, important distinctions would be missed. For instance, Clusters 1 and 2 are both characterized by having low values for the emotional intensity yet differ on inertia, emotional cross-lag, and variability. Clusters 3, 4, and 5 have similar values for the white noise variance per emotion yet are different on the other characteristics. This demonstrates the need for the multivariate model with the ingredients as employed in this article.

A relevant future research question is how much measurements are needed to obtain accurate parameter estimates in the VAR(1)-BDM given certain settings. Research on related models (Krone, Albers, & Timmerman, 2016b, 2017; Schuurman et al., 2015) suggest that at least 50 to 100 measurements are required to estimate the parameters of an individual accurately. However, in the present study we mainly focus on clusters of nine to 12 individuals each. Parameter estimation for the clusters can draw power from all individuals, thus requiring fewer measurements per person for accurate estimates.

References

Albers, C. J. (2017). *Robustness of discrete time models against violation of equidistant time point assumption*. Manuscript in preparation.

Barrett, L. F., & Barrett, D. (2000). The Experience Sampling Program. Retrieved from http://www.experience-sampling.org

Barrett, L. F., & Gross, J. J. (2001). Emotional intelligence: A process model of emotion representation and regulation. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions and social behavior. Emotions: Current is-*

sues and future directions (pp. 286–310). New York, NY: Guilford Press.

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion, 15,* 713–724. http://dx.doi.org/10.1080/02699930143000239

Berger, J. O., & Yang, R.-Y. (1994). Noninformative priors and Bayesian testing for the AR(1) model. *Econometric Theory, 10,* 461–482. http://dx.doi.org/10.1017/S026646660000863X

Bolger, N., & Laurenceau, J.-P. (Eds.) (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Bos, F. M., Schoevers, R. A., & aan het Rot, M. (2015). Experience sampling and ecological momentary assessment studies in psychopharmacology: A systematic review. *European Neuropsychopharmacology, 25,* 1853–1864. http://dx.doi.org/10.1016/j.euroneuro.2015.08.008

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). The regulation of negative and positive affect in daily life. *Emotion, 13,* 926–939. http://dx.doi.org/10.1037/a0032400

Bringmann, L. F., Hamaker, E. L., Vigo, D. E., Aubert, A., Borsboom, D., & Tuerlinckx, F. (2017). Changing dynamics: Time-varying autoregressive models using generalized additive modeling. *Psychological Methods, 22,* 409–425. http://dx.doi.org/10.1037/met0000085

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., . . . Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment, 23,* 425–435. http://dx.doi.org/10.1177/1073191116645909

Bringmann, L. F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., . . . Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE, 8*(4), e60188. http://dx.doi.org/10.1371/journal.pone.0060188

Brose, A., de Roover, K., Ceulemans, E., & Kuppens, P. (2015). Older adults' affective experiences across 100 days are less variable and less complex than younger adults'. *Psychology and Aging, 30,* 194–208. http://dx.doi.org/10.1037/a0038690

Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. R. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology, 79,* 644–655. http://dx.doi.org/10.1037/0022-3514.79.4.644

Chaubert, F., Mortier, F., & Saint André, L. (2008). Multivariate dynamic model for ordinal outcomes. *Journal of Multivariate Analysis, 99,* 1717–1732. http://dx.doi.org/10.1016/j.jmva.2008.01.011

Cook, J., Tyson, R., White, J., Rushe, R., Gottman, J., & Murray, J. (1995). Mathematics of marital conflict: Qualitative dynamic mathematical modeling of marital interaction. *Journal of Family Psychology, 9,* 110–130. http://dx.doi.org/10.1037/0893-3200.9.2.110

De Haan-Rietdijk, S., Gottman, J. M., Bergeman, C. S., & Hamaker, E. L. (2016). Get over it! A multilevel threshold autoregressive model for state-dependent affect regulation. *Psychometrika, 81,* 217–241. http://dx.doi.org/10.1007/s11336-014-9417-x

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space models*. Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199641178.001.0001

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48,* 1–18. http://dx.doi.org/10.18637/jss.v048.i04

Erbas, Y., Ceulemans, E., Lee Pe, M., Koval, P., & Kuppens, P. (2014). Negative emotion differentiation: Its personality and well-being correlates and a comparison of different assessment methods. *Cognition and Emotion, 28,* 1196–1213. http://dx.doi.org/10.1080/02699931.2013.875890

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology, 69,* 153–166. http://dx.doi.org/10.1037/0022-3514.69.1.153

Frijda, N. H. (2007). *The laws of emotion*. Mahwah, NJ: Erlbaum.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman and Hall/CRC.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing, 24,* 997–1016. http://dx.doi.org/10.1007/s11222-013-9416-2

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science, 7,* 457–472. http://dx.doi.org/10.1214/ss/1177011136

Gross, J. J. (2015). The extended process model of emotion regulation: Elaborations, applications, and future directions. *Psychological Inquiry, 26,* 130–137. http://dx.doi.org/10.1080/1047840X.2015.989751

Grühn, D., Lumley, M. A., Diehl, M., & Labouvie-Vief, G. (2013). Time-based indicators of emotional complexity: Interrelations and correlates. *Emotion, 13,* 226–237. http://dx.doi.org/10.1037/a0030363

Hamaker, E. L., Ceulemans, E., Grasman, R. P. P. P., & Tuerlinckx, F. (2015). Modeling affect dynamics: State of the art and future challenges. *Emotion Review, 7,* 316–322. http://dx.doi.org/10.1177/1754073915590619

Hamaker, E. L., & Grasman, R. P. P. (2012). Regime switching state-space models applied to psychological processes: Handling missing data and making inferences. *Psychometrika, 77,* 400–422. http://dx.doi.org/10.1007/s11336-012-9254-8

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

Harvey, A. C. (1990). *Forecasting, structural time series models, and the Kalman filter*. Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781107049994

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57,* 97–109. http://dx.doi.org/10.1093/biomet/57.1.97

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin, 141,* 901–930. http://dx.doi.org/10.1037/a0038822

Jahng, S., Wood, P. K., & Trull, T. J. (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. *Psychological Methods, 13,* 354–375. http://dx.doi.org/10.1037/a0014173

Kashdan, T. B., & Farmer, A. S. (2014). Differentiating emotions across contexts: Comparing adults with and without social anxiety disorder using random, social interaction, and daily experience sampling. *Emotion, 14,* 629–638. http://dx.doi.org/10.1037/a0035796

Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review, 30,* 865–878. http://dx.doi.org/10.1016/j.cpr.2010.03.001

Koval, P., Brose, A., Pe, M. L., Houben, M., Erbas, Y., Champagne, D., & Kuppens, P. (2015). Emotional inertia and external events: The roles of exposure, reactivity, and recovery. *Emotion, 15,* 625–636. http://dx.doi.org/10.1037/emo0000059

Koval, P., Kuppens, P., Allen, N. B., & Sheeber, L. (2012). Getting stuck in depression: The roles of rumination and emotional inertia. *Cognition and Emotion, 26,* 1412–1427. http://dx.doi.org/10.1080/02699931.2012.667392

Krone, T., Albers, C. J., & Timmerman, M. E. (2016a). Bayesian dynamic modelling to assess differential treatment effects on panic attack frequencies. *Statistical Modelling, 16,* 343–359. http://dx.doi.org/10.1177/1471082X16650777

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Krone, T., Albers, C. J., & Timmerman, M. E. (2016b). Comparison of estimation procedures for multilevel AR(1) models. *Frontiers in Psychology, 7,* 486.

Krone, T., Albers, C. J., & Timmerman, M. E. (2017). A comparative simulation study of AR(1) estimators in short time series. *Quality & Quantity, 51,* 1–21. http://dx.doi.org/10.1007/s11135-015-0290-1

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science, 21,* 984–991. http://dx.doi.org/10.1177/0956797610372634

Kuppens, P., Oravecz, Z., & Tuerlinckx, F. (2010). Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology, 99,* 1042–1060. http://dx.doi.org/10.1037/a0020962

Kuppens, P., & Verduyn, P. (2015). Looking at emotion regulation through the window of emotion dynamics. *Psychological Inquiry, 26,* 72–79. http://dx.doi.org/10.1080/1047840X.2015.960505

Larsen, R. J. (2000). Toward a science of mood regulation. *Psychological Inquiry, 11,* 129–141. http://dx.doi.org/10.1207/S15327965PLI1103_01

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. In H. T. Reis (Eds.), *New directions for methodology of social and behavioral sciences* (Vol. 15, pp. 41–56). San Francisco, CA: Jossey-Bass.

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis, 100,* 1989–2001. http://dx.doi.org/10.1016/j.jmva.2009.04.008

Lewis, M., Haviland-Jones, J. M., & Feldman Barret, L. (Eds.). (2008). *Handbook of emotions* (3rd ed.). New York, NY: Guilford Press.

Lodewyckx, T., Tuerlinckx, F., Kuppens, P., Allen, N., & Sheeber, L. (2011). A hierarchical state space approach to affective dynamics. *Journal of Mathematical Psychology, 55,* 68–83. http://dx.doi.org/10.1016/j.jmp.2010.08.004

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). Berkeley, CA: University of California Press.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics, 21,* 1087–1092. http://dx.doi.org/10.1063/1.1699114

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2011). A hierarchical latent stochastic differential equation model for affective dynamics. *Psychological Methods, 16,* 468–490. http://dx.doi.org/10.1037/a0024375

Pe, M. L., Kircanski, K., Thompson, R. J., Bringmann, L. F., Tuerlinckx, F., Mestdagh, M., . . . Gotlib, I. H. (2015). Emotion-network density in major depressive disorder. *Clinical Psychological Science, 3,* 292–300. http://dx.doi.org/10.1177/2167702614540645

Pe, M. L., & Kuppens, P. (2012). The dynamic interplay between emotions in daily life: Augmentation, blunting, and the role of appraisal overlap. *Emotion, 12,* 1320–1328. http://dx.doi.org/10.1037/a0028262

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Röcke, C., Li, S.-C., & Smith, J. (2009). Intraindividual variability in positive and negative affect over 45 days: Do older adults fluctuate less than young adults? *Psychology and Aging, 24,* 863–878. http://dx.doi.org/10.1037/a0016276

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177. http://dx.doi.org/10.1037/1082-989X.7.2.147

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion, 23,* 1307–1351. http://dx.doi.org/10.1080/02699930902928969

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods, 21,* 206–221. http://dx.doi.org/10.1037/met0000062

Schuurman, N. K., Houtveen, J. H., & Hamaker, E. L. (2015). Incorporating measurement error in n = 1 psychological autoregressive modeling. *Frontiers in Psychology, 6,* 1038. http://dx.doi.org/10.3389/fpsyg.2015.01038

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Scott, S. B., Sliwinski, M. J., Mogle, J. A., & Almeida, D. M. (2014). Age, stress, and emotional complexity: Results from two studies of daily experiences. *Psychology and Aging, 29,* 577–587. http://dx.doi.org/10.1037/a0037282

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4,* 1–32. http://dx.doi.org/10.1146/annurev.clinpsy.3.022806.091415

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and applied multilevel analysis*. London, UK: SAGE.

Stan Development Team. (2014). Rstan: The R interface to Stan (Version 2.4). Retrieved from http://mc-stan.org/rstan.html

Suls, J., Green, P., & Hillis, S. (1998). Emotional reactivity to everyday problems, affective inertia and neuroticism. *Personality and Social Psychology Bulletin, 24,* 127–136. http://dx.doi.org/10.1177/0146167298242002

Terui, N., Ban, M., & Maki, T. (2010). Finding market structure by sales count dynamics-multivariate structural time series models with hierarchical structure for count data. *Annals of the Institute of Statistical Mathematics, 62,* 91–107. http://dx.doi.org/10.1007/s10463-009-0244-2

Tugade, M. M., Fredrickson, B. L., & Feldman Barrett, L. (2004). Psychological resilience and positive emotional granularity: Examining the benefits of positive emotions on coping and health. *Journal of Personality, 72,* 1161–1190. http://dx.doi.org/10.1111/j.1467-6494.2004.00294.x

Vansteelandt, K., Van Mechelen, I., & Nezlek, J. B. (2005). The co-occurrence of emotions in daily life: A multilevel approach. *Journal of Research in Personality, 39,* 325–335. http://dx.doi.org/10.1016/j.jrp.2004.05.006

Verduyn, P., & Lavrijsen, S. (2015). Which emotions last longest and why: The role of event importance and rumination. *Motivation and Emotion, 39,* 119–127. http://dx.doi.org/10.1007/s11031-014-9445-y

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross-validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research, 11,* 3571–3594.

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models*. New York, NY: Springer-Verlag.

Wichers, M., Wigman, J. T. W., & Myin-Germeys, I. (2015). Micro-level affect dynamics in psychopathology viewed from complex dynamical system theory. *Emotion Review, 7,* 362–367. http://dx.doi.org/10.1177/1754073915590623

(*Appendices follow*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix A

R Code

```
require (rstan)
##take note: Need to install Rtools first (see http://mc-stan.org/)
rstan_options (auto_write=TRUE)
options (mc.cores = 5)

#create Data
data <- List ()
data$N <- 1 #subjects
data$T <- 20 #maximum Time Points
data$I <- 2 #number of Emotions
data$T_N <- array(data=20, dim=1) #time Points per Individual
#Missing Value Indicator
data$M <- array (data=c(1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0,
                        0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 0), dim=c(N, T, I))

#Observed Scores
data$Y <- array (data=c(1, 3, 2, 5, 3, 5, 5, 4, 3, 3, 1, 1, 2, 6, 6, 5, 5, 4, 1, 6, 1, 1, 1,
1, 1, 2, 3, 3, 2, 1, 1, 1, 5, 6, 3, 5, 2, 4, 1, 5),
     dim=c (N, T, I))

#call Model File From Working Directory
Mod <- stan_model (file="Stancode_full_model.stan")

#sample Model
fit <- Sampling (
Mod
, Data = Data #The Data List
, Iter = 300 #This Number Should Be Large Enough to Reach Convergence (i.e., 20.000 or such)
, Chains = 3 #number of MCMC Chains
, Verbose = F
, Refresh = 300 #update on Every 300 Iterations
, Seed = 2016 #seed to Replicate
)
```

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Appendix B

Stan Code

```
data{
 int <lower=2> T;                                  //maximum Length of Series Across All Individuals
 int <lower=1> I;                                  //number of Emotions
 int <lower=1> N;                                  //number of Individuals
 matrix [T, I] Y [N];                              //observed Scores in N Matrices of T*I
 int M[N, T, I];                                   //m==1 if Missing, m==0 Is non-missing
 int T_N[N];                                       //length of Individual Time Series
}


parameters {
  matrix <lower=-1, upper=1> [I, I] phi [N];       //cross-lag Matrix, Values Restricted to (-1, 1)
  corr_matrix[I] Omega [N];                        //correlation Matrix Q
  vector<lower=0>[I] tau [N];                      //scale Matrix Q
  row_vector [I] mu [N];                           //mean Vector
  matrix [T, I] Z [N];                             //latent Scores y*
  matrix <upper=0> [I, I] X [N];                   //constant for Prior of phi
  vector<lower=0>[I] Lambda [N];                   //vector Matrix Q
}
transformed parameters{
  matrix [I, I] H[N];                              //covariance Matrix of White Noise
  matrix [I, I] Q[N];                              //covariance Matrix of Innovation


  for (n in 1:N) {
  //quad_form=diag (tau) *omega*diag (tau)', Diag (tau) Is a Diagonal Matrix of Vector tau
  H[n] <- quad_form_diag(omega[n], tau[n]);
  Q[n] <- diag_matrix (lambda [n]);
}
  }

model{
  for (n in 1:N) {
    for (t in 2:T_N[n]) {
    //system equation: Estimate Latent Score Theta
    Z[n, t] ~ multi_normal (Z[n, t-1] *phi [n], Q[n]);
       //link to Vector Y_t of Individual n if Y Is Observed for All emotions:
       if (sum(M[n, t])==0) {
       //observation equation: Estimate y (identity Link Function omitted)
       Y[n, t] ~ multi_normal (mu[n] + Z[n, t], H[n]);
       }
    }
  }
```

(*Appendices continue*)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

```
#prior
  for (n in 1:N) {
  mu[n] ~ Normal (0, 2);
  tau[n] ~ Cauchy (0, 2.5);
  omega[n] ~ lkj_corr (2);
  lambda [n] ~ Gamma (3, 3);
    for (i in 1:I) {
       for (j in 1:I) {
       //symmetrized Reference Prior for phi
       increment_log_prob (X[n, i, j] - log(1−(phi [n, i, j] *phi[n, i, j]))/2);
       }
    }
  }
}
```

Received July 22, 2016
Revision received September 6, 2017
Accepted September 11, 2017 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.