Psychological Methods
2011, Vol. 16, No. 4, 468–490

1082-989X/11/$12.00 DOI: 10.1037/a0024375

A Hierarchical Latent Stochastic Differential Equation Model for Affective Dynamics

Zita Oravecz, Francis Tuerlinckx, and Joachim Vandekerckhove
University of Leuven

In this article a continuous-time stochastic model (the Ornstein–Uhlenbeck process) is presented to model the perpetually altering states of the core affect, which is a 2-dimensional concept underlying all our affective experiences. The process model that we propose can account for the temporal changes in core affect on the latent level. The key parameters of the model are the average position (also called home base), the variances and covariances of the process, and the regulatory mechanisms that keep the process in the vicinity of the average position. To account for individual differences, the model is extended hierarchically. A particularly novel contribution is that in principle all parameters of the stochastic process (not only the mean but also its variance and the regulatory parameters) are allowed to differ between individuals. In this way, the aim is to understand the affective dynamics of single individuals and at the same time investigate how these individuals differ from one another. The final model is a continuous-time state-space model for repeated measurement data taken at possibly irregular time points. Both time-invariant and time-varying covariates can be included to investigate sources of individual differences. As an illustration, the model is applied to a diary study measuring core affect repeatedly for several individuals (thereby generating intensive longitudinal data).

Keywords: stochastic differential equation, Bayesian statistics, hierarchical model, affective dynamics

Supplemental materials: http://dx.doi.org/10.1037/a0024375.supp

Affective experience colors our lives. Feelings like joy, sadness, anger, and love are the affective tides of our being. A crucial feature of affect is that it changes and evolves over time. Hence, a genuine understanding of our affective system must be based on understanding the affective dynamics. Prominent researchers in the field consider this a challenging goal that deserves systematic study. For instance, Scherer (2000) believed that the study of the time course of emotional experience could bring about a paradigm shift in emotion psychology, and Boker (2002) claimed that understanding the dynamics of emotion is a challenging goal deserving a comprehensive investigation. Referring specifically to research into the patterns of change that are particular to human emotions, Davidson (2003) even coined a particular term: *affective chronometry*.

To make such an endeavor possible, the use of dynamic systems theory has been put forward by a number of researchers (e.g., Lewis, 2005; Scherer, 2000; Shoda, LeeTiernan, & Mischel, 2002; Witherington & Crichton, 2007). However, there are relatively few recent studies in which dynamical systems models have been applied to investigate affective dynamics, although there are some notable recent exceptions (e.g., Boker, 2002; Boker & Laurenceau, 2006; Chow, Ram, Boker, Fujita, & Clore, 2005; Hamaker, Zhang, & van der Maas, 2009; Hoeksma, Oosterlaan, Schipper, & Koot, 2007). Our contribution in the present article is the development of a specific substantively grounded dynamic systems model with individual differences for the crucial parameters. Additionally, we present the necessary methods to apply it to real data.

Specifically, we will focus on the dynamics of *core affect*, which has a central role in current emotion theory (Barrett, Mesquita, Ochsner, & Gross, 2007; Russell, 2003). Core affect lies at the heart of our affective experience and consists of hedonic (pleasure–displeasure) and arousal (deactivated–activated) dimensions. According to Russell (2003), core affect does not always actively come to the surface of our consciousness, but it is at all times consciously accessible, so that people can provide information about their current state. It is assumed that at least part of the momentary emotional experience relies on the current core affect state.

When we construct a model for temporal fluctuations in the core affect state over time, three important factors need to be taken into account in order to have a plausible, realistic model. First, because the complexity of the subject makes any description necessarily incomplete, our model should allow for a fair degree of randomness. Second, the model should contain an aspect of control and regulation. Third, although a single model can underlie the core affect dynamics, quantitative individual differences can show up in different places and need to be accounted for.

Central Aspects of Affective Changes: Randomness, Regulation and Individual Differences

Our emotional life in general, and core affect in particular, is constantly influenced by external and internal factors (Denissen, Butalid, Penke, & van Aken, 2008; Russell, 2003; Russell &

This article was published Online First August 8, 2011.

Zita Oravecz, Francis Tuerlinckx, and Joachim Vandekerckhove, Department of Psychology, University of Leuven, Leuven, Belgium.

Correspondence concerning this article should be addressed to Zita Oravecz, Department of Psychology, University of Leuven, Tiensestraat 102 Box 3713, B-3000 Leuven, Belgium. E-mail: zita.oravecz@psy.kuleuven.be

Barrett, 1999; Watson, Wiese, Vaidya, & Tellegen, 1999). Listing all these minor and major impacts would be a cumbersome, if not an impossible task, and measuring them would be an even greater challenge. External factors may include environmental effects like the weather, food intake, drugs, ionization in the air, physical activity, social company, and so on. Internal effects, on the other hand, involve important physiological and psychological processes (e.g., the level of certain critical hormones, perception of one's own emotion). Because most of these effects exhibit continuous change over time, the resultant core affect state will also be subject to incessant variation. This suggests that, at least in theory, subsequently visited states in the core affect space should form something like a random trajectory (i.e., a kind of random walk in two dimensions). Unfortunately, when we attempt to explain such a core affect trajectory of a person by linking it to external and internal effects, we may realize that the obtainable information is very limited: It would be overly optimistic to believe that we are able to find all events influencing the core affect at a given moment. As a result, we are left with describing the summed influence of many contributing unknown factors. The current core affect state can then be considered the culmination of numerous, simultaneously occurring small and large impacts. Consequently, we can assume a considerable degree of uncertainty in the core affect dynamics, and therefore we will introduce stochastic models to model the inherent randomness. These models assume that noise drives the changes in the true score of the variable. Note that in the model to be presented, measurement noise will also be added, so that we will be able to distinguish between two sources of random variation.

A second aspect that we need to consider is the degree of control present in the emotional life. At the substantive level, our model should be anchored in the principles of emotion regulation: the conscious or unconscious efforts people make to exert some influence on their emotions (Frijda, 2007; Gross, 2007; Gross & John, 2003; Hemenover, 2003). Many theories based on homeostatic principles (see, e.g., Carver & Scheier, 1990; Chow et al., 2005; Forgas & Ciarrochi, 2002; Hemenover, 2003; Larsen & Prizmic, 2004) suggest that each person has an "ideal point" in the two-dimensional core affect space to which he or she is drawn back to a varying extent. If an individual's current position is further from the ideal point, the "traction" or restoring force will become stronger. On a metaphorical level, this assumption might conjure up the picture of an elastic band connecting the ideal point with the current position: The force exerted by the elastic will be stronger for positions farther from the ideal point but will be very small close to the ideal point.

Finally, individual differences are present at different levels and locations in the emotion system (see, e.g., Kuppens, Stouten, & Mesquita, 2009; Kuppens, Van Mechelen, & Rijmen, 2009). To take the individual differences into account, we could define a separate model per individual. However, it may happen that the number of observations for a certain individual is low and the within-subject estimate might become overly noisy. For this reason, we prefer hierarchical modeling (see also Gelman & Hill, 2007). In a hierarchical model, some of the parameters of the model can differ across individuals (these parameters are also called random effects), but given the hierarchical structure of the model, the parameters of individuals that contribute less information to the sample (e.g., because they have fewer measurements) can nevertheless be reliably estimated. In addition, constraining parameters to be equal across participants is not problematic, and one may regress the random effects on covariate information.

Introducing a Dynamical Approach Based on the Ornstein–Uhlenbeck (OU) Model

On the basis of these important aspects of affective changes, we turn to the branch of mathematics that deals with random phenomena, namely the theory of stochastic processes. Stochastic processes have proven very useful in describing probabilistically governed change in other domains (e.g., statistical physics, ecology), and they seem to be appropriate for modeling complex human phenomena as well. In the present article, we focus on a dynamical model built on a particular stochastic differential equation (SDE), which leads to the OU (Uhlenbeck & Ornstein, 1930) process. This stochastic process combines elements of stochastic variability and deterministic control in an elegant way. Moreover, the OU process is continuous in time (it is the continuous-time analogue of a first-order autoregressive model), which is appropriate for modeling core affect because it does not cease to exist between observations.[1] Also, due to the continuous-time property, the measurements can be taken at person-specific time points, with varying numbers of observations per person. Our approach also allows for the data to be unbalanced and unstructured. The aforementioned complications (unequally spaced measurements, unbalancedness and unstructured data) are very common in data stemming from diary studies. Therefore, from a data-analytical perspective, such flexibility is a key advantage in several applications in the field of emotion psychology. In our model, the OU process represents the change in the true core affect position for a single person over time. However, it is reasonable to assume that the subjectively reported or observed core affect positions will typically be perturbed by measurement error. Incorporating measurement error in the OU model results in a model representation that belongs to the general class of state-space models (see, e.g., Fahrmeir & Tutz, 2001; Jazwinski, 1970; Oud & Jansen, 2000; Ringo Ho, Shumway, & Ombao, 2006). This stochastic process may be appropriate to describe the core affect dynamics in a single individual, but an extension is needed to take into account individual differences in the parameters governing the process. Therefore the key parameters of the state-space OU model will be allowed to differ across persons. In addition, it will be possible to link them to person-specific covariate information. Because of the hierarchical extension, we denote our model as the hierarchical OU (or HOU) model.[2]

[1] Moreover, simplifying continuous time to discrete may compromise the possibilities of inference on the dynamics of change (for an example, see Delsing, Oud, & De Bruyn, 2005).

[2] Readers who are interested in an accessible introduction to the topic of SDE modeling in general and the OU process in particular are referred to Tuma and Hannan (1984). More general and detailed introductions to stochastic processes are in Cox and Miller (1972), Gardiner (2004), Karlin and Taylor (1981), Lawler (2006), and Ross (1996). In this article, we explain the HOU model in an informal and intuitive way so that no previous background in SDE or stochastic processes is required.

The structure of the remainder of the article is as follows: In the next section, the basic OU model for a single individual is introduced first. It is followed by a hierarchical extension allowing for the inclusion of individual differences. Then we describe briefly how statistical inference can be carried out for this model within a Bayesian framework. Following that, the HOU model will be applied to data from a diary study. After a discussion of model fit, the Conclusion section ends the article.

Modeling the Affect Dynamics of a Single Individual: The OU Model With Measurement Error

In this section, we describe the OU process with measurement error as a model for the measured state of a single person. Let us start with some notation. The true or latent position in a two-dimensional latent space at time t will be denoted by the vector $\boldsymbol{\Theta}(t)$ defined as $\boldsymbol{\Theta}(t) = \Theta_1(t), \Theta_2(t)^{\mathrm{T}}$, and the superscript T indicates the transpose operation. In the core affect application, $\Theta_1(t)$ refers to the position on the first dimension (pleasantness) and $\Theta_2(t)$ to the position on the second dimension (arousal). We will define the model for two dimensions here and we refer specifically to core affect, but generalizations to more dimensions and other application areas are possible. In the model formulation, it is assumed that the true core affect changes continuously throughout time, but the measurements are taken at a finite number of time points: $t_1, t_2, \ldots, t_s, \ldots t_n$, where n stands for the number of measurements (Time 0 can be defined arbitrarily by, for instance, setting it equal to the first time point: $t_1 = 0$). We define the vector $\mathbf{Y}(t_s) = (Y_1(t_s), Y_2(t_s))^{\mathrm{T}}$ as the observed pleasantness and arousal scores at time point t_s. The general model can then be written as follows:

$$\begin{cases} d\boldsymbol{\Theta}(t) = \mathbf{B}(\boldsymbol{\mu} - \boldsymbol{\Theta}(t))dt + \boldsymbol{\sigma} d\mathbf{W}(t) \\ \mathbf{Y}(t_s) = \boldsymbol{\Theta}(t_s) + \boldsymbol{\varepsilon}(t_s) \end{cases}, \tag{1}$$

where $\boldsymbol{\mu}$ is a vector with two components and $\boldsymbol{\sigma}$ and $\mathbf{B}$ are positive-definite 2×2 matrices. The measurement error is represented by $\boldsymbol{\varepsilon}(t_s)$, which is a random draw from a bivariate normal distribution with mean $(0,0)^{\mathrm{T}}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. The component $\mathbf{W}(t)$ stands for the standard bivariate Wiener process. The interpretation of these parameters is elaborated below.

At this point, it is important to note that the model in Equation 1 consists of two parts. The first equation describes the change in the true core affect position and is therefore a *transition equation*; it represents the dynamical aspect of the model. The second equation maps the true process onto the observed variable and is called the *observation equation*.

In the remainder of the section, we explain, step by step, the model in Equation 1, by first defining the one-dimensional version of the transition equation. Accordingly, we will introduce the properties of the full, two-dimensional form. In the end, the role of the observation equation is clarified.

A Unidimensional SDE

Let us take the transition equation of Equation 1, transform it into an equation for a unidimensional variable $\Theta(t)$ (e.g., we can consider either the pleasantness or the activation dimension at once, but we cannot see how they influence each other). The result is a linear first-order SDE that describes the dynamics of the OU stochastic process in one dimension:

$$d\Theta(t) = \beta(\mu - \Theta(t))dt + \sigma dW(t), \tag{2}$$

where we assume that $\beta > 0$.[3] The right hand side of Equation 2 can be divided into two parts: The first part of the sum is deterministic, and the second one is stochastic. Considering only the deterministic part, it can be deduced that the instantaneous change in $\Theta(t)$, that is, $d\Theta(t)$, depends on how far the current state $\Theta(t)$ is from the point μ. If $\Theta(t)$ is below μ (i.e., $\mu - \Theta(t) > 0$), the first derivative is positive, and consequently $\Theta(t)$ will increase. The opposite holds when Θ is above μ. Hence, $\Theta(t)$ will always change in the direction of μ and never the other way. Because the process settles itself at μ, this parameter is called a steady state or attractor. However, we will use the term *home base* to refer to μ, as in the context of emotions, one may think of μ as an ideal point to which one is drawn. The parameter β controls the magnitude of the "drawing" effect: If β is large ($\beta \gg 1$), the difference between the actual state and μ tends to be magnified; therefore a faster change will occur in the direction of μ. With small βs (i.e., β close to zero), the change becomes substantially slower. Based on this property, the parameter β is often called the dampening force or centralizing tendency. If we considered only this deterministic part as a model for the core affect dynamics, we would encounter a major disadvantage: The model assumes a gentle but certain return to the attractor or home base, and then the process remains there. This appears unrealistic, because over the course of time, many effects will lead to a divergence between the home base and the actual state, as has been pointed out in the introduction. Therefore a realistic model should incorporate an element of randomness, like the second, stochastic part of the right side of Equation 2. In this, parameter $W(t)$ stands for a unidimensional Wiener process or Brownian motion.[4] However, it is not the Brownian motion as such that is added but rather the quantity $\sigma dW(t)$, where σ is the scale of the stochastic term. The random variable $dW(t)$ can be interpreted loosely as the change in a standard Brownian motion

[3] Note that we put all SDEs in the differential form. This differs from the prime notation often encountered with deterministic differential equations. However, when the stochastic term is added, not all derivatives are defined in the traditional way (see, e.g., Jazwinski, 1970), and therefore we prefer to keep the differential form.

[4] In 1828 the English botanist Robert Brown observed that when part of a pollen of grain is suspended in water, it exhibits an irregular "animated" motion, and the phenomenon was named after him. The random Brownian motion was explained by Einstein (1905) by supposing that the movements are the result of frequent impacts on the particle by the molecules of the water. Because these impacts are incessant, complicated, and highly numerous, the resulting movement path of the particle requires a probabilistic description; a deterministic one is not feasible. Norbert Wiener (1923) provided a rigorous mathematical formalization of the Brownian motion, and therefore the Brownian motion is often called the Wiener process.

process in very small time interval.[5] It can also be written as $dW(t) = \xi(t)dt$, where $\xi(t)$ is called a white noise process, the simplest stochastic process in continuous time on the real line (Gardiner, 2004). It is a mathematical model for a continuous-time process with independent realizations.

In summary, it follows that in the model of Equation 2, the change in $\Theta(t)$ is a function of two factors. First, there is stochastic innovation driving the change process, which is represented by $dW(t)$. This stochastic innovation term incorporates the multiple smaller and larger "impacts" that the emotional system undergoes at a given moment. Second, the control exerted is captured by the centralizing tendency parameter β.

The solution of the unidimensional OU process involves integrating over Equation 2 and solving a stochastic integral. The derivation of this solution, together with a brief overview of the most important properties of the stochastic integral, can be found in Appendix A (see also Tuma & Hannan, 1984). The general solution results in an expression for $\Theta(t)$, given that the process was at Θ_0 at Time 0. However, for our purposes, it is more useful to condition on the position d time units before, that is, $\Theta(t - d)$, such that it becomes possible to model a chain of subsequent measurements. With $\Theta(t - d)$ as initial value, the solution of Equation 2 becomes

$$\Theta(t)|\Theta(t-d) \sim N\left(\mu + e^{-\beta d}(\Theta(t-d) - \mu), \frac{\sigma^2}{2\beta}(1 - e^{-2\beta d})\right). \tag{3}$$

This conditional normal distribution will turn out to be very convenient when estimating the parameters of the model, because it allows us to construct the likelihood.

From Equation 3, we can see that the position at time t, that is, $\Theta(t)$, depends on the already introduced parameters and the previously measured position $\Theta(t - d)$. Figure 1 displays some solution curves for this unidimensional OU stochastic process. The initial values or starting points, that is, $\Theta(t - d)$, are different, but the home base, the centralizing tendency, and the scale of the stochastic term remain the same ($\mu = 0$, $\beta = 1$, and $\sigma = 0.1$). It can be seen that for an initial value of $\Theta_0 = 0$, there is no change at all over time. We can clearly see how the stochastic disturbance term has a profound influence on the trajectories: The return to the baseline shows a noisy pattern.

If we let d go to infinity in Equation 3 (i.e., we condition on a position a very long time ago), then we see that the distribution of $\Theta(t)$ does not depend on $\Theta(t - d)$ anymore:

$$\Theta(t) \sim N\left(\mu, \frac{\sigma^2}{2\beta}\right), \tag{4}$$

assuming that $\beta > 0$. This fact indicates that the initial state $\Theta(t - d)$ is forgotten as $d \to \infty$. Parameter $\sigma^2/2\beta$ is the stationary variance of the process. Because the stationary variance, which we will denote $\gamma = \sigma^2/2\beta$, is easier to interpret than the instantaneous variance σ^2, as it corresponds directly to the total intraindividual variance, and because it offers some computational advantages in the bivariate model, from now on we use a reparameterized version of the process in which σ^2 is replaced by $2\beta\gamma$.

The Two-Dimensional SDE

With a two-dimensional model, we can incorporate modeling pleasantness together with activation in one model, and we can also investigate how they influence each other. For example, we can see whether the two dimensions tend to change together, in the same direction or the opposite, and so on. We have already shown the SDE of the OU process in two dimensions in the first line of Equation 1. This equation can also be solved to arrive to a conditional distributional representation of the two-dimensional OU process. The derivation of this solution can be found in Appendix B. Here we present only the solution, which is

$$\boldsymbol{\Theta}(t)|\boldsymbol{\Theta}(t-d) \sim N_2(\boldsymbol{\mu} + e^{-\mathbf{B}d}(\boldsymbol{\Theta}(t-d) - \boldsymbol{\mu}), \boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}^{\mathrm{T}}d}), \tag{5}$$

where N_2 refers to the bivariate normal distribution. As in the unidimensional case, the two-dimensional process converges to a stationary distribution:

$$\boldsymbol{\Theta}(t) \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Gamma}), \tag{6}$$

provided that all eigenvalues of $\mathbf{B}$ are positive. The latter condition also ensures that the process is stable (Oud & Singer, 2008). Equation 6 can be considered as a two-dimensional version of Equation 4. An informal justification of Equation 6 can be obtained by letting d go to infinity in Equation 5. The matrix $\boldsymbol{\Gamma}$ is the stationary covariance matrix.

The Parameters of the Two-Dimensional OU Process

In this section, we study the interpretation of the parameters (i.e., $\boldsymbol{\mu}$, $\boldsymbol{\Gamma}$, and $\mathbf{B}$) of the two-dimensional OU process in detail. This will be done mainly by making use of simulated trajectories and varying the parameter of interest so that we can visualize its effect. Such a study will increase the understanding of the process and will set the stage for allowing for individual differences.

To simulate the trajectories, first a set of 200 time differences (i.e., t_1, $t_2 - t_1$, etc.) was sampled uniformly between 0.5 and 1. Next, the two-dimensional process was simulated with the conditional distribution from Equation 5. The vector $\boldsymbol{\Theta}(0)$ was drawn from the stationary distribution. The specific parameter values for the parameters are given below.

The home base. As in the unidimensional case, $\boldsymbol{\mu}$ is again the ideal point to which the process is drawn (also called the home base), with the only difference that $\boldsymbol{\mu}$ is now a two-dimensional vector. From Equation 6, one can also infer that the home base is the average of the stationary distribution.

In Figure 2, two trajectories are simulated with different home bases, but all other parameters are constant. As expected, it can be noticed that with different home bases, the visits are concentrated around different areas. Stated otherwise, the stationary distributions are just simple translations of one another.

[5] We use $dW(t)$ notation instead of dW/dt, because the latter would be undefined, as the path of a Wiener process is not differentiable with respect to time.

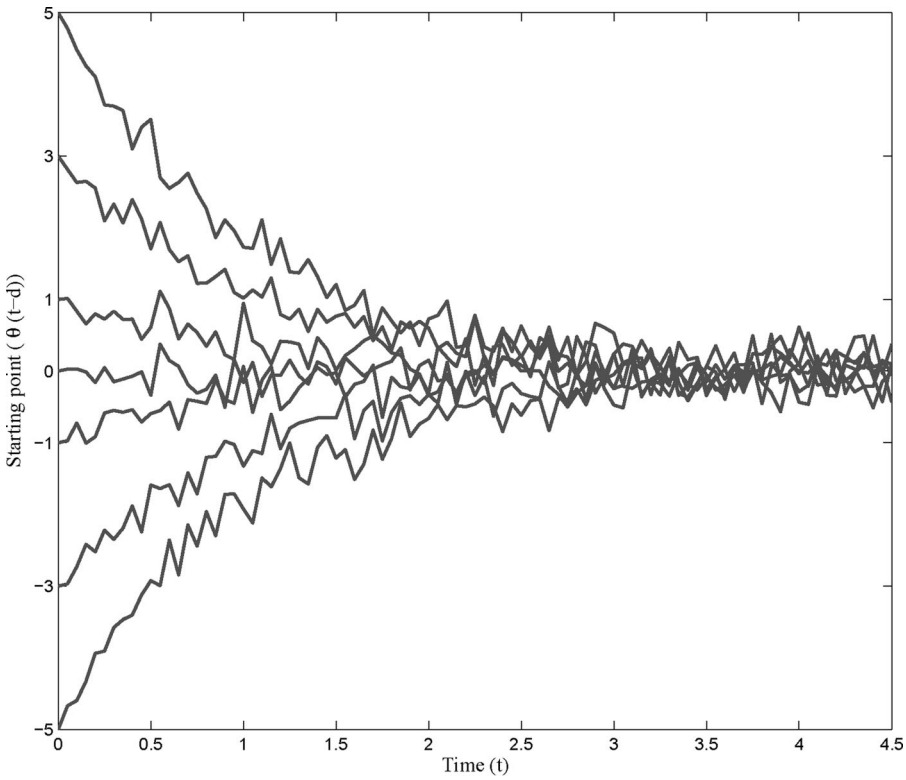


Figure 1. Solution curves for Equation 3.

Intraindividual variation: Fluctuations around the home base. The positive-definite matrix $\mathbf{\Gamma}$ is the covariance matrix of the stationary distribution. To simplify the notation of the separate elements in $\mathbf{\Gamma}$, the matrix is decomposed as follows:

$$\mathbf{\Gamma} = \begin{pmatrix} \gamma_1 & \rho_\gamma \sqrt{\gamma_1 \gamma_2} \\ \rho_\gamma \sqrt{\gamma_1 \gamma_2} & \gamma_2 \end{pmatrix}. \quad (7)$$

The parameters γ_1 and γ_2 correspond to the variance in the first and second dimensions, respectively. The parameter ρ_γ is the cross-correlation, and it quantifies the strength of the linear dependency between the two dimensions in the stationary distribution. If ρ_γ is close to 1, then positive (negative) displacements on one dimension go together with positive (negative) displacements on the other dimension.

In Figure 3, two realized OU processes with different $\mathbf{\Gamma}$ matrices are shown. In Figure 3A, low variance values were used for the simulation, yielding small changes, and therefore the process tends to stay near the home base. In contrast, because of the larger variances used for Figure 3B, the simulated process covers a wider area (i.e., higher volatility and thus more dramatic changes). Moreover, in Figure 3A, the cross-correlation is set to 0, whereas in Figure 3B it is equal to 0.5. As can be seen, increasing the correlation leads to displacements that coincide as the shape of the trajectory clearly suggests. (Again, the other parameters were kept fixed across the two simulations.)

Regulation: The centralizing tendency. The matrix $\mathbf{B}$ is the matrix equivalent of the scalar β in the unidimensional process, and as such it governs the strength and the direction with which the process is pulled back to the home base $\boldsymbol{\mu}$. As stated before, it is required that $\mathbf{B}$ be a positive-definite matrix such that there is always an adjustment toward the home base (and hence the process is stable). If $\mathbf{B}$ were not positive definite, the process would become "explosive," meaning that it would be pushed away from the home base. The HOU model cannot capture such a phenomenon, and constantly being pushed away from the ideal point does not seem to be very realistic to model affective dynamics.[6] Besides positive definite, we will also require that $\mathbf{B}$ be symmetric, and we decompose it in a similar way as was shown for $\mathbf{\Gamma}$:

$$\mathbf{B} = \begin{pmatrix} \beta_1 & \rho_\beta \sqrt{\beta_1 \beta_2} \\ \rho_\beta \sqrt{\beta_1 \beta_2} & \beta_2 \end{pmatrix}. \quad (8)$$

The symmetry constraint is specific to our model formulation. One consequence is that the effects of the first process on the second and vice versa are equal. The reason why we made this simplifying assumption is that in a later modeling stage, we allow for individual differences in the elements of $\mathbf{B}$. Without this assumption, it is extremely difficult to satisfy a basic condition in the model, namely that $\mathbf{B}_p\mathbf{\Gamma}_p + \mathbf{\Gamma}_p\mathbf{B}_p^T$ has to be positive definite, because this sum represents the instantaneous covariance matrix, as shown in Appendix B (see also Dunn & Gipson, 1977). For this reason we sacrificed the additional level of complexity in the basic

[6] Another possibility could be that $\mathbf{B} = 0$, and in that case the resulting process is simply the Wiener process or Brownian motion. Then we would assume only random fluctuations in affect, which also seems unlikely.

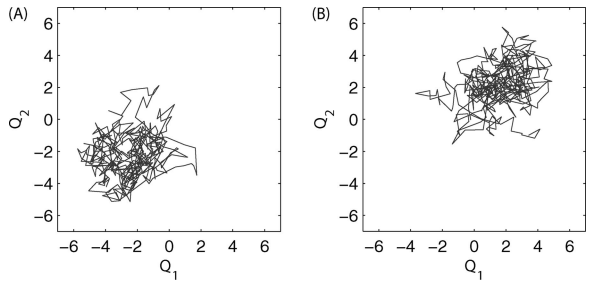


Figure 2. Two simulated two-dimensional Ornstein–Uhlenbeck processes with different home bases. For both plots, $\mathbf{\Gamma} = 2I$ and $\mathbf{B} = 0.1I$, with I being the 2×2 identity matrix. In Figure 2A, $\boldsymbol{\mu} = (-2, -2)^{\mathrm{T}}$. In Figure 2B, $\boldsymbol{\mu} = (2, 2)^{\mathrm{T}}$.

model (i.e., asymmetry) to allow for the study of individual differences later on.[7]

The parameters β_1 and β_2 correspond to the centralizing tendency in the first and second dimensions, respectively. The parameter ρ_β represents a common cross-centralizing tendency. To get a clear idea of the interpretation of **B**, we make use of three distinct ways of visualizing its effects: via simulations, via autocorrelation functions, and via orbital portraits. The result for three **B**s can be found in Figure 4. In each row a different visualization method is used. In each column the parameter settings remain constant, but across columns the **B**s differ systematically. In Figures 4A, 4D, and 4G, we have chosen $\beta_1 = \beta_2 = 0.01$ and $\rho_\beta = 0$ (low beta and no cross-effects). In Figures 4B, 4E, and 4H, we have set $\beta_1 = \beta_2 = 0.5$ and again $\rho_\beta = 0$ (high beta and no cross-effects). Finally, in Figures 4C, 4F, and 4I, $\beta_1 = 0.01$, $\beta_2 = 0.1$, and $\rho_\beta = -0.4$ (mixed betas and cross-effects). For all plots, the home base is located at the origin, and $\mathbf{\Gamma} = 2I$ (with I being the 2×2 identity matrix).

Figures 4A–4C contain three simulated trajectories. From Figure 4A, it can be seen that a low β leads to a low centralizing tendency. The simulated process tends to stay close to the previous observation and is not strongly attracted by the home base. In contrast, the simulated trajectory in Figure 4B is based on a large β, and this corresponds to a large centralizing tendency: The process fluctuates to a large extent around the mean. Note that for both simulations, the ultimate (i.e., stationary) covariance matrix is equal by definition. In Figure 4C, one can see that the small β_1

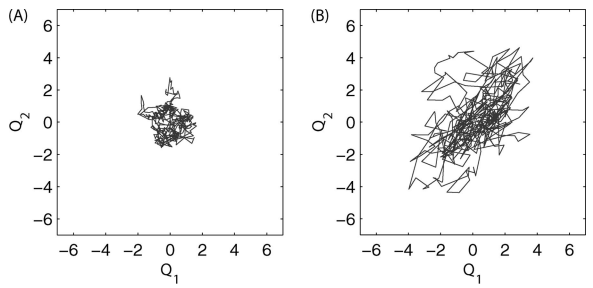


Figure 3. Two simulated two-dimensional Ornstein–Uhlenbeck processes with different stationary covariance matrices $\mathbf{\Gamma}$. For both plots, $\boldsymbol{\mu} = (0, 0)^{\mathrm{T}}$ and $\mathbf{B} = 0.1I$. In Figure 3A, $\gamma_1 = \gamma_2 = 0.5$ and $\rho_\gamma = 0$. In Figure 3B, $\gamma_1 = \gamma_2 = 4$ and $\rho_\gamma = 0.5$.

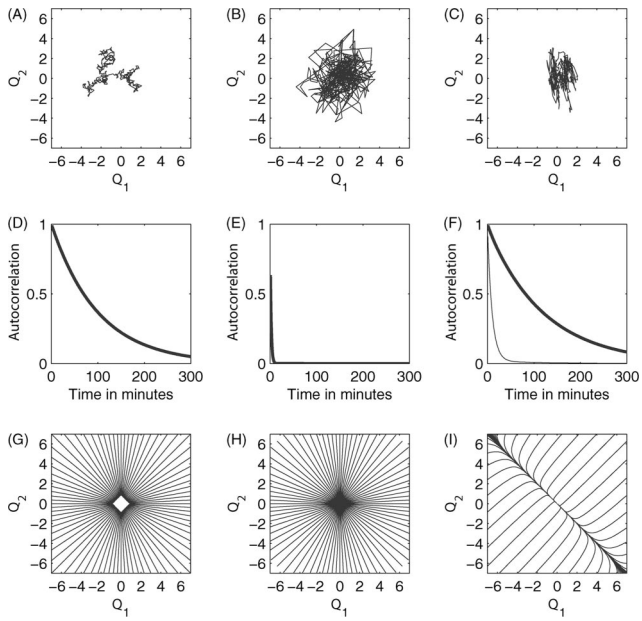


Figure 4. Shows the effect of different **B** matrices.

leads again to a slowly moving process, whereas this does not hold for the second dimension because β_2 is large (and therefore the centralizing tendency in the direction of the second dimension is also large). Hence, this third simulated trajectory combines the features of the previous two.

In Figures 4D–4F, the autocorrelation functions for the two processes are shown. The autocorrelation function value at time point t can be computed with the matrix exponential (see Appendix A) as follows: $e^{-\mathbf{B}t}$. The result of the matrix exponential is a matrix itself (of the same size as **B**), and the autocorrelation function value at time t for the first (second) dimension is then the first (second) diagonal element. If we let time t vary from 0 to 300 min, we can plot the continuous autocorrelation functions. For the first two settings (first two columns), the two autocorrelation functions in the two dimensions are the same, because $\beta_1 = \beta_2$. All autocorrelation functions are exponentially decaying. It can be seen from Figure 4D that if β is low (i.e., a low centralizing tendency), the autocorrelation function shows a slower decay than if β is high (see Figure 4E). For instance, for the process in Figure 4D, even after 100 min, the autocorrelation is still around 0.4 (i.e., if two subsequent states are 100 min separated, they correlate 0.4). In Figure 4E, the autocorrelation is practically 0 after 15 min (about 0.015). The fact that a low β corresponds with a large autocorrelation is not surprising if one looks at the simulated trajectory in Figure 4A: The process tends to stay close to the previous observation, indicating a large autocorrelation. On the other hand, a large value for β (as in Figures 4B and 4E) leads to a fast decaying autocorrelation with time. Therefore the simulated trajectory connects almost independently sampled points from a

[7] We tried to remove the symmetry constraint by not taking care of the Dunn–Gipson condition. However, the numerical algorithm to estimate the model's parameters (see below) failed in that case.

bivariate normal distribution. In Figure 4F, the β values for the two dimensions differ, and thus two autocorrelation functions appear.

Figures 4G–4I contain so-called orbital portraits. In deterministic systems of linear differential equations, the matrix **B** controls the type of deterministic trajectories traced out in the state space, that is, the space defined by the axes $\Theta_1(t)$ and $\Theta_2(t)$. A collection of such deterministic trajectories or orbits is called a portrait. Although we are working with a stochastic model, it is useful to look at the orbital portraits in the corresponding deterministic case. Note that these orbital portraits also correspond to the expected trajectories resulting from the two-dimensional SDE model: Given that a point on one of the orbits is the previous observation, the conditional mean is also located somewhere on the orbit but closer to the home base. In Figures 4G–4I, we find three orbital portraits corresponding to the three **B** matrices. It can be seen that an isotropic matrix (i.e., of the form β_1, as used in the first and second column) gives rise to a so-called star node: The expected trajectories toward the home base are straight lines (in fact, Figures 4G and 4H are similar because their **B** is not qualitatively different). In Figure 4I, the orbital portrait of a model with unequal elements on the diagonal of **B** and $\rho_\beta = 0.4$ leads to a so-called improper node. In the latter situation, the adjustment toward the home base falls along a curved trajectory. Orbital portraits might reveal interesting patterns in affect regulation in the core affect, as they display the way that the centralizing tendency force acts to restore the balance of the dynamic system.[8]

A time-varying home base $\boldsymbol{\mu}(t)$. Hitherto we have assumed that the home base is constant over time. However, it seems reasonable to expect that the ideal point to which the process is attracted is subject to changes throughout time. For example, several studies indicate diurnal patterns in how active and how pleasant people feel throughout the day (Caminada & De Bruijn, 1992; Haug & Fähndrich, 1990; Rusting & Larsen, 1998). To take such structural changes into account, we extend the model as follows:

$$\boldsymbol{\Theta}(t) \sim N_2(\boldsymbol{\mu}(t), \boldsymbol{\Gamma})$$

and

$$\boldsymbol{\Theta}(t)|\boldsymbol{\Theta}(t-d) \sim N_2(\boldsymbol{\mu}(t) + e^{-\mathbf{B}d}(\boldsymbol{\Theta}(t-d) - \boldsymbol{\mu}(t)), \boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}^{\mathrm{T}}d}). \quad (9)$$

In our application below, $\boldsymbol{\mu}(t)$ is assumed to be a polynomial (e.g., quadratic) function of time of the day. However, in principle one could take any other function of time (e.g., spline based) or one could let measured time-varying covariates have an effect on the home base (e.g., important life events).

It is important to note that we assume the systematic variation throughout time (as expressed by $\boldsymbol{\mu}(t)$) and the stochastic dynamics of the model (i.e., the adjustment to the time-varying home base) to be two aspects of the model. We have simply changed the constant value $\boldsymbol{\mu}$ to the time-varying variant $\boldsymbol{\mu}(t)$ without affecting the stochastics of the model. Stated differently, if the time-varying home base is subtracted from a simulated trajectory, the result is a simulated trajectory from a model with a constant home base. A different type of process would result if the time-varying home base were inserted in Equation 1, because then it would be part of the intrinsic dynamics of the model.

The Observation Equation

A final aspect of the model for the within-person dynamics concerns the measurement error. Almost any measurement in psychology will be affected with measurement error to some extent, and therefore it is important to take it into account. By adding measurement error to the latent process, we have discussed thus far, we will step down to the observational level. In Equation 1, the first part serves as a transition equation in our model by describing the changes in the true score vector, that is, $\boldsymbol{\Theta}(t)$. To link the underlying dynamical change $\boldsymbol{\Theta}(t_s)$ to the observed data $\mathbf{Y}(t_s)$, we use the following observation equation:

$$\mathbf{Y}(t_s) = \boldsymbol{\Theta}(t_s) + \boldsymbol{\varepsilon}(t_s), \quad (10)$$

for observations at time points $t_1, t_2, \ldots, t_s, \ldots, t_n$. Although the underlying process is assumed to be continuous in time, it is impossible to make continuous observations in the contexts we consider. Therefore the observations will necessarily be restricted to a discrete set of time points. Jazwinski (1970) called such models continuous-time models with discrete-time sampling.

The measurement errors $\boldsymbol{\varepsilon}(t_1), \ldots, \boldsymbol{\varepsilon}(t_2), \ldots, \boldsymbol{\varepsilon}(t_n)$ are assumed to follow a bivariate normal distribution:

$$\boldsymbol{\varepsilon}(t_s) \overset{\text{iid}}{\sim} N_2(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}), \quad (11)$$

with $\mathbf{0}$ as mean and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ as covariance matrix. In the remainder of the article, we will assume a diagonal matrix for $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ (i.e., uncorrelated errors for the two dimensions). By establishing the observation equation, we have completed our description of the OU model for a single person, as it was given in Equation 1, and now we move on to the hierarchical part of the model.

The HOU Model

In the previous sections, we have described a model for the affective dynamics of a single individual. From the interpretation of the parameters, it becomes clear that they all capture an important aspect of the dynamics. As argued in the introduction, individual differences are common in affective processes. More specifically, individual differences are expected in each of the parameters of the single-person OU process (with the exception of the measurement error variance). For instance, in a recent study, Kuppens, Van Mechelen, Nezlek, Dossche, and Timmermans (2007) found that people differ consistently in the mean level of pleasantness and activation (corresponding to individual differences in the home base) but also in their variabilities on these dimensions (corresponding to individual differences in the diagonal elements of $\boldsymbol{\Gamma}$). Moreover, Gross and John (2003) provided evidence for individual differences in the regulation of affect (corresponding to the **B** matrix).

In order to describe and explain individual differences, a natural approach is to use a hierarchical model (see, e.g., Gelman & Hill, 2007; Snijders & Bosker, 1999). In a hierarchical model (or multilevel or mixed model), the parameters that are subject to individual differences are random effects, sampled from a popu-

[8] Many more orbital portraits are possible (e.g., saddle points, spiral points), if one allows **B** to be a nonsymmetric positive-definite matrix.

lation distribution that is characterized by a set of parameters. In traditional hierarchical models, the random effects are usually allowed only for the mean structure. In contrast, in the model and application we consider, it is meaningful to allow for individual differences in the variability and centralizing tendency parameters as well. Hence, the variance and centralizing tendency parameters will be assumed to be sampled from a population distribution as well. Note that this allows us in a next step to regress the random person-specific parameters on individual difference covariates.

In the hierarchical model, it is assumed that each parameter comes from a specific population distribution. For reasons of computational and interpretational convenience, we assume normal distributions for all random effects. However, not all parameters are defined on the real line (e.g., the diagonal elements of the covariance matrix of the stationary distribution can assume only positive values), and for such parameters we will use appropriate transformations that map them on the real line and then proceed with a normal distribution.

Let us introduce some new notation: A specific person p ($p = 1, \ldots, P$) is measured n_p times at the following sequence of time points: $t_{p1}, t_{p2}, \ldots, t_{ps}, \ldots, t_{p,np}$. The index s denotes the sth measurement occasion of that individual. As mentioned in the introduction, a great strength of our model is that we do not require that measurements occur at regular time intervals nor that the measurement occasions be identical across participants. For notational convenience, we will use p and s as the only indices when denoting parameters or data that are related to the specific observation at t_{ps}. This way, for example, the measured position for person p in the two-dimensional space at time t_{ps} is denoted as $\mathbf{Y}_{ps}$ instead of the more cumbersome notation $\mathbf{Y}(t_{ps})$.

The model for a single person p for whom the observed data are a function of an underlying OU process and some measurement error can now be written as follows:

$$\mathbf{Y}_{ps} = \mathbf{\Theta}_{ps} + \boldsymbol{\varepsilon}_{ps}, \tag{12}$$

where $\mathbf{Y}_{ps}$ stands for the observed random vector, $\mathbf{\Theta}_{ps}$ for the latent state (or true score), and $\boldsymbol{\varepsilon}_{ps}$ for the measurement error with the same distributional assumption as presented in Equation 11. As expressed in Equation 9, the conditional distribution of $\mathbf{\Theta}_{ps}$ given $\mathbf{\Theta}_{p,s-1}$ is normally distributed as follows (for $s > 1$):

$$\mathbf{\Theta}_{ps} | \mathbf{\Theta}_{p,s-1} \sim N_2(\boldsymbol{\mu}_{ps} + e^{-\mathbf{B}_p(t_{ps}-t_{p,s-1})}(\mathbf{\Theta}_{p,s-1} - \boldsymbol{\mu}_{ps}), \; \mathbf{\Gamma}_p - e^{-\mathbf{B}_p(t_{ps}-t_{p,s-1})}\mathbf{\Gamma}_p e^{-\mathbf{B}_p^{\mathrm{T}}(t_{ps}-t_{p,s-1})}). \tag{13}$$

For the first observation, $\mathbf{\Theta}_{p1}$, it is assumed that $\mathbf{\Theta}_{p1} \sim N_2$ ($\boldsymbol{\mu}_{ps}$, $\mathbf{\Gamma}_p$). Although Equations 12 and 13 are identical to Equations 10 and 9, respectively, we present them again because they are now defined with a slightly different but more convenient notation. Note that the presence of the indices p in Equation 13 reflects that all driving parameters of the OU process are allowed to be person specific. In the next sections we develop the hierarchical extension for each of the parameters.

Model for the Person-Specific and Time-Varying Home Base

As can be deduced from the notation (i.e., the indices s and p), the home base $\boldsymbol{\mu}_{ps}$ consists of a time-varying and person-specific aspect. Corresponding to these two aspects, background measurements may be available. Regarding the person-specific aspect, it is assumed that k covariates are measured and x_{jp} denotes the score of person p on covariate j ($j = 1, \ldots, k$). All person-specific covariate scores are collected into a vector of length $k + 1$, denoted as $\mathbf{x}_p = (x_{p0}, x_{p1}, x_{p2}, \ldots, x_{pk})^{\mathrm{T}}$, with $x_{p0} = 1$. Regarding the time-varying aspect, suppose that we measure for person p the scores on m time-varying covariates that are collected in a vector $\mathbf{z}_{ps} = (z_{ps1}, \ldots, z_{psm})^{\mathrm{T}}$, where the presence of the index s indicates that the scores may change from one time point to another. In order to avoid collinearity problems, no intercept is introduced in the vector $\mathbf{z}_{ps}$.

The regression of $\boldsymbol{\mu}_{ps}$ onto the two types of covariates and allowing for a person-specific random deviation is defined as follows:

$$\boldsymbol{\mu}_{ps} = \mathbf{\Delta}_\mu \mathbf{z}_{ps} + \mathbf{A}_\mu \mathbf{x}_p + \mathbf{E}_{p\mu}, \tag{14}$$

with $\mathbf{E}_{p\mu} \sim N_2(\mathbf{0}, \mathbf{\Sigma}_{\boldsymbol{\mu}})$. The matrices $\mathbf{\Delta}_\mu$ and $\mathbf{A}_\mu$ are parameter matrices of dimension $2 \times m$ and $2 \times (k + 1)$, respectively, containing the regression weights for the covariates. Furthermore, the covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\mu}}$ is defined as follows:

$$\Sigma_{\boldsymbol{\mu}} = \begin{pmatrix} \sigma^2_{\mu_1} & \sigma_{\mu_1\mu_2} \\ \sigma_{\mu_1\mu_2} & \sigma^2_{\mu_2} \end{pmatrix}. \tag{15}$$

Because the vector $\boldsymbol{\mu}_{ps}$ is bivariate, it may be illuminating to write the component regression models in more detail:

$$\begin{pmatrix} \mu_{ps1} \\ \mu_{ps2} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\delta}^{\mathrm{T}}_{\mu_1} \\ \boldsymbol{\delta}^{\mathrm{T}}_{\mu_2} \end{pmatrix} z_{ps} + \begin{pmatrix} \boldsymbol{\alpha}^{\mathrm{T}}_{\mu_1} \\ \boldsymbol{\alpha}^{\mathrm{T}}_{\mu_2} \end{pmatrix} x_p + \begin{pmatrix} \mathbf{e}_{p\mu_1} \\ \mathbf{e}_{p\mu_2} \end{pmatrix}. \tag{16}$$

One can consider Equation 14 as a decomposition of the person and time-specific home base $\boldsymbol{\mu}_{ps}$ into three components. The first component allows for the home base to fluctuate over time, but it is constant over all persons (i.e., $\mathbf{\Delta}_\mu \mathbf{z}_{ps}$). The time-varying covariates in the vector $\mathbf{z}_{ps}$ can be anything for which it is meaningful to assume that it relates to the change of home base across time. The most straightforward covariates are time and functions of time itself, which are illustrated in the application section in this article. However, if one has information on other variables that may be related to home base and change through time, they can be incorporated in as well. A second component allows the home base to differ because of the effect of person-specific covariates (i.e., $\mathbf{A}_\mu \mathbf{x}_p$). These person-specific effects leave the time-varying part unaffected. If no person-specific covariates have been collected, then only the bivariate intercept is present (i.e., the two regression weights of the constant 1). The last part, $\mathbf{E}_{p\mu}$, represents the contribution of a bivariate person-specific random effect that will make the home base different from one person to another. It follows a bivariate normal distribution with 0 mean vector and covariance matrix $\Sigma_{\boldsymbol{\mu}}$ (see Equation 15). The covariance matrix is the residual covariance matrix, representing the variability and associations of home base intercepts that exist in the population between the individual means of the stationary distribution, after taking into account the effects of the person covariates and of the time-varying covariates. If only the intercept is present in the covariate vector, then the model just describes the population mean vector of the home bases and the variability in the population.

The so far presented fixed and random effects in the home base (i.e., mean structure) are strongly related to the so-called growth curve models (see, e.g., Pan & Fang, 2002). The particular novel contribution of our HOU model is that also the variance and centralizing tendency parameters will be turned into random effects. In this way, individual differences in intraindividual variation and regulatory dynamics can be described and explained. Such an extension is covered in the next sections. However, it should be mentioned that only the home base is allowed to change as a function of time-varying covariates (and thus time), whereas the other parameters cannot be affected by time-varying covariates.

Model for the Person-Specific Stationary Variances and Cross-Correlation

As was clear from Figures 3A and 3B with simulated OU process, changing the stationary variance matrix has a profound effect on the appearance of the simulated trajectories. Such interindividual differences may even be seen in real data. An example is shown in the application section (in Figure 5). A convenient way to model the individual differences in the stationary covariance matrix $\mathbf{\Gamma}_p$ starts with the decomposition presented in Equation 7, in which $\mathbf{\Gamma}_p$ is split into two variances (i.e., γ_{1p} and γ_{2p}) and a cross-correlation (i.e., ρ_{γ_p}). We will discuss these in turn.

First, the person-specific intraindividual variance γ_{1p} is assumed to be drawn from a population distribution. The most straightforward choice would be to assume a normal distribution, but γ_{1p} is constrained to be positive (while the normal distribution has support on the whole real line). However, we may put a normal population distribution on a transformation of γ_{1p}. The most convenient choice is a logarithmic transformation:

$$\log(\gamma_{1p}) = \mathbf{x}_p^{\mathrm{T}}\boldsymbol{\alpha}_{\gamma_1} + e_{p\gamma_1},$$

with $e_{p\gamma_1} \sim N(0, \sigma^2_{\gamma_1})$ and $\mathbf{x}_p^{\mathrm{T}}$ the vector of covariates with $k + 1$ components (of which the first one is the constant 1). The vector $\alpha_{\gamma 1}$ contains the (fixed) regression coefficients for the covariates. The parameter $\sigma^2_{\gamma_1}$ is the residual variance in the random log variance of the first dimension, after having taken the covariate effects into account. If only the intercept is present in the model, $\sigma^2_{\gamma_1}$ reflects the total amount of variance present in the population in the log variance of the first dimension. A similar logic applies to the modeling of γ_{2p}. The population distributions of γ_{1p} and γ_{2p} are assumed to be log normal on the original scale.

The final parameter of the matrix $\mathbf{\Gamma}_p$ is the person-specific cross-correlation parameter ρ_{γ_p}. Because ρ_{γ_p} is a correlation bounded between -1 and 1, Fisher z transformation is implemented in order to be able to use a normal distribution to model the random effects. The Fisher z transformation

$$F(\rho_{\gamma_p}) = \frac{1}{2}\log\frac{1+\rho_{\gamma_p}}{1-\rho_{\gamma_p}}$$

to get a function value on the whole real line is

$$\frac{1}{2}\log\frac{1+\rho_{\gamma_p}}{1-\rho_{\gamma_p}} = \mathbf{x}_p^{\mathrm{T}}\boldsymbol{\alpha}_{\rho_\gamma} + e_{p\rho_\gamma},$$

with $e_{p\rho_\gamma} \sim N(0, \sigma^2_{\rho_\gamma})$. The density of the original ρ_{γ_p} can be derived by applying the transformation of variables technique (see, e.g., Mood, Graybill, & Boes, 1974). Again, $\boldsymbol{\alpha}_{\rho_\gamma}$ contains $k + 1$ regression coefficients, $\mathbf{x}_p$ the k covariate values for person p with 1 for the intercept, and $\alpha^2_{\rho_\gamma}$ represents the variation in the population in terms of cross-correlation.

It should be noted that for reasons of simplicity, the population distributions for the log variance parameters and Fisher z-transformed cross-correlations are modeled unidimensionally. That is, we do not allow for correlations among $e_{p\gamma_1}$, $e_{p\gamma_2}$, and $e_{p\rho_\gamma}$.

Model for the Person-Specific Centralizing Tendencies and Cross-Centralizing Tendencies

A crucial part of the model is the regulatory mechanism that is included. It is parameterized by the matrix $\mathbf{B}_p$, which is decomposed (see Equation 8) into two centralizing tendencies, one for each dimension (i.e., β_{1p} and β_{2p}) and a standardized cross-centralizing tendency parameter (ρ_{β_p}). As with the stationary variance matrix, all three elements of $\mathbf{B}_p$ are assumed to be person specific. This way some people might show only a mild level of regulation or almost no regulation, whereas others might have a very strong regulatory force. The centralizing tendency is a much less straightforward property of observed core affect trajectories than, for instance, the home base or the variabilities, because it affects the trajectories in more subtle ways. In a graphical representation of the trajectory, it might be covered by the effects of measurement noise. Nevertheless, when fitting the model to the data, one can clearly see there is a considerable amount of interindividual variability in this aspect of the model (see the application section below). The two centralizing tendencies of each dimension will be discussed first and then the cross-centralizing tendency.

Because the diagonal elements of $\mathbf{B}_p$ are required to be positive, a similar approach for β_{1p} and β_{2p} as for stationary variances is taken to model the corresponding population distributions:

$$\log(\beta_{1p}) = \mathbf{x}_p^{\mathrm{T}}\boldsymbol{\alpha}_{\beta_1} + e_{p\beta_1}, \tag{17}$$

with $e_{p\beta_1} \sim N(0,\sigma^2_{\beta_1})$ and $\mathbf{x}_p$ the vector of covariates with $k + 1$ components (of which the first one is the constant 1). The vector $\boldsymbol{\alpha}_{\beta_1}$ contains the (fixed) regression coefficients for the covariates. The parameter $\sigma^2_{\beta_1}$ is the residual variance in the random log-

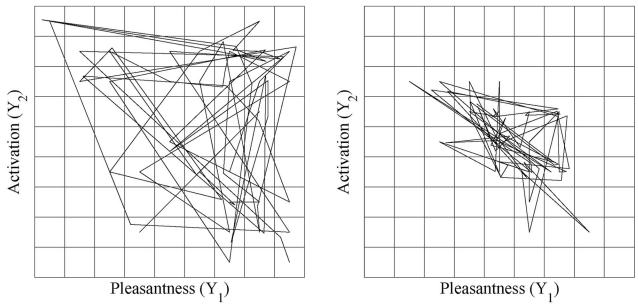


Figure 5. Person trajectories in the core affect grid: in the left plot, participant nr.1 (62 self-reports); in the right plot, participant nr.3 (62 self-reports). The visits falling in the same cell are jittered for a clearer graphical representation.

centralizing tendency specific to the first dimension, after having taken the covariate effects into account. If only the intercept is present in the model, $\sigma^2_{\beta_1}$ reflects the total amount of variance present in the population in the log-centralizing tendency of the first dimension. The model for β_{2p} is analogous. The population distributions for β_{1p} and β_{2p} are assumed to be log normal on the original scale.

The standardized off-diagonal element (ρ_{β_p}) controls the trajectory of the return to the home base (see Figures 4G–4I). It is assumed to have the same population distribution as ρ_{γ_p}; that is, after taking the Fisher z transformation, it is normally distributed:

$$\frac{1}{2}\log\frac{1+\rho_{\beta_p}}{1-\rho_{\beta_p}} = \mathbf{x}_p^{\mathrm{T}}\boldsymbol{\alpha}_{\rho_\beta} + e_{p\rho_\beta},$$

with $e_{p\rho_\beta} \sim N(0,\sigma^2_{\rho_\beta})$. We can interpret the parameters in the same manner as for Equation 17.

Bayesian Inference for the HOU Model

Although the hierarchical extension is both substantively interesting (because all parameters may differ across persons) and straightforward (one has to assign population distributions to the individual difference parameters), statistical inference for such models is not a trivial task. For the model presented here, statistical inference with maximum likelihood would involve a high-dimensional integration over the numerous random effect distributions. Because most of these integrals have no closed-form solutions, these would have to be approximated by finite sums, which is computationally prohibitive. An additional problem with the likelihood method for this case lies in the nonlinear function of latent variables (multiplication of latent variables, exponentiation, etc.; see, e.g., Klein & Moosbrugger, 2000; Schumacker & Marcoulides, 1998). The latent variables (parameters that are allowed to vary in the population) are normally distributed, but because of the nonlinear functions, the marginal distribution of the data Y (after integrating out the latent variables) will not be normal anymore. In frequentist mixed model inference, it is precisely the marginal likelihood that is maximized (Verbeke & Molenberghs, 2000). However, these difficulties are avoided in the Bayesian paradigm in which the explicit integration over the random effects is avoided, because the inference is based on the full joint posterior distribution of the parameters (and not on the marginal). Integration in the Bayesian context typically occurs only to obtain summary measures of the posterior distribution and is based on posterior samples from a Monte Carlo procedure (Klein & Moosbrugger, 2000). In sum, choosing for a Bayesian framework to perform the statistical inference carries some obvious pragmatic value. But additionally, the Bayesian approach is more appropriate for investigating problems in behavioral sciences than the classical statistical inference framework. Parameters in the Bayesian framework have a probability distribution, which offers an intuitively appealing way of thinking about uncertainty and the knowledge one has about the parameters. Moreover, the Bayesian framework is a coherent method for making decisions. A lot of recent methodological work in psychology makes use of the Bayesian framework (see, e.g., Gallistel, 2009; Klein Entink, Kuhn, Hornke, & Fox, 2009; Rouder, Speckman, Sun, Morey, & Iverson, 2009; Rouder, Tuerlinckx, Speckman, Lu, & Gomez, 2008; Smith & Batchelder, 2010).

An advantage of Bayesian statistical inference is that we can apply algorithms to sample from the posterior density of the parameters. The posterior density represents the probability distribution of the parameters given the data, and it is directly proportional to the product of the likelihood of the data (given the parameters) and the prior distribution of the parameters. Formally, $p(\boldsymbol{\xi}|\mathbf{Y}) \propto p(\mathbf{Y}|\boldsymbol{\xi})p(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ stands for the vector of all parameters in the model and where $\mathbf{Y}$ stands for the data.[9] The prior distribution incorporates prior knowledge about the parameters, and if there is none, it is best as vague or diffuse as possible. The more data one acquires, the less influential the prior becomes on the posterior. Because the presented model yields a high-dimensional posterior (due to the large number of parameters), we will opt for Markov chain Monte Carlo (MCMC) methods to draw values from the posterior. Practically speaking, these algorithms perform iterative sampling: Values are drawn from approximate distributions, and they are improved in each step, in such a way that they converge to the targeted posterior distribution. After a sufficiently large number of iterations, one obtains a Markov chain with the posterior distribution as its equilibrium distribution, and the generated samples can be considered as draws from the posterior distribution (it is said that the Markov chain has converged to its equilibrium distribution). More details about the Bayesian methodology and MCMC can be found in Gelman, Carlin, Stern, and Rubin (2004) and Robert and Casella (2004).

For our model we have implemented a specific MCMC algorithm, the Gibbs sampler. In this algorithm, alternating conditional sampling is performed: The parameter vector is divided into subparts (a single element or a vector), and in each iteration the algorithm draws a new sample from the conditional distribution of each subpart given all the other parameters and data; these conditional distributions are the so-called full conditionals. More details on the sampling algorithm, as well the derived full conditionals of each parameter and simulation studies testing the accuracy of the algorithm, can be found in the supplemental materials.

In the estimation algorithm, several of these Markov chains are initiated from different starting values, thereby offering a way to check for convergence of the algorithm (because one has only draws from the posterior after the Markov chain has converged to its equilibrium distribution). The particular convergence check statistic that is used is the Gelman–Rubin $\hat{R}$ statistic (for more information, see Gelman et al., 2004).

Sampling from the posterior distribution was performed with a custom-written MATLAB program.[10] To decrease the computation time, we translated some of the more computationally demanding subroutines into C and applied parallel computing where possible (different sample chains are independent and can be computed on separate processors). The computation time for the example shown in the next section (six chains of 10,000 iterations each) was around 90 min on a computing node with an AMD Opteron 250 processor and 2 GB RAM.

[9] The normalization constant, $p(\mathbf{Y})$, does not depend on the parameter and is therefore not considered.

[10] The MATLAB codes are available on request from the first or second author.

An Experience Sampling Study of the Core Affect

In this section, the HOU is applied to longitudinal core affect measurements. In the present study (for more information on the design of the study, see Kuppens et al., 2007), 80 students from the University of Leuven were paid to provide self-reports about their position in the core affect space over 1 week. Such a study is called an experience sampling study (Csikszentmihalyi & Larson, 1987; Larson & Csikszentmihalyi, 1983). The average age of the participants was 22 years ($SD = 5$), and 60% of them were women. In practice, participants received special booklets containing the so-called affect grid (Russell, Weiss, & Mendelssohn, 1989). The participants carried a preprogrammed wristwatch that beeped nine times a day (at semirandom moments), and upon beeping, they were supposed to indicate their emotional position in the grid. Ideally, each participant completed 63 core affect assessments during the study.

The time difference between two measurements was semirandom. In an introductory session, the participants provided information about their daily routine, more specifically about the time they wake up and go to sleep. Their awake time was divided into equal intervals, and a random beep was scheduled into each of them. As a result of this procedure, the wristwatch did not beep while the participants were sleeping. Occasionally, a participant failed to notice the beeping wristwatch. The most frequent reason for missing a beep was that they simply did not hear the signal (e.g., they were taking a shower without wearing the wristwatch). We assume that the core affect is not an influencing factor for skipping a measurement, and thus the missing data mechanism is assumed to be ignorable (Little & Rubin, 2002), and such an occasion was treated as if there had not been an observation at that particular time. The missingness will create an unbalanced data structure (not all participants have an equal number of measurements), but this does not present special problems for the HOU model. On average we acquired 60 measurements ($SD = 3$) per person.

Because the HOU parameters can be regressed on covariates, interindividual differences can be analyzed and related to stable traits. In this study the five dimensions of the five-factor model of personality (Big Five) were measured (with the Dutch version NEO Five-Factor Inventory; see Hoekstra, Ormel, & De Fruyt, 1996). The NEO Five-Factor Inventory consists of 60 items divided equally into five scales assessing Neuroticism, Extraversion, Openness (to experience), Agreeableness, and Conscientiousness. All items are rated on a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). The items are then summarized into five averaged scores per person, corresponding to the five dimensions.

The data were subjected to an exploratory analysis; some results of this analysis are presented here. For instance, the observed trajectories of two randomly sampled individuals are shown in Figure 5. The trajectories are obtained by connecting subsequent measurements. From the figure, it can be seen that the average position is different for the two persons. Moreover, another observation is that there are different levels of intraindividual variation in core affect. Such observations suggest that there is substantial interindividual variability that can be captured, both in location and in intraindividual variability. Combining the data from all participants in a heat map or three-dimensional histogram over the core affect grid (graph not shown because of space constraints) reveals that the most frequently visited area ranges from the central part of the core affect grid (neutral point) to the center of the upper right quadrant (corresponding to higher activation and pleasantness values).

In Figure 6, an estimated velocity plot is drawn of the core affect grid based on the data of all the participants: The thick black lines are the two-dimensional escape velocity vectors for the corresponding cell (the gray lines are explained later). The length of a vector is proportional to the speed of escape from the cell, and the direction indicates the area toward which one moves (the beginning of the vector is always the middle of the square). One can observe that at the average position in the grid, the average speed is very small (i.e., short vectors), but the farther we move from the average position, the higher the escape speed becomes. Most vectors point more or less directly to the central location, and with increasing distance from the central point, the vector length increases. This observation corresponds to the implied OU model assumption about the centralizing tendency, which increases as the distance from the home base extends. However, at the border cells, we notice more irregularities. These are due to sampling variability because there are much fewer measurements in these outer cells.

In the next two sections, the fitted model is discussed. First, we take up the model fit issue: Several models will be compared by means of the deviance information criterion (DIC), and the best model will undergo a series of statistical tests to evaluate how well it fits the data. In the next part, we discuss the interpretation of an empty model and the best fitting model. The first model is an empty model, containing only the time-varying home base effect but no predictors on the level of the individual differences. As can be judged from the results below, the time effect on the home base needs to be taken into account. The results from the empty model can then be used to interpret the size of the individual differences for the different parameters. In the second, and best fitting, model, the individual difference covariates will be regressed onto a set of person-specific covariates (i.e., the Big Five).

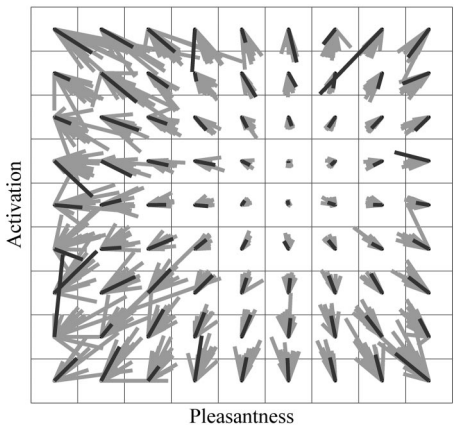


Figure 6. Estimated vector field of the core affect grid. It shows the angle and speed of leaving core affect grid positions. The black lines are calculated based on the observed data, and the gray lines are calculated based on simulated data.

Application of the HOU Model to the Experience Sampling Study: Model Fit

Computational Aspects

A series of different HOU models were estimated for the experience sampling data with a Bayesian procedure. Estimation was carried out by sampling six chains of 10,000 values from the posterior distribution via the Gibbs sampler. The six Markov chains started out from sufficiently different values: These initial values of the chains are strongly randomly perturbed values derived from the data (e.g., the sample average for the home bases). The first 5,000 iterations were discarded (the so-called burn-in period), to avoid any residual influence of the random starting values on our chain. Inference will be based on a total of 30,000 draws. In the left plot of Figure 7, the first 300 iterations of all six chains for the parameter α_{γ_1} in a model without person-specific covariates are shown (in this model α_{γ_1} is then the population mean of the intraindividual log variance parameter γ_{1p} of pleasantness). It can be seen that the chains start from fairly different values, but they all move quickly to the same region, which is an indication of convergence. It should also be noted that this happens already within the first 300 iterations. The right plot displays the smoothed estimated posterior density of the same parameter.

As mentioned above, an appropriate way to check convergence is to calculate the $\hat{R}$ value (Gelman et al., 2004), which roughly equals the ratio of the between- and within-chain variances (from the left plot in Figure 7, it can be seen that after iteration 200, the between- and within-chain variances are approximately equal). As a rule of thumb, the chains are considered converged if the $\hat{R}$ value is below 1.1. We use $\hat{R}$ to assess convergence but also checked the chains visually, and there were no problems (all $\hat{R}$s < 1.1). Convergence was fast for all parameters, typically within the first 500 iterations.

To test goodness of fit, we did not calculate traditional indices like R^2 measures. The reason is that it is not fully clear how such measures should be calculated even in the case of simple linear multilevel models (for different alternatives, see, e.g., Gelman & Hill, 2007; Snijders & Bosker, 1999). Because the HOU model is a hierarchical nonlinear model, it is far more complicated than the simple multilevel cases, so that finding an appropriate R^2 type of measure seems infeasible. Moreover, we do not regard it a good indication of model fit, since sometimes high R^2 values can be associated with relatively poorly fitting models (Ramsey & Schafer, 2002). Our chosen strategies for testing model fit incorporate relative goodness-of-fit testing and graphical comparison of the observed data and replicated data sets based on the model parameters.

Testing Model Fit With the DIC (Relative Goodness of Fit)

To compare different models, we will make use of the DIC statistic (Spiegelhalter, Best, Carlin, & van der Linde, 2002). As its frequentist counterparts (i.e., Akaike information criterion, Bayesian information criterion), it simultaneously takes into account two important features of the model: the complexity (based on the number of parameters) and the fit (typically measured by a deviance statistic). The DIC formula is the sum of the effective number of parameters and the posterior mean of the deviance (defined as -2 times log-likelihood). Theoretically, the model with smaller DIC would better predict a replicate data set of the same structure.

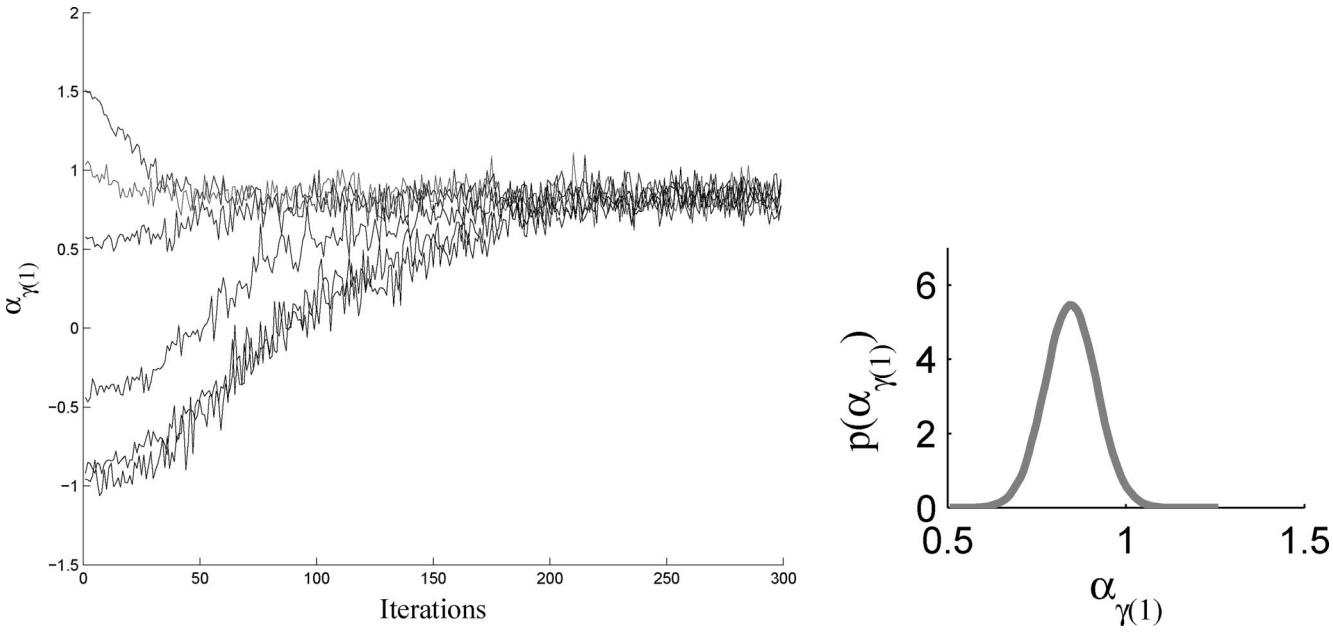


Figure 7. Illustration of the sampling of the parameters. The left plot shows the posterior sampling of the population mean of the variability in the pleasantness dimension, and the right plot displays its estimated posterior density.

For model selection purposes, we have constructed four alternative models, which all have the time-varying home base included. The first model is an empty model with all parameters differing across individuals but no person-specific covariates. In the second model, the covariate information is introduced. The third model is a variant of the first but without individual differences in the centralizing tendency (i.e., an equal **B** matrix for all persons). In the fourth model, the cross-effects for the variance and centralizing tendency were set equal to 0 (i.e., $\rho_{\gamma_p} = 0$ and $\rho_{\beta_p} = 0$ for all persons). The home bases and stationary variances were allowed to differ among persons in all models and were not altered, because the exploratory data analysis suggested that these parameters tend to differ substantially between individuals.

The results are displayed in Table 1 (a lower DIC value suggests a better fit). The best fitting model has all parameters person specific and covariates included. As one can see, the model without person-specific centralizing force shows a relatively poor fit (compared with the empty model). The decrease in the fit index is less dramatic when the person-specific cross-effects are taken out, but nevertheless it goes down (compared with the first model and suggesting that the two dimensions are connected).

Testing Model Fit With Graphical Posterior Predictive Checks

The idea behind posterior predictive checks is that if the model fits, then replicated data generated under the model should look similar to observed data (Gelman et al., 2004). On the basis of the estimated model parameters, we can replicate observations and see how well they resemble the original data. With graphical model checking, we display the data alongside simulated data. If they look similar, we can say that the observed data look plausible under the posterior predictive distribution. Systematic discrepancies indicate poor model fit. In what follows, we simulated new data sets from the empty model (instead of the model with covariates) because that gives a simpler simulation and test procedure.

Comparing observed and replicated trajectories. First, let us look at the data from individual persons. Plotting a specific person's data in the two-dimensional core affect space and connecting subsequent points with a line results in an observed person-specific core affect trajectory. Based on the fitted HOU model, such trajectories can also be simulated from the model (keeping the same time differences as in the observed data). If the model fits the data, the replicated trajectories should closely resemble the observed trajectories with respect to the spatial characteristics.

Table 1
Deviance Information Criterion (DIC) Values for the Fitted Models

Number	Model type	DIC
1	Fully person-specific model without covariates	−6524
2	Fully person-specific model with covariates	−6834
3	Model without person-specific centralizing tendency	−1639
4	Model without cross-effects	−5940

In Figure 8, we plotted the observed data of six students in the first column. In the following four columns, generated data are displayed based on the estimated person-specific parameters and population values (for the measurement error and the time-varying coefficients). It is important to stress that the replicated trajectories cannot follow exactly the same path as the observed trajectory because the HOU model is inherently stochastic. However, it is important that the key characteristics of the trajectories that are captured by the model parameters are similar. For demonstrational purposes, we selected the six persons according to their HOU parameter values. With the first two people (first two rows), the differences in the home bases are demonstrated: The first person is below population average with respect to person-specific home base values (4.65 and 4.95), where the second one is above (6.82 and 6.19). The third and the fourth rows correspond to individuals with low (0.81 and 1.11) and high (5.64 and 5.58) intraindividual variance, whereas the fifth and the sixth rows display participants with low (0.0063 and 0.0103) and high (0.0333 and 0.0199) centralizing tendencies. The graphs in Figure 8 suggest that the important characteristics of the observed trajectories are preserved very well in the replicated ones.

Comparing observed and simulated escape velocities for the core affect grid. Figure 6 displays the escape velocities from the core affect grid positions with length proportional to the speed of escape from the cell and direction indicating to which area it tends toward. We can also calculate such velocity vectors based on simulated data. In Figure 6, the black lines correspond to the vectors calculated from the observed data, and the gray lines correspond to those of the simulated data.

It can be seen that for the majority of the cells, the observed velocity vectors (in black) fall nicely within the range of velocity vectors predicted by the model (in gray). A few exceptions occur, mostly on the left side (low pleasantness), where there is somewhat more deviation between the observed and the replicated data. This is probably due to the fact that there are only a couple of observations with very low pleasantness values, and the observed data vectors could therefore be calculated based only on a handful of data points.

Application of the HOU Model to the Experience Sampling Study: Interpretation of Two Models

In this section, we discuss more in detail the two best fitting models from the previous section. Both models allow all OU parameters to differ across individuals and include a time-varying home base; the difference lies in the presence of time-invariant (i.e., person-specific) covariates.

Model 1: Empty Model With a Time-Varying Home Base but No Time-Invariant Predictors

The first model we investigate is an empty (or unconditional) model that contains a quadratic time effect for the home base but no time-invariant individual difference covariates. In the right plot of Figure 7, an example is given of a smoothed histogram of the samples of α_{γ_1} from its posterior. Like this one, most parameters have marginal posteriors that are very close to normal, except for the posteriors of the two measurement error variances, which are a little bit more

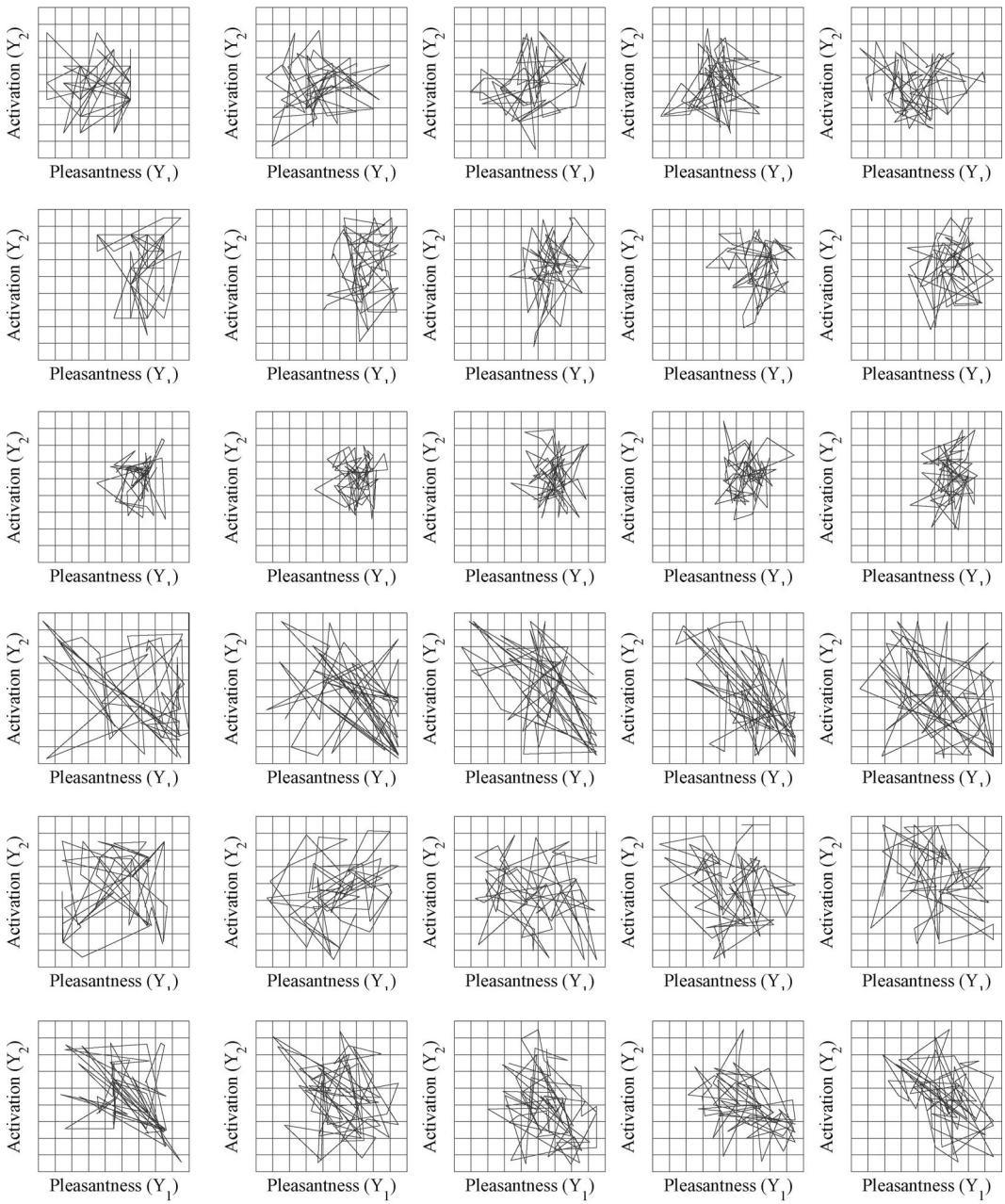


Figure 8. Observed and simulated person-specific data trajectories. Each row corresponds to a different individual. The first column depicts the observed data for these individuals, the following four columns display simulated trajectories based on Ornstein–Uhlenbeck estimates.

skewed to the right. The estimated posterior means, posterior standard deviations, and associated credibility intervals for all population parameters are shown in Table 2. In the following three subsections we interpret the most interesting findings from this model.

The time-varying home base. Both home bases vary as a quadratic function of time, where time is measured in hours and the scores were centered around noon, so that $z_{ps} = 0$ is at 12 noon for all persons. This means that the person-specific intercepts (contained in the vector $\boldsymbol{\mu}_p$) can be considered as the expected pleasantness and activation scores for the different individuals in the middle of the day (i.e., at noon). Moreover, the parameters $\alpha_{\mu 1}$ and $\alpha_{\mu 2}$ are the population means of these intercepts (because $\mathbf{x}_p = 1$). We can summarize the findings about the home base in the core affect space by depicting how it changes with time for each individual separately and compare it to an averaged trajectory. In the case of the pleasantness dimension, as we can see in the left plot of Figure 9, there is a steady increase during the day that is almost linear. We can observe that during the course of the day, the students start feeling more and more pleasant. Because the quadratic time effect for pleasantness is negligible, we may say that with every 2 hr the average pleasantness increases with

Table 2
Summary of the Results Estimated With the Hierarchical Ornstein–Uhlenbeck Model Without Covariates

Model parameter	Description	Posterior mean	95% PCI		Posterior *SD*
			LL	*UL*	
	Pleasantness				
α_{μ_1}	Average home base	5.7552	5.5921	5.9160	0.0821
$\sigma^2_{\mu_1}$	Variance of the average home base	0.3928	0.2606	0.5733	0.0802
$\delta_{L\mu_1}$	Linear time effect	0.0715	0.0426	0.1004	0.0148
$\delta_{Q\mu_1}$	Quadratic time effect	−0.0012	−0.0048	0.0023	0.0018
$s(\alpha_{\gamma_1})$	Average intraindividual variability	2.7885	2.3995	3.2700	0.2213
$s(\sigma^2_{\gamma_1})$	Variance of the intraindividual variability	3.3927	1.7887	6.2642	1.1847
$s(\alpha_{\beta_1})$	Average centralizing tendency	0.0187	0.0149	0.0242	0.0024
$s(\sigma^2_{\beta_1})$	Variance of the centralizing tendency	0.0002	0.0000	0.0006	0.0001
$\sigma^2_{1\varepsilon}$	Measurement error	0.1291	0.0817	0.1903	0.0277
	Activation				
α_{μ_2}	Average home base	5.1761	5.0209	5.3316	0.0790
$\sigma^2_{\mu_2}$	Variance of the average home base	0.3388	0.2236	0.4968	0.0700
$\delta_{L\mu_2}$	Linear time effect	0.2936	0.2606	0.3273	0.0170
$\delta_{Q\mu_2}$	Quadratic time effect	−0.0329	−0.0369	−0.0288	0.0021
$s(\alpha_{\gamma_2})$	Average intraindividual variability	3.2714	2.8796	3.7370	0.2174
$s(\sigma^2_{\gamma_2})$	Variance of the intraindividual variability	3.1039	1.7271	5.5096	0.9782
$s(\alpha_{\beta_2})$	Average centralizing tendency	0.0209	0.0165	0.0275	0.0028
$s(\sigma^2_{\beta_2})$	Variance of the centralizing tendency	0.0003	0.0001	0.0008	0.0002
$\sigma^2_{2\varepsilon}$	Measurement error	0.1069	0.0633	0.1678	0.0267
	Cross-effects				
$\sigma_{\mu_1\mu_2}$	Covariance between the home bases	0.0568	−0.0447	0.1680	0.0534
α_{ρ_γ}	Average Fisher *z*-transformed cross-correlation	0.0227	−0.0460	0.0912	0.0349
α_{ρ_β}	Average Fisher *z*-transformed off-diagonal of **B**	−0.0494	−0.1329	0.0357	0.0431

Note. The $s(.)$ notation stands for a scale transformation for that model parameter. The reason for this notation is that the population distributions for γ and β are normal on the log scale. However, in this table we transformed these values back to the original scale of γ and β, but because no specific notation was introduced for the corresponding population parameters on the original scale, we simply indicate them by adding an $s(.)$ operator to the log-scale notation. PCI = posterior credibility interval; *LL* = lower limit; *UL* = upper limit.

0.14 units. There is substantial interindividual variability with respect to the intercept of pleasantness: The estimated population standard deviation equals $\sqrt{0.3928} \approx 0.63$, which is considerably larger than the average increase. On the whole, we can conclude that a participant's mood is pleasant and that it increases slightly during the day in a linear way.

Unlike the home base of pleasantness, the home base of activation clearly shows a quadratic evolution during the day, as we can

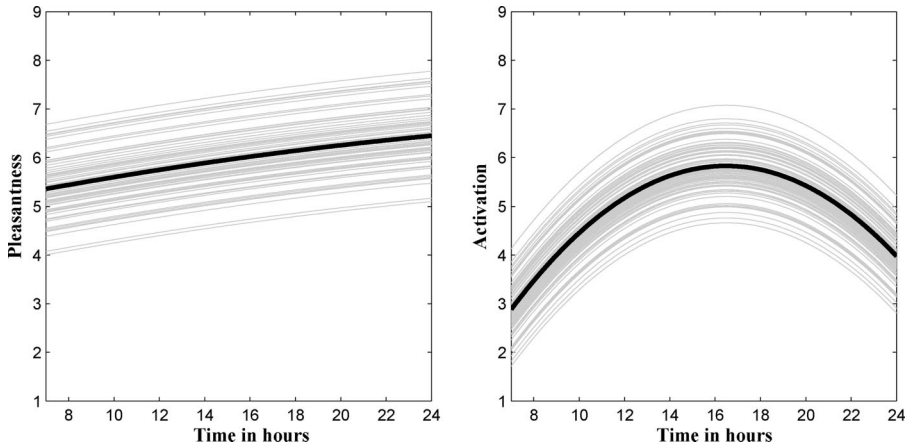


Figure 9. Home base of the pleasantness (left plot) and the activation dimensions (right plot) evolving with time. The thick line depicts the averaged trajectory over persons, and the thin lines correspond to the person-specific values.

see in the right plot of Figure 9. This pattern is not surprising: In the literature on diurnal variation, it is often found that daily variation in the self-reported arousal level shows an inverted U-shaped function (for a summary on these findings, see Caminada & De Bruijn, 1992). In our student population, the average arousal home base starts low in the morning: At 7 a.m., for example, it is on average 2.97. In the next 5 hr, the activation level increases by more than 2 points, so that at noon the intercept is $\alpha_{\mu_2} = 5.17$. The activation home base reaches its apex around 4:50 p.m., and afterward it decreases when evening comes. It may appear that this daily maximum falls late during the day, but we have to keep in mind that the examined population consists of college-aged students (who tend to be awake later than people in the general population). Also, the individual differences in the intercept of the time-varying activation component of home base are shown. Based on the spread of the person-specific lines around the population mean, it appears that the interindividual variability is substantial here as well: The estimated population standard deviation equals $\sqrt{0.3388} \approx 0.58$. The source of the individual variation with respect to the person-specific intercepts of both dimensions may be explored by time-invariant covariates, as we see later when estimating the second model.

To get a better idea of the individual differences in the home base intercept, we will use some simulated trajectories. In Figure 10, several simulated OU processes based on the estimated OU population values from Table 2 are displayed (but without the structural time-varying home base, which we left out when simulating the trajectories). Figures 10A–10C contain three simulated trajectories in which the home base is varied. In Figure 10B, a trajectory was simulated with the mean home base in the population. In Figure 10A, one population standard deviation was subtracted from each of the two components of the mean home base vector (i.e., 0.63 and 0.58 for pleasantness and activation, respectively). In Figure 10C, the above-mentioned two population standard deviations were added to the two components of the mean home base vector. It can be seen that this variation in the home base has a profound effect on the simulated curves: The majority of the visits in Figure 10C are concentrated in a slightly more pleasant and activated area (compared with Figure 10B), whereas in Figure 10A the OU process tends to be drawn more into a slightly unpleasant and deactivated area.

Stationary variance matrix and measurement error variance. Looking at Table 2, we can see that overall there is substantial intraindividual variation in the core affect space and that it is slightly larger with respect to arousal (i.e., 2.79) than with respect to valence (i.e., 3.27). If we compare the intraindividual variabilities of the two dimensions with the measurement error variances, it turns out that the latter are relatively small in both dimensions: $\sigma^2_{1\varepsilon} = 0.13$ and $\sigma^2_{2\varepsilon} = 0.11$, suggesting that the variability in the data is mainly due to the intraindividual variations of the OU process.

As emphasized before, an important aspect of the model is that it allows for interindividual variability in the intraindividual variation. The existence of such individual differences was already visible during data exploration in a simple plot of the core affect measurements of two participants (see Figure 5). Now from Table 2, we can see that in terms of HOU model parameters, in both dimensions the population variance of the intraindividual variation is quite large (i.e., 3.39 and 3.10 for pleasantness and activation, respectively). The size of these individual differences was illustrated again by means of simulation. In Figures 10C–10F, we graphically display three simulated OU processes in which the intraindividual variability is systematically changed. Figure 10E is a trajectory simulated with the means of the population distributions of the parameters; in Figures 10D and 10F, a standard deviation was subtracted from (added to) the intraindividual log variances (we performed the calculations on the scale in which the statistical inference is done). The figures suggest that individual differences are indeed quite substantial. A further step after discovering such variation is to try to tie the person-specific intraindividual variabilities onto predictors.

As part of the intraindividual stationary variance matrix $\mathbf{\Gamma}_p$, we also estimate the cross-correlation between the measurements. On average, this correlation between the changes in the two dimensions is not substantial: α_{ρ_γ} equals 0.0227, on the Fisher z-transformed scale, which corresponds roughly to the same value on the normal scale, as the transformation is close to linear around zero. However, with the person-specific correlation values, there is considerable variability among them: Although their mean value is almost zero, the population standard deviation is 0.27 (not shown in Table 2 because of space limitations), which suggests sizable interindividual differences. In the next subsection, we also look into the explanation of this individual variability with the second model.

The centralizing tendency. Taking into account individual differences in the centralizing tendency matrix $\mathbf{B}_p$ is one of the most interesting parts of the model. First, let us look at the population means and variances of the diagonal elements of $\mathbf{B}_p$ (i.e., β_{1p} and β_{2p}). There does not seem to be any substantial difference in these values between the two dimensions. Both population variances again show the existence of some individual variability, which we displayed again in the same manner as before in Figures 10G–10I. The construction of Figures 10G–10I is similar as before: Figure 10G shows an OU process with one population standard deviation subtracted from the mean centralizing tendencies; Figure 10H is simulated with the population means; in Figure 10I, the trajectory is based on the mean centralizing tendencies plus one population standard deviation. Comparing Figures 10G–10I, we see that in Figure 10G, with the lower centralizing tendency, when the process moves away from the home base, the return is not fast, whereas in Figure 10I, the visits are more concentrated around the population home base, leading to a density in the home base area. In core affect terms, this means that when our mood changes in terms of pleasantness or activation, with low centralizing tendency we will be staying longer in that state, and the adjustment to our comfort zone is slower. As before, finding connections between the different levels of regulation and stable personality traits might lead to interesting new discoveries; an attempt for that will be made by adding covariates to the model in the next subsection.

With the third element of the centralizing tendency matrix $\mathbf{B}_p$, the cross-centralizing tendency $\rho_{p\beta}$, it turns out that the population mean of the off-diagonal element $\alpha_{\rho_{p\beta}}$ equals -0.05, and applying the inverse Fisher z transformation yields roughly the same value. This small off-diagonal value suggests that there is on average no strong dependence between the two dimensions in the adjustment. Described in terms of orbital portraits, the current population would have a very similar picture for its expected dynamics around

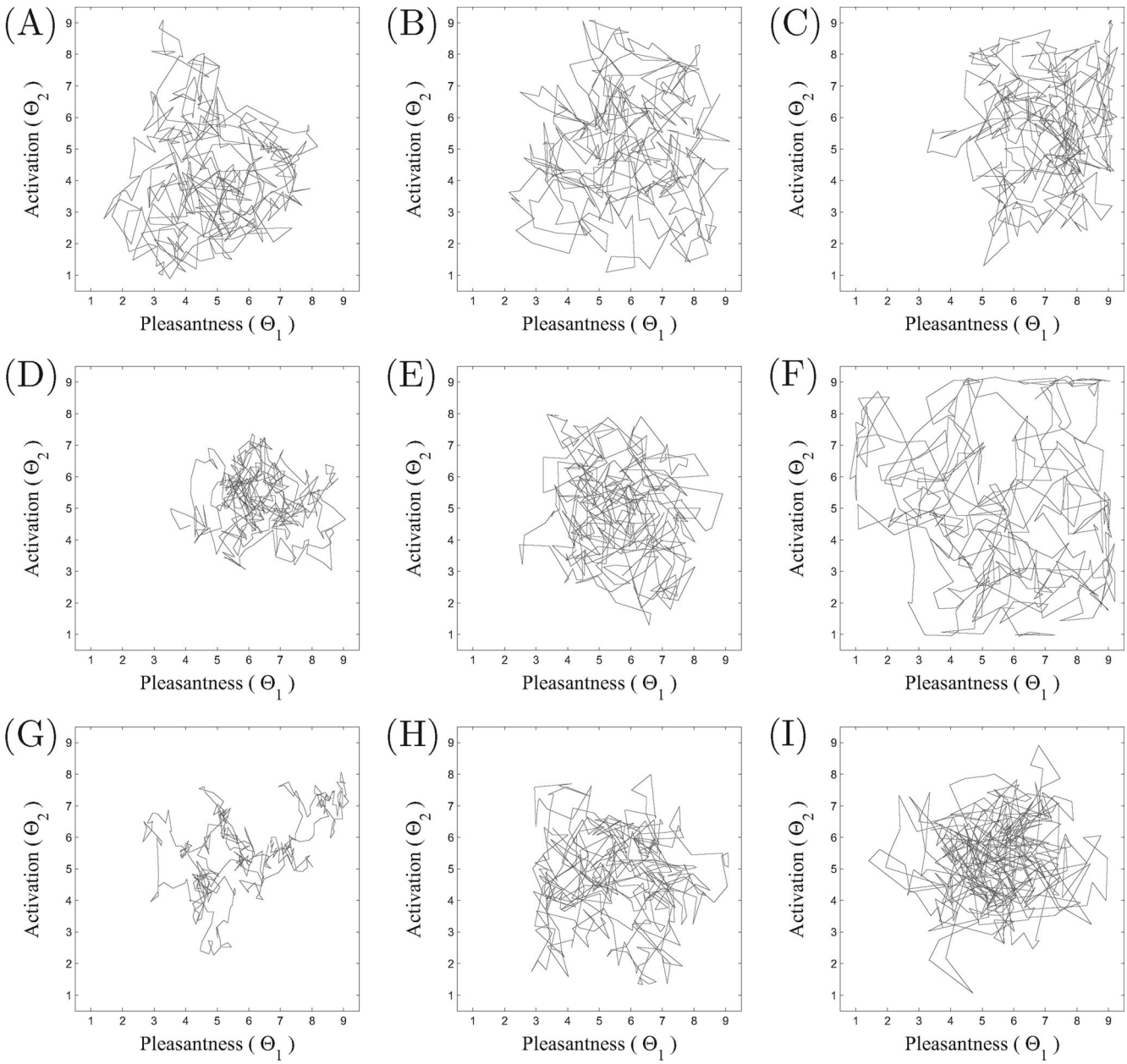


Figure 10. Simulated Ornstein–Uhlenbeck processes based on the estimated population values. B, E, and H show processes based on the estimated OU parameters in Table 2. The home bases (A–C), intraindividual variances (D–F), and centralizing forces (G–I) are manipulated by first subtracting (first column) and then adding (third column) one population standard deviation to these parameters in both dimensions.

the home base as it is shown in Figure 4G. Also, there is some degree of individual difference in the cross-centralizing tendency (population standard deviation equals 0.21, not displayed in Table 2 because of space limitations).

As mentioned before, we can also look at the centralizing tendency from another perspective, as is graphically done in Figure 11: The centralizing tendency values can be converted into autocorrelation functions (see, e.g., Oravecz, Tuerlinckx, & Vandekerckhove, 2009). The reason for this relationship can be understood intuitively: If the adjustment to the average level in the process is large, the autocorrelation function will decrease quickly (because the next value does not influence the current one substantially); whereas if the adjustment is small, there will be a strong relation between the current and next observation (implying a high autocorrelation). In Figure 11, the person-specific autocorrelation functions are shown together with the autocorrelation functions for the two dimensions based on the average β value. We can see a high degree of individual variability here. Although the autocorrelation

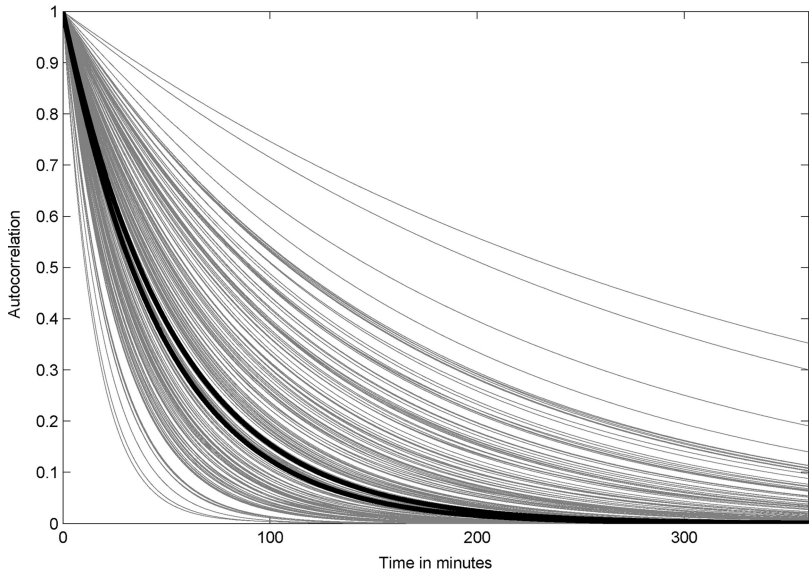


Figure 11. The autocorrelation based on the estimated population means for the centralizing tendencies of the pleasantness and activation (black lines) with all person-specific autocorrelation functions (gray lines). The higher black line corresponds to pleasantness, and the lower corresponds to activation.

is not substantial after 2 hr for many students, some of them do tend to stay in the same core affect region for a longer period. With the help of the covariates, we can see whether these people share some other features as well. This question is investigated in the next subsection.

Model 2: Predicting Core Affect Characteristics From Person-Specific Covariates

In the second estimated model, all the person-specific parameters (i.e., the home bases, intraindividual variances, the cross-correlation, the centralizing tendencies, and the cross-centralizing tendency) were regressed on person-specific predictors. In the presented experience sampling study, these predictors were the five personality dimensions of the Big Five. We discuss only the regression coefficients for which the 95% posterior credibility interval did not contain 0. Table 3 summarizes these coefficients.

With respect to the home base, we have the rather unsurprising finding that neurotic individuals tend to have a lower level of pleasantness. A similar finding can be found in Kuppens et al. (2007), although they use the raw sample average of the pleasantness ratings as outcome (and not a model-based parameter). From Table 3, it can also be deduced that students who score high on neuroticism tend to show higher variability with respect to their pleasantness level—a finding that is consistent with the research of Kuppens et al.

A more surprising result is that neuroticism is positively associated with a higher degree of association between the two dimensions. On the other side, conscientiousness is inversely related to the cross-correlation, which is also new. This means that for highly conscientious people, when their mood becomes more pleasant, their arousal levels drop, and vice versa. For neurotic individuals with a positive cross-correlation coefficient, higher activation levels are associated with more pleasant feelings. These findings are,

Table 3
Summary of the Regression Coefficients With a 95% Posterior Credibility Interval Not Containing Zero

Model parameter	Description	Covariate	Posterior mean	95% PCI		Posterior *SD*
				LL	*UL*	
			Pleasantness			
$\alpha_{\mu_1 N}$	Home base	Neuroticism	−0.32	−0.58	−0.07	0.13
$\alpha_{\gamma_1 N}$	Variability	Neuroticism	0.26	0.01	0.51	0.12
			Cross-effects			
$\alpha_{\rho_\gamma N}$	Cross-correlation	Neuroticism	0.13	0.01	0.25	0.06
$\alpha_{\rho_\gamma C}$	Cross-correlation	Conscientiousness	−0.18	−0.30	−0.05	0.06
$\alpha_{\rho_\beta A}$	Off-diagonal of **B**	Agreeableness	−0.25	−0.45	−0.07	0.09

Note. Model parameters refer to the regression weights. For example $\alpha_{\mu_1 N}$ is the regression weight for neuroticism relating to the home base in the pleasantness dimension (μ_1).

to the best of our knowledge, new and have not yet been described in the literature.

The current study did not find any remarkable connection between the Big Five personality dimensions and the centralizing force in the core affect space. However, the results did show an effect on the off-diagonal element of the centralizing tendency matrix: Agreeable persons seem to have a lower off-diagonal value. This might mean that when their regulation in one dimension increases, the level of centralizing tendency in the other one decreases. Also, the orbital portrait of agreeable individuals would be more like the improper node as in Figure 4I.

Conclusion

In this article, we have introduced a hierarchical model for analyzing change in longitudinal variables. The model is based on the stochastic OU process, which can represent latent states that change over time. Although the model comprises many attractive features, it has not been introduced in psychology so far for modeling affective dynamics. The most important assets of the model are the following. First, the dynamics of the OU process provide a sound theoretical framework: We can account for the observed changes with an underlying dynamical concept. The OU process-based parameterization offers a reasonable description of the dynamics of change and is especially fit for modeling core affect variation. By conceptualizing the parameters of the OU process as random effects, we are able to account for interindividual differences. Furthermore, we may attempt to explain this variability by introducing covariate information. Because we model two longitudinal variables simultaneously, we can also investigate cross-effects. Finally, the time-varying nature of the home base parameter allows us to readily include explanatory variables that are functions of time.

Our approach has many links to other types of models. For example, because we incorporate both structural and random effects, there is a clear connection with mixed or multilevel models (e.g., Diggle, Heagerty, Liang, & Zeger, 2002; Verbeke & Molenberghs, 2000). But several things set the HOU model apart from the traditional mixed models. First, we do not focus exclusively or even primarily on the mean structure but rather on the dynamical aspects of the model. The former is usually the point of attention in mixed modeling. Second, our model is derived from a stochastic process, and it has substantive roots in emotion theory. Because of the specific assumptions from which it is derived, its applicability is not as general as for a linear mixed model (see below). Third, all parameters of the model, including the variance and (auto)correlation parameters, are allowed to vary randomly over persons, whereas traditional mixed models typically allow individual differences only in the mean structure. As a consequence, individual differences at different locations in the model can be investigated with the HOU model.

The HOU model also shares some similarities with structural equation models (Bollen, 1989) because of the presence of a measurement (or observation) model and a structural (or transition) model. However, the HOU model is based on a continuous-time stochastic process that cannot as such be represented in a structural equation model. Moreover, our type of data can be highly unbalanced and very unequally spaced. In addition, there may be many more measurements for each person than there are persons. Such situations are typically hard to handle for structural equation models, but the HOU model does not have any problem with them. Besides these constraints, it is not conventional in structural equation models to allow all driving parameters to vary randomly.

SDE models have been used before in different areas in the behavioral sciences. Oud (2007) and Singer (2007) demonstrated the use of a stochastic second-order differential equation to model oscillatory patterns in the data. Compared with their approach, ours assumes a more simple first-order SDE on the latent level. However, we allow for individual differences in all aspects of the model, also in the dynamical part, which is rather exceptional. Moreover, our model is fitted with methods from the Bayesian framework, whereas Oud and Singer use classical techniques.

The HOU model can be applied to other areas beyond emotion psychology, but it is not a data-analytical panacea. First, the theoretical assumptions about the modeled psychological construct should be in line with the specific assumptions of the stochastic model. For instance, it should be reasonable to assume the existence of a centralizing tendency or a regulatory mechanism such that the process reverts to the mean. Such an assumption may not be realistic for various learning or developmental processes. In addition, there is a considerable computational cost that comes with fitting the model, because we make use of computationally intensive MCMC techniques. Therefore, if one is interested mainly in individual differences in the mean structure, traditional methods (mentioned above) should be considered first.

Concerning the issue of regulation, its parameter $\mathbf{B}$ also has some constraints. Because our focus was also on modeling interindividual differences, we sacrificed the asymmetry property of $\mathbf{B}$ to be able to allow for interindividual differences. However, in other areas, like autoregressive cross-lagged panel designs (Oud & Delsing, 2010), an asymmetric $\mathbf{B}$ is estimated. However, in such approaches, modeling interindividual differences in $\boldsymbol{\Gamma}$ or $\mathbf{B}$ is generally not considered.

Overall, we find that the OU diffusion process is an intuitively appealing way of describing the continuous change of certain phenomena over time, certainly for constructs related to emotion and mood. Future challenges that may lead to new model extensions may consist of measuring important impacts on the modeled processes and possibly discovering physiological connections. The HOU model allows such information to be entered into the model in a time-dependent fashion as well, which makes it especially useful when the emphasis lies on the dynamical aspects of the psychological constructs.

References

Arnold, L. (1974). *Stochastic differential equations: Theory and applications.* New York, NY: Wiley.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58,* 373–403. doi:10.1146/annurev.psych.58.110405.085709

Boker, S. M. (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. *Multivariate Behavioral Research, 37,* 405–422. doi:10.1207/S15327906MBR3703_5

Boker, S. M., & Laurenceau, J.-P. (2006). Dynamical systems modeling: An application to the regulation of intimacy and disclosure in marriage.

In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 195–218). New York, NY: Oxford University Press.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley.

Brown, R. (1828). A brief account of microscopical observations made in the months of June, July, and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philosophical Magazine, 4,* 161–173.

Caminada, H., & De Bruijn, F. (1992). Diurnal variation, morningness–eveningness, and momentary affect. *European Journal of Personality, 6,* 43–69. doi:10.1002/per.2410060105

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97,* 19–35. doi:10.1037/0033-295X.97.1.19

Chow, S.-M., Ram, N., Boker, S. M., Fujita, F., & Clore, G. (2005). Emotion as a thermostat: Representing emotion regulation using a damped oscillator model. *Emotion, 5,* 208–225. doi:10.1037/1528-3542.5.2.208

Cox, D. R., & Miller, H. D. (1972). *The theory of stochastic processes.* London, England: Chapman & Hall.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Disease, 175,* 526–536. doi:10.1097/00005053-198709000-00004

Davidson, R. J. (2003). Darwin and the neural bases of emotion and affective style. In P. Ekman, J. J. Campos, R. J. Davidson, & F. B. M de Waal (Eds.), *Annals of the New York Academy of Sciences* (Vol. 1000, pp. 316–336). New York: New York University Press.

Delsing, M. J. M. H., Oud, J. H. L., & De Bruyn, E. E. J. (2005). Assessment of bidirectional influences between family relationships and adolescent problem behavior: Discrete vs. continuous time analysis. *European Journal of Psychological Assessment, 21,* 226–231. doi:10.1027/1015-5759.21.4.226

Denissen, J. J. A., Butalid, L., Penke, L., & van Aken, M. A. G. (2008). The effects of weather on daily mood: A multilevel approach. *Emotion, 8,* 662–667. doi:10.1037/a0013497

Diggle, P. J., Heagerty, P., Liang, K.-Y., & Zeger, S. L. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford, England: Oxford University Press.

Dunn, J. E., & Gipson, P. S. (1977). Analysis of radio telemetry data in studies of home range. *Biometrics, 33,* 85–101. doi:10.2307/2529305

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen [On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat]. *Annalen der Physik, 332,* 549–560.

Fahrmeir, L., & Tutz, G. (2001). *Multivariate statistical modelling based on generalized linear models* (2nd ed.). New York, NY: Springer-Verlag.

Forgas, J. P., & Ciarrochi, J. V. (2002). On managing moods: Evidence for the role of homeostatic cognitive strategies in affect regulation. *Personality and Social Psychology Bulletin, 28,* 336–345. doi:10.1177/0146167202286005

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

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review, 116,* 439–453. doi:10.1037/a0015251

Gardiner, C. W. (2004). *Handbook of stochastic methods: For physics, chemistry and the natural sciences* (3rd ed.). New York, NY: Springer-Verlag.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). New York, NY: Chapman & Hall.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge, England: Cambridge University Press.

Gross, J. J. (Ed.). (2007). *Handbook of emotion regulation.* New York. NY: Guilford Press.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85,* 348–362. doi:10.1037/0022-3514.85.2.348

Hamaker, E. L., Zhang, Z., & van der Maas, H. L. J. (2009). Using threshold autoregressive models to study dyadic interactions. *Psychometrika, 74,* 727–745. doi:10.1007/s11336-009-9113-4

Haug, H. J., & Fähndrich, E. (1990). Diurnal variations in depressed patients in relation to severity of depression. *Journal of Affective Disorders, 19,* 37–41. doi:10.1016/0165-0327(90)90007-U

Hemenover, S. H. (2003). Individual differences in rate of affect change: Studies in affective chronometry. *Journal of Personality and Social Psychology, 85,* 121–131. doi:10.1037/0022-3514.85.1.121

Hoeksma, J. B., Oosterlaan, J., Schipper, E., & Koot, H. (2007). Finding the attractor of anger: Bridging the gap between dynamic concepts and empirical data. *Emotion, 7,* 638–648. doi:10.1037/1528-3542.7.3.638

Hoekstra, H. A., Ormel, J., & De Fruyt, F. (1996). *NEO-PI–R, NEO-FFI Big Five persoonlijkheidsvragenlijsten: Handleiding* [NEO-PI–R, NEO-FFI Big Five personality questionnaire: Manual]. Lisse, the Netherlands: Swets & Zeitlinger.

Jazwinski, A. H. (1970). *Stochastic processes and filtering theory.* New York, NY: Academic Press.

Karlin, S., & Taylor, H. M. (1981). *A second course in stochastic processes.* New York, NY: Academic Press.

Klein, A., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika, 65,* 457–474. doi:10.1007/BF02296338

Klein Entink, R. H., Kuhn, J.-T., Hornke, L. F., & Fox, J.-P. (2009). Evaluating cognitive theory: A joint modeling approach using responses and response times. *Psychological Methods, 14,* 54–75. doi:10.1037/a0014877

Kuppens, P., Stouten, J., & Mesquita, B. (2009). Individual differences in emotion components and dynamics: Introduction to the special issue. *Cognition & Emotion, 23,* 1249–1258. doi:10.1080/02699930902985605

Kuppens, P., Van Mechelen, I., Nezlek, J. B., Dossche, D., & Timmermans, T. (2007). Individual differences in core affect variability and their relationship to personality and adjustment. *Emotion, 7,* 262–274. doi:10.1037/1528-3542.7.2.262

Kuppens, P., Van Mechelen, I., & Rijmen, F. (2008). Toward disentangling sources of individual differences in appraisal and anger. *Journal of Personality, 76,* 969–1000. doi:10.1111/j.1467-6494.2008.00511.x

Larsen, R. J., & Prizmic, Z. (2000). Affect regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 40–61). New York, NY: Guilford Press.

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social and Behavioral Science, 15,* 41–56.

Lawler, G. F. (2006). *Introduction to stochastic processes.* New York, NY: Chapman & Hall/CRC.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic modeling. *Behavioral and Brain Sciences, 28,* 169–245. doi:10.1017/S0140525X0500004X

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York, NY: Wiley.

Mood, A. M., Graybill, F. A., & Boes, D. C. (1974). *Introduction to the theory of statistics.* New York, NY: McGraw-Hill.

Oravecz, Z., Tuerlinckx, F., & Vandekerckhove, J. (2009). A hierarchical Ornstein–Uhlenbeck model for continuous repeated measurement data. *Psychometrika, 74,* 395–418. doi:10.1007/S11336-008-9106-8

Oud, J. H. L. (2007). Comparison of four procedures to estimate the damped linear differential oscillator for panel data. In K. van Montfort,

J. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 19–40). Mahwah, NJ: Erlbaum.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244). Heidelberg, Germany: Springer-Verlag.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika, 65,* 199–215. doi:10.1007/BF02294374

Oud, J. H. L., & Singer, H. (2008). Continuous time modeling of panel data: SEM versus filter techniques. *Statistica Neerlandica, 62,* 4–28. doi:10.1111/j.1467-9574.2007.00376.x

Pan, J.-K., & Fang, K.-T. (2002). *Growth curve models and statistical diagnostics.* New York, NY: Springer.

Ramsey, F. L., & Schafer, D. W. (2002). *The statistical sleuth: A course in methods of data analysis* (2nd ed.). Pacific Grove, CA: Duxbury.

Ringo Ho, M.-H., Shumway, R., & Ombao, H. (2006). The state-space approach to modeling dynamic processes. In T. A. Walls & J. L. Schafer (Eds.), *Models for intensive longitudinal data* (pp. 148–175). New York, NY: Oxford University Press.

Robert, C. P., & Casella, G. (2004). *Monte Carlo statistical methods* (2nd ed.). New York, NY: Springer.

Ross, S. M. (1996). *Stochastic processes* (2nd ed.). New York, NY: Wiley.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16,* 225–237. doi:10.3758/PBR.15.6.1201

Rouder, J. N., Tuerlinckx, F., Speckman, P., Lu, J., & Gomez, P. (2008). A hierarchical approach for fitting curves to response time measurements. *Psychonomic Bulletin & Review, 15,* 1201–1208. doi:10.3758/PBR.15.6.1201

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110,* 145–172. doi:10.1037/0033-295X.110.1.145

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* 805–819. doi:10.1037/0022-3514.76.5.805

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57,* 493–502. doi:10.1037/0022-3514.57.3.493

Rusting, C. L., & Larsen, R. J. (1998). Diurnal patterns of unpleasant mood: Associations with neuroticism, depression, and anxiety. *Journal of Personality, 66,* 85–103. doi:10.1111/1467-6494.00004

Scherer, K. R. (2000). Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes. In M. D. Lewis & I. Granic (Eds.), *Emotion, development, and self-organization: Dynamic systems approaches to emotional development* (pp. 70–99). New York, NY: Cambridge University Press.

Schumacker, R. A., & Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling.* Mahwah, NJ: Erlbaum.

Shoda, Y., LeeTiernan, S., & Mischel, W. (2002). Personality as a dynamical system: Emergence of stability and distinctiveness from intra- and interpersonal interactions. *Personality and Social Psychology Review, 6,* 316–325. doi:10.1207/S15327957PSPR0604_06

Singer, H. (2007). Stochastic differential equation models with sampled data. In K. van Montfort, J. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 73–106). Mahwah, NJ: Erlbaum.

Smith, J. B., & Batchelder, W. H. (2010). Beta-MPT: Multinomial processing tree models for addressing individual differences. *Journal of Mathematical Psychology, 54,* 167–183. doi:10.1016/j.jmp.2009.06.007

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: Sage.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit [With discussion]. *Journal of the Royal Statistical Society: Series B. Statistical Methodology, 64,* 583–616.

Tuma, N. B., & Hannan, M. T. (1984). *Social dynamics: Models and methods.* New York, NY: Academic Press.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of Brownian motion. *Physical Review, 36,* 823–841.

Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data.* New York, NY: Springer-Verlag.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76,* 820–838. doi:10.1037/0022-3514.76.5.820

Wiener, N. (1923). Differential space. *Journal of Mathematics and Physics, 58,* 131–174.

Witherington, D. C., & Crichton, J. A. (2007). Frameworks for understanding emotions and their development: Functionalist and dynamic systems approaches. *Emotion, 7,* 628–637. doi:10.1037/1528-3542.7.3.628

Appendix A

Properties of the Stochastic Integral and the Solution of the Stochastic Differential Equation for One-Dimensional Ornstein–Uhlenbeck Process

First, we discuss some properties of the standard Brownian motion process:

1. For every t, $W(t)$ has a normal distribution.
2. $E(W(t)) = 0$ and $\mathrm{Cov}(W(s), W(t)) = \min(s,t)$.
3. It has independent increments: For every ordered sequence $t_1 < t_2 < t_3 < t_4$ of four time points, $W(t_2) - W(t_1)$ is independent of $W(t_4) - W(t_3)$.

Note that a direct consequence of the second property is that $\mathrm{Cov}(W(t), W(t)) = \mathrm{Var}(W(t)) = t$, and consequently a Brownian motion process is not stationary because its variance changes over time. However, the increments of the Brownian motion process are stationary; hence, the distribution of $W(t + h) - W(t)$ does not depend on t, only on the time difference h, and is thus identical for all t.

The solution of a stochastic integral relies on a specific calculus, mostly Itô calculus, although other possibilities exist (see, e.g., Arnold, 1974). To define the stochastic integral, suppose $G(t)$ is an arbitrary function of time and $W(t)$ is the standard Brownian motion. The stochastic integral $\int_{t_0}^{t} G(u)dW(u)$ is defined (Gardiner, 2004) as a limit of the partial sums:

$$S_n = \sum_{i=1}^{n} G(\tau_i)[W(t_i) - W(t_i - 1)],$$

where it holds that $t_{i-1} < \tau_i < t_i$ and $t_n = t$ so that

$$\lim_{n \to \infty} S_n = \int_{t_0}^{t} G(u)dW(u).$$

In principle, the function $G(\cdot)$ could be deterministic or (nonanticipatory) stochastic, but we consider in this article only the first option.

Here we highlight three important properties. For the full description of properties, we refer to Arnold (1974) and Tuma and Hannan (1984). The first one is the following:

$$\mathrm{E}\left[\int_{t_0}^{t} G(u)dW(u)\right] = 0,$$

which is to say that the expected mean of the stochastic integral is 0. This property can be understood easily by taking the expectation of $\lim_{n\to\infty} S_n$. The second property states that

$$\mathrm{E}\left[\left(\int_{t_o}^{t} G(u)dW(u)\right)\left(\int_{t_0}^{s} G(u)dW(u)\right)\right] = \int_{t_0}^{s} G^2(u)du,$$

where $t_0 \le s \le t$, which is an ordinary integral of time. As a third property, we mention that the distribution of the stochastic integral with respect to a Wiener process is normally distributed (and the mean and variance of this distribution are derived above).

Let us consider now the specific case of the solution of the stochastic differential equation for the Ornstein–Uhlenbeck process:

$$d\Theta(t) = \beta(\mu - \Theta(t))dt + \sigma dW(t).$$

First, we integrate over this equation, which results in

$$\Theta(t) = \mu + e^{-\beta t}(\Theta_0 - \mu) + \sigma e^{-\beta t}\int_0^t e^{\beta u}dW(u),$$

where the last term in the solution is a stochastic integral (it is an integral with respect to the Brownian motion process $W(t)$).

We will often not condition on the position at Time 0 but on the position d time units before, that is, $\Theta(t - d)$, so that the solution then becomes

$$\Theta(t) = \mu + e^{-\beta d}(\Theta(t - d) - \mu) + \sigma e^{-\beta t}\int_{t-d}^{t} e^{\beta u}dW(u). \tag{A1}$$

From Equation A1, we can see that the position at time t, that is, $\Theta(t)$, depends on the already introduced parameters and the previously measured position $\Theta(t - d)$.

Making use of the properties of the stochastic integral as introduced above, the conditional distribution of $\Theta(t)$ given $\Theta(t - d)$ can be derived from Equation A1, and is as follows:

$$\Theta(t)|\Theta(t - d) \sim N\left(\mu + e^{-\beta d}(\Theta(t - d) - \mu), \frac{\sigma^2}{2\beta}(1 - e^{-2\beta d})\right).$$

(*Appendices continue*)

Appendix B

Solution of the Stochastic Differential Equation for Two-Dimensional Ornstein–Uhlenbeck Process

Based on the extensive treatment of the unidimensional case, the two-dimensional process will not appear entirely novel. As established before, $\boldsymbol{\Theta}(t)$ represents the position in a two-dimensional space at time t. The stochastic differential equation describing the change in the vector $\boldsymbol{\Theta}(t)$ is then as follows:

$$d\boldsymbol{\Theta}(t) = \mathbf{B}(\boldsymbol{\mu} - \boldsymbol{\Theta}(t))dt + \boldsymbol{\sigma}d\mathbf{W}(t). \tag{B1}$$

The vector $\boldsymbol{\mu}$ now stands for the home base in a two-dimensional space. The adjustment to $\boldsymbol{\mu}$ is no longer determined by a single scalar β but the matrix $\mathbf{B}$. The $d\mathbf{W}(t)$ represents the already introduced white noise in two dimensions. The matrix $\boldsymbol{\sigma}$ controls the variances and covariances of the two driving white noise processes. The instantaneous covariance matrix $\boldsymbol{\Sigma}$ can be derived from $\boldsymbol{\sigma}$ as follows: $\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}$.

The solution of the two-dimensional stochastic differential equation in Equation B1 is very similar to the unidimensional solution (assuming we condition on $\boldsymbol{\Theta}(t - d)$):

$$\boldsymbol{\Theta}(t) = \boldsymbol{\mu} + e^{-\mathbf{B}d}(\boldsymbol{\Theta}(t - d) - \boldsymbol{\mu}) + \boldsymbol{\sigma}e^{-\mathbf{B}d}\int_{t-d}^{t} e^{\mathbf{B}u}d\mathbf{W}(u),$$

where d denotes the time difference and $e^{-\mathbf{X}}$ is the matrix exponential defined as

$$e^{-\mathbf{X}} = I - \mathbf{X} + \frac{\mathbf{X}^2}{2!} - \frac{\mathbf{X}^3}{3!} + \frac{\mathbf{X}^4}{4!} - \frac{\mathbf{X}^5}{5!} + \cdots.$$

We follow the reparameterization that was already introduced in the case of a unidimensional process. Instead of using the Cholesky decomposition of the instantaneous covariance matrix (i.e., $\boldsymbol{\sigma}$), we prefer the parameterization based on the stationary covariance matrix $\boldsymbol{\Gamma}$ (see Gardiner, 2004):

$$\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}} = \mathbf{B}\boldsymbol{\Gamma} + \boldsymbol{\Gamma}\mathbf{B}^{\mathrm{T}},$$

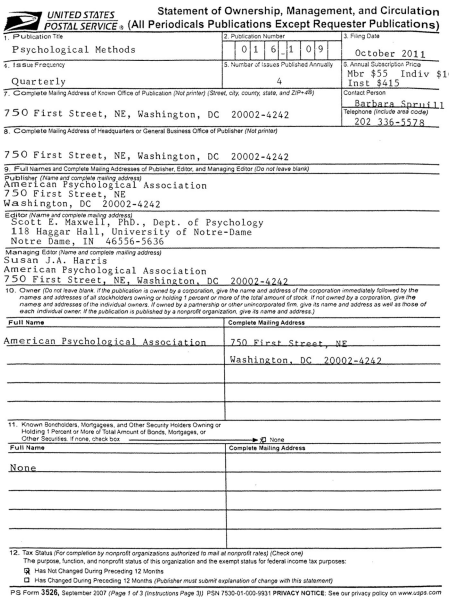
where Σ is the instantaneous covariance matrix and $\boldsymbol{\sigma}$ its Cholesky decomposition. Then the conditional distribution of $\boldsymbol{\Theta}(t)$ given $\boldsymbol{\Theta}(t - d)$ equals

$$\boldsymbol{\Theta}(t)|\boldsymbol{\Theta}(t - d) \sim N_2(\boldsymbol{\mu} + e^{-\mathbf{B}d}(\boldsymbol{\Theta}(t - d) - \boldsymbol{\mu}), \boldsymbol{\Gamma} - e^{-\mathbf{B}d}\boldsymbol{\Gamma}e^{-\mathbf{B}^{\mathrm{T}}d}),$$

where N_2 refers to the bivariate normal distribution. Also in this two-dimensional case, the process converges to a stationary distribution:

$$x\boldsymbol{\Theta}(t) \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Gamma}).$$

Received November 7, 2008
Revision received September 29, 2010
Accepted February 26, 2011 ■



UNITED STATES POSTAL SERVICE ® Statement of Ownership, Management, and Circulation (All Periodicals Publications Except Requester Publications)

1. Publication Title: Psychological Methods
2. Publication Number: 0 1 6 - 1 0 9
3. Filing Date: October 2011
4. Issue Frequency: Quarterly
5. Number of Issues Published Annually: 4
6. Annual Subscription Price: Mbr $55 Indiv $107 Inst $415
7. Complete Mailing Address of Known Office of Publication (Not printer) (Street, city, county, state, and ZIP+4®): 750 First Street, NE, Washington, DC 20002-4242
Contact Person: Barbara Spruill
Telephone (Include area code): 202 336-5578
8. Complete Mailing Address of Headquarters or General Business Office of Publisher (Not printer): 750 First Street, NE, Washington, DC 20002-4242
9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor (Do not leave blank)
Publisher (Name and complete mailing address): American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242
Editor (Name and complete mailing address): Scott E. Maxwell, Phd., Dept. of Psychology, 118 Haggar Hall, University of Notre-Dame, Notre Dame, IN 46556-5636
Managing Editor (Name and complete mailing address): Susan J.A. Harris, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242
10. Owner (Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)

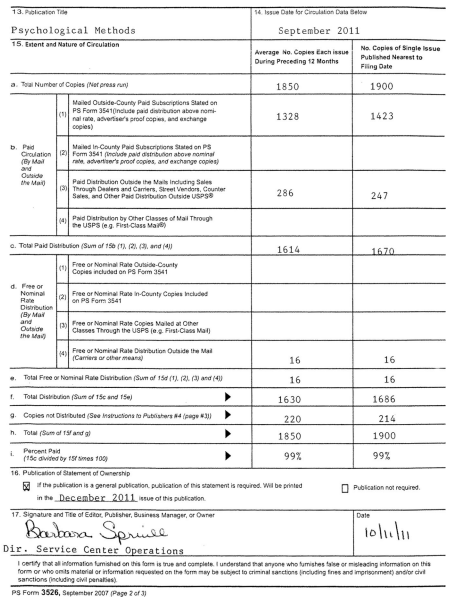
Full Name	Complete Mailing Address
American Psychological Association	750 First Street, NE Washington, DC 20002-4242

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address
None	

12. Tax Status (For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:
☒ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months (Publisher must submit explanation of change with this statement)

PS Form 3526, September 2007 (Page 1 of 3 (Instructions Page 3)) PSN 7530-01-000-9931 PRIVACY NOTICE: See our privacy policy on www.usps.com



13. Publication Title: Psychological Methods
14. Issue Date for Circulation Data Below: September 2011

15. Extent and Nature of Circulation			Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies (Net press run)			1850	1900
b. Paid Circulation (By Mail and Outside the Mail)	(1)	Mailed Outside-County Paid Subscriptions Stated on PS Form 3541(Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)	1328	1423
	(2)	Mailed In-County Paid Subscriptions Stated on PS Form 3541 (Include paid distribution above nominal rate, advertiser's proof copies, and exchange copies)		
	(3)	Paid Distribution Outside the Mails Including Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Paid Distribution Outside USPS®	286	247
	(4)	Paid Distribution by Other Classes of Mail Through the USPS (e.g. First-Class Mail®)		
c. Total Paid Distribution (Sum of 15b (1), (2), (3), and (4))			1614	1670
d. Free or Nominal Rate Distribution (By Mail and Outside the Mail)	(1)	Free or Nominal Rate Outside-County Copies included on PS Form 3541		
	(2)	Free or Nominal Rate In-County Copies Included on PS Form 3541		
	(3)	Free or Nominal Rate Copies Mailed at Other Classes Through the USPS (e.g. First-Class Mail)		
	(4)	Free or Nominal Rate Distribution Outside the Mail (Carriers or other means)	16	16
e. Total Free or Nominal Rate Distribution (Sum of 15d (1), (2), (3) and (4))			16	16
f. Total Distribution (Sum of 15c and 15e)			1630	1686
g. Copies not Distributed (See Instructions to Publishers #4 (page #3))			220	214
h. Total (Sum of 15f and g)			1850	1900
i. Percent Paid (15c divided by 15f times 100)			99%	99%

16. Publication of Statement of Ownership
☒ If the publication is a general publication, publication of this statement is required. Will be printed in the December 2011 issue of this publication. ☐ Publication not required.
17. Signature and Title of Editor, Publisher, Business Manager, or Owner: Barbara Spruill, Dir. Service Center Operations — Date: 10/11/11
I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

PS Form 3526, September 2007 (Page 2 of 3)