

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: www.tandfonline.com/journals/hmbr20

Dynamics Between Asynchronously Measured Variables: A Multilevel Approach to Momentary Affect and Morning Sleep Reports

Sophie W. Berkhout, Noémi K. Schuurman, Koen Niemeijer, Peter Kuppens & Ellen L. Hamaker

To cite this article: Sophie W. Berkhout, Noémi K. Schuurman, Koen Niemeijer, Peter Kuppens & Ellen L. Hamaker (2025) Dynamics Between Asynchronously Measured Variables: A Multilevel Approach to Momentary Affect and Morning Sleep Reports, Multivariate Behavioral Research, 60:6, 1253-1273, DOI: 10.1080/00273171.2025.2551370

To link to this article: https://doi.org/10.1080/00273171.2025.2551370

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC

Published online: 15 Sep 2025.

Submit your article to this journal

Article views: 654

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=hmbr20

MULTIVARIATE BEHAVIORAL RESEARCH
2025, VOL. 60, NO. 6, 1253–1273
https://doi.org/10.1080/00273171.2025.2551370

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Dynamics Between Asynchronously Measured Variables: A Multilevel Approach to Momentary Affect and Morning Sleep Reports

Sophie W. Berkhout[a], Noémi K. Schuurman[a], Koen Niemeijer[b], Peter Kuppens[b], and Ellen L. Hamaker[a]

[a]Department of Methodology and Statistics, Utrecht University, Utrecht, The Netherlands; [b]Quantitative Psychology and Individual Differences Katholieke Universiteit Leuven, Leuven, Belgium

ABSTRACT

The reciprocal relations between sleep and affect have been a common focus in psychological research. Researchers studying affective processes often collect data multiple times a day over several days. Subjective sleep quality, on the other hand, is generally measured once at the beginning of the day. This difference in measurement frequency creates a challenge when analyzing these data, because standard dynamic models are not equipped for this. Furthermore, many of the popular approaches are based on the assumption of stationarity, meaning that processes are assumed to continue throughout the night in the same way as throughout the day. In this paper, we introduce a dynamic structural equation model that incorporates reciprocal relations between momentary affect and daily measures of sleep, tackling both of these challenges and also incorporating individual differences in these relations. To demonstrate the practical applicability of this model, we make use of an empirical example of positive and negative affect. Furthermore, we aim to give researchers the means to adapt or build on this model to align it with different research questions and other asynchronously measured variables.

KEYWORDS

Dynamic Modeling; Measurement Frequency; Multilevel Modeling; Experience Sampling Method; Affective Inertia

Introduction

Many people have the experience that a good night's sleep makes them feel happier, more energized, and better able to cope with ordinary challenges the following day. Moreover, people often associate negative affect—like distress and nervousness—with difficulties of falling and/or staying asleep. This dynamic interplay between sleep and affect is not merely a commonly held belief, but it has been studied extensively by psychological researchers as well (e.g., Goldstein & Walker, 2014; Konjarski et al., 2018; Sikka & Gross, 2023; van der Helm & Walker, 2009). In a comprehensive review of 29 studies that together span two decades, Konjarski et al. (2018) underlined the reciprocal relationship between sleep quality and affect. Their findings consistently showed that better sleep quality is associated with increased positive affect and decreased negative affect the following day. Conversely, while the following effects are somewhat inconsistent across studies, elevated positive affect is associated with better sleep quality the following night, and elevated negative affect is associated with worse sleep quality the subsequent night (Konjarski et al., 2018; Neubauer et al., 2021). The interrelatedness of sleep and affect is further demonstrated by the major role of disrupted sleep in affective disorders such as depression (Baglioni et al., 2016; Benca et al., 1992; Palmer et al., 2024).

However, the measurements of sleep and affect are often asynchronous, which poses major data-analytic challenges that hinder the further study of their reciprocal relation. In this paper, we focus on subjective sleep quality often measured once per day through self-report in the morning with questions like "How well did you sleep?". In contrast, momentary affect is often measured throughout the day with the experience sampling method (ESM) consisting of repeated self-reports at (semi-)random occasions. Due to the asynchronous nature of these measurements, regressing these variables on one another is not straightforward. As an ad hoc solution, many studies of sleep and affect either ask participants to report their overall daily affect once per day (e.g., Flueckiger et al.,

CONTACT Sophie W. Berkhout s.w.berkhout@uu.nl Utrecht University, Padualaan 14, 3584 CH Utrecht, The Netherlands.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

2017; Fortier et al., 2015; Fuligni & Hardway, 2006; Kalmbach et al., 2014) or aggregate the affect variable into daily averages (e.g., Bouwmans et al., 2017; Cousins et al., 2011; Garcia et al., 2014; van Zundert et al., 2015; Wild-Hartmann et al., 2013), thus analyzing dynamic relations between sleep and daily rather than momentary affect. Other studies resort to conducting multiple analyses for separate time-of-day effects, for example, morning, afternoon, and evening (e.g., Könen et al., 2016; Neubauer et al., 2021; Talbot et al., 2012; Wrzus et al., 2014). Both strategies ensure that the frequency of momentary affect measurements is reduced—either through aggregation or partitioning—to one measurement per day, such that it matches the frequency of sleep measurements. As a result, existing analysis techniques can be readily used such as a multilevel regression analysis based on having days nested in persons (e.g., Bouwmans et al., 2017; Könen et al., 2016; Takano et al., 2014; van Zundert et al., 2015; Wild-Hartmann et al., 2013; Wrzus et al., 2014).

While this approach helps to gain some insights into the underlying dynamics of sleep and affect, it is based on reducing the frequency of momentary affect measures to a single score per day. Consequently, we can no longer link the results we obtain to the broader literature that focuses on momentary affective dynamics, including the large and growing body of literature on the role of affective inertia. Affective inertia has been considered an important feature of self-regulation (Koval et al., 2015, 2021; Kuppens et al., 2010; Suls et al., 1998), and individual differences in affective inertia are for instance related negatively to psychological well-being (Brose et al., 2015; Houben et al., 2015; Koval et al., 2021; Kuppens et al., 2010; Shin et al., 2022; van de Leemput et al., 2014; Wichers & Groot, 2016; but see Dejonckheere et al., 2019), and positively to neuroticism (Bringmann et al., 2016; Koval et al., 2015; Suls & Martin, 2005; Waugh et al., 2017). It would thus be desirable to have a model that can accommodate the difference in measurement frequencies between sleep and affect, and that allows us to investigate and compare the dynamic relations in both directions, while also incorporating individual differences in these relations. This would enable us to investigate, for instance, whether individuals who have a stronger than average relation from sleep to affect also tend to have a stronger than average relation from affect to sleep, but also whether individuals with stronger affective inertia tend to also have stronger sleep inertia. While these relations between individual effects may also be estimated in a two-step approach by estimating the random effects using separate univariate models in a first step and then correlating these estimates in a second step, this approach would be less efficient and would underestimate the uncertainty of the correlations. Note further that while the present work focuses on subjective sleep quality measured through self-report, there are other sleep measures such as duration or sleep quality measured through a wearable device. Each of these sleep constructs may be characterized by different dynamic processes and be differently related to affect (Konjarski et al., 2018).

In this paper, we propose a new dynamic multilevel model for daily and momentary measures, which does not require aggregation or partitioning of the latter to match the frequency of sleep measurements. To this end, we first introduce a univariate model that distinguishes between daytime and nighttime inertia in the momentary variable to capture how the momentary process evolves from day to night. Then, we integrate the daily variable to develop a bivariate model that enables the study of reciprocal relations between the daily and momentary variables and individual differences therein. These approaches can be considered multilevel extensions of the single-level models for the analysis of data from a single person proposed by Berkhout et al. (2024) and Berkhout et al. (2025). To motivate and illustrate our developments, we first introduce an empirical example with data from Niemeijer et al. (2023) in Section "Empirical example: sleep quality and positive/negative affect", where we will use the momentary measurements of positive and negative affect, daily measures of sleep quality, and baseline measures of neuroticism and chronotype. Neuroticism has been related to emotional inertia before (Bringmann et al., 2016; Koval et al., 2015; Suls & Martin, 2005; Waugh et al., 2017), and some studies suggest a relation between chronotype and affect (Carciofo, 2023; Jankowski, 2014; Miller et al., 2015; Stolarski et al., 2016). In Section "A new dynamic multilevel model" we explain why the standard models are unsuitable for analyzing these data, and we build a new dynamic multilevel model, starting from the existing multilevel autoregressive model often used to study individual differences in affective inertia. In a first step, we extend this model so that it can accommodate separate daytime and nighttime inertia. In a second step, we extend this new daytime-nighttime inertia model to include the less frequently measured sleep variable, allowing for reciprocal relations between sleep and affect. We discuss how to prepare the data to estimate the new models in Mplus (Muthén & Muthén, 2017) in Section "Model

implementation", and present the results from applying the two extensions to positive and negative affect separately in Section "Results". We end with a discussion covering limitations and future directions in Section "Discussion".

We preregistered the empirical illustration to ensure transparency and to strengthen the methodological integrity of our analyses. The preregistration—consisting of our research questions, hypotheses, and planned analyses—is available at https://osf.io/geyw4. Furthermore, the annotated Mplus code used for the analyses is available in the supplementary materials at https://osf.io/h3dkj/.

Empirical example: sleep quality and positive/negative affect

We use an empirical example to motivate and illustrate the models that we develop. The data set for this example was provided by the Research Group of Quantitative Psychology and Individual Differences at KU Leuven, and was first described in Niemeijer et al. (2023). The study was approved by the Katholieke Universiteit Leuven Social and Societal Ethics Committee (G-2020-2200-R3[AMD]). Over the course of three weeks, data were collected from 104 participants (87.5% women, 11.5% men, 1% other) with ages ranging from 18 to 52 (mean 20.7, *SD* 4.08).

Variables

We consider both positive and negative affect, which were collected using an ESM design with ten measurements per day for 21 days. Prior to the data collection, participants indicated a 12-h period in which they would receive the measurement prompts every day, for example, from 9 AM to 9 PM. At every measurement occasion, participants answered the questions "To what extent do you experience positive feelings?" (PA) and "To what extent do you experience negative feelings?" (NA) on a scale ranging from 0 (no positive/negative feelings) to 100 (many positive/negative feelings). Additionally, we focus on sleep quality, which was included in the ESM design at the first recorded measurement occasion of every day. Participants answered "How well did you sleep?" on a scale from 0 (very bad) to 100 (very well). Finally, we also consider baseline measures of neuroticism and chronotype. For neuroticism, participants completed the Big Five Inventory (BFI-2) where neuroticism was computed as the average of the corresponding scale (Denissen et al., 2008; Soto & John, 2017), and for chronotype participants answered "Are you a morning or evening person?" on a scale from 1 (morning) to 7 (evening).

Measurement intervals and frequencies

The ESM design used a semi-random sampling scheme to determine the timing of the measurement prompts within the 12-h sampling period every day. The sampling period was divided into ten equal blocks of 76.5 min in which the timing of the measurement prompt was sampled randomly, with a minimum of 15 min between prompts and a 30-minute response window; The full details of the ESM design are provided in the supplementary materials. There are three aspects of the study design that result in unequal time intervals between the ESM observations. First, due to the random sampling of measurement prompts within this sampling period and the 30-minute response window, the intervals between observations within days are unequal by design. Second, as the sampling period does not cover the nighttime when participants are asleep, the intervals at night between the last measurement prompt of the day and the first measurement prompt of the next day are much longer than the daytime intervals. Third, missing values impact the length of the intervals between observed values, and subsequent missing values can make the intervals between observations even longer.

The measurement frequency of the ESM observations (ten times per day) is necessarily much higher than that of the sleep quality observations (once per day). This becomes clear from Figure 1, which contains both the momentary PA and the daily sleep quality scores of a random participant plotted against time, for the 21-day study period. While there are 21 daily measurements for sleep quality, there are 181 measurements for the affect variable for this participant. Additionally, while the intervals between observations for the sleep quality variable are more or less equal (i.e., about 24 h), the intervals for the positive affect variable vary both within the day, but also between the first observation of a day and the last observation of the previous day. We will explain how these features of the data can be handled when we discuss our modeling approach for these data.

Research questions and hypotheses

We focus on the following research questions: (1) Is there nighttime affective inertia? (2) Is there a difference between daytime and nighttime affective inertia? (3) Are there individual differences in the daytime

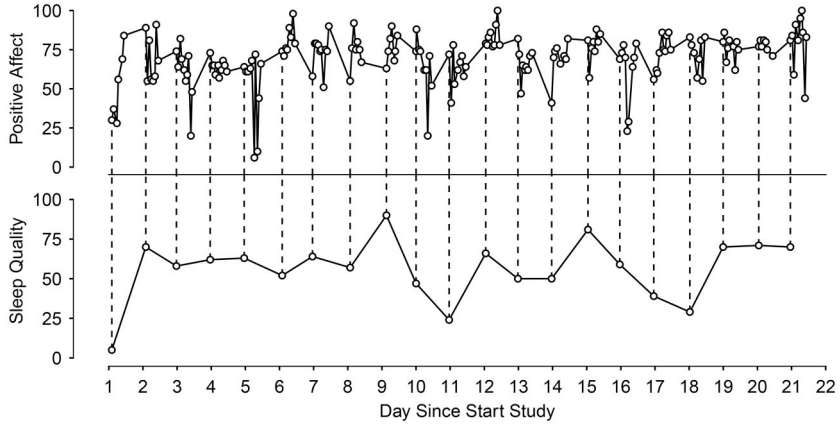

Figure 1. Example of positive affect and sleep quality measurements for a random participant over the 21-day study period. The dashed line indicates the first measurement occasion of the day.

Table 1. Hypotheses of the empirical example.

Label	Hypothesis
$\mathcal{H}_1$	Both PA and NA have positive daytime and nighttime inertia.
$\mathcal{H}_2$	The daytime inertia is larger than the nighttime inertia for PA and NA.
$\mathcal{H}_3$	There are individual differences in the daytime and nighttime inertia, average affect, and innovation variances for PA and NA.
$\mathcal{H}_4$	There is a relation between the daytime and nighttime inertia (undirectional).
$\mathcal{H}_5$	Sleep quality positively relates to PA and negatively relates to NA at the beginning and end of the day.
$\mathcal{H}_6$	PA positively and NA negatively relate to sleep quality the subsequent night.
$\mathcal{H}_7$	There are individual differences in average sleep quality, sleep inertia, the reciprocal effects between sleep quality and affect, and the innovation variances.
$\mathcal{H}_8$	Neuroticism positively predicts NA averages, daytime and nighttime inertia, while chronotype predicts (undirectional) reciprocal effects between sleep quality and affect for NA and PA.

and nighttime inertia? (4) How are these related to each other and individual differences in average affect and innovation variances? (5) Is sleep quality related to affect at the beginning and/or end of the subsequent day? (6) Is affect at the end of the day related to sleep quality the following night? (7) Are there individual differences in sleep inertia, average sleep quality, and the reciprocal effects between sleep quality and affect, and how are these related? Finally, by bringing in baseline measures of individual characteristics, we can consider the research question: (8) Can between-level predictors explain individual differences in the random effects?

For each of these eight questions, we preregistered our hypotheses; an overview is given here in Table 1 (for more details, see https://osf.io/geyw4). Note that for some parts of the research questions we did not specify hypotheses and therefore the results should be interpreted as exploratory. For example, we had no hypotheses on other relations between the individual differences of the random effects except for the relation between the daytime and nighttime autoregression ($\mathcal{H}_4$). We also had no hypotheses on whether neuroticism and chronotype predict other random effects. Finally, while we include an autoregressive relation for sleep, we did not specify a research question or hypothesis for this.

A new dynamic multilevel model

The specific combination of sleep quality assessed upon awakening and momentary affect measures obtained throughout the day presents a major challenge that existing modeling approaches were not designed to handle.

Why standard models fall short

A widely used model for handling multivariate time series data in psychological research is the vector autoregressive (VAR) model, and multilevel extensions of this (Asparouhov et al., 2018; Bringmann et al., 2013; Epskamp et al., 2018). In the VAR(1) model each variable is regressed on itself and the other variable(s) at the previous time point. Related models are higher-order VAR models that allow for longer lags (i.e., the current observation is not only regressed on the preceding one, but also further backward in time), and structural VAR models that allow for lag zero regressions (Beltz & Gates, 2017; Gates & Molenaar, 2012). What makes VAR models particularly attractive is that they allow for the study of reciprocal relations, such as daily dynamics between spouses or the reciprocal relationship between momentary stress and anxiety. Being familiar with that literature, the VAR model might

seem like a logical choice for capturing the interplay between sleep and affect.

However, the data we are concerned with here are characterized by just one measurement occasion per day when both variables—sleep and affect—are observed; this occurs at the first observed measurement of the day, while at all subsequent time points only momentary affect is recorded. If we nevertheless insist on using a VAR model, we have to either reduce the number of affect measurements to one per day, which would mean a problematic impoverishment of the data, or we should adopt an estimation procedure that can handle missing data. Well-known solutions from the latter category are continuous-time modeling (Boker, 2002; Driver et al., 2017; Voelkle et al., 2012) or a Kalman filter approach (Cipra & Romera, 1997; Kim & Nelson, 1999), which have been implemented in existing software and can thus be readily used. However, they rely on the implicit assumption that the missing values are missing only because they were not observed. Using such approaches therefore requires accepting the premise that the sleep process continues to fluctuate and interact with affect throughout the day; it is based on the idea that sleep during the day is a latent process, rather than non-existent. Clearly, this assumption does not align with the phenomenon we are investigating. While nighttime sleep may influence a person's affect not only at the start of the day but also throughout the day via delayed effects, sleep itself is not an ongoing—latent—process during waking hours. While a VAR model (or a continuous time equivalent) can be applied to these data and will provide us with numerical results, from a substantive point of view the results from such a model are meaningless here.

Another limitation of standard models like the VAR(1) or its univariate AR(1) counterpart is their reliance on the assumption of stationarity. This means that when applying an AR(1) model to study the dynamics of momentary affect, we are implicitly assuming that this process continues overnight in the exact same manner as throughout the day. Hence, we subscribe to the belief that for the affective process there is no inherent difference between being awake or asleep, that external influences and regulatory mechanisms operate in the same way across these conditions, and that the psychological function of affect remains unchanged—despite the fundamental differences between these two states of consciousness. This assumption is highly implausible and most of us would not want to endorse it. Moreover, recent empirical studies are refuting it. If the assumption were correct, the overnight autoregression (from the last measurement on a day to the first measurement the following day), would approach zero (Berkhout et al., 2024). The reason for this is the exponential decay that characterizes autoregressive processes, which means that the autoregressive effects decrease when the time interval between observations increase (Brockwell & Davis, 1993; Inoue, 2008; Takemura, 2016). The intervals at night are typically long enough for the decay to tend toward zero. Yet, various researchers have found overnight affective autoregressions that are nonzero (Berkhout et al., 2024; Frérart et al., 2023; Minaeva et al., 2021), strongly suggesting that the affective process during nighttime is not a simple continuation of the affective process during waking hours.

These two issues underscore the need for an alternative to model the reciprocal dynamics between sleep and momentary affect. This alternative should not merely be able to handle differences in measurement frequency, but it should do justice to the fact that the phenomena we are interested in do not behave like stationary ongoing processes during the periods when we are not observing them. The reality is that the measurements are absent for good reason: We are not tracking sleep throughout the day, because there is no sleep process taking place during that time; and we are not measuring momentary affect during sleep, because individuals are in an altered state of consciousness then.

To build a model that can handle these specific particularities, we begin with a model for momentary affect only. In this model, which is a multilevel extension of one of the $N=1$ models presented by Berkhout et al. (2024), individuals can have distinct daytime and nighttime autoregressions that need not be related to each other. Subsequently, we present an extension, based on one of the $N=1$ models proposed by Berkhout et al. (2025), which allows us to include both daily sleep measurements and momentary affect measurements in a single model, thereby offering new opportunities for studying their dynamics.

A new model for daytime and nighttime autoregression in momentary affect

Our modeling approach is based on a decomposition of the affect observations into two parts: (a) a trait component represented by average affect, which remains stable within a person for the duration of the study, but varies across individuals representing individual differences in their general tendencies; and (b) a dynamic state component that represents the person-specific temporal deviation from their mean. We denote the momentary affect variable as $A_{b,d,i}$, with the three subscripts referring to beep b, day d,

and individual *i*. As we will incorporate the sleep variable (*S*) later on, we will insert a subscript *A* here to identify the parameters associated with variable *A*. The decomposition may be mathematically expressed as

$$A_{b,d,i} = \mu_{A,i} + \tilde{A}_{b,d,i}, \quad (1)$$

where $\mu_{A,i}$ is the average affect of individual *i*, and $\tilde{A}_{b,d,i}$ is the deviation from the mean of individual *i* on day *d* and beep *b*. This basic decomposition is also shown in the left panel of Figure 2.

Within-person dynamics

At the within-person level, each observation depends on the previous observation through an autoregressive effect. But in contrast to a regular AR(1) model, we allow for a different nighttime autoregressive effect compared to the daytime autoregressive effect. Hence, the affective state at the first measurement occasion of the day (i.e., $\tilde{A}_{1,d,i}$) is regressed on the last affective state at the preceding day (i.e., $\tilde{A}_{B,d-1,i}$), through

$$\tilde{A}_{1,d,i} = \alpha_{A,i}\ \tilde{A}_{B,d-1,i} + \zeta_{1,d,i}, \quad (2)$$

where $\alpha_{A,i}$ represents the nighttime autoregression of individual *i*. The innovation $\zeta_{1,d,i}$ captures everything about morning affect that is not predictable from last night's affect. The innovations are assumed to be normally distributed with a mean of zero and residual variance $\psi_{A_1,i}$. Within the day (so from the second beep to the last beep), the affect dynamics are modeled through

$$\tilde{A}_{b,d,i} = \phi_{A,i}\ \tilde{A}_{b-1,d,i} + \zeta_{b,d,i} \quad \text{for } b = 2, \ldots, B, \quad (3)$$

with $\phi_{A,i}$ as the daytime autoregression, and $\zeta_{b,d,i}$ as the innovation, which comes from a normal distribution with a mean of zero and residual variance $\psi_{A_o,i}$.

Hence, in this model, there are distinct parameters that represent the daytime and nighttime autoregression, but there are also distinct parameters for the innovation variance at the first beep versus that at the other beeps. The latter is preferable over having a single parameter for this, because that would imply that differences in the variances at different times during the day must be the result of differences in the autoregression; such a restriction (across time, and/or across persons) may bias the estimation of the autoregressive effect (Jongerling et al., 2015).

The subscript *i* for the autoregressions and the innovation variances indicates that each person may have different values for these parameters. This is visualized by having filled circles for them at the within-person level shown in the upper right panel of Figure 2.

Between-person differences

At the between-person level, we may further model the individual differences in the parameters outlined above; this is illustrated in the lower right panel of Figure 2, where the the person-specific average affect (i.e., $\mu_{A,i}$), nighttime and daytime autoregression (i.e., $\phi_{A,i}$ and $\alpha_{A,i}$), and (log-transformed) innovation variances (i.e., $\psi_{A_1,i}$ and $\psi_{A_o,i}$) are represented as open circles. The innovation variances are log-transformed at the between-person level to ensure non-negative values for these variance terms.

Individual differences in these model features can be understood as follows. Differences in $\mu_{A,i}$ imply that people differ in the general tendency of their affective state. Differences in $\phi_{A,i}$ and $\alpha_{A,i}$ are of interest, because individuals with higher affective inertia are considered to have more regulatory weakness compared to

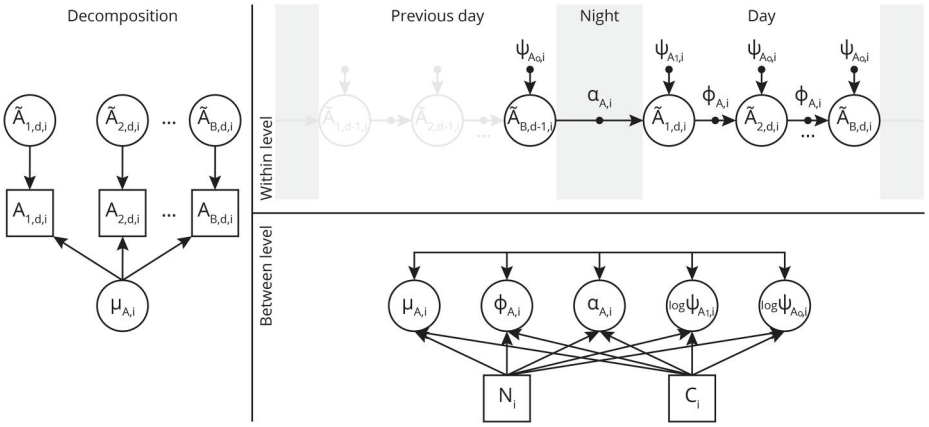

Figure 2. Graphical representation of the decomposition, within-person level, and between-person level specification for the daytime and nighttime autoregressive model. On the within-person level, all components related to day *d* are highlighted in black. $A_{1,d,i}$ is the observation of variable *A* at beep 1, day *d*, and of person *i*, $\mu_{A,i}$ is the average of person *i*, $\tilde{A}_{1,d,i}$ is the temporal deviation from this average, $\phi_{A,i}$ is the daytime autoregression, $\alpha_{A,i}$ is the nighttime autoregression, $\log(\psi_{A_1,i})$ is the innovation variance of the first beep, $\log(\psi_{A_o,i})$ is the innovation variance of the other beeps, N_i is the neuroticism score, and C_i is the chronotype score of person *i*.

individuals with lower inertia; that is, individuals with low inertia are likely to recover faster from disturbances away from their stable state $\mu_{A,i}$. Finally, individual differences in $\psi_{A_1,i}$ and $\psi_{A_o,i}$ may be understood as stemming from differences between individuals in the degree to which they are exposed to variation in external factors, and differences between individuals in their response to these factors (Jongerling et al., 2015). The latter can be thus considered a measure of sensitivity or reactivity to unobserved factors that impact the affective state of a person.

We include baseline measures of neuroticism (N_i) and chronotype (C_i) as between-level predictors for all random effects. Past research has shown that individuals with high levels of neuroticism show more emotional inertia (Bringmann et al., 2016; Koval et al., 2015; Suls & Martin, 2005; Waugh et al., 2017). Additionally, some studies indicate that chronotype is also associated with affect, where evening persons show a lowered affect compared to morning persons (Carciofo, 2023; Jankowski, 2014; Miller et al., 2015; Stolarski et al., 2016). For example, the random effect of average affect can be modeled as

$$\mu_{A,i} = \gamma_{\mu_A} + \theta_{\mu_A}\ N_i + \omega_{\mu_A}\ C_i + \upsilon_{\mu_A,i}, \tag{4}$$

where θ_{μ_A} captures the effect of neuroticism, ω_{μ_A} the effect of chronotype, and $\upsilon_{\mu_A,i}$ the part of a person's mean that is not predictable from the person's score on neuroticism and chronotype. The other random effects are modeled similarly, as shown in Table 2. Neuroticism and chronotype are grand mean centered so the intercept γ_{μ_A} represents the overall mean here; similarly, the intercepts in the expressions for the other random effects represent the average across individuals of that effect (sometimes referred to as the fixed effect).

We are also interested in how much the random effects correlate with each other, either before or after accounting for neuroticism and chronotype. For instance, we may be interested in the between-person correlations of (the residuals of) daytime autoregression and nighttime autoregression. These correlations can be obtained from the covariance matrix T of the multivariate normal distribution that bounds (the residuals of) the random effects, that is,

$$T = \begin{bmatrix} \tau^2_{\mu_A} & & & & \\ \tau_{\alpha_A,\mu_A} & \tau^2_{\alpha_A} & & & \\ \tau_{\phi_A,\mu_A} & \tau_{\phi_A,\alpha_A} & \tau^2_{\phi_A} & & \\ \tau_{\psi_{A_1},\mu_A} & \tau_{\psi_{A_1},\alpha_A} & \tau_{\psi_{A_1},\phi_A} & \tau^2_{\psi_{A_1}} & \\ \tau_{\psi_{A_o},\mu_A} & \tau_{\psi_{A_o},\alpha_A} & \tau_{\psi_{A_o},\phi_A} & \tau_{\psi_{A_o},\psi_{A_1}} & \tau^2_{\psi_{A_o}} \end{bmatrix}. \tag{5}$$

The variances (τ^2) on the diagonal capture the degree to which individuals differ on (the residual of) a specific parameter, and the covariances (τ) indicate how the (residuals of) random effects are related to each other. For example, individuals with more average positive affect may be likely also to have a larger (positive) autoregression, which would imply $\tau_{\mu,\phi} > 0$.

A new model for daily sleep and momentary affect dynamics

We can now extend the previous model—which allows for different daytime and nighttime affect dynamics—to include the less frequently measured sleep variable in it. The affect variable is decomposed in the same way as before (see Equation 1). Because the sleep quality variable S is measured only once per day, it is given only two indices, for day d and individual i. Like affect, it is also decomposed into a stable individual mean $\mu_{S,i}$, and a person- and day-specific deviation from this mean $\tilde{S}_{d,i}$, that is,

$$S_{d,i} = \mu_{S,i} + \tilde{S}_{d,i}. \tag{6}$$

This decomposition of both affect and sleep is shown in the left panel of Figure 3.

Table 2. Random effect expressions of the daytime and nighttime autoregressive model (1) and the reciprocal sleep-affect model (2).

Model	Parameter	Expression
1 & 2	Average affect	$\mu_{A,i} = \gamma_{\mu_A} + \theta_{\mu_A}\ N_i + \omega_{\mu_A}\ C_i + \upsilon_{\mu_A,i}$
1 & 2	Nighttime autoregression	$\alpha_{A,i} = \gamma_{\alpha_A} + \theta_{\alpha_A}\ N_i + \omega_{\alpha_A}\ C_i + \upsilon_{\alpha_A,i}$
1 & 2	Daytime autoregression	$\phi_{A,i} = \gamma_{\phi_A} + \theta_{\phi_A}\ N_i + \omega_{\phi_A}\ C_i + \upsilon_{\phi_A,i}$
1 & 2	Innovation variance first beep	$\log(\psi_{A_1}) = \gamma_{\psi_{A_1}} + \theta_{\psi_{A_1}}\ N_i + \omega_{\psi_{A_1}}\ C_i + \upsilon_{\psi_{A_1},i}$
1 & 2	Innovation variance other beeps	$\log(\psi_{A_o}) = \gamma_{\psi_{A_o}} + \theta_{\psi_{A_o}}\ N_i + \omega_{\psi_{A_o}}\ C_i + \upsilon_{\psi_{A_o},i}$
2	Innovation variance last beep	$\log(\psi_{A_B}) = \gamma_{\psi_{A_B}} + \theta_{\psi_{A_B}}\ N_i + \omega_{\psi_{A_B}}\ C_i + \upsilon_{\psi_{A_B},i}$
2	Average sleep quality	$\mu_{S,i} = \gamma_{\mu_S} + \theta_{\mu_S}\ N_i + \omega_{\mu_S}\ C_i + \upsilon_{\mu_S,i}$
2	Sleep quality autoregression	$\phi_{S,i} = \gamma_{\phi_S} + \theta_{\phi_S}\ N_i + \omega_{\phi_S}\ C_i + \upsilon_{\phi_S,i}$
2	Spillover evening affect to sleep	$\beta_{S,i} = \gamma_{\beta_S} + \theta_{\beta_S}\ N_i + \omega_{\alpha_S}\ C_i + \upsilon_{\beta_S,i}$
2	Spillover sleep to morning affect	$\beta_{A,i} = \gamma_{\beta_A} + \theta_{\beta_A}\ N_i + \omega_{\alpha_A}\ C_i + \upsilon_{\beta_A,i}$
2	Spillover sleep to evening affect	$\delta_{A,i} = \gamma_{\delta_A} + \theta_{\delta_A}\ N_i + \omega_{\delta_A}\ C_i + \upsilon_{\delta_A,i}$
2	Innovation variance sleep quality	$\log(\psi_{S,i}) = \gamma_{\psi_S} + \theta_{\psi_S}\ N_i + \omega_{\psi_S}\ C_i + \upsilon_{\psi_S,i}$

Note. The γ's are the fixed effects, the θ's are the effects of neuroticism, the ω's are the effects of chronotype, and the υ's are person-specific deviations.

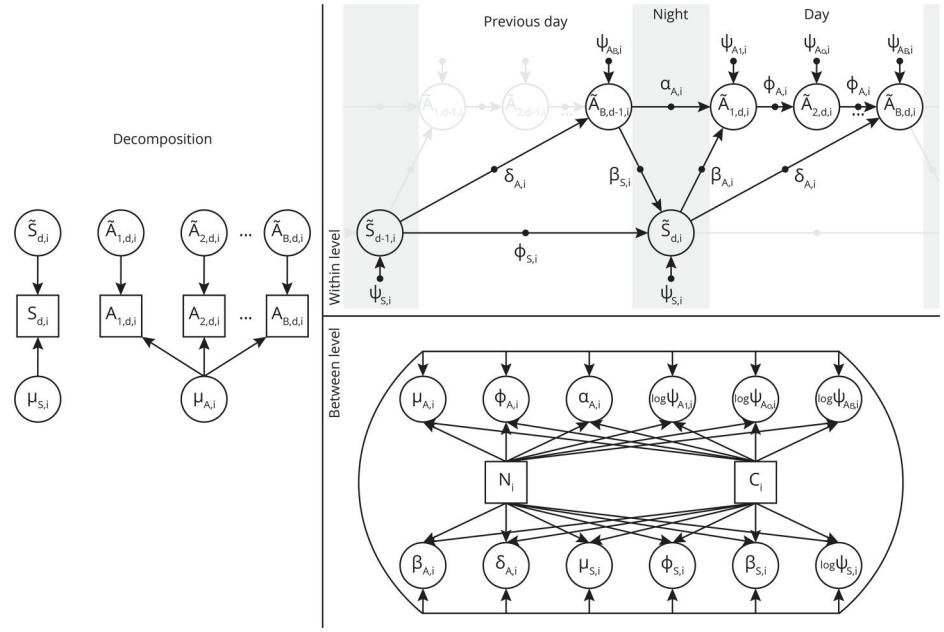

Figure 3. Graphical representation of the decomposition, within-person level, and between-person level specification for the reciprocal sleep-affect model. On the within-person level, all components related to day *d* are highlighted in black. $A_{1,d,i}$ is the observation of variable *A* at beep 1, day *d*, and of person *i*, $\mu_{A,i}$ is the average of person *i*, $\tilde{A}_{1,d,i}$ is the temporal deviation from this average at occasion *t* of person *i*, $\phi_{A,i}$ is the daytime autoregression, $\alpha_{A,i}$ is the nighttime autoregression, $\log(\psi_{A_1,i})$ is the innovation variance of the first beep, $\log(\psi_{A_o,i})$ is the innovation variance of the other beeps, $\log(\psi_{A_B,i})$ is the innovation variance of the last beep, $\beta_{A,i}$ is the cross-regression from *S* to the first beep of *A*, $\delta_{A,i}$ is the cross-regression from sleep quality to the last beep of *A*, μ_S is the average of *S*, $\phi_{S,i}$ is the autoregression of *S*, $\beta_{S,i}$ is the cross-lagged regression from *A* to *S*, and $\log(\psi_{S,i})$ is the innovation variance of *S*, N_i is the neuriticism score, and C_i is the chronotype score of person *i*.

Within-person dynamics

At the within-person level, we assume sleep quality is related to the affect variable *A* at the beginning of the day. Additionally, we also relate sleep to affect at the end of the day, because this may be the time of day that poor sleep has the largest effect. This means we need separate expressions for the first and last beep in comparison to the other beeps during the day. Hence, momentary affect at the first beep (i.e., $b = 1$) is now modeled as

$$\tilde{A}_{1,d,i} = \alpha_{A,i}\ \tilde{A}_{B,d-1,i} + \beta_{A,i}\ \tilde{S}_{d,i} + \zeta_{1,d,i}, \tag{7}$$

showing that it depends on affect at the end of the previous day, and on sleep during the night in between these affective measurements. The latter effect is captured by the cross-regression $\beta_{A,i}$. The subsequent occasions throughout the day are not directly dependent on sleep, although they are indirectly affected by it through the autoregression in affect, that is,

$$\tilde{A}_{b,d,i} = \phi_{A,i}\ \tilde{A}_{b-1,d,i} + \zeta_{b,d,i} \quad \text{for } b = 2, \ldots, B-1. \tag{8}$$

Finally, we also included an effect of last night's sleep quality on momentary affect at the end of the day (i.e., when $b = B$), because the effect of poor sleep on affect may be amplified at the end of the day due to physical and emotional exhaustion. The last measurement of each day is therefore modeled as

$$\tilde{A}_{B,d,i} = \phi_{A,i}\ \tilde{A}_{B-1,d,i} + \delta_{A,i}\ \tilde{S}_{d,i} + \zeta_{B,d,i}, \tag{9}$$

where the cross-regression $\delta_{A,i}$ captures this delayed effect of sleep on affect. The three expressions above each come with their own innovation variance, $\psi_{A_1,i}$, $\psi_{A_o,i}$, and $\psi_{A_B,i}$ respectively.

As sleep may also depend on affect, we model it as

$$\tilde{S}_{d,i} = \phi_{S,i}\ \tilde{S}_{d-1,i} + \beta_{S,i}\ \tilde{A}_{B,d-1,i} + \xi_{d,i}, \tag{10}$$

where ϕ_S is the autoregressive effect of sleep quality (from day to day), β_S is the cross-lagged effect of affect at the last beep of the preceding day on sleep quality the following night, and $\xi_{d,i}$ are the innovations that are normally distributed with a mean of zero and residual variance $\psi_{S,i}$.

Hence, this model for the within-person dynamics is characterized by the following individual-specific parameters: a daytime and nighttime autoregression for affect (i.e., $\phi_{A,i}$ and $\alpha_{A,i}$), and a day level autoregression for sleep (i.e., $\phi_{S,i}$); a direct and a delayed effect of sleep on affect the following day (i.e., $\beta_{A,i}$ and $\delta_{A,i}$), and a direct effect of affect at the end of the day on sleep the subsequent night (i.e., $\beta_{S,i}$); and distinct innovation variances for affect at the first, intermediate,

and last beeps of the day (i.e., $\psi_{A_1,i}$, $\psi_{A_o,i}$, and $\psi_{A_B,i}$), and for sleep (i.e., $\psi_{S,i}$). These ten parameters may differ across individuals, as indicated by the subscript i; this is also visualized in the upper right panel of Figure 3 by including these parameters as filled circles. Individual differences in these parameters and the means (obtained from the decomposition of A and S), are further modeled at the between-person level.

Between-person differences

The individual differences may be further studied to see whether and how they are related to each other and to other person characteristics. For instance, evening affect may play an important role in sleep quality the subsequent night for some individuals, but not for others. Similarly to the daytime versus nighttime autoregressive model, we are interested in the extent to which these individual differences may be explained by neuroticism and chronotype, and how they are related to each other.

The twelve random effects that characterize this model are presented in Table 2. The (residuals of these) random effects are allowed to be correlated, in the same way as in the previous model for affect only. This is also visualized in the lower right panel of Figure 3.

Conclusion

We have developed a new dynamic multilevel model that allows researchers to study the dynamics between asynchronously measured sleep and affect in a single model. To this end, we first proposed a daytime and nighttime autoregressive model that allows us to account for different daytime and nighttime autoregression. This is more realistic than assuming a priori that the nighttime process is a simple continuation of the daytime process. The reciprocal sleep-affect model further builds on this by including sleep as a second variable, and considering ways in which sleep may affect momentary affect throughout the day, and how momentary affect at the end of the day may affect subsequent sleep.

Model implementation

The two dynamic multilevel models that we developed above do not seamlessly fall into the structure of existing multilevel software. While flexible Bayesian software like WinBUGS or Stan allows researchers to specify these models from scratch, it requires defining priors for the variances of random effects—a particularly challenging task for autoregressive parameters, which inherently have very small variances (see Schuurman, Grasman, & Hamaker, 2016). Instead, we opted for Mplus (Muthén & Muthén, 2017), which provides a more user-friendly approach through its dynamic structural equation modeling (DSEM) module, making it a more accessible option for researchers. It uses Bayesian estimation with the Markov chain Monte Carlo Gibbs sampler and the Metropolis-Hastings sampler (Muthén & Muthén, 2017). Mplus also provides default, non-informative prior distributions, reducing the need for users to manually specify complex priors. Additionally, its DSEM module efficiently handles missing values by accounting for the time order of observations, assuming the data are missing at random (Asparouhov et al., 2018; Harvey, 1989). Researchers should evaluate the plausibility of this assumption before relying on this approach to handle missing values.

Typically in DSEM analysis, the repeated measures of the same variable are handled in long format, with all the repeated measures of a single person underneath each other. The model then uses a two-level approach with a within-person level for the repeated measures and a between-person level for the different individuals. However, such a model does not allow for the two distinct autoregressions of the first extension we proposed, nor does it easily allow for the particular reciprocal relations between daily sleep measures and momentary affect which characterize the second extension we proposed. Therefore, we will take an approach based on having the repeated measures within a day in wide format represented by different variables, while the repeated days are in long format, represented by different rows in the data file.

To ensure that the observations within the day, which are transformed from long to wide format, are (approximately) equally spaced in time, we begin by actively adding missing values in the long format data set-up. Subsequently, we explain how to reformat the long-format data into the required combination of wide-and-long format. These data preparations were both preregistered, and we explain them in more detail below. An `R` file with annotated code demonstrating these data preparation procedures is provided in the supplementary materials.

Handling unequal daytime intervals

Figure 4 shows the distribution of time interval lengths in the original data in black, with the daytime intervals in the left panel, and nighttime intervals in the right panel. These interval lengths were computed by taking the difference in time between the start time of one observation and the start time of the next

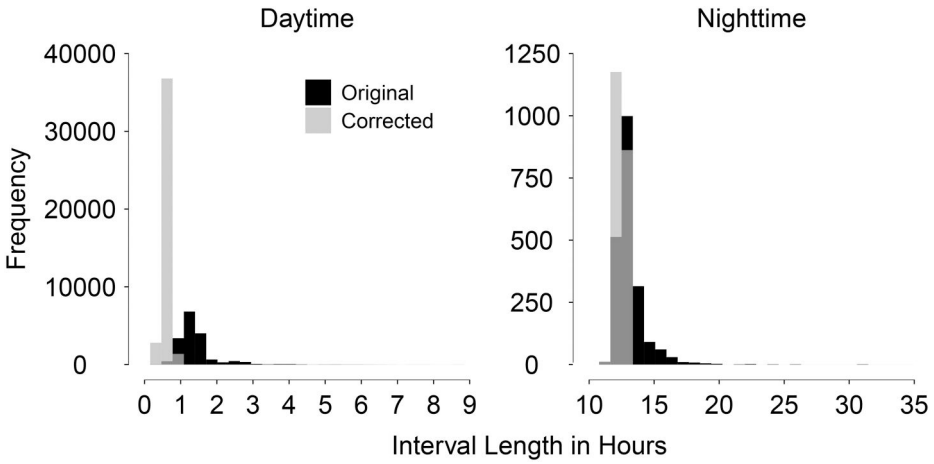

Figure 4. Histograms of interval lengths in hours for the original ESM intervals (black) and the corrected intervals where missing values were added to make the intervals approximately equal (grey) for the daytime intervals (left) and nighttime intervals (right). The dark grey color represents the part of the corrected interval lengths distribution that overlaps with the original intervals distribution.

observation; hence, missing values (i.e., missed beeps) between realized observations lead to longer intervals. The intervals were labeled as nighttime if it was between the last observation of one day and the first observation of the next day.

Although DSEM can handle such differences in time intervals when taking a long format approach (using its TINTERVAL option; Asparouhov et al., 2018; Hamaker, 2023), we will convert the measurements within a day into wide format, which means we need to actively account for the varying intervals. We do this by adding missing values in between observations separated by larger intervals so that the intervals between these observations and the added missing value are then reduced to half the original length. Our approach will ensure that every individual will have the same number of beeps per day, regardless of the number of observed beeps. We begin by adding ten missing values for every person every day. This is done in a sequential manner, adding between observations with the largest interval of that day, which reduces that interval by half, then again finding the largest interval of that day, etc. In the case of missing observations, we increased the number of missing values to be added by the number of missing observations in a day; for example, if there were only six out of the ten intended observations in a day, we added the regular ten missing values, plus an additional four, to give us twenty occasions for that day. We also took into account the interval between the individual's start time of the day and the first observation, and between the last observation and the individual's end time of the day, and added missing values prior to the first or after the last observation of the day if these intervals were the largest. This is relevant because the timing of the first and last observation plays a prominent role in the dynamic models that we want to estimate.

Figure 5 demonstrates this procedure for a hypothetical day with ten observations, where ten missing values are added (indicated by the ten steps and the division of the original interval into equal parts). In the first column, we see the ten beeps, and in the second column we see the time at which the beeps occurred between the starting time of 09:00 and the end time of 21:30. The third column shows the time intervals between two consecutive observed beeps, including the time interval between the start time and the first beep and the interval between the last beep and the end time, which range from 18 to 130 min. Then, the ten steps are shown from left to right, indicating where a missing value is added. The first missing value in step 1 is added to the interval between observations 4 and 5, which breaks up the 130 min into two new intervals of 65 min each. The second missing value is added to the interval between observations 6 and 7 (111 min), the third to the interval between 7 and 9 (89 min), and so on. Two of the original intervals were so long (130 and 111 min) that the divided intervals became the largest interval again at a later step requiring a second missing value to be added within the same interval (at step 6 and 9, respectively). This procedure is continued until ten missing values were added. The result is shown in the far right column, where beeps with actually obtained observations are printed in bold. While the intervals between these twenty beeps are still not equal the range is now between 18 and 44.5 (rather than 130) minutes.

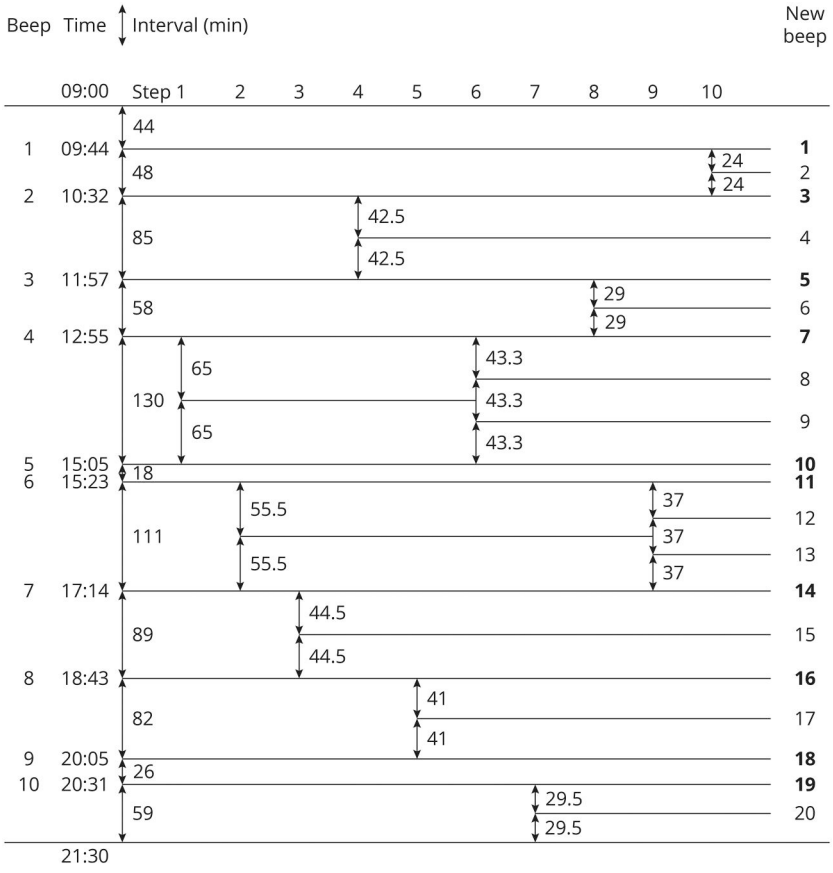

Figure 5. Example of correcting unequal intervals for one day with ten observations by adding ten missing values between observations. The time indicates a hypothetical response time for the beeps. The interval length in minutes is indicated on the right of the double-headed arrows. One missing value is added to the largest interval at each step for ten steps (from left to right), as indicated by the division of the original interval. The new beep numbers highlighted in bold indicate the observed beeps.

Using this procedure for the empirical data, the distribution of the new interval lengths (grey) compared to the distribution of the intervals in the original data (black) is shown in Figure 4. The daytime intervals of the original data range from 16 min to 8 h and 53 min with an average of 1 h and 24 min, while the new daytime intervals range from 16 to 55 min with an average of 37 min. The longest interval in the corrected data is thus roughly eight hours shorter than the longest interval in the original data. The nighttime intervals of the original data range from 11 h and 7 min to 31 h and 24 min[1] with an average of 13 h and 8 min, while the new nighttime intervals range from 11 h and 5 min to 13 h and 7 min with an average of 12 h and 26 min. Thus, the longest nighttime interval in the corrected data is about 18 h shorter than the longest interval in the original data. The percentage of missing values in the positive and negative affect variables was 13% before the correction, and 56% after correction.

[1]As the nighttime intervals were between the first observation of one day and the last observation of the previous day, the interval could cover daytime as well, for example, if a participant does not respond before 8 PM on one day, and had not responded after 1 PM the previous day, this is denoted as a nighttime interval of 31 h.

Handling different measurement frequencies

Table 3(a) provides an example of momentary affect and daily sleep measures from multiple individuals represented in long format, with each measurement occasion placed on a new row. This necessarily leads to missing values in the sleep variable that is only measured at the first measurement occasion; yet, it would not make sense to measure this at every beep within a day, and considering it therefore as missing observations seems inappropriate.

To be able to apply the models that we developed in the previous section using Mplus, we need to rearrange the data into a two-level long-and-wide format in which: (a) the beeps are in wide format as distinct variables (i.e., separate columns); and b) the days are in long format, with every day per individual on a separate row. Table 3(b) demonstrates this new long-and-wide format, which will allow us to estimate

Table 3. Data format examples for *I* individuals, *D* days, and three beeps.

(a) Long format

Indiv.	Day	Beep	S	M
1	1	1	S_{11}	M_{111}
1	1	2	–	M_{211}
1	1	3	–	M_{311}
1	2	1	S_{21}	M_{121}
1	2	2	–	M_{221}
1	2	3	–	M_{321}
2	1	1	S_{11}	M_{112}
⋮	⋮	⋮	⋮	
I	*D*	3	S_{DI}	M_{3DI}

(b) Long-and-wide format

Indiv.	Day	S	M_1	M_2	M_3
1	1	S_{11}	M_{111}	M_{211}	M_{311}
1	2	S_{21}	M_{121}	M_{221}	M_{321}
1	3	S_{31}	M_{131}	M_{231}	M_{331}
2	1	S_{12}	M_{112}	M_{212}	M_{312}
⋮	⋮	⋮	⋮	⋮	
I	*D*	S_{DI}	M_{1DI}	M_{2DI}	M_{3DI}

Note. Panel A shows the typical long-format data, whereas Panel B shows the required long-and-wide data format for estimation in the software Mplus. Indiv. is the individual, S is the daily sleep variable, and M is the momentary affect variable. This wider data format depicts each beep of the affect variable as a separate column. S_{11} is the sleep variable on the first day for the first individual, M_{111} is the first beep of the affect variable on the first day for the first individual, M_{211} is the second beep, and M_{31} is the third beep. An empty cell is represented by '-' which indicates these data are not available.

parameters on a beep, day, and individual level, and across beep and day level, using DSEM in Mplus.

Estimation

Due to the scale of the sleep quality, positive, and negative affect variables (0–100), they have a large variance which led to convergence problems. Therefore, we rescaled these variables by dividing them by 10 which should otherwise not influence the results. The Mplus code is provided in the supplementary materials available at https://osf.io/h3dkj/.

For the random effects estimates of the regression coefficients, we use the averages of the individually standardized parameter estimates, which were standardized using person-specific within-person variances (Hamaker et al., 2023; Schuurman, Ferrer, et al., 2016). As the between-person level is rather complex with many parameters, we may use a step-by-step approach when estimating the model to check the convergence and robustness of parameter estimates. In this step-by-step approach, we start with a relatively basic model and add random effects and between-level predictors sequentially. First, we estimated a model with only random average affect, then we included nighttime and daytime autoregression with covariances, next we included random innovation variances, and finally we added the between-level predictors. The parameter estimates were similar across steps, which indicates that the models converged and that the parameter estimates are robust. We further checked model convergence, by examining the trace plots of the MCMC samples which showed that the chains overlapped well indicating that the models converged. Below, we discuss the results of the full model (last step), and we provide the parameter estimates per step in the supplementary materials.

Results

Below, we first show the results of the univariate extension that allows for separate daytime and nighttime autoregression for PA and NA, and afterward, we show the results of the bivariate extension that includes reciprocal relations between PA/NA and sleep quality. The results of the tested hypotheses are summarized in Table 4, where $\mathcal{H}_1$ to $\mathcal{H}_4$ show the results of the daytime and nighttime autoregressive model and $\mathcal{H}_5$ to $\mathcal{H}_8$ show the results of the reciprocal sleep-affect model.

Daytime versus nighttime autoregression

Individuals reported an average PA of 55.11 (95% CI [52.40, 57.84]) and an average NA of 24.09 95% CI [21.34, 26.87] on a scale from zero to 100. Table 4 shows the averaged individually standardized parameter estimates for the two autoregressive effects (α_A and ϕ_A) as obtained with the daytime and nighttime autoregressive model ($\mathcal{H}_1$). Since the daytime autoregressions are only constrained to be equal on the between-person level, Mplus provides 19 separate averaged individually standardized estimates per daytime autoregression (A_2 on A_1, A_3 on A_2, etc.), so we took the median of these separate estimates to get a single daytime autoregression estimate. The averaged daytime autoregression for both PA and NA is large (PA: 0.65, 95% CI [0.64, 0.66]; NA: 0.64, 95% CI [0.62, 0.65]), which means that PA and NA tend to persist over time while controlling for individual differences in means, that is, the autoregression in the temporal deviation from the individual means. This parameter should be considered the average carryover over a 37-minute interval, as that is the average time length between data points (within the same day) after restructuring the data. Additionally, there is a smaller but still substantial nighttime autoregression (PA: 0.21, 95% CI [0.15, 0.27]; NA: 0.25, 95% CI [0.19, 0.65] which means that PA and NA also persist from evening to the next morning. This autoregression pertains to a time interval of 12 h and 26 min.

Table 4. Evidence for the hypotheses of the empirical example.

		Positive affect			Negative affect		
$\mathcal{H}$	Parameters	Median	95% CI	Evidence	Median	95% CI	Evidence
$\mathcal{H}_1$	α_A	0.21	(0.15, 0.27)	√	0.25	(0.19, 0.30)	√
	ϕ_A	0.65	(0.64, 0.66)	√	0.64	(0.62, 0.65)	√
$\mathcal{H}_2$	$\gamma_{\phi_A} - \gamma_{\alpha_A}$	0.46	(0.40, 0.53)	√	0.42	(0.33, 0.52)	√
$\mathcal{H}_3$	$\tau^2_{\alpha_A}$	0.04	(0.01, 0.08)		0.14	(0.08, 0.23)	√
	$\tau^2_{\phi_A}$	0.02	(0.01, 0.02)	√	0.02	(0.02, 0.03)	√
	$\tau^2_{\mu_A}$	1.92	(1.44, 2.63)	√	2.41	(2.13, 2.69)	√
	$\tau^2_{\psi_{A_1}}$	0.61	(0.42, 0.89)	√	0.67	(0.47, 0.86)	√
	$\tau^2_{\psi_{A_o}}$	0.53	(0.40, 0.72)	√	0.39	(0.27, 0.52)	√
$\mathcal{H}_4$	τ_{ϕ_A, α_A}	0.32	(−0.04, 0.64)		0.34	(0.07, 0.58)	√ [a]
$\mathcal{H}_5$	β_A	0.25	(0.20, 0.29)	√	−0.20	(−0.25, −0.15)	√
	δ_A	−0.02	(−0.08, 0.04)		−0.01	(−0.06, 0.05)	
$\mathcal{H}_6$	β_S	0.03	(−0.02, 0.09)		−0.04	(−0.09, 0.01)	
$\mathcal{H}_7$	$\tau^2_{\mu_S}$	1.57	(1.11, 2.28)	√	1.61	(1.13, 2.34)	√
	$\tau^2_{\phi_S}$	0.07	(0.04, 0.11)		0.06	(0.03, 0.11)	
	$\tau^2_{\beta_A}$	0.07	(0.04, 0.12)		0.05	(0.03, 0.09)	√
	$\tau^2_{\delta_A}$	0.05	(0.02, 0.10)		0.03	(0.01, 0.07)	
	$\tau^2_{\beta_S}$	0.06	(0.03, 0.12)		0.05	(0.02, 0.10)	
	$\tau^2_{\psi_S}$	0.68	(0.46, 1.00)	√	0.68	(0.47, 1.02)	√
$\mathcal{H}_8$	θ_{μ_A}	[b]			0.44	(0.27, 0.58)	√
	θ_{ϕ_A}	[b]			0.12	(−0.09, 0.31)	
	θ_{α_A}	[b]			0.03	(−0.22, 0.27)	
	ω_{β_A}	0.11	(−0.15, 0.36)		−0.25	(−0.50, 0.03)	
	ω_{δ_A}	−0.07	(−0.42, 0.26)		−0.08	(−0.41, 0.28)	
	ω_{β_S}	−0.27	(−0.53, 0.02)		0.23	(−0.08, 0.51)	

Note. $\mathcal{H}_1$ to $\mathcal{H}_4$ show the results of the daytime and nighttime autoregressive model and $\mathcal{H}_5$ to $\mathcal{H}_8$ show the results of the reciprocal sleep-affect model. The parameter notations are explained in Section "A new dynamic multilevel model". The median and credible interval (CI) are based on the individually standardized parameter estimates for α_A, ϕ_A, β_A, δ_A, and β_S; standardized parameter estimates for τ_{ϕ_A, α_A}, θ, and ω; and unstandardized parameter estimates for $\gamma_{\phi_A} - \gamma_{\alpha_A}$ and τ^2. The evidence tick is based on whether the credible interval contains zero except for the residual variance parameters of the random effects (τ^2) which are based on whether the random effects are predicted by or relate to other variables. [a] In the results of the reciprocal sleep-affect model there was no evidence for the relation between the daytime and nighttime autoregression. [b] We did not specify hypotheses on these relations for positive affect.

If we had not separated the nighttime from daytime autoregression and estimated the model without taking the differences of the daytime and nighttime intervals into account, we would implicitly assume that the daytime and nighttime autoregression are the same. This approach implies that the process pauses during the nighttime (Berkhout et al., 2024). We wanted to determine whether separating nighttime from daytime autoregression was more appropriate than not doing so, so we looked at the difference between nighttime and daytime autoregression ($\mathcal{H}_2$). To get an uncertainty estimate for this difference, we subtract the unstandardized fixed average of the nighttime autoregression (PA: $\gamma_{\alpha_A} = 0.19$; NA: $\gamma_{\alpha_A} = 0.22$) from the unstandardized fixed average of the daytime autoregression (PA: $\gamma_{\phi_A} = 0.65$; NA: $\gamma_{\phi_A} = 0.64$) at every MCMC iteration, which gives us $\gamma_{\phi_A} - \gamma_{\alpha_A} = 0.46$ (95% CI [0.40, 0.53]) for PA and $\gamma_{\phi_A} - \gamma_{\alpha_A} = 0.42$ (95% CI [0.33, 0.52]) for NA. This means that the autoregressive effect is 0.46 larger between PA observations during the day compared to overnight, holding all else constant. Thus, the data support separating daytime and nighttime autoregression over constraining them to be equal.

Figure 6 visualizes the between-person results when we use the observed variable neuroticism to predict the random effects for positive affect (left) and negative affect (right). The predictive relations are indicated by one-headed arrows, while the correlations between the residuals of the random effects are represented by non-directed lines; solid lines represent positive relations, whereas dashed lines represent negative parameters. Although both neuroticism and chronotype were included, the latter was left out of the figure since we did not find evidence that it explained any variation in the random effects. The plot shows that neuroticism negatively predicts the average PA per individual (−0.44, 95% CI [−0.57, −0.27]), and positively predicts the average NA (0.47, 95% CI [0.31, 0.60]), indicating that individuals who score higher on neuroticism tend to experience less PA and more NA on average ($\mathcal{H}_8$). Furthermore, neuroticism positively predicts the innovation variance of the other (i.e., not the first) beeps (PA: 0.21, 95% CI [0.02, 0.38]; NA: (0.40, 95% CI [0.21, 0.57]) and only for NA the innovation variance of the first beep (0.39, 95% CI [0.21, 0.53]), indicating that individuals with higher neuroticism have more unexplained variance.

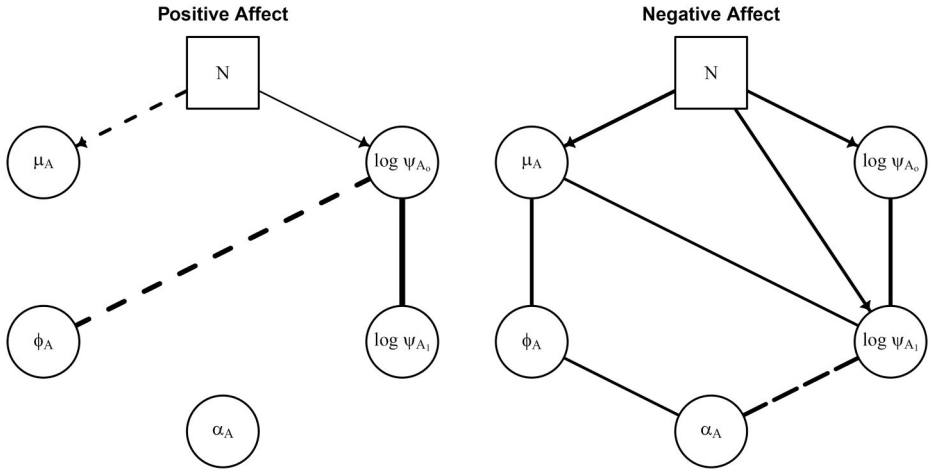

Figure 6. Between-person effects of observed predictor Neuroticism (N) on the random effects (single-headed arrows) and correlations between the random effects (lines without arrows) of the day versus night autoregressive model for positive affect and negative affect. Solid lines represent positive coefficients, dashed lines represent negative coefficients, and the width of the lines indicates the strength of a coefficient. μ_A is the average, ϕ_A is the daytime autoregression, α_A is the nighttime autoregression, $\log(\psi_{A_1})$ is the innovation variance of the first beep, and $\log(\psi_{A_o})$ is the innovation variance of the other beeps.

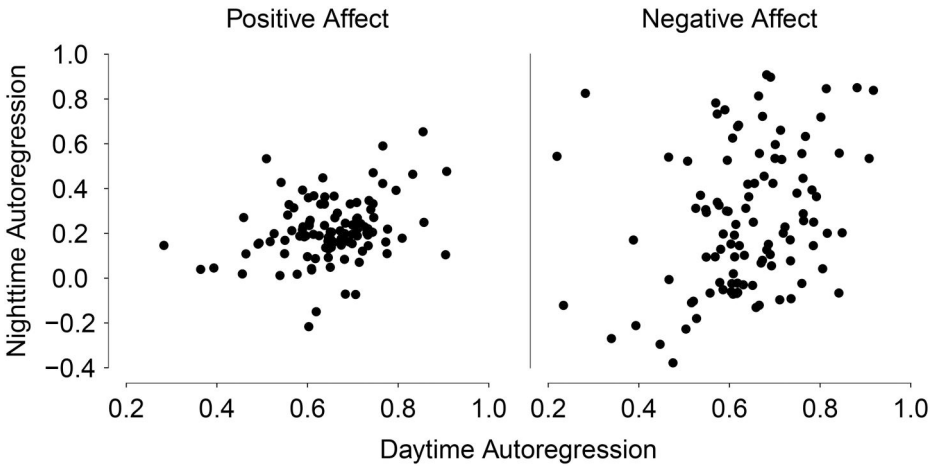

Figure 7. Scatter plot of individually standardized estimates of daytime autoregression (*X*-axis) and nighttime autoregression (*Y*-axis) for positive affect (left) and negative affect (right).

This may reflect their heightened reactivity to external influences.

The (residual) variances of the random effects indicate the degree to which there are individual differences in the estimated individual parameters ($\mathcal{H}_3$). However, we cannot use their credible intervals to make inferences about the population on the existence of individual differences, as the lower bound will always be above zero. We can, however, infer there are individual differences in the parameters when they correlate with other random effects or when it is predicted by a between-level variable. Thus, based on Figure 6, we can conclude there is evidence that there are individual differences in all random effects except for the PA nighttime autoregression.

Interestingly, for PA the daytime autoregression is not correlated with the nighttime autoregression (0.32, 95% CI [−0.04, 0.64]), which indicates that individuals with a higher PA daytime autoregression do not necessarily also have a higher PA nighttime autoregression,[2] and vice versa ($\mathcal{H}_4$). In contrast, the NA nighttime autoregression is positively correlated with the NA daytime autoregression (0.34, 95% CI [0.07, 0.58]). Figure 7 visualizes the relation between the individually standardized daytime autoregression and nighttime autoregression using a scatter plot. While there appears to be a small positive correlation between the parameters of PA, the uncertainty around

[2]This conclusion is based on the credible interval containing zero. There are alternative approaches, such as the probability of direction which is a measure of certainty about the direction (positive or negative) of an effect (Makowski et al., 2019), but they fall beyond the scope of this paper.

the estimate is too large to conclude that there is a correlation. Additionally, the variance of the daytime autoregression estimates of NA seems similar to PA, but the variance of the nighttime autoregression of NA is much larger than of PA.

To assess how well the model fits the data, we look at the proportion of explained variance (R^2). At the within-person level, the nighttime autoregression explains on average 9% of the variance in positive affect and 19% of the variance in negative affect on the first beep. The daytime autoregression explains 43% of the variance in positive and negative affect on the rest of the beeps. At the between-person level, the proportions of explained variance indicate how much variance in the parameter is explained by neuroticism and chronotype. They explain 23% of the individual differences in the average of negative affect; 20% in the average of positive affect; 19% in the NA innovation variance of the first beep; 17% in the NA innovation variance of the other beeps; 6% in the PA innovation variance of the first beep; and 5% or less in the other parameters.

Reciprocal sleep-affect relations

Table 4 shows that last night's sleep quality has a positive effect on the first beep of PA (β_A; 0.25, 95% [0.20, 0.29]) and a negative effect on the first beep of NA (−0.20, 95% [−0.25, −0.15]) which means that better sleep quality leads to a higher PA and lower NA in the morning ($\mathcal{H}_5$). However, there is no delayed direct effect from sleep to PA/NA at the end of the day (δ_A), nor does PA/NA at the end of the day predict sleep the following night (β_S; $\mathcal{H}_6$). Another interesting effect that was not part of the tested hypotheses is the sleep quality autoregression: there is a small positive average autoregression from night to night (0.10, 95% CI [0.05, 0.14]) which means that nights with good sleep quality follow nights with good sleep quality, and nights with poor sleep quality tend to follow nights with poor sleep quality.

Figure 8 visualizes the between-person relations for the reciprocal sleep-affect model, with chronotype left out again due to a lack of evidence. Due to the between-level and correlation effects shown in the figure, we can infer that there are individual differences in the random effects ($\mathcal{H}_7$) except for: the cross-lagged effect from yesterday's affect to subsequent sleep quality (β_S); the delayed cross-regression effect from last night's sleep to affect at the end of the day (δ_A); and the sleep quality autoregression (ϕ_S). Furthermore, for PA only, we cannot infer individual differences in the cross-regression effect from last night's sleep quality to affect at the first beep (β_A) and the nighttime autoregression (α_A). In addition, the figure shows that neuroticism negatively predicts the average individual PA (−0.401, 95% CI [−0.55 − 0.23]) and the average individual sleep quality (−0.28, 95% CI [−0.45, −0.08]), indicating that individuals with higher neuroticism tend to have lower means on PA and sleep quality. Finally, neuroticism positively

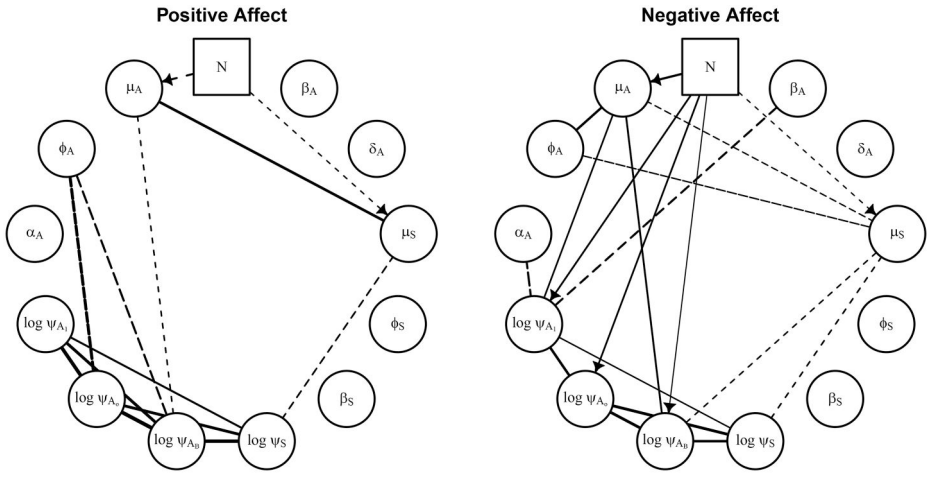

Figure 8. Between-person effects of observed predictor Neuroticism (N) on the random effects (single-headed arrows) and correlations between the random effects (lines without arrows) of the reciprocal sleep-affect model for positive affect and negative affect. Solid lines represent positive coefficients, dashed lines represent negative coefficients, and the width of the lines indicates the strength of a coefficient. μ_A is the average, ϕ_A is the daytime autoregression, α_A is the nighttime autoregression, $\log(\psi_{A_1})$ is the innovation variance of the first beep, $\log(\psi_{A_o})$ is the innovation variance of the other beeps, $\log(\psi_{A_B})$ is the innovation variance of the last beep, $\log(\psi_S)$ is the innovation variance of sleep quality, β_S is the effect from the last beep of the day to sleep quality the subsequent night, ϕ_S is the autoregression of sleep quality, μ_S is the average of sleep quality, δ_A is the effect from last night's sleep quality to the last beep of the day, and β_A is the effect from last night's sleep quality to the first beep of the day.

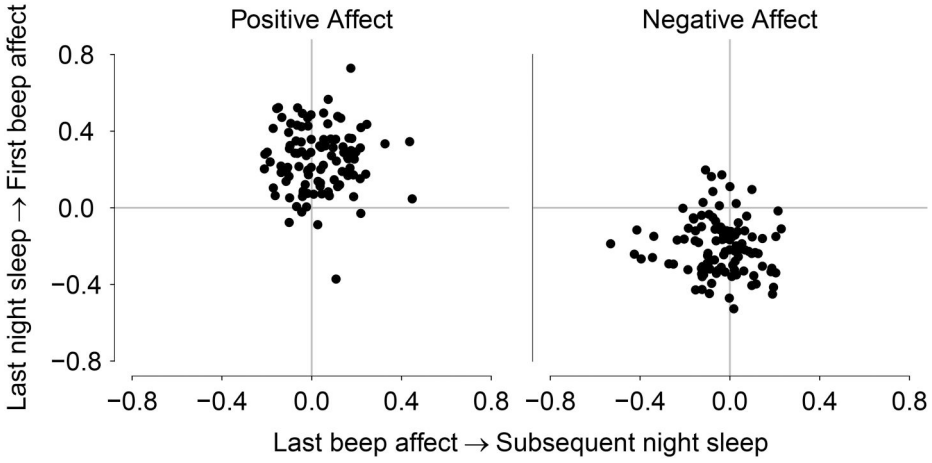

Figure 9. Scatter plot of individually standardized estimates of the cross-lagged regression from yesterday's affect at the last beep to sleep quality the subsequent night (β_S; *X*-axis) and the cross-regression from last night's sleep quality to today's affect at the first beep (β_A; *Y*-axis) for positive affect (left) and negative affect (right).

predicts average NA per individual (0.44, 95% CI [0.27, 0.58]) and the innovation variances ($\theta_{\psi_{A_1}}$: 0.36 95% CI [0.16, 0.53]; $\theta\psi_{A_o}$: 0.35 95% CI [0.17, 0.50]; $\theta_{\psi_{A_B}}$: 0.22 95% CI [0.01, 0.41]) indicating that individuals with higher neuroticism experience more NA ($\mathcal{H}_8$) and have more unexplained variance.

Interestingly, the (residuals of) NA daytime and nighttime autoregression do not correlate with each other, while this was the case in the daytime and nighttime autoregressive model. While many random effects do relate to each other as shown in the figure, we will focus on a couple of these relations between the sleep and affect parameters. The innovation variances of affect and sleep are all positively correlated with each other after accounting for neuroticism. Furthermore, the averages of PA/NA and sleep quality remain correlated, even after accounting for their relation with neuroticism (PA: 0.51, 95% CI [0.31, 0.67]; NA: −0.27, 95% CI [−0.48, −0.03]) indicating that—even within the same level of neuroticism—individuals with higher PA and lower NA have better sleep quality. Another interesting relation is that average sleep is negatively correlated with the NA daytime autoregression (−0.28, 95% CI [−0.50, −0.03]) indicating that individuals with poorer sleep quality have higher NA daytime autoregression.

The results indicate that sleep quality is associated with morning affect, but that this is not reciprocated by evening affect. Additionally, the random effects of the reciprocal relations do not correlate with each other. To assess the reciprocal relations on the individual level, we visualize the reciprocal effects estimated per individual in Figure 9, by plotting the individually standardized effects from affect at the end of the day to subsequent sleep against those from sleep to affect at the start of the day. While for most individuals the effect from sleep to the first beep of affect is positive for PA and negative for NA, the effect from the last beep of affect to sleep is centered around zero. If there were reciprocal effects, we would expect the estimates to be away from zero and either in the upper right corner or lower left corner.

Finally, the proportion of explained variance of the model at the within-person level shows that sleep quality of the previous night and PA and NA at the last beep of the previous day explain on average 12% and 11%, respectively, of the variance in sleep quality. For PA and NA at the first beep, the previous day's PA and NA at the last beep and last night's sleep quality explain 21% and 25% of the variance, respectively. Moreover, the daytime autoregression explains 44% and 42% of the variance in the rest of the PA and NA beeps, and together with the last night's sleep quality they explain 50% and 54% of the variance in the last PA and NA beep, respectively. At the between-person level, neuroticism and chronotype explain 18% and 20% of the variance in individual differences in average PA and NA, respectively; 15%, 14%, and 6% in the NA innovation variances of the first, other, and last beeps, respectively; 12% and 11% of the effect from the previous day's PA and NA, respectively, at the last beep to sleep quality the subsequent night; 8% in average sleep quality; 8% and 6% in the effects from last night's sleep quality to NA at the first and last beep, respectively; and 5% or less in the other parameters.

Conclusion

The results of the daytime and nighttime autoregressive model support the hypotheses that there is nighttime

autoregression and that the daytime autoregression is larger than the nighttime autoregression for both PA and NA ($\mathcal{H}_1$ & $\mathcal{H}_2$). These findings are in line with earlier research on daytime emotional inertia and overnight emotional inertia (e.g., Frérart et al., 2023; Minaeva et al., 2021). The nonzero nighttime autoregressions suggest that emotional processes do not stop or reset at night and the differences between daytime and nighttime autoregression suggest that the processes do not pause at night. Furthermore, they suggest they do not continue like they would during the daytime either: The length of the nighttime intervals was on average 20 times longer than that of the daytime intervals. This implies that if the process follows an AR(1) process that continues at night, we would expect the nighttime autoregression to be ϕ^{20} (for PA $0.65^{20} \approx 0.00018$ and for NA $0.64^{20} \approx 0.00013$). These values are not contained in the credible intervals for the nighttime autoregressions (they are estimated to be considerably larger than zero), which implies that a continuous time version of the AR(1) model is not appropriate for handling the night gaps in these data. For both PA and NA, we found individual differences in average affect, daytime autoregression, and innovation variances ($\mathcal{H}_3$). Furthermore, for NA we found individual differences in the nighttime autoregression but not for PA. With regard to $\mathcal{H}_4$, we find that the nighttime autoregression is correlated with the daytime autoregression for NA, but not for PA after accounting for neuroticism, which may be due to the lack of individual differences in the PA nighttime autoregression.

For the reciprocal sleep-affect model, results support part of $\mathcal{H}_5$, as sleep quality is associated with both PA and NA the next morning, but not at the end of the day. In contrast, we did not find an effect from PA or NA at the end of the day to sleep quality the subsequent night ($\mathcal{H}_6$). This implies that the sleep-affect relation we found is one-sided and not reciprocal, therefore contributing to the inconsistent findings on the relation between affect and sleep quality the subsequent night (Konjarski et al., 2018; Neubauer et al., 2021). We found evidence for individual differences in most of the model parameters; exceptions were the sleep autoregression, the cross-lagged effect from evening affect to sleep, the cross-regression effect from sleep to evening affect, and the PA cross-regression effect from sleep to morning affect ($\mathcal{H}_7$). For $\mathcal{H}_8$, we only found support that neuroticism positively predicts individual differences in average NA. Contrary to our expectations, neuroticism did not predict NA daytime or nighttime inertia. Finally, surprisingly, chronotype did not predict any of the individual differences in the random effects, suggesting that being a morning or evening person does not affect the dynamics of affect and sleep quality.

Discussion

In this paper we introduced a dynamic multilevel model that can capture (1) how a process evolves differently over day and night, and (2) the reciprocal relations between daily and momentary measured variables, in a single modeling approach. Additionally, the model allows researchers to study individual differences in these effects in more detail, through their relation with each other and whether they can be explained by between-level predictors. We proposed a model that allows researchers to study the dynamics between asynchronously measured variables and demonstrated the substantive relevance of these models through an empirical example of daily sleep and momentary affect. We adhered to the preregistered analysis plan (using DSEM in Mplus) to ensure the conducted analyses were confirmatory.

Note that researchers may easily adjust these models to suit different research interests. For example, they may be applied to other ESM variables, such as stress or anxiety, and different between-level predictors may be used, such as depression. Researchers may also specify a different between-level model, for example, where a latent variable accounts for commonalities between the random effects, or a path analysis or mediation model (see Hamaker et al., 2023). Furthermore, one may add or remove cross(-lagged) effects between sleep quality and the momentary variable, for example, in the current study, last night's sleep quality did not affect the last affect beep of the day so this may not be relevant to estimate. Another interesting extension of this model would be to include a day-level component for the affect variable, to study the reciprocal relations between sleep quality and this daily affect component. Our results show that last night's sleep quality relates to affect in the morning, but affect in the afternoon did not relate to sleep quality the subsequent night. It could be that it is not affect at the end of the day that relates to sleep quality, but that the overall affect of the day has an effect, as previous studies showed inconsistent findings of the effect of affect on sleep quality (Konjarski et al., 2018). Berkhout et al. (2025) proposed an $N = 1$ approach to this model and found that a daily component for self-doubt predicted sleep quality the subsequent night. For the current paper, a multilevel approach to

this model was not preregistered and did not converge well in Mplus.

A limitation of the empirical example is that the nighttime autoregression does not represent a period of only sleep, as nighttime is defined as the period between sampling periods, which was on average 12 h and 26 min long. As the average sleep duration of the participants was likely shorter than this, participants were possibly awake for some time before and after sleep during the designated nighttime periods when no momentary affect measures were obtained. Further research may implement a design where the last measurement of the day is collected right before participants go to sleep and the first measurement right after participants wake up so that the nighttime reflects a period of sleep only and the nighttime autoregression can be truly interpreted as the sleep-time autoregression.

Another limitation is that the uncertainty for some parameter estimates was large in our empirical example. While having a design with 103 individuals measured for 21 days with 10 beeps per day (resulting in 21,630 data points, of which 18,868 were observed) can be considered a fairly large dataset in the context of ESM research, the models that are used are rather complex and therefore demanding. For instance, while the daytime versus nighttime autoregressive model has 30 model parameters (in addition to the individual parameters), the reciprocal sleep-affect model has no less than 90 model parameters. The large uncertainty was particularly evident for the between-level covariances and regression coefficients, as reflected in their wide credible intervals. Hence, future research using simulations to determine sample sizes required to have reasonable certainty would be beneficial (e.g., Schultzberg & Muthén, 2018). Alternatively, researchers may reduce the number of parameters to increase certainty. For instance, removing the covariances between the random effects or only incorporating between-level effects of some between-level variables on some random effects that are of substantive interest. Choosing more informative priors, if researchers have good reason (e.g., previous research results) to expect a parameter to lie within a certain range, may also increase certainty. Additionally, for inferences about the existence of individual differences, we only checked whether the random effects correlated with others or if a between-level variable predicted them. If at least one of these criteria is met, this provides evidence for the existence of individual differences. However, if neither criterion is met, this should not be taken as evidence for the absence of individual differences. In such cases, one could turn to descriptive methods to explore meaningful individual differences, such as looking at the distributions, scatter plots such as Figures 7 and 9, or 95% heterogeneity intervals (see e.g., Bolger et al., 2019).

Overall, our findings support the idea that a process may evolve differently during the daytime compared to the nighttime, and that we may find relations between asynchronously measured variables. Furthermore, individual differences in these dynamics may provide interesting insights into how dynamic processes differ from person to person. The proposed dynamic model lays a foundation for future research to further explore the interplay between daily and momentary psychological processes.

Article information

Conflict of Interest Disclosures: Each author signed a form for disclosure of potential conflicts of interest. No authors reported any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The authors affirm having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This research is funded by the European Research Council (ERC) Consolidator Grant awarded to Ellen L. Hamaker (ERC-2019-COG-865468), a FWO postdoctoral mandate grant awarded to Koen Niemeijer, and a C1 grant (C14/23/062) from the KU Leuven Research Council to Peter Kuppens.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas and opinions expressed herein are those of the authors alone, and endorsement by the authors' institutions, the European Research Council, or the KU Leuven Research Council is not intended and should not be inferred.

References

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 359–388. https://doi.org/10.1080/10705511.2017.1406803

Baglioni, C., Nanovska, S., Regen, W., Spiegelhalder, K., Feige, B., Nissen, C., Reynolds, C. F., III., & Riemann, D. (2016). Sleep and mental disorders: A meta-analysis of polysomnographic research. *Psychological Bulletin*, *142*(9), 969–990. https://doi.org/10.1037/bul0000053

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research*, *52*(6), 789–804. https://doi.org/10.1080/00273171.2017.1373014

Benca, R. M., Obermeyer, W. H., Thisted, R. A., & Gillin, J. C. (1992). Sleep and psychiatric disorders. A meta-analysis. *Archives of General Psychiatry*, *49*(8), 651–668. discussion 669–670. https://doi.org/10.1001/archpsyc.1992.01820080059010

Berkhout, S. W., Schuurman, N. K., & Hamaker, E. (2025, January). *How to model ambulatory assessments measured at different frequencies: An N = 1 approach*. Retrieved January 8, 2025, from https://osf.io/6e9w7

Berkhout, S. W., Schuurman, N. K., & Hamaker, E. L. (2024, May). *Let sleeping dogs lie? How to deal with the night gap problem in ESM data*. https://doi.org/10.31234/osf.io/2hec8

Boker, S. M. (2002). Consequences of continuity: The hunt for intrinsic properties within parameters of dynamics in psychological processes. *Multivariate Behavioral Research*, *37*(3), 405–422. https://doi.org/10.1207/S15327906MBR3703_5

Bolger, N., Zee, K. S., Rossignac-Milon, M., & Hassin, R. R. (2019). Causal processes in psychology are heterogeneous. *Journal of Experimental Psychology. General*, *148*(4), 601–618. https://doi.org/10.1037/xge0000558

Bouwmans, M. E. J., Bos, E. H., Hoenders, H. J. R., Oldehinkel, A. J., & de Jonge, P. (2017). Sleep quality predicts positive and negative affect but not vice versa. An electronic diary study in depressed and healthy individuals. *Journal of Affective Disorders*, *207*, 260–267. https://doi.org/10.1016/j.jad.2016.09.046

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., Tuerlinckx, F., & Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, *23*(4), 425–435. https://doi.org/10.1177/1073191116645909

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

Brockwell, P. J., & Davis, R. A. (1993). *Time series: Theory and methods* (2nd ed.). Springer.

Brose, A., Schmiedek, F., Koval, P., & Kuppens, P. (2015). Emotional inertia contributes to depressive symptoms beyond perseverative thinking. *Cognition & Emotion*, *29*(3), 527–538. https://doi.org/10.1080/02699931.2014.916252

Carciofo, R. (2023). Morning affect or sleep inertia? Comparing the constructs and their measurement. *Chronobiology International*, *40*(4), 458–472. https://doi.org/10.1080/07420528.2023.2187211

Cipra, T., & Romera, R. (1997). Kalman filter with outliers and missing observations. *Test*, *6*(2), 379–395. https://doi.org/10.1007/BF02564705

Cousins, J. C., Whalen, D. J., Dahl, R. E., Forbes, E. E., Olino, T. M., Ryan, N. D., & Silk, J. S. (2011). The bidirectional association between daytime affect and nighttime sleep in youth with anxiety and depression. *Journal of Pediatric Psychology*, *36*(9), 969–979. https://doi.org/10.1093/jpepsy/jsr036

Dejonckheere, E., Mestdagh, M., Houben, M., Rutten, I., Sels, L., Kuppens, P., & Tuerlinckx, F. (2019). Complex affect dynamics add limited information to the prediction of psychological well-being. *Nature Human Behaviour*, *3*(5), 478–491. https://doi.org/10.1038/s41562-019-0555-0

Denissen, J. J. A., Geenen, R., Van Aken, M. A. G., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment*, *90*(2), 152–157. https://doi.org/10.1080/00223890701845229

Driver, C. C., Oud, J. H. L., & Voelkle, M. C. (2017). Continuous time structural equation modeling with R Package CTSEM. *Journal of Statistical Software*, *77*(5), 1–35. https://doi.org/10.18637/jss.v077.i05

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research*, *53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Flueckiger, L., Lieb, R., Meyer, A. H., Witthauer, C., & Mata, J. (2017). Day-to-day variations in health behaviors and daily functioning: Two intensive longitudinal studies. *Journal of Behavioral Medicine*, *40*(2), 307–319. https://doi.org/10.1007/s10865-016-9787-x

Fortier, M. S., Guerin, E., Williams, T., & Strachan, S. (2015). Should I exercise or sleep to feel better? A daily analysis with physically active working mothers. *Mental Health and Physical Activity*, *8*, 56–61. https://doi.org/10.1016/j.mhpa.2015.03.001

Frérart, L., Bilsen, L., Dejonckheere, E., & Kuppens, P. (2023). Overnight emotional inertia in relation to depressive symptomatology and subjective sleep quality. *Sleep Advances: A Journal of the Sleep Research Society*, *4*(1), zpac048. https://doi.org/10.1093/sleepadvances/zpac048

Fuligni, A. J., & Hardway, C. (2006). Daily variation in adolescents' sleep, activities, and psychological well-being. *Journal of Research on Adolescence*, *16*(3), 353–378. https://doi.org/10.1111/j.1532-7795.2006.00498.x

Garcia, C., Zhang, L., Holt, K., Hardeman, R., & Peterson, B. (2014). Latina adolescent sleep and mood: An ecological momentary assessment pilot study. *Journal of Child and Adolescent Psychiatric Nursing: Official Publication of the Association of Child and Adolescent Psychiatric Nurses, Inc*, *27*(3), 132–141. https://doi.org/10.1111/jcap.12082

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *NeuroImage*, *63*(1), 310–319. https://doi.org/10.1016/j.neuroimage.2012.06.026

Goldstein, A. N., & Walker, M. P. (2014). The role of sleep in emotional brain function. *Annual Review of Clinical Psychology*, *10*(1), 679–708. https://doi.org/10.1146/annurev-clinpsy-032813-153716

Hamaker, E. L. (2023). The within-between dispute in cross-lagged panel research and how to move forward. *Psychological Methods*. https://doi.org/10.1037/met0000600

Hamaker, E. L., Asparouhov, T., & Muthén, B. (2023). Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (2nd ed.). The Guilford Press.

Harvey, A. C. (1989). *Forecasting, structural time series models, and the Kalman filter*. Cambridge University Press.

Houben, M., Van Den Noortgate, W., & Kuppens, P. (2015). The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, *141*(4), 901–930. https://doi.org/10.1037/a0038822

Inoue, A. (2008). AR and MA representation of partial autocorrelation functions, with applications. *Probability Theory and Related Fields*, *140*(3–4), 523–551. https://doi.org/10.1007/s00440-007-0074-1

Jankowski, K. S. (2014). The role of temperament in the relationship between morningness–eveningness and mood. *Chronobiology International*, *31*(1), 114–122. https://doi.org/10.3109/07420528.2013.829845

Jongerling, J., Laurenceau, J.-P., & Hamaker, E. L. (2015). A multilevel AR(1) model: Allowing for inter-individual differences in trait-scores, inertia, and innovation variance. *Multivariate Behavioral Research*, *50*(3), 334–349. https://doi.org/10.1080/00273171.2014.1003772

Kalmbach, D. A., Pillai, V., Roth, T., & Drake, C. L. (2014). The interplay between daily affect and sleep: A 2-week study of young women. *Journal of Sleep Research*, *23*(6), 636–645. https://doi.org/10.1111/jsr.12190

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

Könen, T., Dirk, J., Leonhardt, A., & Schmiedek, F. (2016). The interplay between sleep behavior and affect in elementary school children's daily life. *Journal of Experimental Child Psychology*, *150*, 1–15. https://doi.org/10.1016/j.jecp.2016.04.003

Konjarski, M., Murray, G., Lee, V. V., & Jackson, M. L. (2018). Reciprocal relationships between daily sleep and mood: A systematic review of naturalistic prospective studies. *Sleep Medicine Reviews*, *42*, 47–58. https://doi.org/10.1016/j.smrv.2018.05.005

Koval, P., Brose, A., Pe, M. L., Houben, M., Erbas, Y., Champagne, D., & Kuppens, P. (2015). Emotional inertia and external events: The roles of exposure, reactivity, and recovery. *Emotion (Washington, D.C.)*, *15*(5), 625–636. https://doi.org/10.1037/emo0000059

Koval, P., Burnett, P. T., & Zheng, Y. (2021). Emotional inertia: On the conservation of emotional momentum. In C. E. Waugh & P. Kuppens (Eds.), *Affect dynamics* (pp. 63–94). Springer Cham.

Koval, P., Sütterlin, S., & Kuppens, P. (2015). Emotional inertia is associated with lower well-being when controlling for differences in emotional context. *Frontiers in Psychology*, *6*, 1997. https://doi.org/10.3389/fpsyg.2015.01997

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*(7), 984–991. https://doi.org/10.1177/0956797610372634

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework. *Journal of Open Source Software*, *4*(40), 1541. https://doi.org/10.21105/joss.01541

Miller, M. A., Rothenberger, S. D., Hasler, B. P., Donofry, S. D., Wong, P. M., Manuck, S. B., Kamarck, T. W., & Roecklein, K. A. (2015). Chronotype predicts positive affect rhythms measured by ecological momentary assessment. *Chronobiology International*, *32*(3), 376–384. https://doi.org/10.3109/07420528.2014.983602

Minaeva, O., George, S. V., Kuranova, A., Jacobs, N., Thiery, E., Derom, C., Wichers, M., Riese, H., & Booij, S. H. (2021). Overnight affective dynamics and sleep characteristics as predictors of depression and its development in women. *Sleep*, *44*(10), zsab129. https://doi.org/10.1093/sleep/zsab129

Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

Neubauer, A. B., Kramer, A. C., Schmidt, A., Könen, T., Dirk, J., & Schmiedek, F. (2021). Reciprocal relations of subjective sleep quality and affective well-being in late childhood. *Developmental Psychology*, *57*(8), 1372–1386. https://doi.org/10.1037/dev0001209

Niemeijer, K., Mestdagh, M., Verdonck, S., Meers, K., & Kuppens, P. (2023). Combining experience sampling and mobile sensing for digital phenotyping with m-path sense: Performance study. *JMIR Formative Research*, *7*(1), e43296. https://doi.org/10.2196/43296

Palmer, C. A., Bower, J. L., Cho, K. W., Clementi, M. A., Lau, S., Oosterhoff, B., & Alfano, C. A. (2024). Sleep loss and emotion: A systematic review and meta-analysis of over 50 years of experimental research. *Psychological Bulletin*, *150*(4), 440–463. https://doi.org/10.1037/bul0000410

Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 495–515. https://doi.org/10.1080/10705511.2017.1392862

Schuurman, N. K., Grasman, R. P. P. P., & Hamaker, E. L. (2016). A comparison of inverse-wishart prior specifications for covariance matrices in multilevel autoregressive models. *Multivariate Behavioral Research*, *51*(2-3), 185–206. https://doi.org/10.1080/00273171.2015.1065398

Schuurman, N. K., Ferrer, E., de Boer-Sonnenschein, M., & Hamaker, E. L. (2016). How to compare cross-lagged associations in a multilevel autoregressive model. *Psychological Methods*, *21*(2), 206–221. https://doi.org/10.1037/met0000062

Shin, K. E., Newman, M. G., & Jacobson, N. C. (2022). Emotion network density is a potential clinical marker for anxiety and depression: Comparison of ecological momentary assessment and daily diary. *The British Journal of Clinical Psychology*, *61 Suppl 1*(Suppl 1), 31–50. https://doi.org/10.1111/bjc.12295

Sikka, P., & Gross, J. J. (2023). Affect across the wake-sleep cycle. *Affective Science*, *4*(3), 563–569. https://doi.org/10.1007/s42761-023-00204-2

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Stolarski, M., Jankowski, K. S., Matthews, G., & Kawalerczyk, J. (2016). Wise "birds" follow their clock: The role of emotional intelligence and morningness–eveningness in diurnal regulation of mood. *Chronobiology International*, *33*(1), 51–63. https://doi.org/10.3109/07420528.2015.1115413

Suls, J., Green, P., & Hillis, S. (1998). Emotional reactivity to everyday problems, affective inertia, and neuroticism. *Personality and Social Psychology Bulletin*, *24*(2), 127–136. https://doi.org/10.1177/0146167298242002

Suls, J., & Martin, R. (2005). The daily life of the garden-variety neurotic: Reactivity, stressor exposure, mood spillover, and maladaptive coping. *Journal of Personality*, *73*(6), 1485–1509. https://doi.org/10.1111/j.1467-6494.2005.00356.x

Takano, K., Sakamoto, S., & Tanno, Y. (2014). Repetitive thought impairs sleep quality: An experience sampling study. *Behavior Therapy*, *45*(1), 67–82. https://doi.org/10.1016/j.beth.2013.09.004

Takemura, A. (2016). Exponential decay rate of partial autocorrelation coefficients of ARMA and short-memory processes. *Statistics & Probability Letters*, *110*, 207–210. https://doi.org/10.1016/j.spl.2015.12.023

Talbot, L. S., Stone, S., Gruber, J., Hairston, I. S., Eidelman, P., & Harvey, A. G. (2012). A test of the bidirectional association between sleep and mood in bipolar disorder and insomnia. *Journal of Abnormal Psychology*, *121*(1), 39–50. https://doi.org/10.1037/a0024946

van Zundert, R. M. P., van Roekel, E., Engels, R. C. M. E., & Scholte, R. H. J. (2015). Reciprocal associations between adolescents' night-time sleep and daytime affect and the role of gender and depressive symptoms. *Journal of Youth and Adolescence*, *44*(2), 556–569. https://doi.org/10.1007/s10964-013-0009-3

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Derom, C., Jacobs, N., Kendler, K. S., van der Maas, H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(1), 87–92. https://doi.org/10.1073/pnas.1312114110

van der Helm, E., & Walker, M. P. (2009). Overnight therapy? The role of sleep in emotional brain processing. *Psychological Bulletin*, *135*(5), 731–748. https://doi.org/10.1037/a0016570

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*(2), 176–192. https://doi.org/10.1037/a0027543

Waugh, C. E., Shing, E. Z., Avery, B. M., Jung, Y., Whitlow, C. T., & Maldjian, J. A. (2017). Neural predictors of emotional inertia in daily life. *Social Cognitive and Affective Neuroscience*, *12*(9), 1448–1459. https://doi.org/10.1093/scan/nsx071

Wichers, M., & Groot, P. C., (2016). Critical slowing down as a personalized early warning signal for depression. *Psychotherapy and Psychosomatics*, *85*(2), 114–116. https://doi.org/10.1159/000441458

Wild-Hartmann, J. A. d., Wichers, M., Bemmel, A. L. v., Derom, C., Thiery, E., Jacobs, N., Os, J. v., & Simons, C. J. P. (2013). Day-to-day associations between subjective sleep and affect in regard to future depressionin a female population-based sample. *The British Journal of Psychiatry: The Journal of Mental Science*, *202*(6), 407–412. https://doi.org/10.1192/bjp.bp.112.123794

Wrzus, C., Wagner, G. G., & Riediger, M. (2014). Feeling good when sleeping in? Day-to-day associations between sleep duration and affective well-being differ from youth to old age. *Emotion (Washington, D.C.)*, *14*(3), 624–628. https://doi.org/10.1037/a0035349