Evaluating dynamics in affect structure with latent Markov factor analysis

Vogelsmeier, Leonie V. D. E.; Cloos, Leonie; Kuppens, Peter; Ceulemans, Eva

Published in:
Emotion

DOI:
10.1037/emo0001307

Publication date:
2023

Document Version
Early version, also known as pre-print

Evaluating Dynamics in Affect Structure with Latent Markov Factor Analysis

Leonie V.D.E Vogelsmeier*

Tilburg University

Leonie Cloos*, Peter Kuppens, Eva Ceulemans

KU Leuven

*Shared first authorship

Author Note

This article has been peer-reviewed and accepted for publication by the journal *Emotion*, American Psychological Association, 2023. **This paper is not the copy of record** and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication.

Correspondence concerning this article should be addressed to: Leonie Cloos, Department of Quantitative Psychology and Individual Differences, KU Leuven, Tiensestraat 102, 3000 Leuven, E-Mail: leonie.cloos@kuleuven.be Leonie V.D.E. Vogelsmeier, Department of Methodology and Statistics, Tilburg University, Warandelaan 2 5037 AB Tilburg, E-Mail: l.v.d.e.vogelsmeier@tilburguniversity.edu

Abstract

In intensive longitudinal research, researchers typically consider the structure of affect to be stable across individuals and contexts. Based on an assumed theoretical structure (e.g., one bipolar or two separate positive and negative affect constructs), researchers create affect scores from items (e.g., sum or factor scores) and use them to examine the dynamics therein. However, researchers usually ignore that the affect structure itself is dynamic and varies across individuals and contexts. Understanding these dynamics provides valuable insights into individuals' affective experiences. This study uses latent Markov factor analysis (LMFA) to study what affect structures underlie individuals' responses, how individuals transition between structures, and whether their individual transition patterns differ. Moreover, we explore whether *the intensity of negative events* and the personality trait *neuroticism* relate to momentary transitions and individual differences in transition patterns, respectively. Applying LMFA to experience sampling data (N = 153; age: mean = 22; standard deviation = 7.1; range = 17 – 66), we identified two affect structures—one with three and one with four dimensions. The main difference was the presence of negative emotionality, and the affect dimensions became more inversely related when the affect structure included negative emotionality. Moreover, we identified three latent subgroups that differed in their transition patterns. Higher *negative event intensity* increased the probability of adopting an affect structure with negative emotionality. However, *neuroticism* was unrelated to subgroup membership. Summarized, we propose a way to incorporate contextual and individual differences in affect structure, contributing to advancing the theoretical basis of affect dynamics research.

Keywords: affect structure, intensive longitudinal data, context, individual differences, affect dynamics

Evaluating Dynamics in Affect Structure with Latent Markov Factor Analysis

Affect is a universally experienced phenomenon. However, feelings, emotions, or moods are individual experiences shaped by a person's interactions with and reactions to a constantly changing environment. The variation or dynamics of individuals' positive and negative affect over time can be studied using Experience Sampling Methodology (ESM; Scollon et al., 2003), where individuals complete questionnaires multiple times per day over a prolonged period of time. The resulting intensive longitudinal data allow for the investigation of *affect dynamics,* which provide insight into how affect changes across time and interacts with other momentary phenomena or person-specific attributes (Dejonckheere et al., 2017; Diener et al., 1995; Mneimne et al., 2018; Mroczek & Almeida, 2004). To this end, researchers create affect scores (most typically in terms of one bipolar or two positive and negative affect constructs) and use them to examine the dynamics therein.

While intensive longitudinal designs are increasingly implemented in contemporary affect research (Hamaker et al., 2015), there is little consensus regarding the theory and measurement of affect (Cloos et al., 2022; Ekkekakis, 2013; Kuppens, 2019). Specifically, there are a number of theoretical perspectives that differ in the number and nature of underlying affect dimensions or factors. The most fundamental component of affect in virtually all theories is valence, a characteristic that contrasts pleasant (positive emotions), with unpleasant (negative emotions) affective experiences (Barrett, 2006). Some theories consider valence to take the form of a bipolar dimension, with positive and negative affect as the opposite ends (positive and negative emotions are mutually exclusive). The presence of positive emotions thus implies a lack of negative emotions, making it improbable that someone feels good and bad at the same time. Yet, not all theories consider valence as a bipolar dimension. The approach proposed by Watson and Tellegen (1985), for instance, states that positive and negative affect form two separate, independent factors. From this theoretical perspective, positive and negative emotions are not mutually exclusive, and thus mixed emotions can occur (Larsen et al., 2001; Larsen & McGraw, 2014). Still other theories propose additional dimensions on top of (bipolar) valence to describe the affective space. Core affect theory posits that aside from valence, arousal (ranging from calm to excited) is a second important dimension that characterizes affect, distinguishing experiences of high arousal (e.g., fear and excitement) from like-valenced experiences of low arousal (e.g., sadness and relaxation) (Barrett & Russell, 1999; Russell & Feldman Barrett, 2009). Osgood (1966) adds a third dimension, namely dominance or control, which, for instance, distinguishes anger (high dominance) from fear (low dominance) (see also Russell & Mehrabian, 1977). Fontaine and colleagues (2007) added still another dimension, unpredictability. In turn, basic emotion theories (Fox et al., 2018; Panksepp, 2004) define the affect space in terms of discrete emotion dimensions such as anger, sadness, fear, happiness, and so on, with each their unique set of characteristics. Appraisal theories, in contrast, identify a number of appraisal dimensions that reflect how individuals evaluate stimuli and events (e.g., event importance, goal (in)congruence, responsibility, coping potential) and what discrete emotion may emerge in this situation (Ellsworth & Scherer, 2003; Smith & Ellsworth, 1985). Finally, psychological construction theories (Barrett & Russell, 2014) hypothesize that at their core emotions are described by underlying valence and arousal dimensions, but if individuals need to meaningfully relate their experience to a specific context, they can effortlessly label them into separate categories (although, as described below, there are individual differences). The important element to take away from this overview is that the diversity of affective experience has been described with a variety of theoretical models that each put forward different dimensions that differ in terms of quantity and quality (for a more thorough discussion of affect theories, see Moors, 2022).

Complementing this theoretical diversity, research has also documented variability in the structure of affect across different individuals and contexts. In terms of individual differences, for instance, some people perceive positive and negative emotions on one bipolar valence dimension, while others experience two rather independent dimensions (Dejonckheere et al., 2018). Moreover, it has been documented that individuals differ in the extent to which they describe their emotional experiences on a general dimension of feeling good versus bad (e.g., bipolar valence) or make more fine-grained distinctions between different emotions of the same valence (reflecting more discrete dimensions) (Barrett et al., 2001). Research on concepts such as emodiversity or emotion differentiation, complexity, and granularity has further provided insight into the variety of subtilities with which people reflect on and report their affective experiences (Barrett, 2004). For example, research on emotional differentiation has looked at the intra-class correlation (ICC) of same-valenced emotions across time or stimuli for a single individual, indicating to what extent an individual experiences a one-dimensional affect structure rather than distinct emotions (Barrett, 1998). As another example, research on emotional complexity often examines individual differences in the number and nature of factors underlying an individual's emotion ratings, reflecting that different affect structures can underlie the structure of affect across individuals (Brose et al., 2015; Lindquist & Barrett, 2008). When it comes to determinants of these individual differences, the personality trait of neuroticism may play a role in an individual's affect structure. Neuroticism has been associated with the experience of more intense negative emotions, stronger reactivity to negative events, and higher emotion variability

(Hisler, Krizan, DeHart, & Wright, 2020).[1] Although neuroticism has not been linked to a specific affect structure, the structure for individuals with higher levels of neuroticism would likely include a factor that captures high-intensity negative emotions.

In terms of contextual modulation, research has identified several factors that can impact the structure of affect, such as stress and relatedly negative, uncontrollable and undesirable events, and cognitive demands (Lindquist & Barrett, 2008; Potter et al., 2000; Zautra et al., 2002). Zautra and colleagues (1997) proposed a model in which the structure of affect is context-dependent and found that with higher rates of negative events, positive and negative affect become more bipolar. This is supported by recent evidence from intensive longitudinal studies, showing that personally relevant events (which take up a lot of cognitive resources, e.g., exams) shift the affect structure to a bipolar model (Dejonckheere et al., 2021).

In sum, the structure of affect, and the number and quality of dimensions describing it, is not a universal constant but varies across theories, individuals, and contexts. For the study of affect, it is therefore important to take into account this variability, as different affect structures can underlie the data depending on the individual and context. It is difficult to determine a priori *which* affect structures applies to *whom* and *when* because the affect structure may depend on many different characteristics of the individuals and the contexts. Embracing dynamics in the affect structure underlying affect responses in intensive longitudinal data can help us understand for *whom*, *when*, and *how* they affect structure changes and *which* substantive theory applies to repeated measures. Because of the lack of a priori information on which we could build specific

[1] However, the latter association has been challenged by the finding that negative affect variability can be explained by the higher intensity (or average) of negative affect that is common among individuals who score high on neuroticism (Kalokerinos et al., 2020).

hypotheses regarding the affect structures, in this study, we explore which theoretical models are present in the data and whether they change over time. More specifically, we use latent Markov factor analysis (LMFA; Vogelsmeier, Vermunt, van Roekel, et al., 2019) to assess *which* affect structures apply for *whom* and *when* in intensive longitudinal data to identify differences between individuals in the frequency of transitioning between affect structures, and to test hypotheses about reasons for transitions between affect structures and differences between individuals in these transitions. Firstly, LMFA reveals how many and which affect structures underlie the individual observations by classifying observations with the same underlying affect structure into the same latent (and thus unobserved) state[2]. Secondly, LMFA shows when (i.e., at which measurement occasions) individuals transition between affect structures by disclosing individuals' transitions between affect structures. To understand why (i.e., in which contexts) individuals transition to a certain latent state and thus affect structure, one can look at relationships between the affect structure associated with a certain state membership and timepoint-specific covariates. Finally, LMFA reveals similarities and differences between individuals in how the affect structure changes by grouping those with similar transition patterns into latent subgroups (Vogelsmeier et al., 2020, 2023). To understand why individuals belong to a certain subgroup, one can investigate relationships between the subgroup membership and *individual-specific covariates* (e.g., personality traits).[3]

[2] Latent states are akin to latent classes. However, in LMFA, individuals can transition between latent classes over time, making the term state more applicable. Note that the number of states and thus measurement models is determined by model selection.

[3] Despite the rather exploratory nature of LMFA, we recommend preregistering hypotheses about the anticipated covariate effects and aiming to confirm findings in other data (with a sample from the same population, similar study characteristics, sets of items, etc.).

Research Questions and Hypotheses

In the present study, we investigate which affect structures underlie the data of a two-week experience sampling method (ESM) study (Cloos et al., 2022). We preregistered several research questions and hypotheses osf.io/jvf7r before analyzing the data.

Research Question 1 (RQ1)

How many and which affect structures underlie the data?

As explained in the introduction, the different theoretical conceptualizations of affect can apply to different individuals and timepoints. Therefore, we assume that more than one affect structure will underlie the data.

Research Question 2 (RQ2)

Do individuals generally differ in whether and how frequently they transition between the affect structures?

As explained in the introduction, the affect structure depends on individual differences. Some individuals may consistently have the same affect structure underlying their responses. Others may transition between two or more affect structures across time. Some of those who transition may do so only occasionally. Others may frequently transition between different affect structures. We, therefore, expect to identify latent subgroups of individuals that differ in which affect structure primarily applies to their data and in their patterns of transitioning to other affect structures.

In case there is more than one affect structure (**RQ1**) and more than one latent subgroup (**RQ2**), we will investigate whether a timepoint-specific covariate relates to the transitions between affect structures and whether an individual-specific covariate predicts the subgroup memberships.

Research Question 3 (RQ3)

Are the momentary transitions between affect structures related to the occurrence of a timepoint-specific negative event*, and does this relation differ across the latent subgroups?*

As explained in the introduction, negative events may influence which affect structure applies to the responses of individuals. Therefore, we expect *negative event intensity* to be related to transitions between affect structures (at least in one subgroup if more than one subgroup is found by answering **RQ2**). For example, an intense negative event could increase the chance that an individual switches to an affect structure with a bipolar affect dimension.

Research Question 4 (RQ4)

Is the latent subgroup membership (and thus the transition pattern) related to neuroticism?

We expect *neuroticism* to be related to the latent subgroup membership. As explained in the introduction, the intensity and frequency of experiencing negative emotions and more intense emotional responses after negative events are characteristic of individuals with high neuroticism levels. Therefore, on the one hand, *neuroticism* may be related to a latent subgroup in which individuals mainly transition to (and stay) in a state representing an affect structure with a relatively high predominance of negative emotions. Such transitions may be particularly the case after the occurrence of a negative event. Thus, if *neuroticism* relates to the subgroup membership and *negative event intensity* is related to momentary transitions between affect structures (see **RQ3**), subgroups characterized by high levels of *neuroticism* may be more likely to switch affect structures after an intense negative event.

On the other hand, since inter-individual differences in *neuroticism* have been related to emotional variability or instability, *neuroticism* may be related to a latent subgroup in which individuals frequently transition between affect structures (which would indicate emotional variability).

Method

Participants

Participants were required to be Dutch-speaking and own a personal smartphone. The sample consisted of students and volunteers recruited through the university's participant management system, advertisement, and social media platforms. Our final sample included 153 individuals with a mean age of 22 years (*SD* = 7.1; range = 17 – 66), an approximately equal male (49.7%) to female gender ratio, with two participants not identifying as either. On average participants completed 123 out of 140 notifications (average compliance = 88%; range = 51% – 100%). There were 18,822 valid observations.[4]

Procedure

The study was approved by the Social and Societal Ethics Committee – KUL. Eligible participants were invited to a video conference session where participants, together with a researcher, set up the ESM platform m-Path (www.m-path.io) on their smartphones and were instructed about the ESM procedure, compliance, and informed consent. After completing the session, participants received an anonymized code that was entered into the app to enroll in the study. They were required to agree to the informed consent and answer demographic and baseline questionnaires. Lastly, they selected their preferred 12-hour timeframe. For the next 14 days, participants received ten notifications per day (total = 140 notifications) asking them to

[4] Responses were valid if participants had a compliance above 50%. Responses to questionnaires that had (a) technical errors, (b) > 15 consecutive items answered in the same range of values, (c) > 15 items answered < 500 ms, (d) a response time > 15 min were rated as invalid. Participants with > 30% invalid data were removed ($N = 2$).

answer the 30-item ESM questionnaire. The completion time was unrestricted, but questionnaires expired 30 minutes after the notification or at the next notification. Notifications occurred within a 12-hour timeframe split into ten blocks of 72 minutes (median interval between notifications = 75 minutes). Within each block, one notification was sent at a random moment and followed by a reminder if the questionnaire was not started after 5 minutes. After the 10^{th} notification on the last day, participants were given an end-questionnaire to complete the study. Participants could earn up to 50€ (or eight credits for research participation required for first-year psychology students). The reward was based on overall compliance, so that 75% compliance was sufficient to earn the full reward, but each decrease of 10% was associated with a deduction of 1 credit or 6€.

Materials

The ESM study included baseline measures (*time-invariant*) and repeated measures (*time-variant*). In the following, we report the materials relevant for the present analyses.

Time-Invariant Measures

Neuroticism. The 8-item neuroticism subscale of the Dutch version of the Big Five Inventory (BFI; Denissen, Geenen, van Aken, Gosling, & Potter, 2008) was used to measure *neuroticism*. The measure was taken the day before the two-week ESM assessment period started. Participants were asked to rate eight items on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The *time-invariant* covariate *neuroticism* is the average of the eight items.

Time-Varying Measures

Momentary Affect. In this study, we use participants' responses on eight positive emotion items *happy, energetic, loving, relaxed, satisfied, alert, caring,* and *calm,* and 12 negative emotion items *depressed*, *exhausted*, *ashamed*, *guilty*, *angry*, *anxious*, *gloomy*, *tired*, *shy*, *regretful*, *irritated*, and *concerned*. Items were rated on a Visual Analog Scale (VAS) from 0 (Not at all…) to 100 (Very…), with the slider starting at the midpoint. An example question is *"How angry do you feel at the moment?" 0 = Not at all Angry to 100 = Very Angry*. The questionnaire was constructed based on a review of the 18 most prominent affect measures (osf.io/g35jb). Across the measures, 11 unique discrete emotion subscales were relevant for the momentary assessment. These were: Happiness, Vigor, Love, Calmness, Sadness, Fatigue, Shame, Guilt, Anger, Anxiety/Fear, and Stress. For each discrete emotion, we measured one high-intensity and one low-intensity item, which were selected according to their valence rating on the normed index of affective ratings for words in Dutch (Moors et al., 2013). Note that we excluded two items that formed the construct of Stress (*stressed* and *nervous*) since we define stress as a strictly external stimulus rather than part of the rubric of emotions (Erbas et al., 2018; Smith & Lazarus, 1990).

Negative Event Intensity. *Negative event intensity* was measured by assessing the intensity of negative events. At each beep, participants were asked to *"Think about the most negative event since the last beep?"* and rate *"How intense was this event?"*, with 0 indicating that no negative event took place and 100 indicating the event intensity was very negative. The *time-varying* covariate *negative event intensity* was quantified by the score on this item.

Latent Markov Factor Analysis

To answer RQs 1–4, we will apply LMFA to the ESM data described above. LMFA combines mixture factor analysis (Lubke & Muthén, 2005; McLachlan & Peel, 2000) to determine the number and nature of the affect structures with latent Markov modeling

(Bartolucci, 2006; Collins, 2006) to describe transitions between the affect structures. In the following, we first explain the mixture factor analysis model and then the latent Markov model.

Mixture Factor Analysis Model

The mixture factor analysis model in LMFA is a factor analysis model that allows the parameters of the affect structures to differ across latent states. Two types of state-specific parameters are of particular interest: (1) The item-specific *factor loadings* indicate if and how well the items measure the latent construct. Differences in the loadings across states therefore imply that items measure the latent constructs differently. (2) The item-specific *intercepts* are the expected scores when the factor scores equal zero. Differences in intercepts across states, therefore, mean that the thresholds to endorse items differ.

One can generally choose between exploratory and confirmatory mixture factor analysis. In the latter, however, one would have to specify a priori how many factors underlie the responses and which items load on which factors for each state-specific affect structure. Because the number and nature of factors vary from theory to theory—and because researchers typically do not know which theory of affect applies to whom and when—they cannot formulate such a priori hypotheses. Exploratory factor analysis, in contrast, can be used without a priori hypotheses. This requires estimating and comparing different plausible models using model selection criteria described in the Data Analysis section (Step 1). To identify the exploratory factor models in each state, factor means are fixed to zero and factor variances to 1.[5]

[5] For technical details about mixture factor analysis, for example, model identification, we refer to (Vogelsmeier et al., 2019).

Summarized, the factor analysis part of the model uncovers *which* affect structures apply. The latent Markov model described next reveals for *whom* which affect structure applies and *when*, and allows to test hypotheses about the *why*.

Latent Markov Model

The basic latent Markov model contains two types of parameters to describe the transitions between the latent states. (1) The *initial state probabilities* quantify the probability of starting in a state. (2) The *transition probabilities* indicate the probability of being in a state at the timepoint of a measurement occasion, given the state membership at the previous measurement occasion. Without further specifications, the model assumes that the probabilities of transitioning to other states or staying in a state are constant across contexts and individuals. As explained in the introduction, this assumption is easily violated. However, it is possible to extend this model (and thereby relax the assumptions) in three ways. First, one can relate timepoint-specific covariates to the transition probabilities to understand why (i.e., in which contexts) individuals transition to a particular state (i.e., the transition probabilities are no longer the same across situations but depend on the values of contextual covariates). In this study, we test the effect of the covariate *negative event intensity* on the transitions.

Second, one can allow for subgroup-specific transition patterns; that is, all parameters—including covariate effect(s) on the transition probabilities—are no longer the same for all individuals but depend on the latent subgroup. Finding subgroups of individuals with the same transition pattern reveals the most salient differences and similarities between individuals in how likely they are to transition to certain states and thus affect structures (and in which contexts). The number of latent subgroups can be determined based on theory or model selection criteria, as described in the Data Analysis Section (Step 3). When using latent subgroups in LMFA, a third type of parameter is added; that is, (3) the *subgroup-membership probabilities* and hence, the probabilities of belonging to a certain subgroup.

Third, to understand why individuals belong to which latent subgroup and, thus, why they have a particular transition pattern, one can relate stable individual characteristics to the subgroup-membership probabilities. This study investigates whether *neuroticism* is related to the subgroup-membership probabilities.

When fitting these models and extensions, we must consider that the intervals between measurement occasions differ across individuals and time. Therefore, we will employ the continuous-time version of LMFA, which accounts for unequal intervals during the estimation (for details about the discrete- and continuous-time approaches, see Vogelsmeier, Vermunt, Böing-Messing, et al., 2019). Moreover, for the estimation of LMFA, we use the three-step approach (Vogelsmeier et al., 2023), which is recommended over a full information maximum likelihood approach when covariates are included, as it simplifies model selection.[6] The stepwise approach splits the estimation into three parts: (1) *obtaining and investigating the affect structures*, (2) *obtaining state assignments and classification errors*, and (3) *obtaining and investigating the covariate and subgroup-specific transition model*. Below, in the Data Analysis Section, we describe the three steps and how they answer our research questions.

Data Analysis

Step 1: Obtaining and Investigating the Affect Structures

[6] For details about the difference between the two approaches, see Vogelsmeier et al. (2021).

In step 1, we disregard the (latent subgroup- and covariate-specific) transition pattern and only focus on obtaining the state-specific affect structures. To this end, we treat all observations as independent[7] and estimate and compare models that differ in the number of different affect structures and the number of dimensions that make up the affect structure, and select the best model based on the Bayesian information criteria (BIC; Schwarz, 1978) and the convex hull (CHull) criterion (Ceulemans & Kiers, 2006; Wilderjans et al., 2013). The BIC balances fit and parsimony by penalizing models with more parameters. The CHull criterion identifies models at the higher boundary of the convex hull in a “loglikelihood vs. number of parameters” plot and then, using scree ratios, automatically chooses the best model among all models on the upper boundary by identifying the point at which improvement in fit levels off for an increasing number of parameters. Both criteria provide valuable information and are well suited for selecting the best model, as shown in previous simulation studies (Bulteel et al., 2013; Vogelsmeier, Vermunt, van Roekel, et al., 2019). However, both criteria come with disadvantages. A relevant disadvantage of the BIC is that, in empirical data, it may keep decreasing for increasingly complex models (Bauer, 2007; McNeish et al., 2021). The CHull has the disadvantage that it cannot select the least and most complex model because no scree ratios can be computed (Bulteel, Wilderjans, Tuerlinckx, & Ceulemans, 2013). Generally, it is suggested to consult both the BIC and the CHull and to additionally consider interpretability for the final model choice (Vogelsmeier et al., 2022).

[7] Note that ignoring serial dependence when estimating the factor structure does not impacts the results (Bulteel et al., 2018; Molenaar & Nesselroade, 2009).

We estimate 31 models with 1 – 3 states (and thus affect structures) and 1 – 4 factors per state.[8] Following recommendations on selecting the number of states and factors per state (Vogelsmeier et al., 2022), we begin with just a few states and factors and increase the numbers only if the model fit (according to the BIC and CHull) still considerably improves when comparing models with two states to models with three states on the one hand and when comparing models with three factors to models with four factors on the other hand (and if the estimations with three states and four factors do not already yield convergence issues).

Once the best model is chosen[9], we can answer **RQ1** and examine the state-specific affect structures. Note that, in each state-specific affect structure, we standardize the loadings by dividing them by the state-specific item standard deviations. Standardizing allows us to use rules of thumb to determine which loadings are large enough to consider items as measures of the underlying factors. We consider loadings large enough when they are larger or equal to 0.3. Moreover, we apply oblique rotation to the dimensions that make up the affect structure to enhance the interpretation of the affect structure. The oblique rotation allows factors to be correlated, which is usually more realistic than orthogonal rotation, which assumes zero factor correlations. If factor correlations are (close to) zero, the oblique rotation will indicate this.

Step 2: Obtaining State Assignments and Classification Errors

In step 2, we assign the observations to the affect structure that most likely underlies the responses. The assignment of observations always involves some classification error unless all

[8] This is a deviation from the preregistration, where our list erroneously did not include the one-factor models. Since the one-factor model is one of the dominant structures in the literature (as explained in the introduction), it should also be included in the model selection.

[9] Note that only converged models are considered.

observations can be assigned with 100% certainty, which is unrealistic for empirical data. However, this classification error is not a problem because, when obtaining the transition model in step 3, the analysis automatically accounts for this classification error from step 2.

Step 3: Investigating the Transition Model

In step 3, we keep the affect structures from step 1 fixed and investigate the latent subgroup- and covariate-specific transitions between the affect structures, accounting for the inherent classification error from step 2. To answer **RQ2** (i.e., whether the transition pattern depends on individuals' latent subgroup membership), we include a latent subgroup variable that allows for subgroup-specific transition and initial state probabilities.

To answer **RQ3** (i.e., whether the transitions are related to *negative event intensity* and whether the effect differs across the latent subgroups), we regress the transition probabilities on the timepoint-specific covariate *negative event intensity* and allow the covariate effects to vary across the latent subgroups. Finally, to answer **RQ4** (i.e., whether the subgroup membership is related to *neuroticism*), we regress the subgroup-membership probabilities on the individual-specific covariate *neuroticism*.

In order to (a) determine the number of latent subgroups, (b) whether *neuroticism* predicts the subgroup membership, and (c) whether *negative event intensity* affects the transition probabilities in one or more latent subgroups, we need to perform model selection. To this end, we will apply the following stepwise procedure:

1. We estimate four models with 1 – 8 latent subgroups[10] and *negative event intensity* as a covariate for the transition probabilities. For the seven models with more than one subgroup, the effect is allowed to differ across subgroups.
2. For the three models with more than one subgroup, we evaluate whether the effect of *negative event intensity* differs significantly across subgroups using Wald significant tests. For models where the effect does not differ significantly, we re-estimate the model with subgroup-independent effects of *negative event intensity* so that the covariate has the same effect on the transition probabilities in all subgroups.
3. Using Wald significant tests, we assess whether *negative event intensity* is a significant predictor of the transition probabilities in one of the subgroups (or overall, in case the model has only one subgroup or was modified in the previous step). For models where this is not the case, we remove *negative event intensity* as a predictor and re-estimate the models.
4. We compare the eight final models (excluding non-converged ones) using the BIC and the CHull criterion to select the number of latent subgroups.
5. If the selected model contains at least two subgroups, we re-estimate the model with *neuroticism* as a subgroup-membership predictor and use the Wald test to evaluate whether the effect is significant.

Transparency and Openness

---

[10]Our actual interest pertained to 1 to 4 subgroups (as stated in the preregistration). However, since the CHull method cannot choose the least and most complex models, we used a wider range of models in this study to increase our confidence in the appropriateness of the selected model.

The data was analyzed using Latent GOLD (Vermunt & Magidson, 2016). The syntax, pre-processing R-code (R Core Team, 2022), and other materials are available at: osf.io/q47pf. The data is available upon request. Details on the sample size, data exclusions, and all measures of the study are reported in Cloos and colleagues (2022) and can be accessed via osf.io/485qy.

Results

Step 1: Obtaining and Investigating the Affect Structures

To answer **RQ1** (i.e., how *many* and *which* affect structures underlie the data), we first investigate the model fit results and then inspect the state-specific affect structures and how they differ.

Model Selection

Table 1 shows the loglikelihood value and number of parameters for all models up to two states and thus affect structures. All models in the table converged. We did not consider the models with three affect structures in our model selection since the solutions to these were highly unstable.[11] Interestingly, models with two affect structures always represent the data better than models with only one (regardless of the number of factors), as can be seen by the smaller BIC values for models with two states and the steep increase in the CHull plot.[12] The model with the lowest BIC value had two state-specific affect structures with four factors each; the second-best model had two state-specific affect structures with four and three factors, respectively. Inspecting

[11] Most three-state models did initially not converge. Therefore, we estimated all three-state models five times. Some eventually converged, but the solutions (i.e., the loglikelihood values) were highly unstable, and most solutions could be identified as local optima solutions (the latter can be seen from the fact that loglikelihood values decrease as the number of parameters increases). Such instability and non-convergence indicate that the data are not informative enough. Therefore, we disregarded the three-state models in our model selection.

[12] The output of the CHull analysis is presented in the Online Supplement.

the CHull plot shows that the increase in model fit levels off with the latter ([43]) model[13]. This means that adding more parameters barely improves fit. Therefore, we choose the model with four factors in the affect structure of one state and three factors in the affect structure of the other state.

Table 1

Step-1 Model fit Results Sorted from Lowest (Best) to Highest BIC

Model	Number of Parameters	BIC	Loglikelihood
[44]	241	2942562.74	-1470095.32
[43]	221	2942948.33	-1470386.54
[42]	201	2945107.31	-1471564.45
[33]	201	2950904.25	-1474462.92
[41]	181	2951813.59	-1475016.02
[32]	181	2953061.00	-1475639.73
[31]	161	2959773.96	-1479094.64
[22]	161	2968363.16	-1483389.24
[21]	141	2975113.40	-1486862.78
[11]	121	2997199.68	-1498004.35
[4]	120	3214793.45	-1606806.16
[3]	100	3224907.25	-1611961.49
[2]	80	3246978.76	-1623095.67
[1]	60	3280280.10	-1639844.77

Note. The number of elements in the square brackets refers to the number of states - the numbers themselves refer to the number of factors in a state (with the first element pertaining to state 1, the second element to state 2, and so on).

[13] Note that the order of the measurement models in Table 1 is presented in decreasing complexity in the output. This order is arbitrary (i.e., model [4 3] is the same as model [3 4]. When discussing the results, we chose to order the models by increasing complexity ([34]) to aid the discussion of the results.

Results Pertaining to the Affect Structure

The main difference between the two affect structures is the number of factors and the pattern of loadings (Table 2). We first describe the simpler affect structure containing three factors (i.e., state 1).

Factor Loadings Affect Structure 1. The first factor includes only positive items that are related to feeling *Content* (*relaxed*, *calm*, *happy*, and *satisfied*); that is, these items have loadings above or equal to the chosen threshold of .30. The second factor contains three negative items (*exhausted*, *tired*, *concerned)* and one positive item (*energetic*) with a reverse loading. We refer to this factor as *Fatigue* because the non-energetic items have larger absolute values. The last factor contained four positive items (*loving*, *caring*, *alert*, and *energetic*) and one negative item (*concerned*) that are typical interpersonal emotions or expressions of *Love*.

Factor Correlations Affect Structure 1. *Love* and *Fatigue* were virtually independent (r = -.07), and *Fatigue* and *Content* were only mildly negatively correlated (r = -.22), indicating that *Fatigue* is a distinct emotion category that can be experienced independently of other emotions. In contrast, the two positive factors, *Love* and *Content*, correlate considerably (r = .42).

Item Intercepts Affect Structure 1. All intercepts for the positive items are larger than for the negative ones. Of the negative items, only *irritated*, *concerned*, *exhausted*, and *tired* have an intercept that is not (close to) zero.

Summary Affect Structure 1. The first affect structure contains three factors, of which only *Fatigue* contains significant loadings of negative emotion items. There is no factor that captures pure negative emotionality. This first affect structure is thus characterized by the lack of variability in negative emotions and thus mainly describes *Positive Emotionality*. This notion is supported by the (near-)zero intercept values for the negative emotions that are not part of *Fatigue*.

Table 2

Step 1. Results: State-Specific Factor Loadings, Intercepts, and Correlations

	Positive Emotionality				*Positive & Negative Emotionality*				
	Factor Loadings			Item Int.	Factor Loadings				Item Int.
	Content	Fatigue	Love		Distress/ Content	SCNE	Fatigue	Love	
angry	-0.04	-0.02	0.02	0.01	0.20	**0.50**	-0.04	-0.08	13.86
irritated	-0.10	0.19	-0.00	6.28	0.26	**0.33**	0.05	-0.05	24.24
anxious	-0.01	0.02	0.02	0.00	**0.44**	**0.37**	0.00	0.26	18.26
concerned	-0.21	**0.31**	**0.33**	7.27	**0.40**	0.28	0.15	**0.41**	29.58
exhausted	0.01	**0.85**	0.08	22.83	0.01	0.03	**0.83**	0.08	39.62
tired	-0.01	**0.92**	0.02	29.79	-0.07	-0.02	**0.92**	0.01	43.94
guilty	-0.02	-0.01	0.01	0.01	0.00	**0.58**	0.14	0.12	18.05
regretful	0.01	0.00	0.01	0.01	0.10	**0.66**	0.05	-0.04	18.44
depressed	-0.07	-0.02	0.01	0.01	0.28	**0.48**	0.13	0.01	15.51
gloomy	-0.15	0.12	-0.03	1.67	**0.41**	**0.38**	0.15	-0.01	22.80
ashamed	-0.00	-0.01	0.02	0.01	-0.12	**0.78**	-0.06	-0.06	13.54
shy	-0.01	-0.01	0.01	0.00	-0.21	**0.63**	-0.03	-0.05	11.84
relaxed	**0.79**	0.03	-0.08	71.94	**-0.78**	0.05	-0.01	-0.02	57.15
calm	**0.64**	0.07	-0.12	72.56	**-0.67**	-0.04	0.07	-0.06	58.96
happy	**0.70**	-0.05	0.17	72.71	**-0.72**	-0.01	-0.09	0.25	58.68
satisfied	**0.69**	-0.09	0.11	71.66	**-0.74**	0.00	-0.07	0.18	56.89
loving	0.24	0.09	**0.65**	58.48	-0.27	-0.06	-0.01	**0.68**	54.79
caring	0.01	0.06	**0.72**	42.49	-0.06	-0.06	0.03	**0.73**	47.03
energetic	0.10	**-0.59**	**0.31**	55.65	-0.17	0.04	**-0.57**	0.25	47.75
alert	-0.09	-0.26	**0.52**	50.99	0.12	-0.09	-0.28	**0.38**	47.37

Factor Correlations

Content	1.00			Distress/ Content	1.00			
Fatigue	-0.22*	1.00		SCNE	0.42**	1.00		
Love	0.42**	-0.07	1.00	Fatigue	0.39**	0.27*	1.00	
				Love	-0.29*	0.06	-0.06	1.00

Note. Item Int. = Intercepts; SCNE = Self-Conscious Negative Emotions.

Item-specific factor loadings > |.30| are highlighted in bold.

* mild correlation > |.20|; ** considerable correlation > |.30|.

Factor Loadings Affect Structure 2. The affect structure of the second state has four factors. There is some overlap with the factors of *Positive Emotionality*: The third factor is similar to the *Fatigue* factor in this affect structure and contains two negative items (*exhausted* and *tired)* and one positive item (*energetic*) with a reverse loading. The last factor is similar to the *Love* factor in *Positive Emotionality* and contains three positive and one negative item (*loving*, *caring*, *alert*, and *concerned*). We refer to the latter two factors with the same names as before (*Fatigue* and *Love*). Moreover, on the first factor, we find the same four positive items related to *Content* (*relaxed, calm, happy,* and *satisfied*) as in the first factor of *Positive Emotionality*. This time, however, the factor also contains three negative items that reflect feelings of general *Distress* (*anxious, concerned*, and *gloomy*). The loadings have opposite signs, meaning that higher feelings of *Distress* go along with being less *Content*. Thus, this factor reflects a bipolar dimension of *Distress/Content*. The second factor contained only negative items (*angry, irritated, anxious, guilty, regretful, depressed, gloomy, ashamed,* and *shy*) that imply some level of emotional pain. The self-conscious items *ashamed, regretful, and guilty* displayed the strongest loadings. Therefore, we call this factor *Self-Conscious Negative Emotions*.

Factor Correlations Affect Structure 2. In the second affect structure, *Fatigue* and *Love* are again uncorrelated (r = -.06). However, *Fatigue* is positively correlated with *Distress/Content* (r =.39) and mildly positively correlated with *Self-Conscious Negative Emotions* (r = .27), which means that higher values of *Fatigue* go along with higher values of *Distress* on the *Distress/Content* dimension as well as higher values on *Self-Conscious Negative Emotions* (and vice versa). Furthermore, *Distress/Content* is positively correlated with *Self-Conscious Negative Emotions* (r =.42), which means that higher values of *Self-Conscious*

Negative Emotions go along with higher values of *Distress* on the *Distress/Content* dimension (and vice versa). Moreover, *Love* and *Self-Conscious Negative Emotions* are uncorrelated (r = .06), but *Love* and *Distress/Content* are mildly negatively correlated (r = -.29). The latter means that higher values of *Love* are somewhat associated with lower values of *Distress* on the *Distress/Content* dimension (and vice versa).

Item Intercepts Affect Structure 2. Like in *Positive Emotionality*, all positive item intercepts are larger than the negative item intercepts. However, in this affect structure, all of the negative items have intercepts considerably larger than zero and are thus endorsed more easily than in affect structure 1. This implies that not only positive but also negative emotions may be experienced.

Summary Affect Structure 2. The second affect structure is similar to the first but contains one more factor capturing negative emotionality and non-zero intercepts for negative emotions. The affect structure thus describes *Positive & Negative Emotionality*, such that both positive and negative emotions may be experienced. Moreover, the affect structure is characterized by less distinct factors. Instead, there seems to be a bipolar element because there are mild to considerable correlations between *Distress/Content* and all other factors and between *Fatigue* and *Self-Conscious Negative Emotions*.

Step 2: Obtaining State Assignments and Classification Errors

In step 2, we assign the observations to the state and thus affect the structure that most likely underlies the responses. To clarify the assignment, consider that an observation belongs to state 1 with a probability of .8 and to state 2 with a probability of .2. The assignment of that observation would then be state 1 because the probability of belonging to this state is higher than of belonging to state 2. The assignment of observations always involves some classification error unless all observations are assigned with 100% certainty, which is unrealistic for empirical data. However, this classification error is not a problem because, when obtaining the transition model in step 3, the analysis automatically accounts for this classification error from step 2 (details about how the classification errors are computed and accounted for can be found in Vogelsmeier et al. 2021).

Step 3: Obtaining and Investigating the Transition Model

In order to answer **RQ2** (i.e., whether the transition pattern depends on individuals' latent subgroup membership) and **RQ3** (i.e., whether the transitions between affect structures are related to *negative event intensity* and whether the effect differs across the latent subgroups), we started with the transition-model selection. Following the procedure described in the Data Analysis Section (Step 3), we estimated all models (with 1– 8 latent subgroups) with *negative event intensity* as a covariate. The Wald tests indicated that the relation between *negative event intensity* and the transitions between affect structures significantly differed across subgroups for all seven models with more than one subgroup. Moreover, *negative event intensity* was significantly related to the transitions in at least one of the latent subgroups. Therefore, we compared the eight models, including the covariate *negative event intensity* in all subgroups.

Model Selection

The BIC values for all models are shown in Table 3. All models in the table converged. While the model with six subgroups is preferred according to the BIC, the CHull criterion[14] selects the model with three subgroups. Because of this disagreement, we inspected all models with three to six subgroups. It could be seen that there were considerable differences in the

[14] The output of the CHull analysis is presented in the Online Supplement.

transition probabilities across the three subgroups. More fine-grained differences could be seen when adding a fourth, fifth, and sixth subgroup, but the differences were minor compared to those between the initial three subgroups. Therefore, we choose the most parsimonious model with three latent subgroups. Of all individuals, 52 % are assigned to subgroup 1, 23 % to subgroup 2, and 25 % to subgroup 3.

Table 3

Step-3 Model Fit Results Sorted from Lowest (Best) to Highest BIC

Model	Number of Parameters	BIC	Loglikelihood
6 subgroups	35	10340.72	-4998.11
5 subgroups	29	10348.47	-5031.51
7 subgroups	41	10349.97	-4973.21
4 subgroups	23	10367.77	-5070.69
8 subgroups	47	10385.32	-4961.36
3 subgroups	17	10494.21	-5163.44
2 subgroups	11	11226.42	-5559.08
no subgroups	5	14226.87	-7088.83

Note. The estimation for all models in the table converged.

Results Pertaining to the Transition Model

The transition probabilities for the three latent subgroups for an interval of 75 minutes (corresponding to the median interval in the sample) and two *negative event intensity* values are presented in Table 4.[15] The chosen values are 3 and 24, which correspond to the median and the third quartile in the sample, respectively. Given that the scale of negative event intensity ranges from 0 to 100, these values indicate that individuals rarely experience negative events, and if they do, their experienced intensity is relatively mild. By inspecting the transition probabilities for *negative event intensity* values of 3 and 24, we investigate the transition patterns for *low* and

[15] 75 minutes was the median length of the interval between two measurement occasions. By design the measurements were scheduled with a random interval of 72 minutes. See procedure.

mild negative event intensity, and by comparing the transition probabilities for the two negative event values, we see how the difference between *low intensity* and *mild intensity* impacts the transition probabilities. Looking at the differences in transition patterns across subgroups for *low negative event intensity*, the most striking difference is that participants differ in which affect structure primarily underlies their responses; that is, which affect structure they are most likely to transition to and which they are most likely to remain in. Subgroup 1 (the largest of the three subgroups) shows high probabilities of transitioning to and staying in the *Positive & Negative Emotionality* state. Thus, once participants are in that state, they will likely stay there. In contrast, subgroup 2 is characterized by high probabilities of transitioning to and staying in the *Positive Emotionality* state. Subgroup 3 shows a slightly different pattern. This group distinguishes itself by a transition probability matrix with lower values on the diagonal. In other words, participants frequently transition between affect structures. However, the probability of transitioning to and staying in the *Positive & Negative Emotionality* state is somewhat higher than transitioning to and staying in the *Positive Emotionality* state.

Table 4

Transition Probabilities of the three Latent Subgroups for an Interval of 75 Minutes

		Median *Negative Event Intensity*		3rd Quartile *Negative Event Intensity*	
		Transition to		Transition to	
Sub-group	Transition from	*Positive Emotionality*	*Positive & Negative Emotionality*	*Positive Emotionality*	*Positive & Negative Emotionality*
1	*Positive Emotionality*	0.34	0.66	0.31 (↓)	0.69 (↑)
	Positive & Negative Emotionality	0.01	0.99	0.00 (↓)	1.00 (↑)
2	*Positive Emotionality*	0.85	0.15	0.75 (↓)	0.25 (↑)
	Positive & Negative Emotionality	0.61	0.39	0.49 (↓)	0.51 (↑)
3	*Positive Emotionality*	0.55	0.45	0.41 (↓)	0.59 (↑)
	Positive & Negative Emotionality	0.26	0.74	0.14 (↓)	0.86 (↑)

Note. The largest probability per row and *negative event* value is shaded in gray. Green upward arrows indicate increases in probabilities, and red downward arrows indicate decreases when changing the *stressor* covariate score from *low intensity* = 3 (sample median) to *mild intensity* = 24 (3rd quartile of the sample).

Comparing the transition probabilities between *low* and *mild negative event intensity* indicates that, in all subgroups, the probabilities of transitioning to and staying in the *Positive & Negative Emotionality* state increase, and the probabilities of transitioning to and staying in the *Positive Emotionality* state decrease. The amount of increase and decrease differs in magnitude across subgroups, but the trend is the same: For *low negative event intensity*, the probabilities of transitioning to and staying in the *Positive & Negative Emotionality* state are larger than for *Positive Emotionality*—with one exception in subgroup 2, where staying in the *Positive Emotionality* state is still more likely than transitioning to the *Positive & Negative Emotionality* state.

Finally, to answer **RQ4** (i.e., whether the subgroup membership is related to *neuroticism*), we re-estimated the transition model with three subgroups while including *neuroticism* as predictor for the subgroup memberships. The Wald test indicated that *neuroticism* did not significantly predict subgroup membership.

Discussion

Affect can be defined by a number of underlying constructs that relate to each other differently, depending on the timepoint and individual. This means that the affect structure is inherently dynamic so that different theoretical models can apply to intensive longitudinal data. In this article, we investigate dynamic changes in the affect structure by exploring the different affect structures underlying the data using Latent Markov factor analysis (LMFA). We found two different affect structures that applied to different individuals and timepoints. In the following, we elaborate on the results and their implications for affect research by answering the four research questions.

Which Affect Structures Underlie the Data, and how do They Differ? (RQ1)

In line with our hypothesis, we found more than one underlying affect structure or affect structure. Specifically, we identified two affect structures—one with the three factors *Content*, *Fatigue*, and *Love* that described *Positive Emotionality*, and one with the four factors *Distress/Content*, *Self-Conscious Negative Emotions*, *Fatigue*, and *Love* that described *Positive & Negative Emotionality*. The first difference was the presence or absence of negative emotionality, which was highlighted by (close to) zero item intercepts for negative items in the *Positive Emotionality* affect structure.

The second difference pertained to the degree to which affect dimensions were (in)dependent. Generally, both affect structures aligned with discrete emotion theories (Smith & Lazarus, 1990): They both included distinct emotion factors such that *Fatigue* and *Love* are separate from other negative and positive emotions, respectively. On the *Love* factor, the positive items *loving* and *caring* load together with energy items *alert* and *energetic,* possibly because people show their *love* to others with caring gestures, excitement, and giving them attention. Additionally, the item *concerned* loads on this factor. Although it is often considered a negative emotion, in the context of love, *concern* may rather be interpreted as similar to *caring*.

Furthermore, when negative emotionality was present, these items did not form a single common dimension of negative affect. In the affect structure that described *Positive & Negative Emotionality*, *Self-Conscious Negative Emotions* were separate from *Fatigue* and *Distress*. Yet, the *Positive & Negative Emotionality* affect structure displayed an element of bipolarity in that *Distress* and *Content* were opposite poles of one dimension. Furthermore, the *Self-Conscious Negative Emotions* correlated strongly with this bipolar factor. Thus, the different affect dimensions were more dependent when negative emotions were present.

Both affect structures were more nuanced than the positive or negative affect dimensions often encountered in the literature. This is consistent with recent findings that the affect structure in repeated momentary measures is rather fine-grained. For example, Eisele and colleagues (2021) found that positive affect was distinguished into high and low arousal and negative affect into irritation, high, and low arousal. Other studies looking at day-to-day assessments have found an even higher separation of affect into eight factors (Jacobson et al., 2023). Note, however, that differences in the study characteristics (e.g., sampling frequency, time frame, item selection, level of analysis; Carroll et al., 1999; Eisele et al., 2021; Scott et al., 2020; Watson et al., 1999), will likely influence the emerging structure of affect. For instance, a *Love* factor would not emerge when items such as concerned, loving, and caring are not part of the questionnaire.

The findings of this and other studies thus clearly show that two factors (positive and negative affect) may not accurately capture momentary affect. This also highlights the caveat of simply aggregating item responses into single negative or positive affect scores, as this may not reflect the depth of information in the data (McNeish & Wolf, 2020). For example, if we had computed a negative affect score based on all negative items and studied negative affect dynamics with this general negative affect score, we would have failed to disentangle fatigue and distress, which might have led to misinterpreting the negative affect dynamics with variability due to fatigue. Simple aggregation thus ignores the dynamics in the affect structure and the plausibility of having more than one structure underlying responses within persons, which—considering the results of this and other studies (Hofmann & Meyer, 2006; Vogelsmeier et al., 2023)—seems rather the rule than the exception. Finally, simply assuming rather than assessing (changes in) affect structures could even lead to results about affect dynamics not being comparable across studies. If different studies (unknowingly) measure different constructs (e.g., negative affect instead of fatigue and distress), dynamics in the construct or relations with other constructs will naturally differ and lead to different conclusions.

Do Individuals Generally Differ in Whether and how Frequently they Transition Between the Affect Structures? (RQ2)

In line with our hypothesis, we found that individuals differed in their transition patterns. Individuals were assigned to one of three latent subgroups, which differed mainly in which affect structure primarily underlay their responses. Most individuals belonged to a subgroup in which responses were characterized by the affect structure that included *Positive & Negative Emotionality*; that means that most individuals did endorse both positive and negative emotions. One-quarter of individuals was in a subgroup that reported primarily positive emotions and, thus, endorsed only a limited range of emotions. The remaining quarter of individuals was in a subgroup in which they frequently transitioned between both affect structures. This may imply that context influenced these individuals more than those in the other subgroups. These findings align with the notion that individual characteristics, as well as contextual effects, can bring about affect structure differences in intensive longitudinal data (Adolf et al., 2014). Classifying individuals into groups based on their individual patterns of change allows us to consider both effects at the same time.

Are the Momentary Transitions Between Affect Structures Related to Timepoint-Specific *Negative Event Intensity*, and Does this Differ Across the Latent Subgroups? (RQ3)

In line with our hypothesis, we found that *negative event intensity* was significantly related to the transitions between the affect structures. There were some differences in the magnitude of the *negative event intensity* effect on the transitions across subgroups. Nevertheless, for individuals in all subgroups, the occurrence of a *negative event intensity* increased the probability of transitioning to the measurement model of *Positive & Negative Emotionality,* in which negative emotionality was present and affect dimensions were somewhat dependent or even bipolar.

Like other intensive longitudinal studies, these findings suggest that the affect structure is context-dependent, such that affect becomes more bipolar during times of stress (Dejonckheere et al., 2021). Although individuals did not experience high negative event intensity, this measure explained differences and changes in the affect structure. This shows that context plays a notable role in emotion dynamics (Lapate & Heller, 2020) and should always be considered in substantive and methodological investigations of intensive longitudinal data.

Is the Latent Subgroup Membership Related to *Neuroticism*? (RQ4)

In contrast to our expectations, *neuroticism* did not significantly relate to subgroup membership. This means that neither the group membership nor differences in the reactivity to *negative event intensity* were related to this personality trait. Thus, higher scores on *neuroticism* did not relate to a higher probability of belonging to a subgroup in which individuals experience more negativity or transition more frequently (which were the two plausible outcomes we anticipated). Neither did higher scores on *neuroticism* relate to more frequent transitions between affect structures. Instead, regardless of their *neuroticism* level, most individuals were in the subgroup with *Positive & Negative Emotionality* being the most prominent measurement model.

Although most individuals experienced and reported negative emotionality regardless of their *neuroticism* level, it is important to understand which individual characteristics explain group differences in the (primary) affect structure. Since only a quarter of observations pertained to the affect structure of *Positive Emotionality*, it may be insightful to determine why this affect structure emerged. This could be linked to experiential avoidance but also to covariates that can explain the predominance of positive emotions in daily life, such as extraversion, goal achievement, or life satisfaction. Unfortunately, the investigation of the determinants and effects of experiencing positive emotions in daily life is less prevalent than the investigation of negative emotions (Heininga & Kuppens, 2021). To improve our knowledge about affective processes over time and differences between individuals, it may be worth to start paying attention to the positive side of the coin (Heininga et al., 2019).

Limitations

A limitation of this study are the researchers' degrees of freedom concerning the model selection in LMFA. There are some guidelines (e.g., selecting models with the lowest BIC value), but, like in this study, the empirical results are not always sufficiently straightforward (e.g., BIC scores can keep decreasing for more complex models; McNeish & Harring, 2017), so the researcher's choices always play a role in choosing the final model. This problem is not specific to LMFA but concerns complex analysis methods in general (del Giudice & Gangestad, 2020; Gelman & Loken, 2014). We propose adhering to guidelines as much as possible and stating all decisions transparently. This is important for understanding why similar or different affect structures are found across studies with the same characteristics.

Another limitation pertains to the period of the data collection. The data were collected when many Covid-19 restrictions were still in place. This means that many of the individuals in this study only had limited interactions with others and may not have experienced a wide variety of contexts (e.g., school, work, home, sports). Future studies could look into more specific events and interpersonal interactions that can explain context-related changes in the affect structure.

Finally, the covariate *negative event intensity* comes with two limitations. Firstly, that it was measured with a single item. Secondly, that the median value of *negative event intensity* was 3, on a scale from 0 to 100, and the distribution was highly skewed to the right (mode = 0; mean = 16). There was thus neither a lot of variability nor did participants report very intense negative events. Future studies could look into more specific events and interpersonal interactions that can explain context-related changes in the affect structure in a more nuanced manner.

Conclusion

By applying latent Markov factor analysis to our experience sampling data, we found that individuals transition between two affect structures with a nuanced affect structure. In this way, our results support previous findings suggesting that (1) more than just two dimensions (i.e., positive and negative affect) are needed to accurately capture momentary affect, (2) affect structure, like affect itself, is not stable but can vary over time (at least for some individuals), and (3) the way individuals transition between different affect structures can be influenced by contextual cues such as the occurrence of a stressor. We encourage other researchers to examine the dynamics of the affect structure in their data as well. This will complement current theories of affect by providing detailed insights into which theories of affect apply to momentary measures of affect, for which individuals, and when. By preregistering and testing specific hypotheses about the effects of context- and individual-specific variables on momentary transitions and latent subgroup memberships (capturing differences dynamics across individuals), respectively, we will eventually also be able to learn more about the *why*.

References

Adolf, J., Schuurman, N. K., Borkenau, P., Borsboom, D., & Dolan, C. V. (2014). Measurement invariance within and between individuals: A distinct problem in testing the equivalence of intra- and inter-individual model structures. *Frontiers in Psychology, 5*(883). https://doi.org/10.3389/fpsyg.2014.00883

Barrett, L. F. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion, 12*(4), 579–599. https://doi.org/10.1080/026999398379574

Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology, 87*(2), 266–281. https://doi.org/10.1037/0022-3514.87.2.266

Barrett, L. F. (2006). Valence is a basic building block of emotional life. *Journal of Research in Personality*, *40*(1), 35–55. https://doi.org/10.1016/j.jrp.2005.08.006

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you’re feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition & Emotion*, *15*(6), 713–724. https://doi.org/10.1080/02699930143000239

Barrett, L. F., & Russell, J. A. (1999). The Number of Dimensions. *Current Directions in Psychological Science, 8*(1), 10–14.

Barrett, L. F., & Russell, J. A. (Eds.). (2014). *The Psychological Construction of Emotion*. Guilford Publications.

Bartolucci, F. (2006). Likelihood inference for a class of latent Markov models under linear hypotheses on the transition probabilities. *Journal of the Royal Statistical Society: Series*

B (Statistical Methodology), *68*(2), 155–178. https://doi.org/10.1111/j.1467-9868.2006.00538.x

Bauer, D. J. (2007). Observations on the Use of Growth Mixture Models in Psychological Research. *Multivariate Behavioral Research*, *42*(4), 757–786. https://doi.org/10.1080/00273170701710338

Brose, A., De Roover, K., Ceulemans, E., & Kuppens, P. (2015). Older adults' affective experiences across 100 days are less variable and less complex than younger adults'. *Psychology and Aging*, *30*(1), 194–208. https://doi.org/10.1037/a0038690

Bulteel, K., Mestdagh, M., Tuerlinckx, F., & Ceulemans, E. (2018). VAR(1) based models do not always outpredict AR(1) models in typical psychological applications. *Psychological Methods*, *23*(4), 740–756. https://doi.org/10.1037/met0000178

Bulteel, K., Wilderjans, T. F., Tuerlinckx, F., & Ceulemans, E. (2013). CHull as an alternative to AIC and BIC in the context of mixtures of factor analyzers. *Behavior Research Methods*, *45*(3), 782–791. https://doi.org/10.3758/s13428-012-0293-y

Carroll, J. M., Yik, M. S. M., Russell, J. A., & Feldman Barrett, L. (1999). On the psychometric principles of affect. *Review of General Psychology*, *3*(1), 14–22. https://doi.org/10.1037/1089-2680.3.1.14

Ceulemans, E., & Kiers, H. A. L. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *British Journal of Mathematical and Statistical Psychology*, *59*(1), 133–150. https://doi.org/10.1348/000711005X64817

Cloos, L., Ceulemans, E., & Kuppens, P. (2022). Development, validation, and comparison of self-report measures for positive and negative affect in intensive longitudinal research. *Psychological Assessment*, *35*(3), 189. https://doi.org/10.1037/pas0001200

Cloos, L., Ceulemans, E., & Kuppens, P. (2023, July 14). The Measurement of Affect in Intensive Longitudinal Research. Retrieved from osf.io/485qy

Cloos, L., Vogelsmeier, L. V. D. E., Kuppens, P., & Ceulemans, E. (2023, June 19). Evaluating Dynamics in Affect Measurement with Latent Markov Factor Analysis. Retrieved from osf.io/q47pf

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*(1), 505–528. https://doi.org/10.1146/annurev.psych.57.102904.190146

Dejonckheere, E., Bastian, B., Fried, E. I., Murphy, S. C., & Kuppens, P. (2017). Perceiving social pressure not to feel negative predicts depressive symptoms in daily life. *Depression and Anxiety*, *34*(9), 836–844. https://doi.org/10.1002/da.22653

Dejonckheere, E., Mestdagh, M., Houben, M., Erbas, Y., Pe, M., Leuven, K. U., Koval, P., Brose, A., Bastian, B., & Kuppens, P. (2018). The bipolarity of affect and depressive symptoms. *Journal of Personality and Social Psychology, 114*(2). https://doi.org/10.1037/pspp0000186.supp

Dejonckheere, E., Mestdagh, M., Verdonck, S., Lafit, G., Ceulemans, E., Bastian, B., & Kalokerinos, E. K. (2021). The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion*, *21*(2), 326–336. https://doi.org/10.1037/emo0000697 del Giudice, M., & Gangestad, S. W. (2020). A traveler’s guide to the multiverse: Promises, pitfalls, and a framework for the evaluation of analytic decisions. *Advances in Methods and Practices in Psychological Science, 4*(1). https://doi.org/10.1177/2515245920954925

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

Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology*, *69*(1), 130–141. https://doi.org/10.1037/0022-3514.69.1.130

Eisele, G., Lafit, G., Vachon, H., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2021). Affective structure, measurement invariance, and reliability across different experience sampling protocols. *Journal of Research in Personality, 92*, 104094. https://doi.org/10.1016/j.jrp.2021.104094

Ekkekakis, P. (2013). *The measurement of affect, mood, and emotion: A guide for health-behavioral research.* Cambridge University Press. https://doi.org/10.1017/CBO9780511820724

Ellsworth, P., & Scherer, K. (2003). Appraisal processes in emotion. In R. J. Davidson, K. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572–595).

Erbas, Y., Ceulemans, E., Kalokerinos, E. K., Houben, M., Leuven, K. U., Koval, P., Pe, M. L., & Kuppens, P. (2018). Why I don’t always know what I’m feeling: The role of stress in within-person fluctuations in emotion differentiation. *Journal of Personality and Social Psychology, 115*(2), 179–191. https://doi.org/10.1037/pspa0000126.supp

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*(12), 1050–1057. https://doi.org/10.1111/j.1467-9280.2007.02024.x

Fox, A. S., Lapate, R. C., Shackman, A. J., & Davidson, R. J. (Eds.). (2018). *The Nature of Emotion: Fundamental Questions* (2nd ed.). Oxford University Press.

Gelman, A., & Loken, E. (2014). The Statistical Crisis in Science. *American Scientist*, *102*(6), 460. https://doi.org/10.1511/2014.111.460

Heininga, V. E., Dejonckheere, E., Houben, M., Obbels, J., Sienaert, P., Leroy, B., Van Roy, J., & Kuppens, P. (2019). The dynamical signature of anhedonia in major depressive disorder: Positive emotion dynamics, reactivity, and recovery. *BMC Psychiatry*, *19*(1). https://doi.org/10.1186/s12888-018-1983-5

Heininga, V. E., & Kuppens, P. (2021). Psychopathology and positive emotions in daily life. *Current Opinion in Behavioral Sciences*, *39*, 10–18. https://doi.org/10.1016/j.cobeha.2020.11.005

Hofmann, B. U., & Meyer, T. D. (2006). Mood fluctuations in people putatively at risk for bipolar disorders. *British Journal of Clinical Psychology*, *45*(1), 105–110. https://doi.org/10.1348/014466505X35317

Jacobson, N. C., Evey, K. J., Wright, A. G. C., & Newman, M. G. (2023). Integration of discrete and global structures of affect across three large samples: Specific emotions within-persons and global affect between-persons. *Emotion*, *23*(4), 1202–1211. https://doi.org/10.1037/emo0001022

Kalokerinos, E. K., Murphy, S. C., Koval, P., Bailen, N. H., Crombez, G., Hollenstein, T., Gleeson, J., Thompson, R. J., van Ryckeghem, D. M. L., Kuppens, P., & Bastian, B.

(2020). Neuroticism may not reflect emotional variability. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(17), 9270–9276. https://doi.org/10.1073/pnas.1919934117

Kuppens, P. (2019). Improving theory, measurement, and reality to advance the future of emotion research. *Cognition and Emotion*, *33*(1), 20–23. https://doi.org/10.1080/02699931.2018.1536037

Lapate, R. C., & Heller, A. S. (2020). Context matters for affective chronometry. *Nature Human Behaviour*, *4*(7), 688–689. https://doi.org/10.1038/s41562-020-0860-7

Larsen, J. T., & McGraw, A. P. (2014). The Case for Mixed Emotions: Mixed Emotions. *Social and Personality Psychology Compass*, *8*(6), 263–274. https://doi.org/10.1111/spc3.12108

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, *81*(4), 684–696. https://doi.org/10.1037/0022-3514.81.4.684

Lindquist, K. A., & Barrett, L. F. (2008). Emotional complexity. In M. Lewis, J. M. Haviland-Jones, & L. Feldman Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 513–530). The Guilford Press.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*(1), 21–39. https://doi.org/10.1037/1082-989X.10.1.21

McLachlan, G., & Peel, D. (2000). Mixtures of Factor Analyzers. *Proceedings of The17th International Conference on Machine Learning*, 599–606.

McNeish, D., & Harring, J. R. (2017). The effect of model misspecification on growth mixture model class enumeration. *Journal of Classification*, *34*(2), 223–248. https://doi.org/10.1007/s00357-017-9233-y

McNeish, D., Mackinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2021). Measurement in Intensive Longitudinal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(5), 807–822. https://doi.org/10.1080/10705511.2021.1915788

McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, *52*(6), 2287–2305. https://doi.org/10.3758/s13428-020-01398-0

Mneimne, M., Fleeson, W., Arnold, E. M., & Furr, R. M. (2018). Differentiating the everyday emotion dynamics of borderline personality disorder from major depressive disorder and bipolar disorder. *Personality Disorders: Theory, Research, and Treatment*, *9*(2), 192–196. https://doi.org/10.1037/per0000255

Molenaar, P. C. M., & Nesselroade, J. R. (2009). The Recoverability of P-technique Factor Analysis. *Multivariate Behavioral Research*, *44*(1), 130–141. https://doi.org/10.1080/00273170802620204

Moors, A. (Ed.). (2022). *Demystifying emotions: A typology of theories in psychology and philosophy*. Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781107588882.015

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A. L., De Schryver, M., De Winne, J., & Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, *45*(1), 169–177. https://doi.org/10.3758/s13428-012-0243-8

Mroczek, D. K., & Almeida, D. M. (2004). The effect of daily stress, personality, and age on daily negative affect. *Journal of Personality*, *72*(2), 355–378. https://doi.org/10.1111/j.0022-3506.2004.00265.x

Osgood, C. E. (1966). Dimensionality of the semantic space for communication via facial expressions. *Scandinavian Journal of Psychology*, *7*(1), 1–30. https://doi.org/10.1111/j.1467-9450.1966.tb01334.x

Panksepp, J. (2004). *Affective Neuroscience: The Foundations of Human and Animal Emotions: Vol. Volume 4 of Series in Affective Science*. Oxford University Press.

Potter, P., Zautra, A. J., & Reich, J. (2000). Stressful events and information processing dispositions moderate the relationship between positive and negative affect: Implications for pain patients. *Annals of Behavioral Medicine*, *22*(3), 191–198. https://doi.org/10.1007/BF02895113

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Russell, J. A., & Feldman Barrett, L. (2009). Core affect. In D. Sander & Scherer, K.R. (Eds.), *The Oxford companion to emotion and the affective sciences*. Oxford University Press.

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(3), 273–294. https://doi.org/10.1016/0092-6566(77)90037-X

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience Sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*, *4*(1), 5–34. https://doi.org/10.1023/A:1023605205115

Scott, S. B., Sliwinski, M. J., Zawadzki, M., Stawski, R. S., Kim, J., Marcusson-Clavertz, D., Lanza, S. T., Conroy, D. E., Buxton, O., Almeida, D. M., & Smyth, J. M. (2020). A coordinated analysis of variance in affect in daily life. *Assessment*, *27*(8), 1683–1698. https://doi.org/10.1177/1073191118799460

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of Cognitive Appraisal in Emotion. *Journal of Personality and Social Psychology*, *48*(4), 813–838. https://doi.org/10.1037//0022-3514.48.4.813

Smith, C. A., & Lazarus, R. (1990). Emotion and adaptation. In L. A. Pervin (Ed.), *Handbook of personality: Theory & Research.* (pp. 609–637). Guliford.

Vermunt, J. K., & Magidson, J. (2016). *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. Statistical Innovations Inc. http://www.statisticalinnovations.com

Vogelsmeier, L. V. D. E., Vermunt, J. K., Böing-Messing, F., & De Roover, K. (2019). Continuous-time latent Markov factor analysis for exploring measurement model changes across time. *Methodology*, *15*(1), 29–42. https://doi.org/10.1027/1614-2241/A000176

Vogelsmeier, L. V. D. E., Vermunt, J. K., Bülow, A., & De Roover, K. (2023). Evaluating covariate effects on ESM measurement model changes with latent Markov factor analysis: A three-step approach. *Multivariate Behavioral Research*, *58*(0), 262–291. https://doi.org/10.1080/00273171.2021.1967715

Vogelsmeier, L. V. D. E., Vermunt, J. K., & De Roover, K. (2022). How to explore within-person and between-person measurement model differences in intensive longitudinal data with the R package lmfa. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01898-1

Vogelsmeier, L. V. D. E., Vermunt, J. K., Keijsers, L., & De Roover, K. (2020). Latent Markov latent trait analysis for exploring measurement model changes in intensive longitudinal data. *Evaluation and the Health Professions*, *44*(1). https://doi.org/10.1177/0163278720976762

Vogelsmeier, L. V. D. E., Vermunt, J. K., van Roekel, E., & De Roover, K. (2019). Latent Markov factor analysis for exploring measurement model changes in time-intensive longitudinal studies. *Structural Equation Modeling*, *26*(4), 557–575. https://doi.org/10.1080/10705511.2018.1554445

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. https://doi.org/10.1037/0033-2909.98.2.219

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838. https://doi.org/10.1037/0022-3514.76.5.820

Wilderjans, T. F., Ceulemans, E., & Meers, K. (2013). CHull: A generic convex-hull-based model selection method. *Behavior Research Methods*, *45*(1), 1–15. https://doi.org/10.3758/s13428-012-0238-5

Zautra, A. J., Berkhof, J., & Nicolson, N. A. (2002). Changes in affect interrelations as a function of stressful events. *Cognition and Emotion*, *16*(2), 309–318. https://doi.org/10.1080/02699930143000257

Zautra, A. J., Potter, P. T., & Reich, J. W. (1997). The independence of affects is context-dependent: An integrative model of the relationship between positive and negative affect Personal mastery and locus of control. In K. W. Schaie & Powell Lawton (Eds.), *Annual Review of Gerontology & Geriatrics* (Vol. 17, pp. 75–103). Springer.