Reply to: Context matters for affective chronometry

Egon Dejonckheere*

Merijn Mestdagh*

Peter Kuppens

Francis Tuerlinckx

KU Leuven – Faculty of Psychology and Educational Sciences

Word Count = 1,901

Display Items = 2

Number of References = 16

* Egon Dejonckheere and Merijn Mestdagh contributed equally to this reply.

Correspondence concerning this reply should be addressed to Egon Dejonckheere, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, Leuven, 3000, Belgium.

E-mail: egon.dejonckheere@kuleuven.be.

Reply to: Context matters for affective chronometry

In Dejonckheere, Mestdagh and colleagues¹, we demonstrate that many commonly studied affect dynamic measures in the experience sampling (ESM) literature essentially explain no additional variance in various psychological well-being outcomes, once the explanatory power of basic mean levels of positive (PA) and negative affect (NA), and the variability in these affective states, is accounted for. In an attempt to foster cumulative science, we encourage researchers to control for these static covariates before attesting the incremental value of more complex, time-dynamic measures in the prediction of psychological (mal)adjustment.

In a convincing comment, Lapate and Heller² contend that our non-findings require further contextualization. Backed up by a literature review of experimental and ESM research, they accentuate the added value of emotional recovery over average levels of affect to explain between-person variation in various mental health outcomes: The temporal slope with which an individual returns to an emotional baseline after a contextual stressor conveys meaningful information about that person's psychological well-being. In light of this evidence, the authors conclude that, in the explanation of individual differences in well-being, affective researchers “should not throw the baby out with the bathwater”, but instead unravel the exact contextual conditions under which affective chronometry shows incremental value above and beyond mean levels of affect.

We largely concur with this conclusion, as discussed in our original article¹. Indeed, in reviewing the potential implications of our findings, we explicitly state that our results “do not necessarily renounce the importance of affect dynamics in psychological well-being” (p. 486). Similarly, in formulating guidelines to improve the current modus operandi of our field, we argue that unique relations between affect dynamics and psychological well-being may more likely be uncovered, “when researchers ask participants about their subjective emotional

experiences in relation to specific events” (p.486). Taken together, we share a similar aspiration to not prematurely deny the unique role of affective dynamics in psychopathology or well-being, but to pinpoint the specific study determinants that boost or diminish their predictive value instead.

That being said, while Lapate and Heller interpret the results of their literature review exclusively as evidence for the promising role of context, we would argue that their referenced research³⁻⁹ also varies in other meaningful ways from the traditional ESM protocols described in our article¹. In our perspective, the crucial reason why these cited studies manage to establish the added value of emotional recovery is not limited to context, but should be interpreted against a broader background of typically higher signal-to-noise ratios (SNRs) in the time series these studies investigate:

$$SNR = \frac{Var(\widehat{PA}_t = a\widehat{PA}_{t-1} + \varepsilon_t)}{Var(\omega_t)} \quad (1)$$

Substantively, the SNR of an affective time series is defined as the ratio of meaningful emotional signal to measurement noise. In the study of affect dynamics, this emotional signal statistically refers to the variance of a latent auto-regressive (AR) model of order 1 [i.e., an AR(1) model]¹⁰, which is defined by an AR parameter (a) that captures the degree with which an individual’s latent affective state (e.g., $\widehat{PA}$) changes from one assessment to the next (i.e., inertia¹¹), and an innovation or dynamic error term ($\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$) that roughly corresponds to the intensity of the emotional stimulus that was introduced, and that carries over to the next assessments via this AR relation¹⁰. In contrast, measurement noise refers to the variance in measurement error that is specific for each particular emotional assessment ($\omega_t \sim N(0, \sigma_\omega^2)$), and does not resonate to subsequent assessments (see Supplementary Notes 1 for more information on the computation of the SNR).

Here, we suggest that low SNRs in traditional ESM research (compared to the studies brought up by Lapate and Heller³⁻⁹) may lie at the basis of our initial non-findings. Consequently, when ESM researchers seek to maximize the SNR of the emotional time series they investigate, we believe the added contribution of real-life affect dynamics in well-being may become apparent. As Equation 1 reveals, this could be achieved in multiple ways (see Figure 1 for a graphical visualisation). On the one hand, ESM researchers could focus on increasing the emotional signal (a) by either studying emotional reactions to stronger contextual stimuli (i.e., impacting ε_t as proposed by Lapate and Heller²; as opposed to Figure 1 Panel B), or (b) by increasing the AR relation through assessing emotions with a finer temporal resolution (i.e., impacting a ; as opposed to Figure 1 Panel C). On the other hand, ESM researchers could also aim to decrease the measurement noise associated with emotional assessments (c) by relying on assessment procedures that are more reliable (i.e., impacting ω_t ; as opposed to Figure 1 Panel C). We will illustrate the promise of this overarching framework by comparing the traditional ESM studies covered in our study¹ with the referenced research³⁻⁹ by Lapate and Heller for each of these parameters.

First, an inherent limitation to traditional ESM designs is that we typically have no control over the contextual input (ε_t) of participants' subjective emotional experiences. In fact, in many instances, ESM researchers are completely blind to the exact emotion-eliciting stimuli that underlie the ups and downs in participants' affective time series. Because we track participants' emotions in the complexity of everyday life, it is difficult to anchor their emotional evaluations to specific objective events or stimuli. Consequently, affective assessments in traditional ESM studies are often the product of a complex interplay of diffuse stimuli and short-lived events, producing emotional time series that generally carry a weak emotional signal (see Figure 1 Panel B). This is in marked contrast with the experimental lab studies³⁻⁸ cited by Lapate and Heller, where researchers have full control over the contextual

input participants receive. In these experiments, researchers track various indicators of emotion in response to a set of carefully selected and strong affective stimuli that are identical across participants (e.g., pictures^{3–7}, film clips⁸). Although it remains unclear to what extent emotional responding to standardized, yet artificial lab stimuli generalises to real-life settings¹², anchoring these assessments to specific stimuli yields a stronger emotional signal.

To empirically illustrate that the anchoring of emotional assessments produces a stronger emotional signal, we computed the median SNR for all (unanchored) PA and NA time series in the traditional ESM studies of our meta-analysis (see Supplementary Notes 1 and Supplementary MATLAB code for exact computation procedure). We compared these with the SNRs of a quasi-experimental ESM study that investigated the anchored PA and NA trajectories of 101 first-year university students in specific relation to the release of their exam results (i.e., anchored emotional assessments; e.g., “When you think about your grades right now, how [positive / negative] do you feel right now?”)¹³. As shown in Figure 2, the median SNRs for PA and NA in this latter study were almost six times larger than the ones we observed in the studies from our meta-analysis, which demonstrates that event-related ESM research captures a relatively stronger emotional signal. In contrast, almost all of our traditional ESM studies had median PA and NA SNRs circling around 1, suggesting that participants’ emotional signal was considerably equivocal, and less stipulated by a signal event (i.e., 36% of all participants in our meta-analysis had an emotional SNR smaller than 1). In sum, this comparison suggests that examining real-life perturbations (e.g., the release of students’ exam results) holds promise for establishing unique associations between affect dynamics and well-being in daily life, as anchored PA and NA time series are more directed and pronounced, and may therefore more resemble the signal value found in standardized experiments.

Second, traditional ESM designs investigate the dynamics of participants' real-life emotions on a time scale that is considerably larger than the studies brought up by Lapate and Heller³⁻⁸. This leaves the auto-regressive effect (a) of consecutive emotional assessments to be relatively weak (see Figure 1 Panel C). While our meta-analysis focussed on the incremental value of affect dynamics that were computed from emotion ratings that were typically hours (or days) apart, many of their cited studies evaluated emotional recovery on a second-to-second basis (i.e., virtually continuously^{3,6,8}). As noted in our discussion (p. 485), “the fact that the temporal resolution in typical ESM research may be insufficient to capture meaningful regularities in affective trajectories” could be another reason why independent associations between affective dynamics and psychological adjustment are more difficult to establish in the reality of everyday life¹⁴: Due to the typically large intervals between two consecutive emotional assessments, the auto-regressive relation is simply too small to pick up a meaningful emotional signal that effectively outweighs the inevitable measurement noise in participants' responding.

To demonstrate that differences in the strength of the auto-regressive effect also impact the SNR of an emotional time series, we trimmed the data of the original exam-anchored ESM study¹³ and only considered every fifth emotional assessment. Next, we compared the median SNR for the original versus trimmed PA and NA time series. Although the contextual input was constant across approaches, Figure 2 illustrates that the SNRs for PA and NA were drastically impaired when the auto-regressive effects of PA and NA was reduced. In sum, this comparison suggests that compressing the time interval between consecutive assessments allows for a better detection of the emotion signal underlying participants' responses. However, if a more fine-grained temporal resolution comes with an increase in the number of emotional assessments, this may also build up the burden or reactivity associated with ESM¹⁵. This brings us to the last parameter that defines the SNR.

Finally, traditional ESM differs from the experimental studies discussed by Lapate and Heller³⁻⁷ in the assessment procedure they adopt to deduce emotional states. To track affective fluctuations in daily life, the ESM protocols in our article rely on numerous emotional self-reports over an extended period of time (i.e., weeks or months¹). In contrast, these experiments record temporal changes in various neurological^{3,4} or psychophysiological⁵⁻⁷ indicators of emotions within the timeframe of an hour or less. Furthermore, before analysing these experimental time series, emotional responses are often pooled together across multiple trials to acquire a robust and reliable emotional signal that reduces measurement noise (ω_t). Thus, not only do the cited studies differ in the emotional components they consider (with physiological and experiential measures actually showing little convergence¹⁶), the burden and reactivity related to repeated self-reports versus effortless and unconscious experimental assessments may be an additional reason why affective chronometry is easier to establish in the lab versus daily life: The real-time self-monitoring of emotions in the complexity of everyday life may be more error prone, which conceals participants' true emotional signal (see Figure 1 Panel D).

In sum, we feel that Lapate and Heller's literature review² does not contradict our findings, nor do we believe that our conclusions reported in Dejonckheere, Mestdagh and colleagues¹ refute the results of these earlier studies. Rather, a comparison between the traditional ESM studies in our meta-analysis and the experiments these authors refer to elucidates the importance of maximizing the SNR of the affective time series ESM researchers investigate. While we concur with Lapate and Heller that anchoring emotional assessments to contextual stimuli is essential to uncover unique relations between affective chronometry and psychological well-being, we believe that the current non-findings in ESM should be framed within a broader context of typically low SNRs in traditional ESM protocols. Besides investigating stronger emotional stimuli, pursuing time series with a more

fine-grained temporal resolution and improving practices to reduce measurement error are potential advancements for ESM researchers who aim to understand how real-life affect dynamics are important for people's well-being.

Author Contributions

E.D. and M.M. contributed equally to the manuscript, both drafting parts of this reply. P.K. and F.T critically revised earlier versions of the manuscript. All authors approved the final version.

Competing Interests

The authors declare no competing interests.

Code Availability

All analyses reported in this reply were conducted in MATLAB (R2017a). The code to reproduce our results is provided in the Supplementary MATLAB Code, and is online available from the Open Science Framework (<http://osf.io/zm6uw>).

Data Availability

In this reply, we rely on the original datasets reported in Dejonckheere, Mestdagh, et al. (2019)¹, of which two are publicly available from the Open Science Framework (<http://osf.io/zm6uw>). For the other datasets, restrictions apply to the availability of these data, as they were used under license for that particular study, and so are not publicly available. Finally, the data for Dejonckheere et al. (2019)¹³ can be found on the Open Science Framework (<https://osf.io/yte2w/>).

References

1. Dejonckheere, E. et al. Complex affect dynamics add limited information to the prediction of psychological well-being. *Nat Hum Behav* **3**, 478–491 (2019).
<https://doi.org/10.1038/s41562-019-0555-0>
2. Lapate, R. C. & Heller A. S. *Nat Hum Behav* **X**, xxx–xxx (2020).
3. Heller, A. S. *et al.* Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of fronto-striatal brain activation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 22445–22450 (2009). <https://doi.org/10.1073/pnas.0910651106>
4. Heller, A. S. *et al.* Sustained striatal activity predicts eudaimonic well-being and cortisol output. *Psychol Sci* **24**, 2191–2200 (2013). <https://doi.org/10.1177/0956797613490744>
5. Javaraś, K. N. *et al.* Conscientiousness predicts greater recovery from negative emotion. *Emotion* **12**, 875–881 (2012). <https://doi.org/10.1037/a0028105>
6. Lapate, R. C. *et al.* Prolonged marital stress is associated with short-lived responses to positive stimuli. *Psychophysiology* **51**, 499–509 (2014).
<https://dx.doi.org/10.1111%2Fpsyp.12203>
7. Schaefer, S. M. *et al.* Purpose in life predicts better emotional recovery from negative stimuli. *PLoS ONE* **8**, e80329 (2013). <https://doi.org/10.1371/journal.pone.0080329>
8. McMakin, D. L., Santiago, C. D. & Shirk, S. R. The time course of positive and negative emotion in dysphoria. *The Journal of Positive Psychology* **4**, 182–192 (2009).
<https://dx.doi.org/10.1080%2F17439760802650600>
9. Metalsky, G. I., Joiner, T. E., Hardin, T. S. & Abramson, L. Y. Depressive reactions to failure in a naturalistic setting: a test of the hopelessness and self-esteem theories of depression. *J Abnorm* **102**, 101–109 (1993). <https://doi.org/10.1037//0021-843x.102.1.101>

10. Schuurman, N. K., Houtveen, J. H. & Hamaker E. L. Incorporating measurement error in $n = 1$ psychological autoregressive modelling. *Front Psychol* **28**, 1038 (2015).
<http://dx.doi.org/10.3389/fpsyg.2015.01038>
11. Kuppens, P., Oravecz, Z. & Tuerlinckx, F. Feelings change: accounting for individual differences in the temporal dynamics of affect. *J Pers Soc Psychol* **99**, 1042–1060 (2010).
<https://doi.org/10.1037/a0020962>
12. Rottenberg, J. & Hindash, A. C. Emerging evidence for emotion context insensitivity in depression. *Curr Opin Psychol* **4**, 1–5 (2015).
<https://doi.org/10.1016/j.copsyc.2014.12.025>
13. Dejonckheere, E. *et al.* The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion*, Advanced online publication (2019). <https://doi.org/10.1037/emo0000697>
14. Ebner-Priemer, U. W. & Sawitzki, G. Ambulatory assessment of affective instability in borderline personality disorder: the effect of the sampling frequency. *Eur J Psychol Assess* **23**, 238–247 (2007). <http://dx.doi.org/10.1027/1015-5759.23.4.238>
15. Vachon, H., Rintala, A., Viechtbauer, W. & Myin-Germeys, I. Data quality and feasibility of the experience sampling method across the spectrum of severe psychiatric disorders: a protocol for a systematic review and meta-analysis. *Syst Rev* **7**, 1–5
<https://doi.org/10.1186/s13643-018-0673-1>
16. Mauss, I. B. & Robinson, M. D. Measures of emotion: A review. *Cogn Emot* **23**, 209–237 (2009). <https://doi.org/10.1080/02699930802204677>

Figure Legends

Fig. 1 | Evaluating how different study determinants impact the signal-to-noise ratio. Simulated affective time series for a hypothetical participant after introducing an emotional stimulus at measurement occasion 3. In each graph, the blue line represents the latent emotional signal, while the red dots refer to the actual emotional assessments. **a**, High SNR due to a strong emotional stimulus, a high measurement resolution, and low measurement error. **b**, Low SNR due to weak emotional stimulus (measurement resolution and measurement error are held constant). **c**, Low SNR due to low measurement resolution (emotional stimulus and measurement error are held constant). **d**, Low SNR due to high measurement error (emotional stimulus and measurement resolution are held constant).

Fig. 2 | Comparing PA and NA signal-to-noise ratios in traditional versus event-related ESM studies. SNRs for positive affect (**a**) and negative affect (**b**) were calculated for each subject following the approach outlined in Schuurman and colleagues¹⁰ (see Supplementary Notes 1). The median SNRs of PA and NA are visualised for each dataset, together with a 95% confidence interval derived from 2,000 bootstraps. Blue bars represent the datasets in our meta-analysis with a traditional ESM protocol (see Dejonckheere, Mestdagh and colleagues¹ for actual references). The red bars refers to an event-related ESM study¹³, where emotional assessments were anchored to the release of participants' exam results. Here, we compare the SNR of the original versus trimmed dataset (in which we only consider every 5th emotional assessment). See Supplementary Figures 1 and 2 for participants' individual data points.