MATTERS ARISING
https://doi.org/10.1038/s41562-020-0861-6

nature human behaviour

Check for updates

Reply to: Context matters for affective chronometry

Egon Dejonckheere[1,2] ✉, Merijn Mestdagh[1,2], Peter Kuppens[1] and Francis Tuerlinckx[1]

REPLYING TO R. C. Lapate & A. S. Heller *Nature Human Behaviour* https://doi.org/10.1038/s41562-020-0860-7 (2020)

In Dejonckheere, Mestdagh and colleagues[1], 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 affect (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 to the incremental value of more complex, time-dynamic measures in the prediction of psychological (mal)adjustment.

In a convincing comment, Lapate and Heller[2] contend that our non-findings require further contextualization. Backed up by a literature review of experimental and ESM research, they emphasize 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 affective dynamics baby out with the bath-water", but instead unravel the exact contextual conditions under which affective chronometry shows incremental values above and beyond mean levels of affect.

We largely concur with this conclusion, as discussed in our original article[1]. Indeed, in reviewing the potential implications of our findings[1], we explicitly describe that our results do not necessarily renounce the importance of affect dynamics in psychological well-being. 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. Taken together, we share a similar aspiration to not deny the unique role of affective dynamics in psychopathology or well-being prematurely, but instead to pinpoint the specific study determinants that boost or diminish their predictive value.

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 research[3–9] also varies in other meaningful ways from the traditional ESM protocols described in our article[1]. In our view, 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:

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

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 autoregressive (AR) model of order 1 (that is, 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 (for example, $\widetilde{\mathrm{PA}}$) changes from one assessment (t) to the next (that is, inertia[11]) and an innovation or dynamic error term ($\varepsilon_t \approx N\left(0, \sigma_\varepsilon^2\right)$, assumed normally distributed with variance σ_ε^2) that roughly corresponds to the intensity of the emotional stimulus that was introduced, and carries over to the next assessments via this AR relation[10]. In contrast, measurement noise refers to the variance in measurement error that is specific for each particular emotional assessment ($\omega_t \approx N\left(0, \sigma_\omega^2\right)$, assumed normally distributed with variance σ_ω^2), and does not resonate to subsequent assessments (see Supplementary Note 1 for more information on the computation of the SNR).

Here, we suggest that low SNRs in traditional ESM research (compared with the studies mentioned by Lapate and Heller[3–9]) 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 that 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 Fig. 1 for a graphical visualization). On the one hand, ESM researchers could focus on increasing the emotional signal by either studying emotional reactions to stronger contextual stimuli (that is, impacting ε_t as proposed by Lapate and Heller[2]; as opposed to the contrasting scenario shown in Fig. 1b), or by increasing the AR relation through assessing emotions with a finer temporal resolution (that is, impacting a; as opposed to the contrasting scenario shown in Fig. 1c). On the other hand, ESM researchers could aim to decrease the measurement noise associated with emotional assessments by relying on assessment procedures that are more reliable (that is, impacting ω_t; as opposed to the contrasting scenario shown in Fig. 1d). We will illustrate the promise of this overarching framework by comparing the traditional ESM studies covered in our study[1] with the referenced research[3–9] by Lapate and Heller for each of these parameters.

First, an inherent limitation of 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 blinded 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

[1]Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium. [2]These authors contributed equally: Egon Dejonckheere, Merijn Mestdagh. ✉e-mail: egon.dejonckheere@kuleuven.be

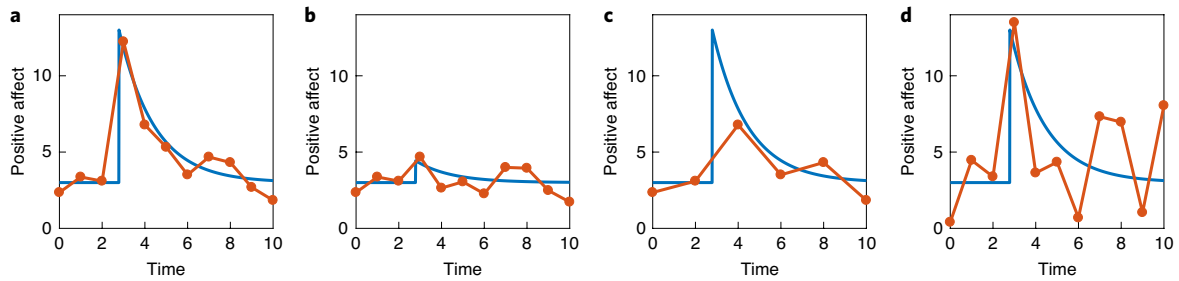


Fig. 1 | Evaluating how different study determinants impact the SNR. Simulated affective time series for a hypothetical participant after introducing an emotional stimulus at measurement time point 3. In each graph, the blue line represents the latent emotional signal and the red dots refer to the actual 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).

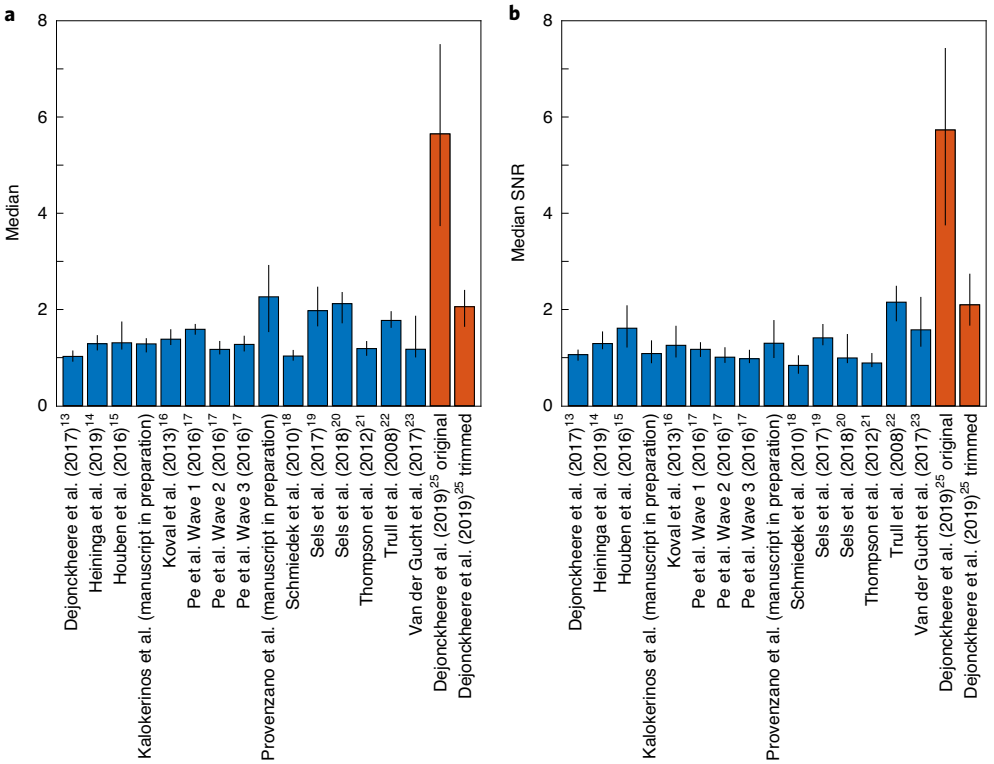


Fig. 2 | Comparing PA and NA SNRs in traditional versus event-related ESM studies. a,**b**, SNRs for PA (**a**) and NA (**b**) calculated for each subject following the approach outlined in Schuurman and colleagues[10] (see Supplementary Note 1). The median SNRs of PA and NA are visualized for each dataset, with the 95% confidence interval derived from 2,000 bootstraps (error bars). Blue bars represent the datasets in our meta-analysis with a traditional ESM protocol[13–23]. The red bars refer to an event-related ESM study[25], in which 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 fifth emotional assessment). See Supplementary Figs. 1 and 2 for participants' individual data points.

carry a weak emotional signal (see Fig. 1b). This is in marked contrast with the experimental lab studies[3–8] 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 (for example, pictures[3–7] or film clips[8]). Although it remains unclear to what extent emotional responses to standardized, yet artificial, lab stimuli generalize to real-life settings[12], 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[13–23] (see Supplementary Note 1 and Supplementary MATLAB[24] code for the 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 (that is, anchored emotional assessments; for example, "When you think about your

grades right now, how [positive/negative] do you feel?")[25]. As shown in Fig. 2, the median SNRs for PA and NA in this latter study were almost six times larger than those observed in the studies from our meta-analysis[13–23], which demonstrates that event-related ESM research captures a stronger emotional signal. In contrast, almost all of our traditional ESM studies[13–23] had median PA and NA SNRs of around 1, suggesting that the participants' emotional signals were considerably equivocal, and less stipulated by a signal event (that is, 36% of all participants in our meta-analysis[13–23] had an emotional SNR smaller than 1). In sum, this comparison suggests that examining real-life perturbations (for example, 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 closely resemble the signal value found in standardized experiments.

Second, traditional ESM designs investigate the dynamics of participants' real-life emotions on a timescale that is considerably longer than the studies cited by Lapate and Heller[3–8]. This leaves the AR relation (*a*) values of consecutive emotional assessments relatively weak (see Fig. 1c). Whereas our meta-analysis[13–23] focused 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 (that is, almost continuously[3,6,8]). As noted in our discussion[1], 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 life[26]. Owing to the typically large intervals between two consecutive emotional assessments, the AR relation is simply too small to detect a meaningful emotional signal that effectively outweighs the inevitable measurement noise in participants' responses.

To demonstrate that differences in the strength of the AR effect also impact the SNR of an emotional time series, we trimmed the data from the original exam-anchored ESM study[25] 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, Fig. 2 illustrates that the SNRs for PA and NA were markedly diminished when their AR effects were reduced. In sum, this comparison suggests that compressing the time interval between consecutive assessments allows for 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 increase the burden or reactivity associated with ESM[27]. This brings us to the last parameter that defines the SNR.

Third, traditional ESM differs from the experimental studies discussed by Lapate and Heller[3–7] in the assessment procedure adopted to deduce emotional states. To track affective fluctuations in daily life, the ESM protocols in our article[13–23] rely on numerous emotional self-reports over an extended period of time (weeks or months[1]). In contrast, those experiments record temporal changes in various neurological[3,4] or psychophysiological[5–7] indicators of emotions within the timeframe of an hour or less. Furthermore, before analysing the 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 convergence[28]), 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 Fig. 1d).

In summary, we feel that Lapate and Heller's literature review[2] does not contradict our findings, nor do we believe that our original conclusions[1] refute the results of these earlier studies. Instead, a comparison between the traditional ESM studies in our meta-analysis and the experiments Lapate and Heller refer to elucidates the importance of maximizing the SNR of the affective time series ESM researchers investigate. Although 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. In addition to investigating stronger emotional stimuli, pursuing time series with a more fine-grained temporal resolution and improving practices to reduce measurement error would constitute potential advancements for ESM researchers who aim to understand how real-life affect dynamics are important for people's well-being.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

We rely on the original datasets reported in Dejonckheere, Mestdagh, et al.[1], of which two are publicly available from the Open Science Framework (http://osf.io/zm6uw). Restrictions apply to the availability of the other datasets, as they were used under license for that particular study, and so are not publicly available. The data for Dejonckheere et al.[25] can be found on the Open Science Framework (https://osf.io/yte2w/).

Code availability

All analyses reported in this reply were conducted in MATLAB (R2017a)[24]. The code for reproducing our results is provided in the Supplementary MATLAB Code, and is available from the Open Science Framework (http://osf.io/zm6uw).

Received: 12 September 2019; Accepted: 13 March 2020;
Published online: 27 April 2020

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).
2. Lapate, R. C. & Heller, A. S. Context matters for affective chronometry. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-020-0860-7 (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. USA* **106**, 22445–22450 (2009).
4. Heller, A. S. et al. Sustained striatal activity predicts eudaimonic well-being and cortisol output. *Psychol. Sci.* **24**, 2191–2200 (2013).
5. Javaras, K. N. et al. Conscientiousness predicts greater recovery from negative emotion. *Emotion* **12**, 875–881 (2012).
6. Lapate, R. C. et al. Prolonged marital stress is associated with short-lived responses to positive stimuli. *Psychophysiology* **51**, 499–509 (2014).
7. Schaefer, S. M. et al. Purpose in life predicts better emotional recovery from negative stimuli. *PLoS ONE* **8**, e80329(2013).
8. McMakin, D. L., Santiago, C. D. & Shirk, S. R. The time course of positive and negative emotion in dysphoria. *J. Posit. Psychol.* **4**, 182–192 (2009).
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).
10. Schuurman, N. K., Houtveen, J. H. & Hamaker, E. L. Incorporating measurement error in $n = 1$ psychological autoregressive modelling. *Front Psychol.* **28**, 1038(2015).
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).

12. Rottenberg, J. & Hindash, A. C. Emerging evidence for emotion context insensitivity in depression. *Curr. Opin. Psychol.* **4**, 1–5 (2015).
13. Dejonckheere, E., Bastian, B., Fried, E. I., Murphy, S. & Kuppens, P. Perceiving social pressure not to feel negative predicts depressive symptoms in daily life. *Depress. Anxiety* **34**, 836–844 (2017).
14. Heininga, V. E. et al. The dynamical signature of anhedonia in major depressive disorder: positive emotion dynamics, reactivity, and recovery. *BMC Psychiatry* **19**, 59 (2019).
15. Houben, M. et al. Emotional switching in borderline personality disorder: a daily life study. *J. Pers. Disord.* **7**, 50–60 (2016).
16. Koval, P., Pe, M. L., Meers, K. & Kuppens, P. Affect dynamics in relation to depressive symptoms: variable, unstable or inert? *Emotion* **13**, 1132–1141 (2013).
17. Pe, M. L., Brose, A., Gotlib, I. H. & Kuppens, P. Affective updating ability and stressful events interact to prospectively predict increases in depressive symptoms over time. *Emotion* **16**, 73–82 (2016).
18. Schmiedek, F., Lövdén, M. & Lindenberger, U. Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study. *Front. Aging Neurosci.* **2**, 1–27 (2010).
19. Sels, L., Ceulemans, E. & Kuppens, P. Partner-expected affect: how you feel now is predicted by how your partner thought you felt before. *Emotion* **17**, 1066–1077 (2017).
20. Sels, L., Ceulemans, E., & Kuppens, P. All's well that ends well? A test of the peak-end rule in couples' conflict discussions. *Eur. J. Soc. Psychol.* https://doi.org/10.1002/ejsp.2547 (2018).
21. Thompson, R. J. et al. The everyday emotional experience of adults with major depressive disorder: examining emotional instability, inertia, and reactivity. *J. Abnorm. Psychol.* **121**, 819–829 (2012).
22. Trull, T. J. et al. Affective instability: measuring a core feature of borderline personality disorder with ecological momentary assessment. *J. Abnorm. Psychol.* **117**, 647–661 (2008).
23. Van der Gucht, K. et al. An experience sampling study examining the potential impact of a mindfulness-based intervention on emotion differentiation. *Emotion* **19**, 123–131 (2018).
24. MATLAB. version 9.2.0 (R2017a) (The MathWorks Inc., 2017).
25. Dejonckheere, E. et al. The relation between positive and negative affect becomes more negative in response to personally relevant events. *Emotion* https://doi.org/10.1037/emo0000697 (2019).
26. 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).
27. 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**, 7 (2018).
28. Mauss, I. B. & Robinson, M. D. Measures of emotion: a review. *Cogn. Emot.* **23**, 209–237 (2009).

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.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-020-0861-6.

Correspondence and requests for materials should be addressed to E.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

nature research

Corresponding author(s): Egon Dejonckheere

Last updated by author(s): Feb 28, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For details on the data collection software used in each individual study, we refer to the sources referenced in Supplementary Table 1 of the original article (Dejonckheere, Mestdagh, et al., 2019).
Data analysis	All analyses reported in this article were conducted in MATLAB (version R2017a).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

We rely on the original datasets reported in Dejonckheere, Mestdagh, et al. (2019), of which two are publicly available from OSF (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. The data for Dejonckheere et al. (2019) can be found from OSF (https://osf.io/yte2w/).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Research sample	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Sampling strategy	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Data collection	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Timing	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Data exclusions	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Non-participation	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Randomization	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Recruitment	See reporting summary of our original article (Dejonckheere, Mestdagh, et al., 2019).
Ethics oversight	See our original article (Dejonckheere, Mestdagh, et al., 2019).

Note that full information on the approval of the study protocol must also be provided in the manuscript.