OSF Preregistration

Preregistration template

(Pre-)registration template for experience sampling methodology (ESM) research

Kirtley, O. J., Lafit, G., Achterhof, R., Hiekkaranta, A. P., & Myin-Germeys, I. (2020, March 12). A template and tutorial for (pre-)registration of studies using Experience Sampling Methods (ESM). https://doi.org/10.17605/OSF.IO/2CHMU

The template is accompanied by an open access tutorial paper: Kirtley, O. J., Lafit, G., Achterhof, R., Hiekkaranta, A. P., & Myin-Germeys, I. (2021). Making the Black Box Transparent: A Template and Tutorial for Registration of Studies Using Experience-Sampling Methods. Advances in Methods and Practices in Psychological Science. https://doi.org/10.1177/2515245920924686

Study Information

1. Title (required)
   Naming before taming? The pivotal role of emotion differentiation to emotion regulation variability in adolescents.

2. Authors (required)

   Tak Tsun Lo, MSSc
   Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

   Dr Dominique Maciejewski
   Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

   Dr Loes Pouwels
   Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

   Prof Dr Jacqueline Vink
   Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

   Dr Maaike Verhagen
   Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

3. Description (optional)

Is knowing what emotions one experiences important for how one regulates them? Emotion differentiation (ED) refers to how well one can distinguish and label emotional experiences (Barrett et al., 2001). According to the theoretical framework of Kashdan et al. (2015), ED precedes and facilitates emotion regulation (ER), which refers to how one influences emotion experiences (Gross, 1998) for hedonic or instrumental goals (Tamir, 2009). However, Kalokerinos et al. (2019) found mostly insignificant or inconsistent relationships between ED at one moment and selection of specific ER strategies (e.g., reappraisal, distraction) at the next moment. Kalokerinos et al. (2019) suspected it was due to examination of separate strategies which fail to capture how ER strategies flexibly interact with changing contexts in daily life. This concurs with Aldao et al. (2015)'s suggestion to study ER beyond separate strategies by collectively considering the variability in the use of ER strategies across occasions. Aldao et al. (2015) proposed ER variability as a foundation to capture ER flexibility (ER variability that reacts to changing contexts), which is a necessary (but not sufficient) condition in achieving adaptive outcomes. The current study reexamines Kashdan et al. (2015)'s theory by investigating the relationship between ED and ER variability in daily lives of adolescents.

In this study, both between-person individual differences and within-person variations will be examined, as they can be different (Molenaar & Campbell, 2009). There are many examples where variables may associate in one direction across a population but oppositely within each individual across different assessments. As an example, there is a negative between-person association between physical activity and heart rate (i.e., people who exercise more generally have lower heart rate) but a positive within-person association between physical activity and heart rate (i.e., there is a higher heart rate in vigorous activity versus one's resting heart rate; see Kievit et al., 2013 for detailed discussions with more examples). The current study will examine ED and ER variability in both trait- and momentary-level, which allows us to better understand between-person differences and within-person variations respectively.

4. Hypotheses (required)

Overall, we hypothesize that higher ED predicts higher ER variability at both trait- and momentary-level in adolescents. Particularly, at momentary-level, we are interested in the time-lagged effect between ED and ER variability, in both temporal directions. Kashdan et al. (2015)'s theoretical framework suggests that ED precedes and facilitates ER (operationalized as ER variability in this study), but not the other way round.

The hypotheses are specified as follow (see section 22 – Main Analyses for details of the hypotheses):

H1. We hypothesize that ED positively associates with ER variability at the trait-level.

H2. We expect a positive association between ED at one moment and ER variability at the next moment.
H3. We expect no significant associations between ER variability at one moment and both the level of and changes in ED at the next moment.

Previous research on emotion differentiation calculated from negative affect was commonly found to have predictive validity on different physical and mental health related outcome measures. ED calculated from positive affect were less often reported, or, when reported, less likely to have significant associations with the concerned outcome measures (e.g. Oh & Tong, 2020; Pond et al., 2012; Willroth et al., 2020). Therefore, all trait- and momentary-level ED in our analyses refer to negative ED, i.e., ED calculated from negative affect. However, positive ED will also be calculated and analyzed exploratorily.

Design Plan

In this section, you will be asked to describe the overall design of your study. Remember that this research plan is designed to register a single study, so if you have multiple experimental designs, please complete a separate preregistration. For large ESM datasets that will be used for multiple analyses by different people (and possibly different labs), for clarity, we recommend one registration per study/paper.

5. Study type (required)

   Observational.

6. Blinding (required for intervention studies)

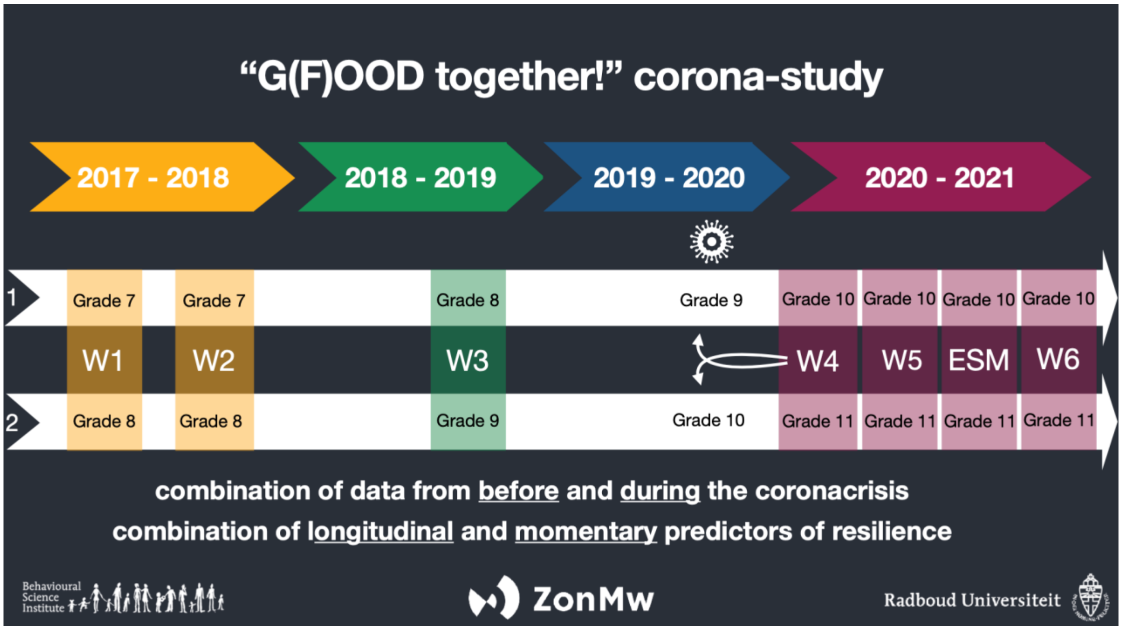
   No blinding is involved in this study.

7. Is there any additional blinding in this study?

   No.

8. Study design (required)

   Data from “G(F)ood Together!” (*G(V)oed voor elkaar* in Dutch, hereafter GVE) will be used for this study. Started in 2017, GVE is primarily a 6-wave longitudinal cohort study in Dutch adolescents and their parents to assess adolescents’ and parents’ health behaviour. GVE has an additional week-long experience sampling method (ESM) study (named *Flits Study* in the GVE project page) that collected data from a sub-sample of 89 parent-child dyads between wave 5 and 6. For the current preregistration, only the adolescent ESM data will be used. Please see the following flowchart for an overview of assessment and data collection procedures on the whole project.



GVE consists of adolescents of both genders aged 12-18 years old over the course of data collection, and their mothers at their mid-40s. Participant characteristics, recruitment procedures, ethics approval process and incentives for the whole project can be found in a published study at https://doi.org/10.3390/nu12030786.

Upon invitation to all W5 participants in the cohort study, 89 parent-child dyads volunteered to participate in the week-long ESM study. Details of the ESM study, including the additional incentives for ESM study participation, can be found at section 12 and the GVE project ESM study (*Flits Study*) page at https://osf.io/9axte/.

9. Randomization

Not applicable.

Sampling Plan

In this section we'll ask you to describe how you plan to collect samples, as well as the number of samples you plan to collect and your rationale for this decision. Please keep in mind that the data described in this section should be the actual data used for analysis, so if you are using a subset of a larger dataset, please describe the subset that will actually be used in your study.

10. Existing data (required)

Registration prior to access of the data.

11. Explanation of existing data (if applicable)

ESM data were collected in GVE as discussed in section 8. Data have not been accessed at time of preregistration.

The current study will use a subset of variables (see variables section for names) from the ESM study within GVE, using data from all participating adolescents (N=89). Data are not publicly available at this stage. We will start data checks and preparation after making this preregistration on the Open Science Framework (OSF),

The codebook for the GVE project ESM study is documented on OSF at https://osf.io/9axte/.

No other analyses have been done on the dataset from this ESM study at this moment.

12. ESM data collection procedure (required)

The ESM data collection took place between June and July 2021. There were 3 periods of data collection. Each participant only participated in one of them. All participants downloaded and installed the SEMA-app (Koval et al., 2019) on their mobile phones a few days before starting the study, which enabled us to deliver momentary assessment prompts. Participants completed a baseline questionnaire on a Sunday and a follow-up questionnaire on a Monday a week later. The ESM period was in between, always starting on a Monday and lasting for 7 consecutive days. Participants completed the ESM assessments 10 times a day. There was no extension of ESM period. Only the data collected in the 7-day ESM period are used in this study.

An overview of the ESM notification scheme and items in questionnaires can be found at https://osf.io/hbmwt/.

A semi-random sampling scheme was used. Participants received 10 notifications per day. The first 9 notifications were given in 9 blocks of 90 minutes between 07.30 a.m. and 09.00 p.m. Each notification was randomly scheduled within the first 30 minutes of the 90-minute block (i.e., notification 1 between 07.30 a.m. and 08.00 a.m.; notification 2 between 09.00 a.m. and 09.30 a.m., etc.). The 10^{th} notification was given randomly between 09.00 p.m. and 09.30 p.m.

Following the initial notification on the phone, the ESM questionnaire was available to the participant for 30 minutes. The end-of the-day questionnaire had a longer available time of 149 minutes. Participants who did not open the momentary questionnaires received 2 reminders from the app 15 minutes and 25 minutes after the initial notification (75 minutes and 145 minutes for the end-of-the-day questionnaire).

At each momentary assessment prompt, participants answered 9 questions regarding their affective state in that specific moment (4 positive affect and 5 negative affect items, see section 17). The items were randomized at each assessment prompt, alternating between positive and negative affect. At every even beep (i.e., the 2nd, 4th, 6th, 8th, and 10th beep) on a day (i.e., assessed 5 times per day), participants answered 1 additional question (after rating affect) on a slider scale regarding the intensity of the most unpleasant event since the last beep, ranging from (0) 'not at all unpleasant' to (10) 'very much unpleasant'. If they gave a score at or above 5 to this unpleasantness rating, they were asked to further rate the extent they used 5 different emotion regulation strategies for the event (see section 17 for details). At the last beep of the day, the ER strategies questions were always asked regardless of the intensity of event rated. Maximum possible number of measurements for affect is 70 and for ER strategy, 35 during the one-week ESM.

There were other measurements which were beyond the scope of this study. Please refer to the codebook documented on OSF at https://osf.io/9axte/ for more details. Overall, the ESM questionnaire consisted of a minimum of 17 (affect and relationships) and a maximum of 32 (unpleasant event, emotion regulation, parenting and miscellaneous) true/false or sliding scale items per assessment. Questionnaire length varied as a function of certain item responses (e.g., ER strategies only asked if they indicated high intensity from an unpleasant event). Depending on the length of questionnaire, the average completion time ranged from one to two minutes.

Prior to enrolment, participants were informed about the ESM incentive scheme. They could receive €5 to €25, depending on the compliance of both mother and adolescent in the same dyad. Additionally, participants entered into a raffle of two €250 rewards (i.e., a voucher for a weekend holiday break) upon participation. Participants had access to their compliance information during the ESM study period.

GVE project members called participants with low compliance to help resolve any technical difficulties arose and encourage their continued participation.

13. Sample size (number of participants) (required)

89 adolescents participated. However, 2 of them dropped out after 1-2 days of participation.

14. Rationale for sample size: Temporal design and number of participants (if applicable)
A simulation study conducted by Schultzberg and Muthén (2018) showed that 100 participants with 25 assessments were warranted to estimate models with a weak to moderate effect with a power of 80%. Taken an average compliance rate of 78% of ESM assessment into account (Rintala et al., 2019) and potential attrition into account, GVE targeted to recruit 100 participants for the ESM. Invitations to join the ESM study were sent to families who were still active in the GVE cohort study at wave 5 (*n* = 257). 89 adolescent participants were successfully recruited.

15. Stopping rule (if applicable)

New participant recruitment was not possible after exhausting the pool of participants who were still active in wave 5 of the GVE cohort study.

Variables

In this section, you can describe all variables (both ESM and non-ESM variables and manipulated and measured variables) that will later be used in your confirmatory analysis plan. In your analysis plan, you will have the opportunity to describe how each variable will be used. If you have variables that you are measuring for exploratory analyses, you are not required to list them, though you are permitted to do so.

16. Measured non-ESM/time invariant variables (if applicable)

Not applicable.

17. Measured ESM/time-variant variables (required)

Positive and negative affect, measured by 9 items: “Right now I feel content/relaxed/joyful/energetic/irritated/worried/depressed/insecure/lonely.” (used in Barrantes-Vidal et al., 2013)

Emotion regulation strategies, measured by 5 items, following a question that prompted to rate the most unpleasant event participants experienced since the last assessment (used in Kirtley et al., 2020; Nittel et al., 2018; Brans et al., 2013):
- (acceptance) “I have accepted my feelings about it”
- (reappraisal) “To feel better, I have changed the way I think about it”
- (expression suppression) “I have avoided expressing my feelings about it”
- (rumination) “I couldn’t stop thinking my feelings about it”
- (sharing) “I talked about it to someone”

Answers to all these affect and ER items were given on an 11-point slider scale ranging from 0 to 10 (“not at all” to “a lot” for affect items; “not applicable at all” to “very applicable” for ER items).

18. Open-ended questions (if applicable)

Not applicable.

19. Indices (if applicable)

We will calculate the following indices:

Momentary negative affect will be calculated by the mean of 5 negative affect items (unless indicated to be excluded – see section 22) at each moment for each person (used in Kalokerinos et al., 2019).
Momentary ER strategy use will be calculated by the mean of 5 ER strategies at each moment for each person (used in Blanke et al., 2020).
Momentary negative ED will be calculated using the momentary ED index (Erbas et al., 2021) from negative affect reported momentarily.
Momentary ER variability will be calculated by considering the within-person variation of use of ER strategies across occasions (Aldao et al., 2015). The exact way to calculate momentary ER variability is left open at the time of this preregistration: We are working on another paper to determine the optimal index for momentary ER variability.
Trait mean negative affect will be calculated by the mean of momentary negative affect (the 5 items) of each person across all observations.
Trait mean ER strategy use will be calculated by the mean of momentary ER strategy use of each person across all observations.
Trait negative ED will be calculated by intraclass correlation coefficient (ICC) for each person on repeatedly sampled ratings of 5 negative affect items (unless indicated to be excluded – see section 22; used in Lennarz et al., 2018).
Trait ER variability will be calculated by the mean of momentary ER variability of each person across all observations.
Lagged momentary negative ED and lagged negative affect will be calculated as well. See section 24 for the details of transformation.

20. Manipulated variables (if applicable)

Not applicable.

Prior knowledge of data (if applicable)

21. List the publications, conference presentations (papers, posters), and working papers (in prep, unpublished, preprints) you have worked on that are based on the data set. Describe which variables you have previously analyzed and which information you used

in these analyses. Limit yourself to variables that are relevant to the current study. If the dataset is longitudinal, include information about which wave(s) of data were previously analyzed. Also, include any knowledge regarding missingness within the dataset or compliance, including at what level e.g., overall compliance, compliance for different types of reports, the mean level of compliance, range of compliance across participants.

There are no previous papers that we (or others) have worked on that were based on this ESM dataset. None of the authors in the current (or the other) registration accessed the data prior to this pre-registration.

The only prior knowledge of data is the estimated percentage of ESM assessment collected out of all possible observations. This was known from the monitoring of participants' compliance during the data collection process (given by SEMA-app) and calculation of reimbursement for each participant.

No summary statistics, statistical distribution of variables, nor associations between variables were computed prior to this pre-registration.

Analysis Plan

You may describe one or more confirmatory analyses in this pre-registration. Please remember that all analyses specified below must be reported in the final article, and any additional analyses must be noted as exploratory or hypothesis-generating.

A confirmatory analysis plan must state upfront which variables are predictors (independent) and which are the outcomes (dependent), otherwise, it is an exploratory analysis. You may describe any exploratory work here, but a clear confirmatory analysis is required.

22. Statistical models (required)

Assumptions check and data preparation

Intraclass correlation coefficients (ICCs) will be calculated on the "person × observation" matrix of each of the 9 affect items and 5 ER strategy items. ICCs, calculated by having between-person variance as the nominator and overall variance as denominator, indicate how much variance is due to between-person differences and how much is due to within-person momentary fluctuations.

Multi-level confirmatory factor analysis (MCFA) will be conducted to check if affect items load on expected factor structures (5 negative affect items as one factor and 4 positive affect items as another) within and between individuals (Eisele et al., 2021). Based on the results of MCFA, items that load insufficiently on overall latent factors will be omitted from the calculation of indices.

Prior to analyzing H1, assumptions of homoscedasticity and linearity will be checked through inspection of standardized residual plots.

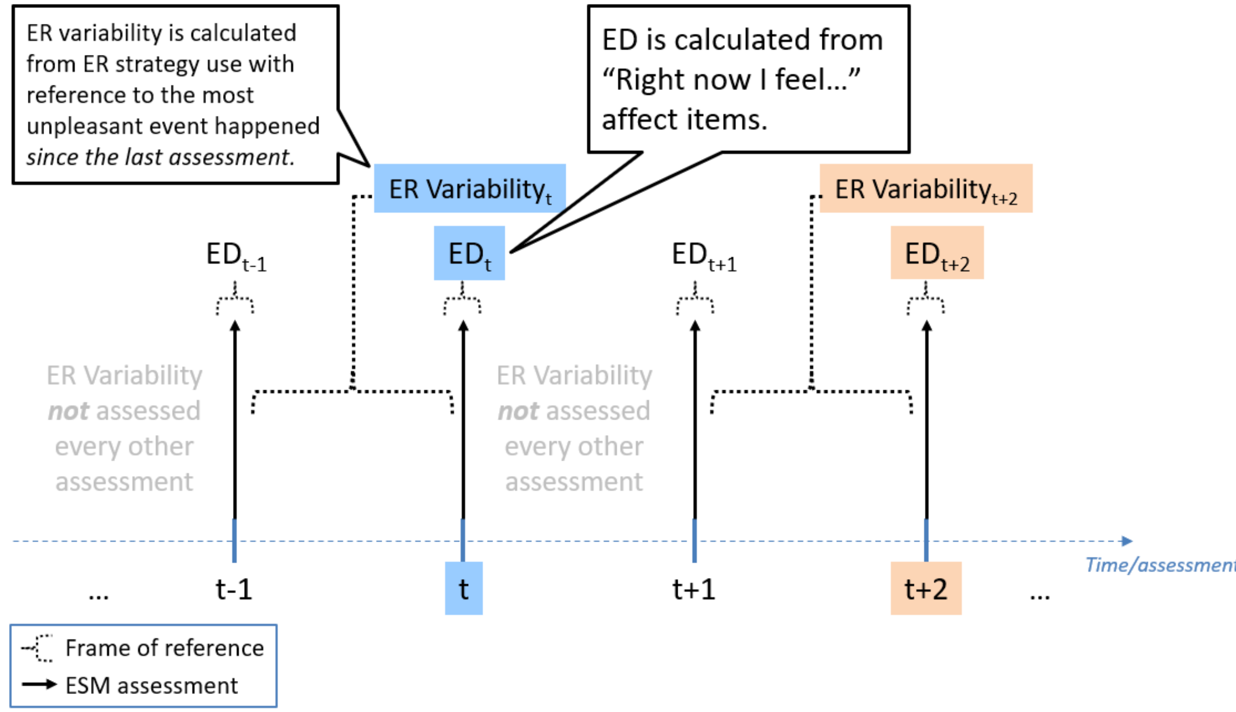
During the analyses of H2 and H3, histograms of random intercepts and slopes will be plotted to check the assumption of normally distributed intercepts and slopes.

Main analyses

For H1, hierarchical regressions will be used to examine associations of trait ED and trait ER variability. The main regression equation is ERvariability ~ ED.

For H2 and H3, multilevel modelling with measurement occasions nested within persons (i.e., momentary ED and momentary ER variability) will be conducted to test the associations between momentary ED and momentary ER variability. Both momentary ED and ER variability are continuous. Please note that in H2, momentary ED is the predictor and momentary ER variability is the outcome variable. In H3, their roles are reversed, that momentary ER variability is the predictor and momentary ED is the outcome variable.

The frame of reference of ER strategies rating (with reference to the most unpleasant event since the last assessment) lied between the last and current assessment. In contrast, affect items assessed “right now” of a certain assessment. Therefore, at the same momentary assessment, momentary ED referred to the moment of assessment, where momentary ER variability referred to the period between the last and current assessment, as illustrated below:



As noted in McNeish and Hamaker (2020), whether or not to use lagged (t – 1) variables in the examination temporal relationships depends on the frame of reference in measurement of the variables concerned. Therefore, the main regression equations for H2 and H3 are as below:

[H2] ER Variability$_t$ ~ ED$_{it-1}$ + (1 + ED$_{it-1}$ | i)

[H3] ED$_t$ ~ ER Variability$_{it}$ + (1 + ER Variability$_{it}$ | i)

For each person *i*, the momentary predictor variable at time point *t* (i.e., ED$_{it}$ and ER Variability$_{it}$) is broken down into (A) a person-mean-centered component (*within*) that varies at each time point (*t*), and (B) a component (*between*) that shows person i's time-invariant between-person difference with the rest of the population. The relationship between the variable and components are represented in the equations below: (How these components are calculated is discussed at section 24)

ED$_{it-1}$ = ED$_{i(within)t-1}$ + ED$_{i(between)}$

ER Variability$_{it}$ = ER Variability$_{i(within)t}$ + ER Variability$_{i(between)}$

Combining the main regression equations and the above equations on variable components we obtain the following equations for our analyses:

[H2] ER Variability$_t$ ~ ED$_{i(within)t-1}$ + ED$_{i(between)}$ + (1 + ED$_{i(within)t-1}$ | i)

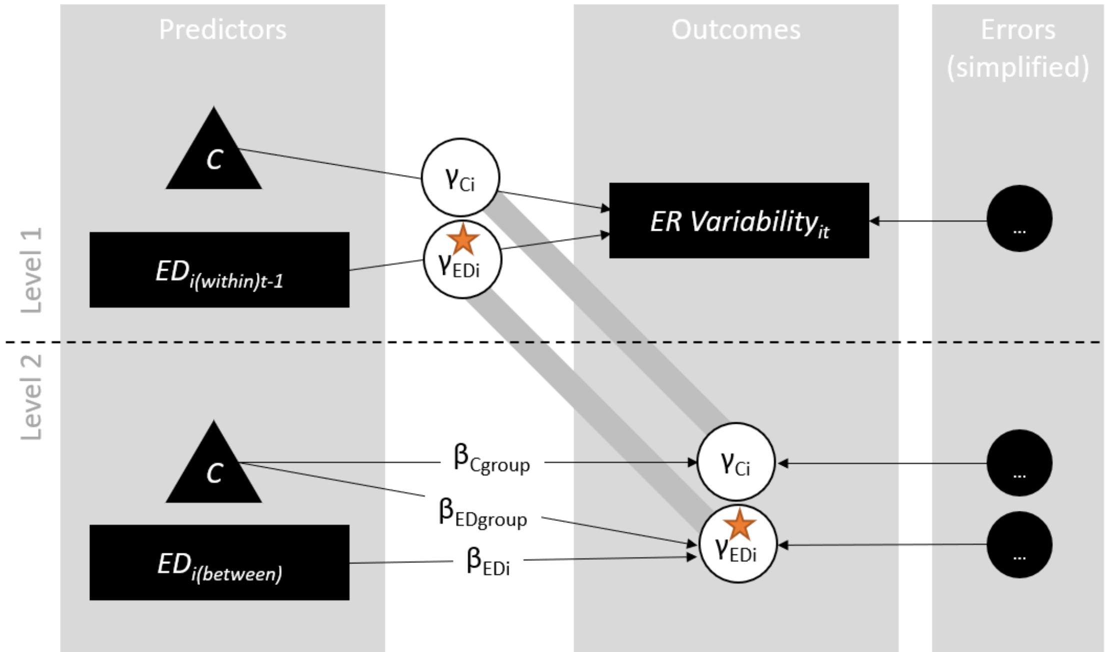
[H3] ED$_t$ ~ ER Variability$_{i(within)t}$ + ER Variability$_{i(between)}$ + (1 + ER Variability$_{i(within)t}$ | i)

The coefficient of the within- component (i.e., $ED_{i(within)t-1}$ and ER $Variability_{i(within)t}$) is the key coefficient of interest in the multilevel modeling. The regression equations above test, with reference to a person's typical level, if there are associations between the predictor at one moment and the outcome variable at the next moment.

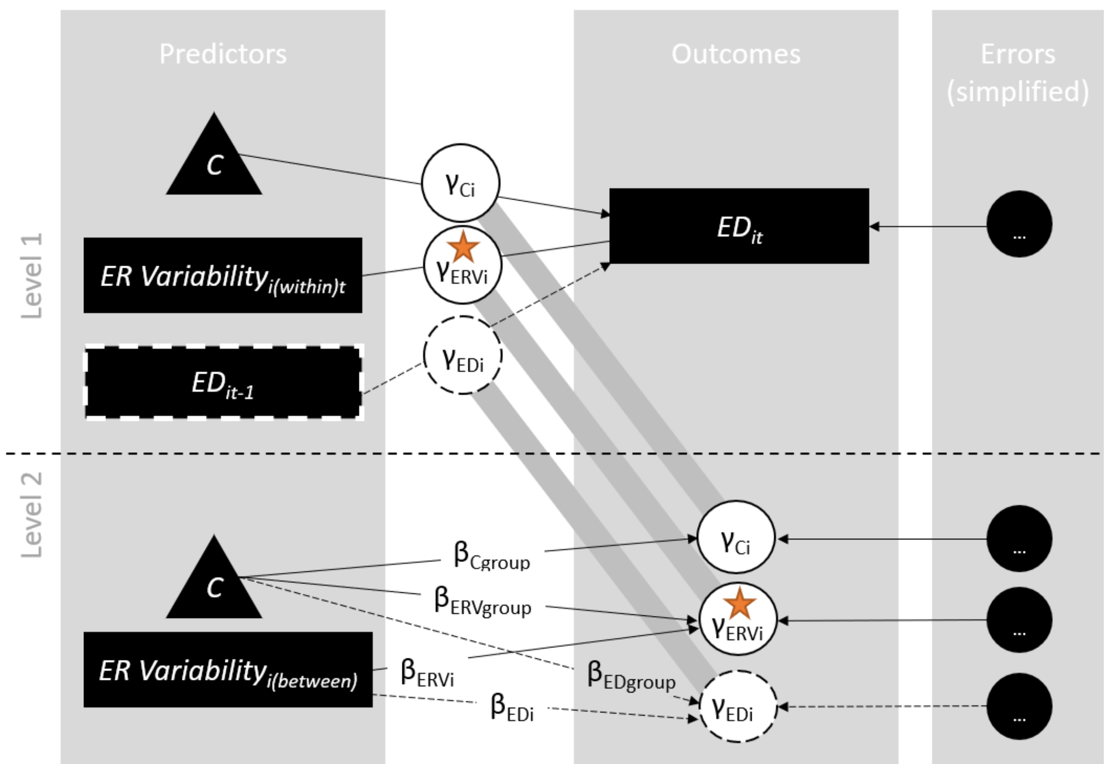
In ESM studies that examined temporal relationships between variables, adding the outcome variable at t – 1 as one of the control variables allows models to estimate the effect of the predictor on *change* in outcome variable at the subsequent moment (e.g. Chan et al., 2019). This additional step of analysis will be done for H3. However, no such additional step will be done for H2. ER variability was only measured every 2 assessments (see section 12 for details) and was dependent on a branching condition, so ER $variability_{t-1}$ is not available for the additional step. Note that in the current study design, participants rated their ER strategies use (thus the calculated ER variability) with reference to an unpleasant event that happened since the last assessment. The event could be the same or different from the previous event they reported. As a result, the target of ER strategies might be different. This makes the inquiry of the change of ER strategy use (thus ER variability) less theoretically meaningful. Instead, we include lagged momentary negative affect as a control variable, because higher negative affect associates with increased ER (Kalokerinos et al., 2017). This was how momentary ER strategies as outcome variables was analyzed in Kalokerinos et al. (2019)'s study.

Below are the graphical illustrations (error terms and control variables not detailed for simplicity) of the multilevel models for H2 and H3, where C = constant, i = individual, group = effect commonly shared by all individuals, t = time point (of concurrent ESM assessment), t – 1 = previous time point, β = between-person coefficient, and γ = within-person coefficient. The illustrations were drawn with reference to Model 7 and Model 9 in Lafit et al., (2021)'s paper, with the between-person difference component placed in level 2. In H3, the additional step of analysis to control for ED_{t-1} is represented in the dashed box. The coefficient of interest is marked by a star.

H2:



H3:



The errors of momentary ED and momentary ER variability are assumed to be Gaussian distributed and serially correlated. The serial correlation will be modeled using an AR(1) process. The model will include a random intercept and random slopes for the predictor variables and control variables where convergence is possible. The model is assumed to be bivariate Gaussian distributed with an uncorrelated correlation structure.

Control variables

Control variables in H1 are gender (0 for male and 1 for female), trait mean negative affect and trait mean ER strategy use.
Control variables in H2 and H3 are gender, time, momentary negative affect, and momentary ER strategy use. For H3, lagged momentary ED will also serve as a control variable.

23. What will you do should your data violate assumptions, your model not converge or some other analytic problem arises?

When analyses in frequentist approaches could not converge, more simple models will be estimated (e.g., with fixed slopes and/or fixed intercepts). Also, Bayesian approaches will be employed to estimate the models that cannot converge in frequentist approaches.

24. Transformations (if applicable)

In the calculation of lagged variables, the first beep of the day will be set as a missing observation.

Prior to multilevel modeling in H2 and H3, momentary ED (as a predictor), momentary ER variability (as a predictor), and other momentary control variables (momentary negative affect and momentary ER strategy) will be separated into within-person and between-person components. The within-person component is obtained by centering on within-person mean. The between-person component is obtained by subtracting the momentary measure with grand mean (mean of all assessments across all participants) and the within-person component (Bolger & Laurenceau, 2013). Time will also be centered with the 35.5th observation as zero.

25. Inference criteria (if applicable)

For H1, the p-value, confidence interval and effect sizes of the regression coefficient will be reported. Effect sizes $f^2 >= 0.02$, $f^2 >= 0.15$, and $f^2 >= 0.35$ are seen small, medium and large respectively (Cohen, 1988).

For H2 and H3, the fixed effect slopes of the predictor variable are of interest. Their estimates, errors, and confidence intervals will be reported. Where possible (i.e., when frequentist modeling converges), p-values will be reported where $p < .05$ is considered significant by convention.

26. Missing data (if applicable)

For H1, the trait-level variables are indices calculated from individual affect and ER strategy items over all time points. Missing data at certain time points will not result in missing data on these trait-level variables.

For H2 and H3, multilevel model estimates random intercepts and slopes according to available data from each person. With the likelihood-based methods missing data are handled well even if data available per person are different (Fitzmaurice et al., 2008; Snijders & Bosker, 2011).

27. Data exclusion (if applicable)

The principle is to include all observations (Jacobson, 2020) unless the data quality is very questionable, for example:
- zero variance across all observations
- items with response time under 500ms which indicated potentially careless responding (McCabe et al., 2012)

28. Exploratory analysis (if applicable)

The current study holds no specific hypotheses regarding that involved positive ED at state or trait level. There are fewer reported findings in positive ED among literature of ED. Therefore, exploratory analyses H1, H2 and H3 will be re-run with trait- and momentary-level positive ED substituting those of negative ED. Trait mean positive affect instead of trait mean negative affect will also be used as a control variable in these exploratory analyses.

These indices will be calculated for exploratory analyses for positive emotions:
Momentary positive affect will be calculated by the mean of 4 positive affect items (unless indicated to be excluded – see section 22) at each moment for each person.
Momentary positive ED will be calculated using the momentary ED index (Erbas et al., 2021) from positive affect reported momentarily.
Trait mean positive affect will be calculated by the mean of momentary positive affect (the 4 items) of each person across all observations.
Trait positive ED will be calculated by intraclass correlation coefficient (ICC) for each person on repeatedly sampled ratings of 5 positive affect items (unless indicated to be excluded – see section 22).
Lagged momentary positive ED will be calculated as well. See section 24 for the details of transformation.

Other

29. Other (if applicable)
R will be used for data preparation and analyses. Specifically, the *lm()* function will be used for H1. The *nlme* and *brms* packages will be used for H2 and H3.

30. Reference

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion Regulation Flexibility. *Cognitive Therapy and Research*, *39*(3), 263–278. https://doi.org/10.1007/s10608-014-9662-4

Barrantes-Vidal, N., Chun, C., Myin-Germeys, I., & Kwapil, T. (2013). Psychometric Schizotypy Predicts Psychotic-Like, Paranoid, and Negative Symptoms in Daily Life. *Journal of Abnormal Psychology*, *122*, 1077–1087. https://doi.org/10.1037/a0034793

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

Blanke, E. S., Brose, A., Kalokerinos, E. K., Erbas, Y., Riediger, M., & Kuppens, P. (2020). Mix it to fix it: Emotion regulation variability in daily life. *Emotion*, *20*(3), 473–485. https://doi.org/10.1037/emo0000566

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. Guilford press.

Brans, K., Koval, P., Verduyn, P., Lim, Y. L., & Kuppens, P. (2013). *The Regulation of Negative and Positive Affect in Daily Life*. 14.

Chan, Y., So, S. H., Mak, A. D. P., Siah, K. T. H., Chan, W., & Wu, J. C. Y. (2019). The temporal relationship of daily life stress, emotions, and bowel symptoms in irritable bowel syndrome—Diarrhea subtype: A smartphone-based experience sampling study. *Neurogastroenterology & Motility*, *31*(3), e13514. https://doi.org/10.1111/nmo.13514

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*.

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

Erbas, Y., Kalokerinos, E. K., Kuppens, P., van Halem, S., & Ceulemans, E. (2021). Momentary Emotion Differentiation: The Derivation and Validation of an index to Study Within-Person Fluctuations in Emotion Differentiation. *Assessment*, 107319112199008. https://doi.org/10.1177/1073191121990089

Fitzmaurice, G., Davidian, M., Verbeke, G., & Molenberghs, G. (2008). *Longitudinal data analysis*. CRC press.

Gross, J. J. (1998). The Emerging Field of Emotion Regulation: An Integrative Review. *EMOTION REGULATION*, 29.

Jacobson, N. (2020). Compliance thresholds in intensive longitudinal data: Worse than listwise deletion: Call for action [Symposium]. *Society for Ambulatory Assessment Conference, Melbourne, Australia*.

Kalokerinos, E. K., Erbas, Y., Ceulemans, E., & Kuppens, P. (2019). Differentiate to Regulate: Low Negative Emotion Differentiation Is Associated With Ineffective Use but Not Selection of Emotion-Regulation Strategies. *Psychological Science*, *30*(6), 863–879. https://doi.org/10.1177/0956797619838763

Kalokerinos, E. K., Résibois, M., Verduyn, P., & Kuppens, P. (2017). The temporal deployment of emotion regulation strategies during negative emotional episodes. *Emotion*, *17*(3), 450–458. https://doi.org/10.1037/emo0000248

Kashdan, T. B., Barrett, L. F., & McKnight, P. E. (2015). *Unpacking Emotion Differentiation: Transforming Unpleasant Experience by Perceiving Distinctions in Negativity*. 7.

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00513

Kirtley, O., Hiekkaranta, A., Kunkels, Y., Eisele, G., Verhoeven, D., Van Nierop, M., & Myin-Germeys, I. (2020). The experience sampling method (ESM) item repository. *Retrieved from: Osf. Io/Kg376*.

Koval, P., Hinton, J., Dozo, N., Gleeson, J., Alvarez, M., Harrison, A., & Sinnott, R. (2019). SEMA3: Smartphone Ecological Momentary Assessment, Version 3. *Computer Software]. Retrieved from Http://Www. Sema3. Com*.

Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the Number of Participants in Intensive Longitudinal Studies: A User-Friendly Shiny App and Tutorial for Performing Power Analysis in Multilevel Regression Models That Account for Temporal Dependencies. *Advances in Methods and Practices in Psychological Science*, *4*(1), 251524592097873. https://doi.org/10.1177/2515245920978738

Lennarz, H. K., Lichtwarck-Aschoff, A., Timmerman, M. E., & Granic, I. (2018). Emotion differentiation and its relation with emotional well-being in adolescents. *Cognition and Emotion*, *32*(3), 651–657. https://doi.org/10.1080/02699931.2017.1338177

McCabe, K. O., Mack, L., & Fleeson, W. (2012). A guide for data cleaning in experience sampling studies. In *Handbook of research methods for studying daily life* (pp. 321–338). The Guilford Press.

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, *25*(5), 610–635. https://doi.org/10.1037/met0000250

Molenaar, P. C. M., & Campbell, C. G. (2009). The New Person-Specific Paradigm in Psychology. *Current Directions in Psychological Science*, *18*(2), 112–117. https://doi.org/10.1111/j.1467-8721.2009.01619.x

Nittel, C. M., Lincoln, T. M., & Lamster, F. (2018). *Expressive suppression is associated with state paranoia in psychosis: An experience sampling study on the association between adaptive and maladaptive emotion regulation strategies and paranoia—Nittel—2018—British Journal of Clinical Psychology—Wiley Online Library*. https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/bjc.12174

Oh, V. Y. S., & Tong, E. M. W. (2020). Negative emotion differentiation and long-term physical health—The moderating role of neuroticism. *Health Psychology*, *39*(2), 127–136. https://doi.org/10.1037/hea0000809

Pond, R. S., Kashdan, T. B., DeWall, C. N., Savostyanova, A., Lambert, N. M., & Fincham, F. D. (2012). Emotion differentiation moderates aggressive tendencies in angry people: A daily diary analysis. *Emotion*, *12*(2), 326–337. https://doi.org/10.1037/a0025762

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2019). Response compliance and predictors thereof in studies using the experience sampling method. *Psychological Assessment*, *31*(2), 226–235. https://doi.org/10.1037/pas0000662

Schultzberg, M., & Muthén, B. (2018). Number of Subjects and Time Points Needed for Multilevel Time-Series Analysis: A Simulation Study of Dynamic Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 495–515. https://doi.org/10.1080/10705511.2017.1392862

Snijders, T. A., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. sage.

Tamir, M. (2009). What Do People Want to Feel and Why?: Pleasure and Utility in Emotion Regulation. *Current Directions in Psychological Science*, *18*(2), 101–105. https://doi.org/10.1111/j.1467-8721.2009.01617.x

Willroth, E. C., Flett, J. A. M., & Mauss, I. B. (2020). Depressive symptoms and deficits in stress-reactive negative, positive, and within-emotion-category differentiation: A daily diary study. *Journal of Personality*, *88*(2), 174–184. https://doi.org/10.1111/jopy.12475