Peer Review History and Correspondence

# Ms Title: Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect

**Author Names:** Daan Alexander Ornée, Albertine J. Oldehinkel, Jojanneke A. Bastiaansen

Submitted: May 4, 2020

**Editor First decision**

July 24, 2020

Dear Daan Alexander Ornée,

Thank you for your extreme patience with us in the review process for this manuscript. I have now received all reviews of your manuscript, “Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect” from qualified researchers. I also independently read the manuscript before consulting these reviews. I agree that your manuscript has important strengths and also that there are some issues that need to be addressed. I therefore encourage you to submit a revised version for further consideration at Collabra: Psychology.

Because both I and the reviewers think your manuscript is already strong, I will review your revision myself and render a decision, in hopes of expiditing the process.

The reviewers did an outstanding job in their reviews. I will highlight issues I think are particularly salient here. In your resubmission, please include a document with a point-by-point response to both the points I list here and the reviewers’ comments, outlining each change made in your manuscript or providing a suitable rebuttal.

First, I would say that I did appreciate your summary of the prior conflicting findings in EMI studies, and I thought your point about the importance of controlling for methods factors in understanding the effects of EMI content on outcomes was well made.

I also really appreciated the carefulness and thoroughness of how you analyzed your data and reported your results. It very much boosts confidence in your findings. I had a few comments and questions about the analyses that I think could help readers understand your findings better. One, I was curious why you didn’t include time of day as a predictor, as well as day of study? Wouldn’t that help the model account for within-day time trends in NA and PA? I also thought you could do a little better in describing the random effects of the time/day trend in the text; you seem to only describe the variance of one of the time slopes, but it’s not clear whether it’s the linear or quadratic slope, and it’s not clear whether both of those time/day trends were estimated with random effects or not. I also wondered if you could go a little further in speculating why PA and NA might be confounded with measurement frequency. What is the psychological process that might account for that?

In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision.

Please ensure that your revised files adhere to our author guidelines, and that the files are fully copyedited/proofed prior to upload. Please also ensure that all copyright permissions have been obtained. This is the last opportunity for major editing, therefore please fully check your file prior to re-submission.

If you have any questions or difficulties during this process, please contact the editorial office at editorialoffice@collabra.org.

We hope you can submit your revision within the next six weeks. If you cannot make this deadline, please let us know as early as possible.

Sincerely,

Kevin King

# Reviewer 1

##### Open response questions

### **Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.**

Dear Editor Thank you for sending me this fascinating paper to review. It is extremely well written, and very clear. While I do have a few suggestions for the authors which I have listed below in roughly the order they occur in the manuscript, these are limited. I hope the authors will accept these in the collegial spirit in which they are intended.

Abstract

1. Note missing “an” in “promise as AN intervention strategy”.
2. Is it possible to spell out ZELF-i in the abstract, or else readers will be a bit confused by the acronym
3. Similarly, spell out EMA.
4. To my mind, one of the main limitations worthy of mention is that there is no neutral control group, which might involve general self-monitoring.
5. The conclusion of the abstract seems out of line with the aims. The aims are to compare the content of Do vs Think modes. The conclusion expands this to claim that Think mode is no different to general self-monitoring. This has not actually been assessed.

Introduction

1. I really enjoyed reading this. It is clearly written and gives a good sense of the previous literature.
2. Having read it through, I still don’t fully understand why a control condition was not possible, lacking an activity component and simply asking for affect ratings. This will need some mention in the discussion section I think.

Methods

1. The pre-registration of the trial is a positive feature and I would praise the team for their honest and straightforward description of why blinding was not possible.
2. My stats knowledge is not sufficient to offer a review on that component of the manuscript.

Discussion

1. I thought the discussion was pithy and to the point. I have no suggestions to make here.

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  |  | ✔ |

# Reviewer 2

##### Open response questions

### **Please write your review here. The author(s) will see this review. Your identity will not be revealed to the authors unless you also include your name (i.e., sign your review) in this box. It is up to you whether to reveal your identity or not, either is fine.**

I write regarding the manuscript “Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect”. I enjoyed reading the manuscript and reviewing the supplementary information submitted by the authors, including the preregistration uploaded to the Open Science Framework. The manuscript focuses on a study of Ecological Momentary Interventions (EMI), which are repeated daily self-assessments with an aim of providing self-correction. The current manuscript asks whether two similar such EMIs differentially change positive or negative affect, and finds that they do not do so in statistically-significant ways.

I commend the authors on an interesting secondary research question that has clear implications both clinically and from a cognitive theory approach. The manuscript is clearly written. My major concern is that this manuscript focuses solely on the overlapping ecological momentary assessment mirrored in both “Do” and “Think” modules (measures of PA and NA—although “physical state” was also measured in both, as mentioned in Figure 1), while ignoring the major connections to the RCT from which the data comes. However, I found the manuscript interesting to read, and believe that it contributes to the literature.

Below, I note some points by which the authors might improve their manuscript.

Introduction

1. The authors devote a significant portion of the introduction discussing two past EMI studies (Kramer et al., 2014; and van Roekel et al., 2017). I am not sure why they focus so deeply on potential explanations for the mixed effects (p. 4. beginning on line 76), as this is not the focus of the current manuscript. I believe this section could be significantly shortened.
2. I appreciate the focus on connecting the “Do” module to Behavioral Activation (p. 5), and on connecting the “Think” module to what I would call cognitive restructuring. (p. 6). However, I believe the manuscript could be bolstered by discussing research on repeated and tailored interventions for depressed mood either in the cognitive bias modification (CBM) literature or positive mental imagery literature. Some such studies have a similar desired outcome to those described by the authors, and similarly take active ingredients from BA or CR. Some examples follow. (I acknowledge that the final citation is to my own work, and leave it to the authors whether they believe it is appropriate to include.)
* <https://doi.org/10.1037/a0024355>
* <https://doi.org/10.1016/j.jbtep.2016.03.012>
* <https://doi.org/10.1016/j.cpr.2010.01.001>
* <https://doi.org/10.1016/j.brat.2018.09.010>
1. I would recommend rewording the phrase “We will therefore test two-sidedly” – perhaps “We will therefore use two-sided tests”.

Methods

1. Were smartphones provided to participants or were they completing the EMI on their own phones? How were feedback reports provided?
2. The preregistration makes clear that data collection was completed at the time of analysis development. This should be explicitly stated in the Analysis section; it seems to be assumed that it’s obvious but it is never explicitly stated. The authors develop an effective solution for analyses (i.e., they asked a condition-blind researcher to re-code group membership before analysis – is that one of the authors?). The preregistration describes a specific analysis and two direction-less hypotheses (“Based on the current literature we cannot make a clear prediction which of the two interventions will have the largest impact on affect compared to the other”), which are mirrored in the introduction.
3. The researchers share their code on the OSF, which is appreciated. Given that data is not available, I wonder whether it might be possible to share the output of an .Rmd file as well (as PDF or HTML), thus allowing one to follow along with the results of the analyses. Certainly, I appreciated being able to review the lme4 models.
4. I recognize that the authors preregistered the comparison of linear or quadratic time effects, but I think the manuscript may benefit from a theoretical background on this decision and what modeling time as quadratic might mean (especially given Figure 2). At the moment, it is only mentioned on p. 10.

Results

1. Table 2 and 3 both have notes that report that “Time was set in days, with five measurements each day”. I believe the wording on Figure 3 is slightly better (“Five measurements represent 1 day”); the first in particular makes it sound like Time in the analyses is Day (rather than measurement).
2. Figure 2 is helpful; I wonder whether adding a web-friendly color might be helpful for online publication. The same goes for Fig 3.
3. I really appreciate the individual predictive lines reported in Figure 3. (The quality in my copy is, unfortunately, somewhat low.) Are the lines predicted based on the models? This is implied but not stated.
4. On p. 15, the authors report that “there were no substantive subgroups of participants that did not read the feedback reports” (297–298). This was not previously mentioned as a possible analysis, although I note that it is described in the supplement. The supplement says that whether participants read the reports is measured by asking them after they are received—is that right? I think that either this should be removed from the manuscript or somewhat elaborated upon.

Discussion

1. The discussion returns to using the terms EMA, EMI, and “EMA intervention”; I think clarity in terms might help! (p. 16) e.g., the second paragraph begins by talking about EMA and ends with EMI. Is this distinction correct?
2. The authors do not report the results of the primary RCT from which this analysis stems. While I understand that such an analysis is likely under review or has been published elsewhere (and was not a preregistered analysis here), I wonder why the authors don’t take this into account – might it be that changes in NA and PA only occurred for participants with improvements in depressive symptoms? At a minimum, I believe that this should be addressed as a discussion point.
3. The authors note that “Because all participants started treatment during or shortly after the EMI…” (p. 17, line 331): This probably should have been mentioned earlier! Does treatment here mean a TAU such as BA/CBT, or does it mean something different? Is the EMI not itself considered a treatment as the authors argued in the introduction?
4. At several points, the authors appear to conclude that their results “could mean that EMA content does not matter” (quote from p. 16). I enjoyed the speculation on what this would mean, but think that the more measured conclusion reached in the abstract is more appropriate. All that we can conclude here is that asking about PA and NA, followed by asking about pleasure or worry, seems not to show differences between conditions.

I appreciated the opportunity to review this manuscript.

Justin Dainer-Best Assistant Professor of Psychology Bard College, USA

##### Rating scale questions

|  | Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree |
| --- | --- | --- | --- | --- | --- |
| The study/studies in this manuscript have strong construct validity (good measures and/or manipulations of the constructs the authors wish to study). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong statistical validity (appropriate statistical tests, assumptions are clear and reasonable, no statistical errors, appropriate statistical inferences, etc.). (Choose “Neutral” if this is not an empirical manuscript) |  |  |  |  | ✔ |
| The study/studies in this manuscript have strong internal validity (any causal claims or implications are well-justified, alternative explanations are thoroughly considered, etc.). (Choose “Neutral” if this is not an empirical manuscript, or no causal claims are made or even vaguely implied.) |  |  |  | ✔ |  |
| The study/studies in this manuscript have strong external validity (authors appropriately constrain their conclusions based on the limits of the generalizability of their findings to other contexts (including from lab to real world), other populations, other stimuli or measures, etc.) |  |  |  |  | ✔ |

**Author response**

Dec 22, 2020

Dear Kevin King,

Thank you for inviting us to revise our manuscript “Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect”. We highly value the thorough and constructive reviews, and are happy that implementing open practices was appreciated, and reciprocated by reviewer 2’s signed report. We addressed the comments provided by you and the reviewers to the best of our abilities, and hope the revised manuscript will be acceptable for publication in your journal. Below you will find a point-by-point response, with changes or additions to the manuscript indicated by blue text, deleted text in red and the page number indicated between brackets (p.x).

Yours sincerely,
Daan Ornée, on behalf of all co-authors.

**Comments from the editor and reviewers:**

**Reviewer 1**

**Dear Editor**

**Thank you for sending me this fascinating paper to review. It is extremely well written, and very clear. While I do have a few suggestions for the authors which I have listed below in roughly the order they occur in the manuscript, these are limited. I hope the authors will accept these in the collegial spirit in which they are intended.**

**Abstract**

1. **Note missing “an” in “promise as AN intervention strategy”.**

We added the word ‘an’, making the (partial) sentence (on p. 2) as follows:

 *…have shown promise as an intervention strategy for depression.*

1. **Is it possible to spell out ZELF-i in the abstract, or else readers will be a bit confused by the acronym**

ZELF-i is the actual name of the study, and not an acronym. ‘Zelf’ is Dutch for ‘self’, hence the name. We realize that the name ‘ZELF-i’ with capital letters suggests an acronym, and can create confusion, therefore we removed or substituted ‘ZELF-i’ in the abstract, so that the first mention will be in the introduction. Hopefully this will create more context and prevent confusion. The relevant text in the abstract (p.2) now reads:

 *The current study consists of two EMA intervention (EMI) modules, enabling us to compare the impact of EMI content on the course of momentary affect during the intervention.* ***Methods****: The intervention, implemented as add-on to regular depression treatment, consists of intensive self-monitoring (5x/day, 28 days) and weekly personalized feedback.*

Furthermore, at the first mention of ZELF-i in the introduction we added the following footnote (p.5):

 *‘Self’ means ‘Zelf’ in Dutch, i stands for intervention*

1. **Similarly, spell out EMA.**

Together with other changes to the use and location of EMI and EMA (see reviewer 2 comment 12), we ensured that the first instance of both the acronyms are spelled out (in the abstract on page 2, in the introduction on page 3).

1. **To my mind, one of the main limitations worthy of mention is that there is no neutral control group, which might involve general self-monitoring.**

The reviewer addresses a complex issue, as there is conceptual debate whether self-monitoring can ever be neutral, regardless of the questions. As the reviewer returns to this point at remark #7, we further address this issue there.

1. **The conclusion of the abstract seems out of line with the aims. The aims are to compare the content of Do vs Think modes. The conclusion expands this to claim that Think mode is no different to general self-monitoring. This has not actually been assessed.**

The reviewer addressed a fair point that the conclusion noted in the abstract drew disproportionate attention towards a specific part of the discussion. We re-wrote both the relevant part in the discussion of the manuscript (p.16) and the conclusion in the abstract (p.2) to match our main results more closely:

*In our sample, the focus of the EMI did not lead to differential effects on momentary affect. This implies that a focus on thoughts and negative affect compared to positive affect and activities may not lead to added adverse effects on mood, which is an often-voiced concern when using EMA in both research and clinical practice.*

**Introduction**

1. **I really enjoyed reading this. It is clearly written and gives a good sense of the previous literature.**

We would like to thank the reviewer for these kind words and are glad we succeeded in presenting a good overview of the literature.

1. **Having read it through, I still don’t fully understand why a control condition was not possible, lacking an activity component and simply asking for affect ratings. This will need some mention in the discussion section I think.**

Our main goal was to examine the effects of two reasonable EMI modules for depression on momentary affect. We agree with the reviewer that it could be interesting to compare such effects to a neutral control condition (to disentangle the content-specific effects from more general effects of self-monitoring). However, it is debatable whether a truly neutral control condition is at all possible in EMA; general self-monitoring is hardly neutral in itself, because participants will most likely reflect on non-neutral events, regardless of the question (see for example Arslan et al., 2020; Bos et al., 2019). Therefore, the limitation is inherent to the design. We have added the following sentence to the discussion section describing this issue (p.15):

*Ideally one would add a neutral EMA control group to disentangle the content-specific effects from more general effects of self-monitoring. However, whether it is possible to create such a truly neutral control condition is subject of debate, since monitoring affect can already be regarded as an intervention in itself (See for example Arslan et al., 2020; or Bos et al., 2019). Hence, the lack of such a group is an inevitable limitation to this design.*

**Methods**

1. **The pre-registration of the trial is a positive feature and I would praise the team for their honest and straightforward description of why blinding was not possible.**

We would like to thank the reviewer for these kind words.

1. **My stats knowledge is not sufficient to offer a review on that component of the manuscript**

# Reviewer 2

I write regarding the manuscript “Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect”. I enjoyed reading the manuscript and reviewing the supplementary information submitted by the authors, including the preregistration uploaded to the Open Science Framework. The manuscript focuses on a study of Ecological Momentary Interventions (EMI), which are repeated daily self-assessments with an aim of providing self-correction. The current manuscript asks whether two similar such EMIs differentially change positive or negative affect, and finds that they do not do so in statistically-significant ways.

I commend the authors on an interesting secondary research question that has clear implications both clinically and from a cognitive theory approach. The manuscript is clearly written. My major concern is that this manuscript focuses solely on the overlapping ecological momentary assessment mirrored in both “Do” and “Think” modules (measures of PA and NA

—although “physical state” was also measured in both, as mentioned in Figure 1), while ignoring the major connections to the RCT from which the data comes. However, I found the manuscript interesting to read, and believe that it contributes to the literature.

Below, I note some points by which the authors might improve their manuscript.

**Introduction:**

1. **The authors devote a significant portion of the introduction discussing two past EMI studies (Kramer et al., 2014; and van Roekel et al., 2017). I am not sure why they focus so deeply on potential explanations for the mixed effects (p. 4. beginning on line 76), as this is not the focus of the current manuscript. I believe this section could be significantly shortened.**

We inserted the discussion of mixed effects in the introduction to ensure the reader has a good understanding why we cannot make a directed hypothesis. Moreover, these studies explain why we expect a difference based on measurement content in the first place. Hence, we are hesitant to shorten this section too drastically. However, to be more concise we have summarized (reducing the wordcount from 104 to 40) the specifics of the previous EMI studies as follows (p.4):

*A second likely source is that the content of the EMA questionnaires and feedback differed between the RCTs. Besides differences in number of PA and NA questions, the first RCT included questions about activities and events, whereas the second RCT added questions regarding stress, worrying and discomfort, which could have created differences in focus between the two interventions.*

1. **I appreciate the focus on connecting the “Do” module to Behavioral Activation (p. 5), and on connecting the “Think” module to what I would call cognitive restructuring. (p. 6). However, I believe the manuscript could be bolstered by discussing research on repeated and tailored interventions for depressed mood either in the cognitive bias modification (CBM) literature or positive mental imagery literature. Some such studies have a similar desired outcome to those described by the authors, and similarly take active ingredients from BA or CR. Some examples follow. (I acknowledge that the final citation is to my own work, and leave it to the authors whether they believe it is appropriate to include.)**
* <https://doi.org/10.1037/a0024355>
* <https://doi.org/10.1016/j.jbtep.2016.03.012>
* <https://doi.org/10.1016/j.cpr.2010.01.001>
* <https://doi.org/10.1016/j.brat.2018.09.010>

We thank the reviewer for pointing out these interesting research areas, which we had not previously encountered in connection to the theoretical basis for our intervention. After reviewing the proposed articles (and other) articles, we added a paragraph with several key points of the mental imagery and memory modification literature in the introduction, mainly in relation to the question why the intervention modules are expected to have an effect on depressive symptoms.

 The results of the cognitive bias modification studies are too inconclusive (as evidenced by Hallion & Ruscio 2011) to derive firm conclusions about possible effects on momentary affect and somewhat less relevant for our study; our intervention simply prompts remembering without specific exercises or assignments for directing attention, contrary to cognitive bias modification interventions. Incorporating (parts of) that literature would, in our view, reduce the focus of the article and unnecessarily broaden the introduction. In sum, we added the following texts discussing self-referential cognition and memory modification in the introduction (p. 5):

*Furthermore, both the broaden and build theory (Fredrickson, 2004) and evidence from positive psychology interventions (Sin & Lyubomirsky, 2009) suggest that repeated focus on what goes well can result in a positive spiral of activities and PA. Similarly, a study by Dainer-Best et al. (2018) shows that repeated focus on positive cues (e.g. achievements) can lead to increased positive self-referential cognition, which in turn is thought to help reduce depressive symptoms.*

(p.6):

*Regardless of focus, both modules could influence affect through various memory processes involved in having participants reflect on the past three hours five times a day. Directly, by automatically triggering emotions associated with recalled events (see for example Holmes & Mathews, 2010 for a review), but possibly also indirectly through increased memory specificity. Depressive symptoms have repeatedly been associated with reduced memory specificity (Williams et al., 2007), and several studies have shown promising results of various memory training interventions on depressive symptomatology (see for example Hitchcock et al., 2016; Raes et al., 2009; or Watkins et al., 2009).These interventions typically entail some form of autobiographical memory recall, adding specificity by focusing on experienced emotions or context, which is comparable to the EMI questionnaires.*

1. **I would recommend rewording the phrase “We will therefore test two-sidedly” – perhaps “We will therefore use two-sided tests”.**

We thank the reviewer for this suggestion and adjusted the sentence in the introduction as follows (p.7):

*We will therefore use two-sided tests to investigate whether there is a differential effect of module type (Do vs. Think) on momentary PA or NA*

**Methods**

1. **Were smartphones provided to participants or were they completing the EMI on their own phones? How were feedback reports provided?**

Based on the reviewer’s comments we clarified the methods section. In the interest of sparsity, we overall try to strike a good balance between providing detailed information and referring to the protocol paper where possible. The altered sections under the subheading Methods-Experience Sampling Intervention (p.8) now read as follows:

*(p.8)*

*Both Zelf-i modules consisted of 28 consecutive days of EMA in which participants filled out brief questionnaires on their own smartphones five times a day. For each measurement a link to the questionnaire, hosted on a secured website for routine outcome monitoring (RoQua,* [*www.roqua.nl*](http://www.roqua.nl)*), was sent by text message after which participants had 30 minutes to complete the survey*.

*(p.9)*

*Personalized feedback reports were automatically generated by RoQua and then emailed as a pdf by a research assistant to the participant after each week of EMA, with each successive report containing richer information. The Do-module reports comprised various graphs showing PA patterns and associations between PA and activities, whereas graphs in the Think-module focused on NA over time and associations with thinking patterns.*

1. **The preregistration makes clear that data collection was completed at the time of analysis development. This should be explicitly stated in the Analysis section; it seems to be assumed that it’s obvious but it is never explicitly stated. The authors develop an effective solution for analyses (i.e., they asked a condition-blind researcher to re-code group membership before analysis – is that one of the authors?). The preregistration describes a specific analysis and two direction-less hypotheses (“Based on the current literature we cannot make a clear prediction which of the two interventions will have the largest impact on affect compared to the other”), which are mirrored in the introduction.**

The researcher who re-coded the group levels was a direct colleague, but not one of the authors, to ensure that discussion of the results could proceed unbiased. We added the suggested statement to the analysis section (p.10), as follows:

 *To reduce experimenter bias, all analyses and data handling procedures were preregistered before any analyses were performed (*[*https://osf.io/4e52q/*](https://osf.io/4e52q/)*), but after data collection was completed.*

1. **The researchers share their code on the OSF, which is appreciated. Given that data is not available, I wonder whether it might be possible to share the output of an .Rmd file as well (as PDF or HTML), thus allowing one to follow along with the results of the analyses. Certainly, I appreciated being able to review the lme4 models.**

At the original time of coding, we were not proficient in the use of Rmarkdown yet. Currently we are, so we transformed the entire analysis code into an Rmarkdown file and added an HTML file to the OSF project, to enable inspection of not only the code, but the results as well. We added the following text at the analyses section (p.10) in the manuscript as follows;

 *All analysis codes and outcomes can be found online (https://osf.io/bg7pr/)*

1. **I recognize that the authors preregistered the comparison of linear or quadratic time effects, but I think the manuscript may benefit from a theoretical background on this decision and what modeling time as quadratic might mean (especially given Figure 2). At the moment, it is only mentioned on p. 10**

We added a short explanation of the quadratic time component in the analysis section (p.11):

 *To test our hypothesis on differential effects over time the interaction term group \* time was included. As depressive symptoms do not necessarily change linearly over time (see for example Dinga et al., 2018), we estimated whether model fit improved with a quadratic effect for time (Table S2). This quadratic effect is technically an interaction of time with time, which indicates whether the effect of time on depressive symptoms differs over the period under inspection (e.g. stronger changes shortly after intervention start than later on).*

**Results**

1. **Table 2 and 3 both have notes that report that “Time was set in days, with five measurements each day”. I believe the wording on Figure 3 is slightly better (“Five measurements represent 1 day”); the first in particular makes it sound like Time in the analyses is Day (rather than measurement).**

As the reviewer pointed out, the description in the notes were not clear. This is particularly evident as the estimates in tables 2 and 3 actually relate to the change over days, not measurements, which was chosen for easier interpretation. We changed the description in the notes of tables 2 and 3 into the following (p.26 & 27):

 *Estimates for Time and Time2 represent the effect of one day, with five measurements each day.*

1. **Figure 2 is helpful; I wonder whether adding a web-friendly color might be helpful for online publication. The same goes for Fig 3.**

We recreated both figures in color, hoping that it will increase readability online. If you would like us to change color schemes please indicate so; we could not find specific color-setting instructions in the author guidelines. After a short inspection online we decided to use the following colorblind-friendly R color codes: #85c0f9 and #f5793a

1. **I really appreciate the individual predictive lines reported in Figure 3. (The quality in my copy is, unfortunately, somewhat low.) Are the lines predicted based on the models? This is implied but not stated.**

We added a short explanation of how the lines were created, following the first mention of Figure 3 in the Results – Momentary affect section (p. 13):

 *This heterogeneity is illustrated by the large spread of the individual predicted regression lines in Figure 3. The lines were created by using the model estimates described in Tables 2 and 3, and adding the individual model residuals.*

Furthermore, we added the word ‘model’ in brackets in the notes of Figure 3:

 *Individual (model) predicted lines for .. etc*

1. **On p. 15, the authors report that “there were no substantive subgroups of participants that did not read the feedback reports” (297–298). This was not previously mentioned as a possible analysis, although I note that it is described in the supplement. The supplement says that whether participants read the reports is measured by asking them after they are received—is that right? I think that either this should be removed from the manuscript or somewhat elaborated upon.**

The reviewer rightfully points out that the location (at the end of the result section) and explanation of not being able to perform this preregistered exploratory analysis is somewhat confusing. We adjusted the sentence in the manuscript (see below) and placed it in the subsection ‘analysis’ in the method section (p.10), where the pre-registration is also discussed.

 *Furthermore, the preregistered exploratory analysis on the number of
 feedback reports read was not possible because there were no
 substantive subgroups of participants that did not read the feedback
 reports (Supplementary materials: Feedback reports).*

Furthermore, we clarified the statement in the subsection ‘feedback reports’ (p.1) of the supplementary materials:

 *Participants indicated in a questionnaire, administered during debriefing
 (T1), which of the three interim feedback reports they read.*

Discussion

1. **The discussion returns to using the terms EMA, EMI, and “EMA intervention”; I think clarity in terms might help! (p. 16) e.g., the second paragraph begins by talking about EMA and ends with EMI. Is this distinction correct?**

We agree that the distinction was not entirely clear in our manuscript. In our revision we tried to be more consistent in our use of terminology, especially in the discussion and conclusion. Throughout the manuscript we have replaced every instance of “EMA intervention” with “EMI”, and made a few minor changes to the use of EMA or EMI.

1. **The authors do not report the results of the primary RCT from which this analysis stems. While I understand that such an analysis is likely under review or has been published elsewhere (and was not a preregistered analysis here), I wonder why the authors don’t take this into account – might it be that changes in NA and PA only occurred for participants with improvements in depressive symptoms? At a minimum, I believe that this should be addressed as a discussion point.**

The reviewer is correct in the assumption that the results of the primary RCT are currently under review elsewhere, and were unknown at the time of both preregistration and analysis of the current manuscript (a preprint has recently become available, see Bastiaansen et al., 2020). More importantly, however, although the question raised is an interesting one, our aim is to investigate the influence of ESM content on momentary affect. We therefore believe it is beyond the scope of the current article to include changes in depressive symptoms as a predictor variable. Because we do agree that the link to the primary RCT results could be of interest for readers, we have added a reference to the preprint in the methods section (p8), next to the reference to the protocol paper.

1. **The authors note that “Because all participants started treatment during or shortly after the EMI…” (p. 17, line 331): This probably should have been mentioned earlier! Does treatment here mean a TAU such as BA/CBT, or does it mean something different? Is the EMI not itself considered a treatment as the authors argued in the introduction?**

The reviewer suggests that the first mention of the role of TAU is on line 331, in the discussion section. This is, however, not entirely accurate, as it was mentioned in the introduction twice; at the first mention of the ZELF-i study and as hypothesis at the end of the introduction. However, we agree that the description is limited. We therefore made the mention of treatment as usual in the ZELF-i study more explicit in the introduction (p5):

 *The recent ZELF-i RCT (Bastiaansen et al., 2018) was set up to
 investigate the effects of two different EMI ~~content~~ modules as an add-
 on to regular depression treatment. ~~on depression and momentary~~ ~~affect~~.*

Furthermore, we added a description of the treatment as usual in the ‘Experience sampling intervention’ subsection of the methods (p9):

 *Treatment as usual was provided for all participants when available,
 irrespective of participation in the ZELF-i study. Most participants
 started a form of psychotherapy during the intervention period.
 This psychotherapy most often consisted of some form of cognitive
 behavioral treatment, in combination with a diverse number of other
 treatments (For further details see the preprint by Bastiaansen et al.,
 2020).*

1. **At several points, the authors appear to conclude that their results “could mean that EMA content does not matter” (quote from p. 16). I enjoyed the speculation on what this would mean, but think that the more measured conclusion reached in the abstract is more appropriate. All that we can conclude here is that asking about PA and NA, followed by asking about pleasure or worry, seems not to show differences between conditions.**

We agree with the reviewer that, in the current form, this point (‘could mean that EMA content does not matter’) is emphasized more strongly than was intended. To ensure that we do not imply this statement as conclusion but rather as part of the larger discussion of possible implications, we changed the phrase in the discussion section into the following (p.13):

*The lack of a diferential effect over time could mean that EMI content doesnot ~~matter~~ have a substantial influence on momentary affect.*

Furthermore, we removed a reference to this point in the limitations section of the discussion (p.15):

*Although ~~the conclusion that EMI content does not matter seems justified, we cannot exclude that our null finding is due to aspects of our study design.~~ the focus of the intervention modules differed, ~~but~~ there was overlap in the first part of each EMA questionnaire; both the Do- and Think-module measured PA and NA.*

**Editor Comments:**

* **First, I would say that I did appreciate your summary of the prior conflicting findings in EMI studies, and I thought your point about the importance of controlling for methods factors in understanding the effects of EMI content on outcomes was well made. I also really appreciated the carefulness and thoroughness of how you analyzed your data and reported your results. It very much boosts confidence in your findings.**

Prior to responding to your comments, we would like to thank you for these kind words, and we are especially glad that our efforts towards reproducible and open science was appreciated.

* **I had a few comments and questions about the analyses that I think could help readers understand your findings better. One, I was curious why you didn’t include time of day as a predictor, as well as day of study? Wouldn’t that help the model account for within-day time trends in NA and PA?**

Thank you for your suggestions. We assumed that your point of ‘day of study’ was intended to distinguish between weekdays and weekends, because the exact day of study (e.g. day 16) is implicit in the time variable used as predictor. Initially we did not include any such predictors because we did not expect (time of) day- predictors to affect the modules’ influence differentially. However, we agree that the model could improve by explaining additional variance, and thereby potentially unmasking specific effects. We ran two additional analyses each for PA and NA (four in total), one with time of day (dummy-coded) and one with weekdays versus weekends as covariate. The overall pattern of time of day and weekdays suggested an improvement of mood during the day and during weekends, but the primary results (the time and time\*group predictors) showed only very minor changes when time of day and weekends were added. We incorporated the results of these post-hoc analyses in the supplementary materials, with the following reference in the sensitivity analysis at the end of the results section (p.13):

*Upon reviewer suggestion we additionally included covariates to account for within-week (weekdays versus weekends) and within-day (time of day) trends of PA and NA. Again no large differences were found with the original models; details can be found in tables S10 to S13.*

* **I also thought you could do a little better in describing the random effects of the time/day trend in the text; you seem to only describe the variance of one of the time slopes, but it’s not clear whether it’s the linear or quadratic slope, and it’s not clear whether both of those time/day trends were estimated with random effects or not.**

We agree with your observations: upon inspection we noticed that the description of random slopes did not distinguish between the linear and quadratic time variable. We therefore added the following sentence in the analyses subsection of the methods (p.11):

*The model included random intercepts and slopes for the linear time variable, effectively allowing participants to vary in their experienced affect at the start (T0) and in trajectories of affect change over time. No random effects were estimated for the quadratic effect,*

Furthermore, we clarified which random effect was described in the momentary affect subsection of the results (p.12):

*The random effects, representing the differences between participants across groups, indicated a large variance in both PA and NA starting levels (S2 PA intercept = 145; S2 NA* *intercept = 212) and linear changes over time (S2 PA slope = 0.38; S2 NA slope = .47).*

* **I also wondered if you could go a little further in speculating why PA and NA might be confounded with measurement frequency. What is the psychological process that might account for that?**

We added the following section at the end of the discussion section (p.15):

*Together with our findings, this suggests that the measurement frequency might play a crucial role in the effects of EMI’s on momentary affect. For example, it is possible that reflecting on your mood a few times a day improves affect by providing insight into what goes well. Reflecting more frequently on the other hand could act as constant reminder of the depressive symptoms and subsequently reduce affect.*

**In summary, I think this is a promising manuscript and, I hope you will revise it for further consideration at Collabra: Psychology. I look forward to receiving your revision.**

**Final Editor decision—Accept**

Jan 19, 2021

Dear Dr. Daan Alexander Ornée,

I have now had a chance to read over your manuscript “Comparison of Two Ecological Momentary Intervention Modules for Treatment of Depression on Momentary Positive and Negative Affect”, along with the letter describing the changes you made. Thank you for your patience with this decision, and for your responsiveness to the concerns that the reviewers and I raised. I am happy to say that your paper is now officially accepted for publication in Collabra: Psychology. Congratulations on this excellent work, I think it will make an important contribution to the literature and I look forward to seeing it published! I hope your experiences with Collabra: Psychology have been positive and that you will continue to consider it as an outlet for your work.

As there are no further reviewer revisions to make, you do not have to complete any tasks at this point. Our managing editor will contact you in case there are any pre-prodution file related questions. You will have an opportunity to check the page proofs before we publish your article. Thank you again for publishing in Collabra: Psychology.

Sincerely,

Kevin King