**Jan XX, 2021**

Dr. Marian Berryhill

Section Editor

***Brain Research***

**We thank you and the reviewers for your careful reading of our manuscript, evinced by the constructive comments and insights provided therein. We would also like to express our gratitude for the opportunity to revise our manuscript, titled "Temporal changes in attentional resources devoted to mind wandering before the awareness: An ERP study." The feedback has helped us to significantly improve the manuscript (ID:** **BRES-D-20-01138).**

**The revisions have been made in consultation with all coauthors, and each author has approved this revised manuscript.**

**We look forward to working with you and the reviewers to move our manuscript closer to publication in *Brain Research*.**

**Sincerely,**

**Kazushi Shinagawa (M.A.)**

**Graduate School of Human Relations,**

**Keio University**

**2-15-45 Mita, Minato-ku, Tokyo Japan**

**+81 (0)3-5427-1517**

**kazushi\_shinagawa@keio.jp**

**Our responses to the reviewers' comments are presented below. The reviewer's comments are presented in italics.**

**-------------------------------------------------------------------------------------------------------**

Reviewer #1

***1. First of all, the online reference according to the authors was Cz, right? This was the same electrode that was used to average the data from the P3 waveform. Despite the fact that authors offline averaged the data to an average reference, in their actual recordings the data was set to 0 at this site. How may this have impacted results? Rolling back the data from 0 using an average reference at least decreased their signal. Why was Cz chosen as the reference channel if it was to be the one used to extract the data for the present results.***

We thank the reviewer for raising this concern. We selected Cz for online reference because it is easy to secure electrode contact, and it can prevent the influence of biological artifacts the most (Yao et al., 2019). The reason for selecting the average in the offline reference is that it is effective for recording the EEG of many electrodes because it can minimize the bias of potential distribution on the scalp (Dien, 1998). The EEG manufacturer recommends these settings. Moreover, we referenced some research using the same environments (Ladouceur et al., 2007; Li et al., 2012). Furthermore, since we focused on the difference between conditions, there is no effect on the main results.

[Added text]

>Methods 5.6 Data acquisition

We referenced previous studies that used a similar environment (Ladouceur et al., 2007; Li et al., 2012).

>Reference

Ladouceur, C. D., Dahl, R. E., Carter, C. S., 2007 Development of action monitoring through adolescence into adulthood: ERP and source localization. Dev. Sci. 10, 874-91. https://doi.org/10.1111/j.1467-7687.2007.00639.x

Li, L., Wang, M., Zhao, Q. J., Fogelson, N.Li, L., et al., 2012 Neural Mechanisms Underlying the Cost of Task Switching: An ERP Study. PLoS One 7, e42233. <https://doi.org/10.1371/journal.pone.0042233>

***2. Also, the authors state that there are fluctuations at least 4.5 seconds prior to the self-report of mind wandering, which is actually different from the beginning of the MW state. As they were using target locked baselines for their probes, how can this be claimed? Did the authors performed other time frequency analysis centered around the time of self-report? Moreover, there seems to be increased frontal negativity - was this analyzed.***

We appreciate the constructive comment of the reviewer. As you pointed out, since we discussed the results from previous studies, there was insufficient data to show that the change was really due to a decrease in the allocation of attention to MW. We reinforced the conclusion by conducting the time-frequency analysis and comparing MW and aware of their MW trials suggested in the next question and added these results to the discussion.

In addition, frontal negativity certainly increases as getting closer to the self-reports by visual inspection. Since medial frontal activities relate to internal thought, the reduction of these regions supports our result. Then, we analyzed this to reinforce our hypothesis. First, we plot the ERP waves to decide the components focused on in the subsequent analysis. However, there was no significant difference by permutation test (On the Cz, significant differences near 300 ms).

[Added text]

> Results 2.4 Post hoc analysis using the estimated change point

Second, we also conducted a two-way (aware/MW, pre/post) repeated measures ANOVA on the averaged power of theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz), gamma(30-40 Hz) bands in the 500-720 ms on Fz, Cz, and Oz electrodes (Fig. 5B). Theta bands on Cz, there was no significant main effect of response type and time (*F*(1, 25) = 0.308, *p*= .581, G = 0.005; *F*(1, 25) = 0.061, *p* = .808, G= 0.001). The interaction between response type and time was significant (*F*(1, 25) = 7.498, *p* = .009, G  = 0.112). A post hoc analysis revealed that when participants were aware of their MW, the power of the theta bands increased compared to the MW condition after the change point(*F*(1, 25) = 4.652, *p* = .041, G  = 0.115).

>Discussion 3.1

Moreover, the present study results showed that the power of theta bands after the change point in the aware condition increased compared to the MW condition (Fig. 5B). Theta bands have been implicated in cognitive control (Cavanagh and Frank, 2014; Cohen and Donner, 2013). Furthermore, the cortical theta band inversely correlated with the activation of the DMN (Wang et al. 2016). In addition to P3, the results from the time-frequency analysis supported our hypothesis that attention moved away from MW and returned to the task before self-report.

>Discussion 3.2

Concerning theta waves in the context of MW studies, Braboszcz and Delorme (2011) found that the power of theta bands increased in the MW condition, while the opposite trend was seen during task concentration. In contrast, Brandmeyer and Delorme (2018) found the opposite. In response to these inconsistencies, a classification study was conducted based on the content details. It revealed that theta waves decreased with self-referential task-irrelevant thinking (Bocharov et al., 2019). Moreover, especially those waves in the medial frontal region were also related to monitoring (Cahn et al., 2013). In the present study, theta band power increase only before self-reports because the focus on self took place just before awareness. It is known that theta waves increase during meditation compared to MW (Brandmeyer and Delorme, 2018). In other words, the increase in theta waves in the present study may be related to the rise in metacognition and cognitive control that led to the subsequent self-report.

>Methods5.6 Data acquisition (filtering)

In addition, we conducted different filtering from ERP analysis for time-frequency analysis. To analyze high frequencies, we used a 0.5-40 Hz bandpass filter.

>Methods 5.6 Data acquisition (epoching)

Moreover, the continuous EEG signal was segmented by the tone onset into 1000 ms epochs with a -200 ms baseline period.

>Methods 5.6 Data acquisition (time-frequency)

In the time-frequency analysis, we focused on 4–40 Hz for the analysis, and applied two cycles to the lowest frequency, and then increased the cycles by 0.5 cycle increments for each step. As a result, 80 ms-720 ms remained as the analysis window.

>Methods 5.9 Statistical analysis

Also, to capture changes in the time-frequency data that are not related to the sound stimulus, we analyzed the averaged power of theta, alpha, beta, and gamma bands in the 500-720 ms on Fz, Cz, and Oz electrodes. Similar to ERP analysis, we conducted a two-way (MW/aware, pre/post) repeated measures ANOVA on these data.

>References

Wang, W., Viswanathan, S., Lee, T., Grafton, S. T., 2016. Coupling between theta oscillations and cognitive control network during cross-modal visual and auditory attention: Supramodal vs modality-specific mechanisms. PLoS One 11, e0158465. <https://doi.org/10.1371/journal.pone.0158465>

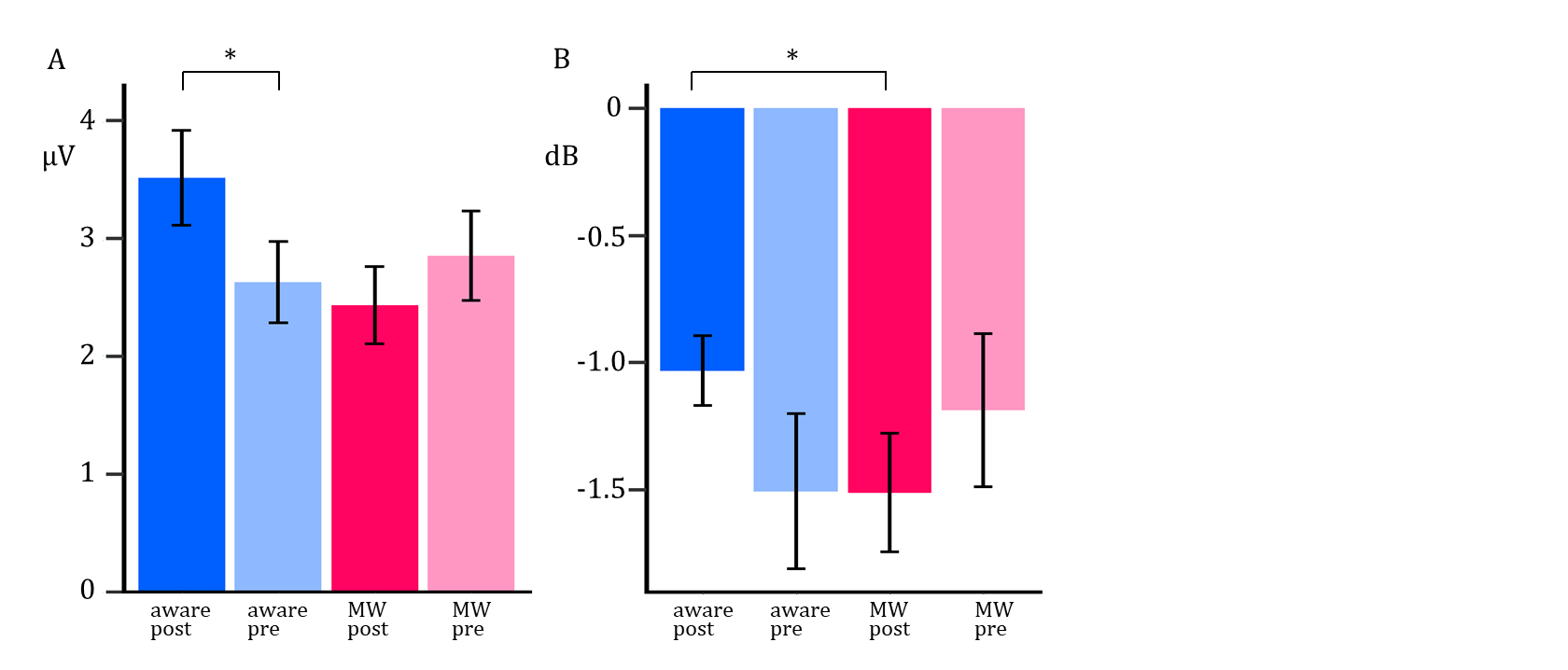
Cavanagh, J. F., Frank, M. J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. 18, 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cohen, M. X., Donner, T. H., 2013. Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. J. Neurophysiol. 110, 2752–2763. https://doi.org/10.1152/jn.00479.2013

Cahn, B.R., et al., 2013. Event-related delta, theta, alpha and gamma correlates to auditory oddball processing during Vipassana meditation. Soc. Cogn. Affect. Neurosci. 8, 100–111. <https://doi.org/10.1093/scan/nss060>

[Added Figure]

> Results 2.4 Post hoc analysis using the estimated change point



**Figure 5. Analysis results around the estimated change point.** **A:** The averaged peak amplitudeof P3. Error bars represent standard error. When participants are aware of their MW, P3 increases with time. B: The averaged power of theta bands on Cz. Error bars represent standard error. Theta bands stronger before participants were aware of their MW as compared to failure to respond.

***3. Also it would be important to show that P3 in a MW trial was distinct from a non MW trial - this would strengthen their arguments. Also the average of failure to respond was 12.44 and according to the cited references, it should be enough to average the P3. The article would also benefit from the inclusion of that data, in my opinion.***

We appreciate the reviewer's supportive suggestion. As you pointed out, it would be important to show that P3 in a MW trial was distinct from a non-MW trial. However, the section that participants non-MW is unclear in this experiment. We added a discussion related to this point in the Limitations section.

In addition, the article would benefit from the inclusion of data from failure response, we conducted an additional analysis. The mean number of failure responses is 12.44, but there are many individual differences. Since many individuals do not have enough trials, it is difficult to perform the same analysis in the old manuscript. Therefore, we performed a 2×2 ANOVA before and after the estimated change points. In addition, since the number of trials between failure condition and self-reports condition is likely to be large, we analyzed as yes when the response time exceeded the mean + 3SD. This is a significant point and is discussed further.

[Revised text]

>Results 2.1 Behavioral data

We, therefore, focused only on the MW reports

→ Therefore, we focused only on the MW reports in the analysis using a statistical model.

[Added text]

>Introduction

We also conducted additional analyses, such as frequency analysis, to examine the estimated change points in more detail to support the hypothesis.

>Results 2.1 Behavioral data

For subsequent analysis using the estimated change point, we integrated the stimulus into the same groups before and after the estimated change point (more details are 2.3 Model parameter inference) and treated them as one group used in the subsequent analysis. That is, -1 and -2 were classified as 'post' groups (after change point), and -3 to -5 were classified as 'pre' groups (before change point). We also defined the self-reports of MW and failure to respond as 'aware' and 'MW' conditions.

> Results 2.4 Post hoc analysis using the estimated change point

In addition to change point analysis, two analyses that focused on the change points were conducted for a more detailed study. First, we conducted a two-way (aware/MW, pre/post) repeated measures ANOVA on ERP data (Fig. 5A). There was no significant main effect of response type and time (*F*(1, 15.261) = 0.579, *p* = .458, G= 0.010; *F*(1, 24.279) = 0.609, *p*= .443, G = 0.010). Interactions between response type and time were significant (*F*(1, 22.276) = 4.814, *p* = .039, G  = 0.081). A post hoc analysis revealed that P3 increased before and after the change point when participants were aware (*F*(1, 26) = 4.675, *p* = .04, G  = 0.12).

>Discussion 3.1 Changes in attentional resources that precede to self-reports of MW

In the analysis including response failures, P3 increased before and after the change point only in the aware condition (Fig 5. A). This result suggests that the increase in attentional allocation is a unique change seen before self-reports since such a change was not seen in an MW condition. In other words, the result supports the hypothesis that continued MW does not lead to awareness and that a decrease in attentional allocation to MW is necessary.

>Methods 5.9 Statistical analysis

We conducted a two-way (aware/MW, pre/post) repeated measures ANOVA on ERP data.

>Limitation

It is also important to compare the results obtained in this study with those of the non-MW interval. However, in this study, the location of the non-MW sections was not clear. Moreover, the points of awareness were captured by self-report, so it may be challenging to capture the moments when participants focused on the task. As mentioned above, it is necessary to consider an experimental design that combines the random probe method.

Reviewer #2:

***1.The manuscript needs to be proofread by a native English speaker. Some sentences are not-well phrased. Many sayings are confusing that I have to infer what it means by the context.***

***For example,***

***p4 "Referred to as mind wandering (MW; Feng et al., 2013), this attentional shift away from a primary task to thoughts unrelated to the task is reportedly mundane…."***

***p9 "A systematic approach by which to assess validation and trustworthiness is to check effective" - does the "validation and trustworthiness" mean "validity and reliability"?***

***p32 subtitle Figure 2 "The number in the legend indicates the number of tones prior to self-reports of MW." - "position of tone" is more appropriate.***

I am very sorry for inconvenience caused. In addition to the parts you pointed out, I have made additional revisions and passed the manuscript through proofreading again.

[Revised text]

p4 "Referred to as mind wandering (MW; Feng et al., 2013), this attentional shift away from a primary task to thoughts unrelated to the task is reportedly mundane…."

→ This thought shift from the primary task to thoughts unrelated to the task is called mind wandering (MW; Feng et al., 2013). It is reportedly common, accounting for approximately 30%–50% of our waking hours (Killingsworth and Gilbert, 2010).

p9 "A systematic approach by which to assess validation and trustworthiness is to check effective" - does the "validation and trustworthiness" mean "validity and reliability"?

→ A systematic approach to assess validity and reliability is to check the effective sample size and R^ for each parameter.

p32 subtitle Figure 2 "The number in the legend indicates the number of tones prior to self-reports of MW." - "position of tone" is more appropriate.

→ The number in the legend indicates the position of tones prior to self-reports of MW

***2. p5 "We hypothesized that attentional resources related to MW also exhibit constant fluctuation, and this spontaneous attentional fluctuation relates to the disruption of MW."***

***It seems to me that you have some misunderstandings of "attentional fluctuation". The attentional fluctuation refers to the attention shifts between the external events (e.g., current task) and the internal thoughts and feelings unrelated to the goings on around it (Smallwood and Schooler, 2015). In other words, mind-wandering is part of this fluctuation process. When the attention shifts to the internal self-generated thoughts, it is a mental phenomenon referred as mind-wandering.***

We apologize for the lack of explanation and the confusion it caused. We used a misleading expression. At this point, we expressed that the amount of attentional resources invested in a particular topic can change spontaneously (fluctuate), and that this is also true for specific thoughts during MW. We think the main reason for the misunderstanding is that the use of ‘fluctuate’ reminds transitions between thoughts, and that the MW mentioned here assumes a specific topic, not a shift between thoughts. Then, to emphasize the element of specific thinking rather than just MW, we changed the wording to change attentional resources to task unrelated thoughts. To clarify the meaning of the sentence, we specifically emphasized the part about within a single task-unrelated thought and avoided the word fluctuate.

[Revised text]

p5 "We hypothesized that attentional resources related to MW also exhibit constant fluctuation, and this spontaneous attentional fluctuation relates to the disruption of MW."

→We hypothesized that attentional resources devoted to task-unrelated thought also always change, and these relate to the disruption of ongoing task-unrelated thoughts.

***3. It is not clear why and how the Markov Chain Monte Carlo model is implemented. Please elaborate it.***

We apologize for the error and thank the reviewer for bringing this to our attention. As you have pointed out, we should specify why we recruited the MCMC; we have therefore added this.

[Added text]

>methods 5.8 Model implementation

We recruited MCMC because this approach is broadly used for estimating the posterior distribution, especially in psychological research (Schoot et al., 2017).

>Reference

Van De Schoot, R., Winter, S. D., Ryan, O., Zondervan-Zwijnenburg, M., Depaoli, S., 2017. A systematic review of Bayesian articles in psychology: The last 25 years. Psychol. Methods. 22, 217. https://doi.org/10.1037/met0000100

***4. If the point is to compare the P3 at different positions prior to the self-reports, why not just do a statistical test for the P3 peak voltage in Figure 2c? It looks that there is a difference between -5 and -1. Please report the statistical results.***

We thank the reviewer for the constructive comment. We analyzed the differences between time points. According to the results, there was a significant difference between timepoints -5 and -1(p = .027). However, since this analysis was applied for within-subject conditions, a correction was made in order to maintain the independence of the data. This modification deletes the information about relevancy between timepoints. That is, it is not considered a point in time, but just factors, so information such as points -4 to -3 is reduced. Therefore, the results obtained by conventional ANOVA do not add new information, but rather may be inappropriate as a model for adapting to the data because of the loss of temporal information. For the above reasons, we have not included it in the manuscript.

***5. I think the main conclusion should be focused more on the self-awareness part than mind-wandering. It makes sense that P3 increases when their self-awareness is enhanced, but it is not directly linked to mind-wandering itself.***

We thank the reviewer for this comment. It is difficult to say that the change is due to the MW. In the original manuscript, the discussion focused on the fact that this was the section that was considered MW and that the component that was seen as an indicator of MW was increasing, but this was insufficient as you pointed out. Therefore, the results of the frequency analysis and the comparison with the MW condition were added to the discussion. (Rev 1 Q2, 3)

In addition, since awareness occurs before self-report, the perspective of heightened self-awareness may well be considered. However, since P3 responds to sounds, it may be an increase in awareness of sound, but it is not possible to mention the self. We added the discussion from the results of the time-frequency analysis. (Rev 1 Q2)

***6. There is no data for an on-task state as a contrast to mind-wandering. Thus, there is no way to observe a P3 change caused simply by mind-wandering.***

Thank you for raising these important concerns. As in the previous question, we thought that the change was due to MW because it is a change from the interval estimated to have MW in the previous study, but it seems to lack data. In the present study, we captured only the point when participants noticed the MW and failed to respond, and we did not know at which point they were on-task.

With the response to Reviewer 1, we have added a discussion of what we can do in this experiment to compare with the case where awareness does not occur and increase reliability. We have also added a discussion of indicators other than ERP concerning previous studies.

In addition, we do not think that the change is simply due to MW, but we aim to elaborate on this process by adding the perspective of attention.

***7. p12 In the discussion, authors mentioned the role of the DMN in mind-wandering and speculated that changes in attentional resources related to the DMN. This cannot be supported by the current EEG findings. Please revise it.***

We apologize for this oversight and thank the reviewer for bringing this to our attention. There are indeed some leaps in the discussion, and I wanted to suggest a future study. As you pointed out, the last part, which is insufficiently supported by previous studies, has been changed to expect future studies.

>Discussion

[Deleted text]

We speculate that a change in attentional resources is related to awareness of MW due to spontaneous fluctuations of the DMN.

[Added text]

It would also be meaningful to capture the amount of attention allocated to MW in terms of the spontaneous activity of the DMN over time. Future research should also focus on the interaction between network changes and the results of this study.

***Minor issues***

***8. Please check the journal requirement if the method should be at the end.***

We confirmed that we followed the rules of the journal again.

***9. p4 'including its adverse effect on…' -- In fact, mind-wandering has both pros and cons (see Smallwood and Schooler, 2015)***

As you indicated, MW has both pros and cons and has been focused on negative effects too much in the previous manuscript. We emphasize two-sided in the new one.

[Added text]

While there are positive aspects to MW,