

Dr. Ljerka Ostojić
Recommender
PCI Registered Reports

November 13, 2021

Dear Dr. Ljerka Ostojić,

Thank you very much for your time and effort. We appreciate the reviewers' thoughtful and constructive comments regarding our Stage 1 manuscript titled, "Is the past farther than the future? A replication and test of the time-expansion hypothesis based on the filling rate of duration."

We are pleased with your prompt action and the reviewers' valuable comments. We apologize for the delayed response. We have substantially revised our Stage 1 manuscript based on the reviewers' comments and our discussion.

We feel that the comments have improved the manuscript. Thanks to this revision, we have modified this manuscript with a more in-depth discussion of its pros and cons. We have also uploaded a copy of the original manuscript with all the changes colored in blue.

Note that we changed the key word, "event frequency" into "the filling rate of duration" in this revision. We use the idea of the filled duration illusion to form the logic and hypotheses of the present study. From this perspective, we believe that not only the number of events, but also the length of each event, contributes to how the durations are filled. Thus we felt that "frequency" was no longer appropriate, which gives importance only to the number of events. The title of our Stage 1 manuscript has also changed to "Is the past farther than the future? A registered replication and test of the time-expansion hypothesis based on the filling rate of duration." We mention this here before our response to the comments from the reviewers as this change is not based on their comments but on our recent discussions.

Comments & Replies to Review by Anonymous Reviewer

1-1

Your work builds off previous research demonstrating that people perceive a specified time interval in the past as being farther away than the same time interval in the future (i.e., the Temporal Doppler Effect). First, you propose a replication of the Temporal Doppler Effect. One explanation for the Temporal Doppler Effect is it results from our subjective experience that we are moving toward the future and away from the past. However, you propose testing an alternative explanation. Specifically, you want to determine if the number of events within the time interval in question explains the temporal asymmetry between the past and the future.

Reply:

We are grateful for your clear understanding of the purpose of our research and your thoughtful comments. Thank you for giving us the opportunity to reply to your comments and revise our manuscript accordingly.

1-2

You have a good start here. You are proposing a replication of the Temporal Doppler Effect and testing an alternative explanation for the effect. Your alternative explanation is reasonable and you provide some literature to support your ideas. However, the Temporal Doppler Effect has already been replicated with additional experiments and your alternative explanation has already been addressed to some extent by Caruso et al. (2013). Caruso et al. (2013) discuss and test event/task "filling in" as an explanation for the Temporal Doppler Effect. Their work related to "filling in" and intervening events as an explanation of the Temporal Doppler Effect should be discussed in the context of the research proposed in your submission. Your proposed method for exploring "filling in" and intervening events is somewhat different than the one described by Caruso et al. (2013) but both methods address "filling in" differences between future and past events.

Reply:

We are very sorry for our negligence in not mentioning Caruso et al.'s (2013) additional experiments on "filling in" and we have now added it to the introduction of the manuscript directly before "Aims of the present study".

1-3

Overall, this submission feels more like the start of a research idea than a fully developed one. There is a contribution here, but more is needed to confidently say the current project is a novel and significant contribution.

Reply:

Thank you for your comment. We believe that our study makes three major contributions. First, this study will extend the filled duration illusion from a short duration (e.g., millisecond timescale) to month and year levels, which is novel, and of great significance. Second, in the field of time perception and time estimation, there are only a few other studies that focus on both past and future orientations. Third, our replication of the Temporal Doppler Effect in the format of a registered report will contribute to open science and improve the robustness of this research field.

1-4

Your discussion of the Temporal Doppler Effect is clear, and your supporting discussion related to our representation of time and space is helpful. However, a more detailed description of the Caruso et al. (2013) Study 3 results is needed.

Reply:

We apologize that our description of Study 3 was not detailed enough, and we have rewritten this to provide a more informative description.

1-5

Much of the literature you discuss centers on previous work related to the Temporal Doppler Effect. This makes sense and your discussion is appreciated. However, a more

developed discussion of time perception as it relates to your research is needed. For example, you briefly discuss research exploring variables that impact time judgments of past events but additional detail would be helpful. A discussion of research exploring variables that are known to affect time judgments of future events is also needed.

Reply:

Thank you for pointing this out. We have added some additional explanations of factors related to past and future psychological distance.

1-6

Your rationale for sample sizes was consistent with current practice. You decided to use quite conservative values for your power analysis but given your focus on replication this seems reasonable.

Reply:

Thank you for your comment. Cashen and Geiger (2004) suggested the use of a power of .95 to reduce the possibility of Type II errors. Moreover, halving the effect size of the previous study is a common practice to make the replication more rigorous.

1-7

In your study, participants will first report the perceived temporal distance to a target day in the past (or future). Next, participants will report how "filled" this time period was with errands and events. Will a participant's errands/events response be driven by their perceived temporal distance report? If so, does this impact your ability to adequately address your research questions/hypotheses? Your submission would benefit from a rethinking of the experimental design or some discussion of these issues.

Reply:

Thank you so much for your careful consideration, but we are afraid that it is hard for us to change the experimental design. There are two main reasons. First, our experiments include a replication of the psychological distance of the past and the future. In order to strictly follow the procedure of the original study, we must address psychological distance first and then assess the filling rate of duration, rather than use a counterbalanced design. Another reason is our hypotheses. We hypothesized that the filling rate of duration will affect participants' perception of psychological distance, and not vice versa. Therefore, if we ask participants to answer the psychological distance after asking them the filling rate of duration, it may affect the estimation of the psychological distance. For these reasons, we will conduct the experiment without counterbalancing.

1-8

For study 2, you identify Cohen's d (0.45) from Caruso et al. (2013) Study 1b when explaining your sample size decision. You note this effect size is larger than the effect size used in your Study 1(0.26) to argue your Study 2 sample size can safely be equal to that of your Study 1. However, the Caruso et al. (2013) effect size that you used for Study 1 was 0.52. You halved the original effect size of 0.52 based on concerns that it was an overestimate. I don't think you can compare the 0.45 effect size (which is still in its original form) to the 0.26 effect size (which you adjusted from the original 0.52). Some revision

here might be needed to clarify your language (larger/smaller) and rationales.

Reply:

Thank you for pointing this out, and we have discussed it again and you are correct that it is hard to compare the original effect size with an adjusted one. So, we conducted a power analysis and to keep the significant numbers consistent with Caruso et al.'s (2013) Study 1b, we modified the effect size in Study 2 to Cohen's $d = .22$, which is half of that in Study 1b. Then, we adjusted the number of participants accordingly. We will recruit 936 participants in Study 1 and 1308 participants in Study 2.

Comments & Replies to Review by Chris Chambers

2-1

In this Stage 1 submission, Zhang et al propose an interesting pair of experiments to test whether a phenomenon termed the temporal doppler effect (TDE) – in which people perceive events in the past to be further from the present than events at an objectively equivalent point in the future – might be explained by a process related to the filled-duration illusion (FDI) – in which periods of time with a greater frequency of events or changes are perceived as being longer in duration.

Over two experiments, the authors propose to replicate the TDE, and further, to test whether the magnitude of the TDE is related to subjective judgments of event frequency.

My overall evaluation of this proposal is positive: the research question is clearly articulated and the hypothesis offers a potentially elegant test of an underlying cause of the TDE – assuming the TDE itself replicates successfully, which the authors will also test.

My main comments concern the justification of the link between the TDE and FDI, whether the proposed studies provide a sufficiently severe and falsifiable test of the overarching hypothesis, and the degree of methodological detail. I summarise these points below and have attached a commented version of the Stage 1 manuscript which expands these points and places them in a specific context (along with some additional suggested revisions).

Reply:

We are pleased with your clear understanding of the main ideas, and constructive and thoughtful comments. Thank you for giving us the chance to reply to your comments. We have revised our manuscript according to your suggestions.

2-2

1. The authors' hypothesis that the TDE is related to (and possibly driven by) same/similar mechanisms to the FDI depends on an auxiliary assumption (tested in H2-1) that people will report a greater frequency of events in the past than they will anticipate in the future. I may have missed this, but I couldn't find any basis for this hypothesis, either based on theory or previous evidence. As noted in my in-line comments in the manuscript, I feel this issue needs more attention and foundation.

Reply:

Thank you for your suggestion; we have added some explanation to make it clear.

2-3

2. It wasn't clear to me whether H2-2 (event frequency related to psychological distance) will be tested through separate correlations in past and future conditions, or through the data collapsed across these conditions. This in turn led me to wonder whether the link between the TDE and event frequency is being tested in the most severe way possible given the current between-subjects design in which past and future judgments are made by separate groups. As noted in my in-line comments, an augmented design could involve the inclusion of past and future judgments in the same participants (in a counterbalanced order) – this would allow for the calculation of a Past vs Future difference score which could then be related to the polarity and magnitude of the TDE. In addition, by selectively analysing the first session of the counterbalanced sequence, the authors could test H1 in a comparable between-subjects manner to Caruso 2013; and by taking into account the difference in event frequency between Past vs Future, the authors could establish a tighter link between the measures. For example, would participants who show a robust opposite trend in event frequency between Past vs Future also show an opposite TDE? If so, this would provide strong support for the authors' overarching proposition. If not, it seems to be it would also provide a more severe and convincing disconfirmation.

Reply:

We are excited to accept this thought-provoking suggestion.

We changed our experimental design so that it could meet both the requirement of the between-subject design in H1, and enable us to test H2-1 and H2-2 using a within-subject design. The sequences that we will ask about are as mentioned below, and will be counterbalanced between participants.

1. past time, past event, future time, future event
2. future time, future event, past time, past event

We have also modified our data analysis plan to fit this suggestion. As a result of the power analysis, a maximum of 936 participants will be recruited for Study 1 and 1308 participants for Study 2. For H1, which follows between-subject design, we will analyze all the data of 936 participants in Study 1 (or 1308 participants in Study 2) using a two-sample t-test. For H2-1, which uses a within-subject design, we will only analyze the first 102 participants' data, which will be extracted from timestamps of the survey form, in both Study 1 and Study 2, using a paired t-test. For H2-2, which will also use within-subject design, we will calculate the Spearman's Correlation coefficient of the first 386 participants' data in Study 1 and Study 2.

2-4

3. Throughout the manuscript, there are various points where greater methodological detail needs to be provided, e.g. concerning sampling, exclusion criteria, and instructions to participants. I was also left wondering whether the survey question should include attention checks to protect against inattentive or random responses. Finally, I felt that the rationale for both the 1-month and 1-year studies needs strengthening, over and above the fact that the studies replicate the approach taken by Caruso 2013.

Reply:

Thank you very much for your suggestions regarding the methodological details. We have

added details and modified the method section.

Regarding rationale for 1-month and 1-year studies, Gan, Miao, Zheng, & Liu (2017) conducted experiments with six time intervals (1 week, 2 weeks, 1 month, 3 months, 6 months, and 12 months). The results showed that the TDE was observed in all these time intervals, and they suggested that the TDE exists regardless of the length of time. In this study, we did not want to examine the effect of the length of time, but we wanted to conduct a strict follow-up replication, so we used 1 month and 1 year, like the original study.

2-5

General note: throughout the manuscript, as well as making comments regarding content, I have included some suggested language edits as tracked changes. These tracked changes are recommendations only and the authors should feel free to decline them where they disagree or feel that the edits reduce accuracy.

Reply:

We are grateful for your comments, including the very detailed points and language edits. They are extremely helpful for non-native English speakers like us. We have modified our content according to all your suggestions. Thank you once again for your thoughtful suggestions.

2-6

I would suggest some rephrasing of these sentences to tighten the language and improve clarity. e.g. “Alternatively, the TDE might be explained by asymmetrical perceptions concerning the frequency of events between the past and future. Previous studies have shown that people tend to perceive durations with more events or changes as longer than those with fewer changes – a phenomenon termed the filled-duration illusion (FDI). If the TDE is related to the FDI, then over a long duration such as a month or a year, the number of events should positively predict the estimated psychological distance. In this research, we aim to create a direct replication of the TDE and test a novel explanation for this phenomenon from the event frequency perspective.”

Reply:

Thank you for your suggestion. We have modified the text accordingly.

2-7

for second-level durations
Suggest explaining this

Reply:

Thank you for pointing this out. This description has been removed during the revision process.

2-8

However, there was no significant difference.

Expand and clarify: no statistically significant difference between what exactly? And if there is was no evidence of a difference, was Caruso’s conclusion (explained in the next two sentences) justified?

Reply:

We apologize for the misleading term; we have modified it to a clearer one as follows:

[p.4]

The results indicated that when the participants perceived moving forward, the past felt more distant than the future. In contrast, the future felt farther than the past when they perceived moving backward, although the effect was not significant. Further, the psychological distance of the future when moving forward was not significantly different from that of the past when moving backward. Similarly, there was no significant difference in the psychological distance of the future when moving backward and the past when moving forward. Consequently, Caruso et al. (2013) proposed that the temporal asymmetry of psychological distance is formed by the perception that people are moving toward the future and moving away from the past; approaching (i.e., future) events are felt psychologically closer while retreating (i.e., past) events are felt farther, despite having the same objective temporal distance.

2-9

To avoid the existence of the anchoring effect

This needs further explanation

Reply:

Thank you for your suggestion. After our discussion, however, we did not describe the study (Aksentijevic & Treider, 2016) in detail because it was focused on the aspect of movement, whereas we wished to examine TDE in terms of the FDI-like effect, such as in the present study.

2-10

The existence of FDI-like phenomenon when changing the duration to hours, days, months or longer, is questionable

Because there is mixed evidence, or because there is no evidence? Suggest clarifying.

Reply:

We apologize for this misleading term and have changed it to a clearer one.

2-11

we assume that a duration's fullness

Do you mean the perceived length of a duration or the extent to which a duration is filled with events? Perhaps clarify here so that the use of fullness is clearer later on

Reply:

Thank you for your comment. We have clarified this in the revised manuscript. After our discussion, we have changed the phrase to “filling rate of duration” rather than fullness, to avoid this vagueness. We hope that this will be clearer than before.

2-12

Therefore, we assume that there is an asymmetry of the number of events between past and future.

Why is this? I don't understand the basis of the assumption that the number of events that has already happened will be greater than the number of events that are scheduled or imagined for the future. In order for the TDE to be explained by the FDI, this is presumably a crucial auxiliary assumption? Is there previous literature supporting this assumption?

Reply:

We are sorry that, to the best of our knowledge, there is no literature that directly supports this hypothesis. As this assumption is the link between the TDE and our hypothesis regarding the filling rate of duration, we think the hypothesis itself is something new to consider and have added some additional details in the manuscript. We hope that this hypothesis is better supported.

2-13

We will replicate the TDE directly by a registered report. To prevent the publication bias of replication research, we choose to do it as a registered report because pre-registration is not enough to prevent the such bias (Ikeda, Xu, Fuji, Zhu, & Yamada, 2019). This will contribute to the robustness and transparency of the TDE research.

I think you could cut these three sentences. There is no need to add a special justification for using the RR format. Instead, I would suggest including a concrete justification for why you are proposing two studies, with 1 month and 1 year timeframes. I'm finding the rationale for the two time-frames somewhat lacking at the moment (aside from the fact that it replicates Caruso 1a and 1b)

Reply:

Thank you for your suggestions, however, we are not willing to cut these sentences because we would like to emphasize the use of the RR format.

The time conditions of 1 month and 1 year are described below. Indeed, the TDE was observed at six intervals (1 week, 2 weeks, 1 month, 3 months, 6 months, and 12 months) in previous replication studies (Gan, Miao, Zheng and Liu, 2017), and it is quite possible that the TDE could be identified in time conditions other than those we are using. However, our main aim was not to test whether the length of time affects the presence or absence of the TDE, but to perform a rigorous follow-up replication in the form of RRs. Therefore, we will not examine conditions other than 1 month or 1 year in this study.

2-14

(H2-1), because all the past events have been experienced, while only the scheduled events can be considered in the future

What is the evidence from previous literature to support the rationale for this hypothesis? It seems to me that this is a crucial auxiliary hypothesis because if the number of events in the past isn't greater than the number of events in the future then is the rationale for H2-2 still justified?

On the other hand, if there are stable individual differences in the difference between the number of (experienced) past and (imagined) future events, it seems to me you could use

this to provide a more severe test of your overarching hypothesis, because participants with a greater difference (i.e. higher value of Past > Future) should exhibit a greater TDE, and participants who (for whatever reason) happen to show a reliable negative difference (i.e. Future > Past) should presumably show an opposite TDE? And if they didn't then maybe this would falsify the hypothesis that the event frequency is the determining factor?

Also: as I understand H2-2, it is a separate correlation for past, and a separate correlation for future? This wasn't clear from the design table either. Suggest clarifying.

Reply:

① Regarding the evidence of H2-1, we have already provided an explanation above, please refer to 2-12.

② Depending on the results, we foresee the following. If the TDE is successfully replicated while the filling rate of duration in the past is not greater than that in the future, the filling rate of duration cannot explain the TDE. On the contrary, if the TDE is not replicated but H2-2 (fuller duration correlates with farther psychological distance) is supported, only the psychological distance of the past and future can be explained by the filling rate of duration.

③ For H2-2, we will conduct a correlation analysis for both future and past data simultaneously. This is because we hypothesize that the relationship between psychological distance and the filling rate of duration would be the same in the past and future. The manuscript and design tables have been revised accordingly.

2-15

Similar to Caruso et al.'s (2013) study, temporal direction is a between-subject factor.

In relation to my previous comment, if the temporal direction is a between-subjects factor, then you wouldn't be able to test the more severe hypothesis I suggested based on the relative difference in event frequency between past and future, because you would only have measures of one or the other in each participant.

It may be worth considering a within-subjects design in which the order of future and past is counterbalanced. You could keep the design as it is now but just add an additional condition for each participant (so if they did future first then they do past second; if they did past first they do future second). That way, you could use the data from the first session in the planned between-subjects analysis to provide a close replication of Caruso 2013, while also having data from future and past sessions within each participant available to be able to measure the difference, and hence, the severe hypothesis that the difference in event frequency between future and past predicts the polarity and magnitude of the individual's TDE.

Reply:

As mentioned in 2-3, we modified our data analysis plan to fit this suggestion. Thank you for your suggestion.

2-16

We will use a Likert scale from 1 (not filled at all) to 10 (all filled up), and participants will

report how many errands and events they have already experienced or will experience in 1 month (1 year in Study 2)

Confirm that this is the same or different methodology as Caruso (eg. Likert with same number of intervals etc) and explain why if there are any differences. Same comment for the psychological distance DV.

Reply:

We apologize for this unclear explanation. We will use the same methodology as Caruso et al. and have modified the content in the manuscript.

2-17

Exclusion criteria include:

Are there any data-based exclusion criteria? E.g. incomplete data (if they only completed part of the questions?)

Reply:

We apologize that we did not elaborate on this in the manuscript. The tool we used (Google Form) allows participants to send a response only if the participants have completed all the answers. We can also make each item mandatory, so that the next page (or the result) cannot be accessed if there is an item that has not been answered. Since we planned to use this feature, we believed that incomplete data could not occur. Therefore, we did not include participants with incomplete data in the exclusion criteria.

2-18

We intend to recruit up to 1000 participants to avoid the risk of insufficient sample size.

Rather than stating an intended total sample size before exclusions, I suggest stating that data collection will continue until the minimum sample size indicated by the power analysis has been met after exclusions, regardless of the exclusion rate. Same comment for Study 2.

Reply:

Thank you for your suggestion, and we are grateful to accept it. We have modified the manuscript.

2-19

and presenting the instructions and questionnaire in Japanese

Will the instructions make clear to participants what is an “event”? I can imagine this being interpreted very differently by different people. Can you provide an English translation of the exact instructions you will use in Supplementary Information?

Reply:

In our study, we will use the word “用事,” which roughly translates to “event” in English, and “出来事,” which translates to “errand.” We agree that the “event” can be interpreted differently by different people in English. However, we think the interpretations of these Japanese words that we will use are common among most Japanese people. The word “用事” includes important activities, such as meetings, rather than everyday routine, such as bathing. And the

word “出来事” refers to something happening accidentally. We apologize for not providing a clear explanation of these Japanese words before. We also added this explanation to the manuscript.

2-20

the day of the study (future condition) and report the target day's psychological distance, using a Likert scale from 1 (大変短い時間である: a really short time from now) to 10 (大変長い時間である: a really long time from now). Next, participants will report how much the past month (past condition) was filled with errands and events or how much the coming month (future condition) will be filled with errands and events using a similar scale from 1 (全く埋まっていなかった: not filled at all) to 10 (すべて埋まっている: all filled up).

Would it be useful to have some form of attention check so that you could exclude responses made randomly (e.g., by bots or by participants just pressing buttons without thinking), or by people who didn't read the question? You could do this by including a question at the end that asks them to recall what the previous questions asked about, or perhaps a purely factual multiple-choice question to test that they are paying attention generally?

Along the same lines, if you are able to measure response times, it may be worth considering discarding data where people respond so quickly that they couldn't have read the question?

Reply:

Thank you for your thoughtful suggestions. We have added an attention check asking “Which year of Reiwa is it now?” to exclude invalid responses. Reiwa is the current regnal era name of Japan, and the year 2021 AD is Reiwa 3. Japanese people are very familiar with this era name; hence, we can use the same 1 to 10 scale for the attention check that has been used for the other questions. We have modified the description in the manuscript.

2-21

($\alpha = .05$)

I'm not sure yet if the disciplinary scope of this submission would be suitable for the journal Cortex (or if you are even considering Cortex yourself), but in case you are thinking of Cortex as a PCI RR-friendly outlet, keep in mind that alpha would need to be set to .02. A threshold of .05 is fine at all other PCI RR-friendly journals.

Reply:

Thank you for this information. We changed the alpha level into .02.

2-22

However, the participants will be limited to those who have not participated in Study 1. How will it be determined that participants did or did not participate in Study 1?

Reply:

We have used the blacklist function provided by Yahoo! Crowdsourcing to prevent those repeat

participants between Study 1, and 2. The manuscript has been modified.

2-23

from 1 month to 1 year.

Confirm here if the inclusion and exclusion criteria are the same.

Reply:

Thank you for your suggestions; we have added the note about the same inclusion and exclusion criteria.

2-24

We will conduct a correlation analysis between psychological distance and event frequency using Spearman's rank correlation coefficient.

Is this one correlation analysis each for future and past conditions, or a combined analysis?

Reply:

We apologize for this vague description. We plan to conduct a correlation analysis combining the past and future data. In other words, we will not distinguish between the future and the past in terms of psychological distance, and will analyze the data simultaneously. This is also because we hypothesize that there is a relationship between the filling rate of duration and psychological distance regardless of past or future. We have modified our manuscript.