*20-Dec-2019*

*Dear Dr Clarke,*

*The Editors assigned to your Stage 1 Registered Report ("Does pre-crastination explain why some observers are sub-optimal in a visual search task?") have now received comments from reviewers. We would like you to revise your paper in accordance with the referee and editors suggestions which can be found below (not including confidential reports to the Editor). Please note this decision does not guarantee eventual acceptance.*

*[...]*

*Kind regards,*

*Anita Kristiansen*

*Editorial Coordinator*

*Royal Society Open Science*

*openscience@royalsociety.org*

*on behalf of Professor Chris Chambers (Registered Reports Editor, Royal Society Open Science) openscience@royalsociety.org*

*Associate Editor Comments to Author (Professor Chris Chambers):*

*Associate Editor: 1*

*Comments to the Author:*

*Two expert reviewers have now appraised the manuscript. Both assessments are broadly positive but raise a number of concerns that span the full range of Stage 1 review criteria and will require careful attention to achieve in-principle acceptance. Reviewer 1 suggests a range of clarifications to the methods and inclusion of greater methodological detail, as well as suggesting some additional or alternative outcome measures. Reviewer 2 queries the rationale for the hypotheses and notes the need for greater justification and elaboration concerning the analyses involving the psychometric measures. Concerning Reviewer 1's question about the inclusion of pilot data in the preregistered experiment, this data should of course not be re-used -- the preregistered study must include new data.*

*Comments to Author:*

*Reviewer: 1*

*Comments to the Author(s)*

*I am intrigued by your proposed study on the pre-crastination effect, visual search strategy, and personality. Your pilot data is promising; it is interesting that pre-crastinators may have a more optimal visual search strategy. I agree that the speed-accuracy trade-off in the visual search task should be further investigated with the entire dataset. My specific comments are attached as a Word file.*

*1. Pilot results (a) Accuracy and reaction time It seems that your most interesting accuracy and reaction time results are for hard and absent trials. I wonder whether you should exclude the easy trials and increase the number of hard and absent trials to increase reliability.*

The reviewer is correct that the easy trials are not particularly interesting - by design, participants should be close to ceiling performance. Unfortunately, reducing the number of trials in which the target is present on the easy half would risk biasing search strategies. At least some of our participants would probably develop a prior that the target is more likely to be present on the hard side, or that it is more likely to be absent, which would change their search behaviour and/or response criteria.

*4. Methods Are you going to include the 30 participants from the pilot experiment in the current experiment?*

No, the data from the pilot experiment will not be reused in the planned main experiment. However, for completeness and to maximise transparency, we plan to include the pilot data in the final manuscript. The pilot and “main” data will clearly be differentiated. For example, we can include a multi-level model in the Exploratory Analysis section (alternatively, in an appendix or supplementary materials) that includes a 3-way random effect (pilot/Aberdeen/Essex). This will allow us to calculate a (hopefully!) robust estimate of the fixed effect and how it varies from one replication to another (due to a range of unspecified site/population and experimenter effects).

But, please note, the main, pre-registered analysis for our experiment will only use data from 200 new participants. We have clarified this point in the text (ref line number).

*(a) Pre-crastination*

*(ii) Procedure I think you referenced the wrong study for your bucket task procedure. You referenced [8], Fournier et al. 2019 Psychological Research, but I think you meant Fournier et al. 2019 Attention, Perception, & Psychophysics. That latter paper is missing in your references. I think there are several other times you mixed up these references throughout the manuscript.*

We have fixed the relevant references, thank you.

*You mention that, “We will closely follow the methods and instructions as described in their appendix.” That appendix, however, has instructions about golf balls and alpha arithmetic problems, which I don’t think applies to your task. I would suggest including all the necessary details of the bucket task in the current manuscript.*

The experiment protocol document followed by the diagram of the bucket task setup will be publically available on the OSF. We now state this in the methods (page 6).

*Do you think three bucket task trials before and after the visual search task is enough to reliably estimate performance? More trials would increase your reliability. It seems like it would be easy enough, in terms of time and physical effort, to have 10 trials (or more?) before and after the visual search task.*

In the interest of being able to predict the results of the planned experiment based on the pilot, we would like to keep the number of trials the same in the experiment as in the pilot. The bucket task is the same on each trial, unlike in the search task in which many aspects of the array vary unpredictably. Most participants were consistent in their bucket choice across all 6 trials. With more trials, participants’ decisions might change. Perhaps this estimate based on the larger number of trials would be more reliable, but equally, repeating the decision more times might change it, for a variety of reasons, and it would be difficult to interpret this change.

*(iii) Data processing I understand that for many of your analyses you will need to divide performance on the bucket task into categorical levels (you mention using two or three levels). Is it possible, for any of your analyses, to use the probability of picking up the closer bucket first instead? It may be beneficial to keep bucket task performance as a continuous variable.*

This is something that we considered and indeed, an earlier version of the manuscript stated that we would consider treating pre-crastination as a continuous variable. But it wasn’t obvious to us how to formally specify the decision criteria for two (or three) levels and a continuous variable ahead of time. One of the strengths of the Registered Report format is that it reduces the number of “researcher degrees of freedom”. We decided to stick to a discrete factor, as this is what previous research has used, and was supported by our pilot data. We plan to return to this issue in the Exploratory Analysis section once we have collected that data.

*(b) Split-half visual search*

*(ii) Procedure I’m confused here about the number of trials. You say there are 10 trials (experimental or total?) but then add that 8 practice trials precede the experimental trials. Do you mean there are 8 practice trials and then 10 test trials? Furthermore, what are easy and hard trials? Do both types always have a target present? If so, I’m unsure how you are splitting 5 target-present trials into easy and hard.*

Thank you for pointing this out. In the pilot, we included 10 “free-view” trials, in which participants passively viewed the split screen array. We included these to examine where participants prefer to look when there is no task. Upon reflection, we do not really need these free-view trials in this experiment, as this is not relevant to our current question. We will take them out of the experiment (and the methods). There are 8 practice trails, half target present, half target absent.

*(iii) Data processing You mention here that you will look at the “proportion of fixations two through five falling on the heterogeneous side of the display”. Yet, in the Planned Analysis and Search Strategy sections, you mention the first five fixations. Which method will you use and why?*

This was a wording error on our part. All (pre-registered) eye-tracking analysis will use fixations two through five. This is based on our previous work using this paradigm and it is where we expect to see the largest individual differences. (In practice, there is negligible difference between “fixations two through five” and “the first five fixations” because, in the vast majority of the trials, the initial fixation will be at the centre of the screen and hence removed from analysis.) We have edited the manuscript to refer to “fixations two-through-five”

*(c) Personality questionnaires*

*(ii) Procedure Have you considered counterbalancing when the questionnaires are completed, with half of the participants completing them at the beginning of the experiment and the other half at the end? I’m trying to think whether completing the questionnaires first could affect your subsequent task performance or, alternatively, whether doing the tasks first might affect your subsequent questionnaires.*

We did consider that the questionnaires might influence the behavioural measures, which is why we decided to leave them to the end. Our main goal in the current experiment is to try and replicate the results we observed in the pilot, so we wanted to keep the two behavioural measures the same. This way, any potential change in the results between the pilot and the proposed experiment would allow us to interpret the pilot results as having been due to chance. If we counterbalanced, and the questionnaires did influence the results, we would effectively have halved our sample size for a clear replication (or need 400 participants, which is not feasible). We will, however, make sure we acknowledge possible order effects when we report the questionnaire results.

*(e) Outcome-neutral conditions I wonder whether the visual difference in pre-crastinators between your pilot participants (43%) and your previous experiment (74%) is caused by the restricted space to navigate and carry buckets in the hallway compared to an open room. It might be best for data analysis if half of the participants precrastinate, so perhaps 43% is better than 74%*

*Thank you. We will make a note of this point in the discussion.*

*Reviewer: 2*

*Comments to the Author(s)*

*The authors propose to examine the role of pre-crastination in sub-optimal visual search performance, as well as assessing the additional role of other personality factors. In general, I thought their proposed methodology was sound, feasible, and well-justified. The clarity and detail of the methodology was also sufficient, and outcome-neutral conditions have been considered well. I have two main comments though that I believe are relevant and should be considered.*

*1) Mainly a methodological point. The authors are proposing to use a number of personality measures (the Big Five Inventory, Barratt Impulsiveness Scale, procrastination scale). It was not clear to me though whether they intend to only relate these questionnaire measures to the likelihood of pre-crastination, visual search performance, or both. As this may introduce a large number of correlations, will the authors apply any corrections for multiple comparisons? Or is this unnecessary given the Bayesian approach proposed?*

The main hypothesis we are interested (and formally testing) is the link between our two behavioural measures (precrastination and visual search). The only prediction we have around the personality measures is that conscientiousness should be positively correlated with precrastination. Taking this a step further, understanding the three-way relationship between precrastination, conscientiousness and efficient search will help us better understand the nature of the relationship between each (e.g., do precrastination and conscientiousness explain independent, or shared, variance in search efficiency?). We have explained this more clearly in the paper (also in response to your second point below). The additional personality measures were an easy addition to our protocol and the analysis of them will be considered “Exploratory Results”, intended to provide useful hypothesis-generating data for others in the field as well as ourselves.

The Bayesian methods that we will be using are less interested in answering “is there an effect of A on B?” and more centred around measuring i) “how large is the effect of A on B”, and ii) “how much uncertainty is there around our estimate?”. As such, I’m not aware of the need to correct for multiple comparisons in a Bayesian framework. That said, using a relatively conservative prior would be a good idea considering the large number of different measures, so we have added this to the analysis plan.

*2) As a more general point, it did seem to me that the rationale leading to the authors’ main hypothesis was vague. The authors outline the literature on individual differences leading to sub-optimal visual search, and then simply state that they hypothesize a role of pre-crastination and review previous literature on this phenomenon. While their proposed study is in my opinion methodologically sound, and already has some pilot data (though it’s unclear how statistically reliable this data is), I am cautious to recommend that this work should be published in principle within the journal unless a more detailed rationale for the hypotheses can be provided. Otherwise the research comes across to the reader as high risk. Alternatively, if it is the case that the idea is just so novel that there is not much to base the predictions on, then stronger emphasis on the applied implications of the work (if hypotheses are supported) would be vital to demonstrate that the potential academic gain associated with the research outweighs the risk.*

We have strengthened the hypothesis/rationale with the following text:

“Our experiment tests the hypothesis that pre-crastination might represent a rational approach to doing tasks, i.e., a reasonable tendency to get the difficult part of the task out of the way first. If precrastination is positively related to efficient search, this would be evidence in favour of a ``rational'' interpretation of precrastination, as opposed to the alternative interpretation that it reflects an irrational impulse to grab the closest task-related object (in which case, a negative correlation with search efficiency would be more likely). In addition, the experiment has the potential to assess whether the already-established link between pre-crastination and conscientiousness could generalize, and be a potential explanation for the wide range of different search strategies previously observed. This would have important implications for theories of visual search, which tend to make assumptions about how ''people'' (as a group) direct their eyes during search; if personality traits can be directly predictive of eye movement strategies, this will challenge the field to more carefully refine its assumptions. In addition to the clear theoretical contribution our results could have, they could also illustrate that efficient/optimal tendencies in one circumstance can translate to inefficient choices in others. Finally, we have noticed in many fields that there is a lack of correlational studies relating behavioural measures to each other (as opposed to relating behavioural measures to personality traits).”

*Minor comments*

*- Page 3, line 40 – the authors refer to a study showing a relationship between pre-crastination and conscientiousness, though the citation is missing.*

Thank you, this citation has been fixed.