Recommender’s decision

Dear Dr Constant and colleagues,
Thank you for submitting your Stage 2 report and sorry for the delay in getting back to you. I've now received two reviews back from the original reviewers and I've read your manuscript as well. I agree with both of the reviewers that your Stage 2 report is extremely impressive and stays true to the Stage 1 report. The exploratory analyses are valuable additions to the manuscript as well, particularly your analysis of other N2pc time windows used in previous work - together, it's a really thorough treatment of the N2pc that'll be a very valuable contribution to the literature and I was excited to see how the results came out.
As you will see, the reviewers have some very minor suggestions. Dr Hickey made an interesting point about what the N2pc reflects. If you do address this point in the manuscript, it would be best to do so in the discussion rather than the introduction, where in-principle acceptance has already been granted.
I look forward to receiving your revised submission.
Best wishes
Maxine

Dear Dr. Sherman,

Thank you and the reviewers very much for your kind and careful reviews, and useful comments. We have replied to and addressed all the points raised. In the tracked-change version, we have marked in red additions and in crossed-out blue deletions compared to the first version of the Stage 2 report and it is also available on the OSF repository. We have also updated the preprint’s clean version.
We would like to point out that the change we made to the introduction in response to Dr. Hickey’s first comment (pp. 8-9) may appear large, but is mostly restructuring, we hope that this is fine.

Best regards,
Martin Constant

Response to Reviews

Reviewer 1 (Reny Baykova)

Constant et al. have replicated the second EEG experiment reported in Eimer (1996), which investigated the N2pc component of the visual event-related potential (ERP). The N2pc is commonly interpreted as a marker of attention, and the original study found that it can be elicited by different stimulus properties, namely shapes and colours. In addition, Eimer (1996) also found that the amplitude of the component was greater for shapes. Constant et al. conducted a replication of the original experiment as part of the EEGManyLabs project, and a total of 22 separate labs were involved in the replication. The N2pc effect using shapes was replicated, as well as the difference in

amplitude between shapes and colours. The effect using colours was not replicated using the strict replication of the original pipeline and the ICA pipeline. It was, however, replicated using the collapsed localiser pipeline and in the additional exploratory analyses which didn't use a fixed time window.

Overall, this is a very ambitious project that was conducted with great care. The study provides a valuable contribution to the literature on the N2pc and ERPs more generally. I commend the authors on the impressive amount of work they undertook with this project, and I have only a few minor comments.

1. Transparency and openness statement

I would suggest adding links to the data, analysis scripts and stage 1 report in the transparency and openness statement so that they are easier to find.

Response: Thank you for this suggestion, we have added this information to the transparency and openness statement. To avoid too much redundancy, we have removed the "Data sharing protocol" subsection. Please also note that the links to the data changed, because we completed data curation during the review time, and, as recommended by our repository, we have also added the citations to the data, so that they are more findable.

2. Tables containing Bays factors

Apologies if this is a common notation I'm not familiar with, but I found the notation "BF-0" initially confusing. I believe BF10 is the more commonly known notation, but if you prefer to go with BF-0 that's completely fine, I'd just suggest adding a note to the tables to describe what it means.

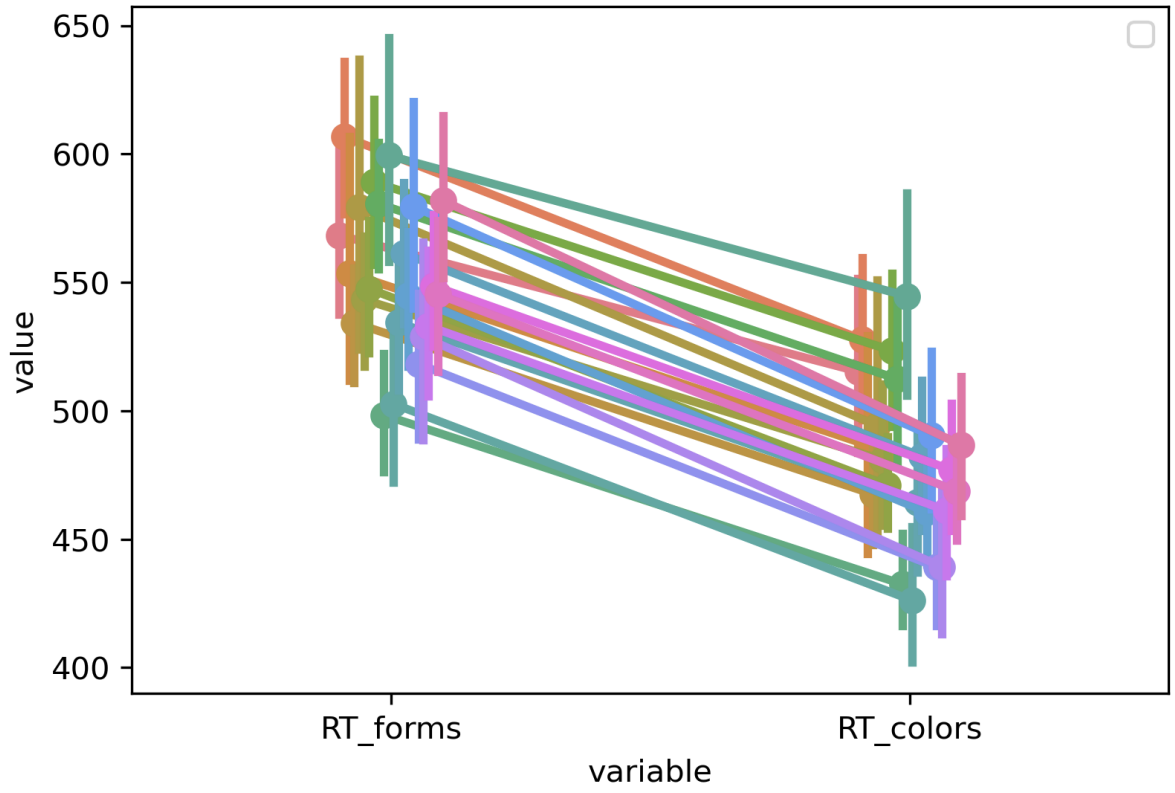
Response: Thank you for raising this issue. This notation is indeed less known than BF10. It is notably used in JASP and BayesFactor to indicate that the test is one-tailed (the minus indicates that we are testing for negativity, if we were testing for positivity, BF+0 would've been used). As this is conveying a bit more information about the nature of the test than BF10, we have kept this notation and added a note to Table 2 (i.e., first table where it is used).

3. Figure 17

Since each dot represents the average reaction time for each lab, I would suggest adding error bars to show the variability of the data (unless it all gets too messy).

In the description of the figure, I would suggest specifying that this shows reaction times on correct trials and trials kept after rejection.

Response: While we agree that showing dispersion information would be ideal, the figure indeed becomes very messy with error bars (see below a first attempt at making this). We have added to the figure note which trials were used to produce it.



4. Figure 18

Similar to my comments for Figure 17, if it isn't too messy, it would be nice to add error bars, and describe what data was used in the figure description

Response: We have also added to this figure note which trials were used to produce it. Both figures show the behavioral data for the distractor conditions (which are the ones for which ERPs are analyzed).

5. Discussion

At the beginning of the discussion, you say that the colour manipulation is more influential because more studies use colour stimuli compared to shapes. Could you just provide some evidence to back this up?

Response: The relevant sentences were: “This non-replicated result is arguably the more influential of the two, because far more N2pc studies use color patches than line patterns (W and M) as stimuli.” (Discussion) and “Our results robustly replicate the N2pc to form stimuli, but a direct replication of the more influential N2pc to color stimuli technically failed.” (Abstract). There are indeed many highly influential studies reporting N2pc to a colored object, such as Woodman & Luck (1999); Hickey et al. (2006, 2009); Kiss et al. (2012); Burra & Kerzel (2013); Jolicoeur et al. (2005), whereas we are not aware of any other study using Ms and Ws. However other shapes (e.g. diamonds and circles) are often employed in N2pc studies, including some of these influential studies. To us, these shapes are very different from the Ms and Ws used by Eimer (1996), but it is also difficult to argue that the colors employed in these studies are in some way more similar to those of Eimer (1996). Thus, this claim is somewhat subjective and hard to back up by the literature. We consequently decided to remove it.

Reviewer 2 (Clayton Hickey)

The paper does exactly what was described in the pre-registration, plus a lot of additional exploratory analysis that is very interesting. The results are not trivial and provide new insight on a seminal paper in cognitive neuroscience.

I have some comments; all are minor.

P2 – ‘Furthermore, the N2pc might reflect engagement at the location of the relevant stimulus rather…’ This is a bit hard to distinguish from the preceding hypotheses and needs some clarification. That is, an account of the N2pc that posits target enhancement is an account of the N2pc as engagement at the location of the relevant stimulus, not the shift of attention proper. Similarly, an account that posits distractor suppression is an account as engagement, not the shift of attention.

Response: We agree that this was non-ideally put. An account positing a mixture of engagement and suppression implies the presence of engagement, of course. To address this point, we restructured our discussion of alternative interpretations. Now the engagement + suppression account comes after the engagement account, because it goes beyond it (pp. 8-9).

The N2pc is commonly used and discussed as an ‘index of the deployment of attention’ or of ‘attentional shifting’. This is not incorrect, but perhaps misleading… the N2pc is (mostly) an indirect index of these constructs. It reflects the large-scale change in striate and extrastriate cortical activity that is the consequence of attention. This was a salient observation in the early papers from Luck, Eimer and others that has become somewhat lost. The first proposal of the N2pc as a reflection of a control mechanism emerged in MEG work from Max Hopf’s group, where the involvement of posterior parietal cortex in generating the mN2pc was tentatively interpreted as reflecting the mechanistic implementation of attention, rather than its consequence on sensory activity.

Response: The first occurance of this claim in our paper is: "The N2pc is most often used as a marker of shifts of attention." (p. 7)
To account for the possibility that readers misunderstand what "a marker of" means, we specified: "The N2pc is most often used as a marker of shifts of attention, which can be valid even if it reflects some process that is a consequence of an attention allocation rather than the allocation proper" (p. 7). We refrained from adding this explanation to the abstract for brevity.

p. 35 – '…the belief in fixed component timing is still widely held.' Needs some sort of support for the authors' impression / intuition / belief. I don't have this same impression… people have been looking at effects on component latency in visual search, and on N2pc latency specifically. There are papers cited in the MS that demonstrate that, like Brisson et al 2007. Other 'early' latency effect papers looking at search or N2pc that come to mind:

Smulders, Kok, Kenemans, Bashore (1995, Acta)
Wijers, Mulder & Mulder (1997, Psychopys)
Robitaille & Jolicoer, (2006, Neuroreport)
Kiesel, Miller, Jolicoeur, & Brisson (2008, Psychophys)
Hickey, van Zoest, & Theeuwes (2010, EBR)

More broadly, there were reports of P3 latency effects from Donchin and Kok in the '80s, probably the '70s too. I remember very early work from Falkenstein in the '90s with latency effects on Ne/ERN, either in the initial report or soon after. It could be more justifiable to say that the shift in N2pc timing observed here is larger than was expected across by the large group of authors contributing to this paper.

Response: This is an interesting point. You are right that variation in latency has been observed early on. This makes it even more astonishing that many labs use a fixed time window and that using a fixed analysis window (either predetermined or based on a collapsed localizer) is also commonly advised (e.g., Kappenman & Luck, 2016; Luck & Gaspelin, 2017). Arguably, this advice from influential ERP researchers is very strong evidence that the fixed-timing belief is commonly held. We added some discussion of this discrepancy accordingly (pp. 76-79). Even prior to this project, we were aware that the N2pc can occur in very different time windows (see Liesefeld et al., 2017, 2022).

Spelling etc:

p. 8 – 'This copy was high-pass filtered at 2 Hz (passband edge)…' for completeness, it would be good to have the same level of detail as for other filter descriptions.

Response: Thank you for noticing that this was missing. We now describe the filter parameters.

p. 9 – '…that is, ERP amplitude at electrodes P07/9 are lower…' to '…are more negative…'

Response: Indeed, this is more correct, we have changed accordingly.

p. 36 – The brightness of monitors has broadly increased over the last 40 years, and this could have something to do with the difference between the original report and these replications. CRT monitors were pretty dim by default, and and if you increased brightness the monitors degraded and became dimmer over time. If it was an early LED panel – unlikely as no one used them for vision research – then it's even more probable, as the original fluorescent backlights were awful.

Response: This is an interesting speculation and we have included it accordingly: "A systematic difference between original and replication studies might be that screens were generally dimmer back in the days when the original Eimer study was conducted." - Footnote: "We thank Clayton Hickey for pointing this out to us."