The Missing N1 or Jittered P2: Electrophysiological Correlates of Pattern-Glare in the Time and Frequency Domain

Austyn J. Tempesta¹, Claire E. Miller¹, Vladimir Litvak², Howard Bowman¹,³, & Andrew J. Schofield⁴,¹

1. School of Psychology, College of Life Sciences, University of Birmingham, Edgbaston, Birmingham, UK, B15 2TT.
2. Wellcome Centre for Human Neuroimaging, University College London, 12, Queen Square, London, UK, WC1N 3AR
3. School of Computing, University of Kent, Canterbury, Kent, UK, CT2 7NF.
4. School of Psychology, College of Health and Life Sciences, Aston University, Birmingham, UK, B4 7ET.

Supplementary Material

S1 | Factor analysis

Factor analysis seeks to derive independent factors from a collection of variables. Thus, the correlations between our factors were zero. Rotation helps to clarify which variables load onto (belong to) each factor, as seen in the Rotated Component Matrix (Table S1) and helps with factor naming. Our factors scores were calculated in SPSS (IBM, NY) using the regression method in which all variables contribute to all factors according to the Component Score Coefficient Matrix (Table S2). For the headache and discomfort factors, the contribution from variables not obviously associated with these terms was minimal. The headache variables did contribute to the visual stress factor, but were pitted against each other such that intensity and frequency contributed positively onto this factor while duration contributed negatively. Since the headache variables correlate relatively strongly, the effect of these positive and negative weightings will be to cancel out the overall effect of headache on the visual stress factor. Although correlations between factors were zero by design, the correlation between discomfort ratings for the medium stimuli alone and headache frequency was 0.3 (p = .068) which while not significant in our sample is similar to previous measures.

Table S1

| Component | Rotated Component Matrix |
|-----------|--------------------------|
|           | 1           | 2           | 3           |
| VSQ       | 0.636       | 0.035       | 0.062       |
| CHi       | 0.845       | 0.212       | -0.074      |
| aura      | 0.789       | 0.122       | 0.292       |
| H-duration| -0.470      | 0.764       | 0.040       |
| H-intensity| 0.366     | 0.716       | -0.116      |
| H-frequency| 0.305      | 0.795       | 0.041       |
| Discomfort| 0.119       | -0.028      | 0.977       |
Table S2

| Component | 1   | 2   | 3   |
|-----------|-----|-----|-----|
| VSQ       | 0.303 | -0.055 | -0.021 |
| CHi       | 0.401 | 0.018 | -0.172 |
| aura      | 0.337 | -0.012 | 0.187 |
| H-duration | -0.332 | 0.51 | 0.135 |
| H-intensity | 0.108 | 0.372 | -0.128 |
| H-frequency | 0.048 | 0.433 | 0.036 |
| Discomfort | -0.065 | 0.013 | 0.934 |

S2 | Direct Comparison of Stimulus Responses

In EEG work, one often attempts to avoid differences in stimulus properties, particularly those that are low-level, such as spatial frequency. This is because one is typically seeking to identify differences in higher-level properties, such as, attentional state, affective salience or linguistic properties, and associated EEG features could be contaminated by differences in low-level features. However, this is not our situation; we are specifically seeking to observe differences in the brain’s electrical response to changes in low-level stimulus features, i.e. spatial frequency.

Other researchers also looking at the brain’s electrical response to aggravating stimuli, have sought to avoid making comparisons across stimuli with different spatial frequencies (Fong, et al., 2020). A consequence of this choice is that effects of different spatial frequencies have to be judged informally, without a statistical test to quantify the confidence in an observed difference; effectively one is left considering the difference of evidences (e.g. comparing p-values), rather than assessing the evidence for a difference. There are, of course, many ways to investigate scientific questions, but since the question of interest for us is how early brain responses change to stimuli that have different effective strengths for different groups, we have chosen to directly compare the EEG generated by stimuli of different spatial frequencies.

Additionally, our parametric regression onto the three factors we identified cannot be impacted by fixed baseline differences between gratings, such as Thick stripes generating an overall higher amplitude than Thin or Medium, since that difference would be consistent across participants. As a result, it would not create a differential pattern “down” the regressor, i.e. it could not generate a non-zero correlation between dependent variable and regressor. Thus, fixed baseline differences in brain responses to different gratings cannot impact our analysis of factors.

S3 | Further Justification of Window Selection

The second ROI that we analysed focussed on the time region of the N1 component, which follows the P1. As discussed in the method the spatial and temporal parameters of this component are less well defined in the literature. Consequently, we applied an orthogonal contrast approach to identify the ROI (Brooks et al., 2017). We used our analysis of the factor intercept in the MUA to produce an ROI mask for further analysis. This approach will not inflate the false-positive rate for the following reasons.
All our three factors are, by construction, orthogonal to the intercept. This can be seen from the fact that they are all de-meaned and consequently, their dot-product with the column vector of all ones (i.e. the intercept regressor), which corresponds to taking the sum of the factor, equals zero. Additionally, the contrast vector for the one-sample t-test (the statistic of interest for us) for each regressor in our design (of which there are four: intercept and three factors) has a zero dot-product with any other. This is because each is a unit vector on one dimension, e.g. [0, 1, 0, 0]. Accordingly, the contrasts being performed are also orthogonal.

There are two further properties that support the parametric contrast orthogonality property, which is the gold-standard demonstration that the false positive rate is not inflated (Bowman, et al., 2020). The first of these is that there are no correlations running down the dependent variable, which holds because every participant is an independent sample. Regarding the last property, which considers trial count asymmetry across conditions, Brooks et al. (2017) showed that such asymmetries have a negligible effect unless they are great. For example, Figure 3, Panel C of Brooks et al. (2017) considers a t-test on an N170 component overlaid with noise at the human frequency spectrum. The N170 has a similar timeframe to that which we are considering in this paper. Essentially, there is very little evidence of an inflation of false positive rate when one condition has twice as many trials as the other, and indeed little until there is an 8 times asymmetry.

Since our main inference is parametric regression, the point of interest for us is whether differences in trial counts between participants could bias the mean/intercept ROI selection. Importantly, in our data, there is trial-count variability between participants, but that variability is not great. For example, the mean trial-count is 288.94 (considering onsets 2–8), while the standard deviation is 37.41. This standard deviation is very small relative to the mean, with a ratio of 0.13 (37.41/288.94). Additionally, the mean count is 1.4 times the lowest trial count and the highest trial count is 1.25 times the mean count. These levels of trial-count asymmetries between participants are very unlikely to bias ROI selection on the mean/intercept.

Note, at an earlier stage in our processing pipeline, there was more trial-count variability. That is, before we removed crown electrodes trial-count asymmetry was a greater issue. This is because presence of the noisy crown electrodes, lead to a higher rate of trial rejection due to artefacts. Indeed, this is why we performed a weighted average when calculating the overall window of analysis in time by identifying the portion of the grand average ERP waveform that deviated from baseline.

S4 | Justification of Prior Precedent for P1 Window

We consulted the literature when choosing our P1 window for the ROI analysis. The P1-contra and P1-ipsi components are associated with electrodes O1 and O2 on the 10/20 cap layout respectively (see Figure 4 of Hillyard et al., 1998). The posterior P1 is described as having an onset in the interval 80–100ms (Hillyard et al., 1998), with a peak between 100-130 ms (Luck, 2014) or 90-110ms (Slotnick, et al., 1999). Finally, Mangun (1995) described the P1 component as a positive going component that typically begins around 70–90ms with a peak around 80-130ms, with maximum amplitude over the peristriate cortex (Mangun et al., 1993).

We converted electrode locations for the 10/20 system (Kungl, Bovenschen, & Spangler 2017) to our 128-electrode Biosemi cap. Electrodes O1 on the 10/20 cap is associated with electrodes B6 and B7 on the biosemi cap O2 with electrodes A9 and A10. These electrodes have the following (x,y) co-ordinates in our SPM images noting that these are slightly skewed relative to the midline due to electrode removals: B6= 34, -94.625 mm,
B7= 34, -73.125 mm, A9= -38.25, -89.25 mm, A10= -38.25, -94.625 mm. Oz lies in the rectangle created by these electrodes; that is Oz= -4.25, -94.625 mm. We therefore placed a box centred at the midpoint of these electrodes: -2.125 mm, -83.875 mm. The bounding box just reaching the above electrode locations on the SPM MIP would measure 72.25 mm x 21.5 mm. However a larger box is required to fully encompass these electrodes so, based on the electrode spacing, we extended it by 10mm in all directions to give a final box size of 92 mm x 42 mm. Finally, constraints in SPM lead us to centre our ROI on location 0, -84 mm.

For the time domain, we wanted to have a window big enough to include the entire P1 effect. We decided to use a window as close to 70-130 ms as possible. We did this in order to encompass the time window precedents from all the literature (both onset and peak), which were described above. The dimensions of the ROI box go in both directions from the box location (which acts as the midpoint of the box). We had chosen to centre our time window at 100 ms, choosing 61 ms as the length of the time window (30.5 ms in both directions). However, because the closest time point SPM could give us was 101 ms, we had to adjust the window to fit the region of analysis. We ended up adjusting the time window to 62 ms, which would fully cover the 70-132 ms needed to investigate the P1.

S5 | Unweighted ERPs

In the results, we displayed the discomfort and headache factor ERPs scaled by factor weights to better visualise the parametric regressor inferences in the MUA. Figure S1, provides unweighted median split ERPs on factors (Discomfort and Headache) for significant electrodes (A8, A20, and A28).
Figure S1. Unweighted median splits for ERPs from statistically significant electrodes for discomfort (A20, left column, & A8, middle column, for Onset 1) and headache (A29, right column, Onsets 2-8). Positive plotted up. Lines indicate start and end of statistically significant time window (See our analysis section for specific time windows).
Throughout our analyses, we drew a logical distinction between onset 1 (the first appearance of the stimulus in each trial) and the remaining onsets (subsequent repeats of the same stimulus), reasoning that the first onset has an element of uncertainty not present in subsequent repeats, whereas the remaining onsets better capture the build-up of effects (such as habituation) over repeated presentation. However, the grouping of onsets 2-8 and the comparison, if only informal, of these data with those from onset 1 alone is statistically unequal, as many more epochs contribute to onsets 2-8 than onset 1. In particular, it is possible that effects seen in onsets 2-8 are present in onset 1 but are too weak to produce statistically significant peaks or clusters in the SPM analysis. To address this issue, we plot (Fig S2) scalp maps for each onset individually estimated at the times of the statistically significant effects for each factor.

For discomfort (Fig. S2a), there is clear activity within the ROI at 97 ms for onset 1, which is absent in the other onsets except for a weak effect, which was not statistically significant, in onset 7.

For the headache factor (Fig S2b) there is clear activity in the ROI at 173 ms in onsets 3-8. Recall that the statistically significant effect for onsets 2-8 as a group was found in a small cluster of voxels located towards the occipital pole close to electrode A29 peaking at this time. This effect is seen most clearly in onsets 5-8 although the activity in onsets 3 & 4 may be a precursor to it. Thus, we see a build-up of activity – that is a difference between those high and low on the factor – as the number of onsets increases; consistent with a habituation-like effect. There is a small amount of positive activity in the posterior region for onset 1 but it is very weak and does not extend to the location of the cluster found for onsets 2-8.
Figure S2. (a) Scalp maps for onsets 1 – 8 (one per onset, indicated by numbers in top left of each sub-panel) for the discomfort factor at 97ms. (b) Scalp maps of onsets 1-8 (one per onset, numbers as (a)) for the headache factor at 173ms. Colours indicate t-values.

S7 | Comparisons with Fong, Law, Braithwaite & Mazaheri (2020)

Fong et al. (2020) used a very similar paradigm to ours, presenting the three pattern glare stimuli, at spatial frequencies very close to our Thick, Medium and Thin, to those with self-reported migraine. However, their stimulus timings were different from ours with a key difference being that they only presented single onsets in each trial, whereas we repeated our stimuli several times per trial.

In comparing our results, we first note a difference in nomenclature. Fong et al. (2020) observe a negative-going deflection just prior to 100 ms in response to high frequency stimuli and report this as N1, with the subsequent positive deflection termed P2 and a second negative deflection N2 at around 200ms. They then carry this labelling through to the ERP for the low frequency stimulus, which starts with a positive deflection around 100 ms, which
they label P2. P1 is absent in all ERPs, with N1 also absent in the ERP for low frequency gratings. However, a good deal of work on visual ERPs has found a positive going occipital peak at around 100 ms, which is called variously the P1 or P100 followed by a negative deflection between 150-200 ms post stimulus, which is typically called the N1 component; see, for example, Figure S1. Our ERP for Thick stimuli follows this typical P1-N1 pattern, as indeed does Fong et al.’s trace for their low-frequency stimulus. We thus prefer to label the positive peak at 100 ms as P1 and the subsequent deflections as N1 and P2 respectively. We note however that, in common with Fong et al., our Thin (high frequency) stimulus tended to produce a negative deflection at 100 ms as can be seen in our Figure 2. Thus, our results have some superficial similarities to those of Fong et al. with a transposition of nomenclature. For the avoidance of doubt, we will refer to stimulus features in terms of milliseconds post stimulus in the comparisons below.

Owing to our initial data-driven ROI, we do not consider ERP components outside the 56-256 ms window. We will therefore not comment on the reduced late negativity found by Fong et al (2020) at around 400 ms, although we do note that those high on our headache factor have more positive ERPs between 300-400 ms than those low on the same factor. Fong et al. also found significant between-group differences between migraineurs and controls at around 200 ms, with the patient group having more negative ERPs for high-frequency stimuli in this region. This was to some extent also visible, but not significant, in the traces for medium frequency stimuli. At first sight, this result is at odds with our finding of more positive traces at 200 ms for those high on our headache factor than those low on the same factor (see our Figure 5); this resulted in a significant relationship between the PGI and headache. However, we found this only for Onsets 2-8 and since Fong et al. did not use repeated stimulus presentations, our result could represent a habituation effect that they were unable to detect due to differences in methodology. Further, although we found no significant results in this region, an examination of our ERPs for the first onset of our medium stimuli, shows a greater negativity for those high on the headache factor at around 150 ms not unlike that found by Fong et al. (see Figure S3). We conclude that the apparent difference between the studies to a certain extent rests on stimulus presentation methods and our use of repeated onsets to address habituation effects.

Figure S3. Unweighted ERPs derived from a median split on headache parameters at Oz (electrode A23) for the first onset of each stimulus. Thick red line is the unweighted mean of those high on the headache factor, blue - low. Thin lines show individual ERPs for those in each quartile top-bottom: red, green, cyan, blue. Positive is up. The vertical dashed line in each plot shows the timing of the effect of interest.
Fong et al. (2020) also posit that those with migraine and those controls with strong pattern-glare have a “phantom” positive deflection at 200 ms for the medium stimulus. We notice a similar effect; our medium stimuli tend to produce greater positivity at 200 ms, which we describe as an early or jittered P2, but could be related to the proposed phantom P200 component.

Additionally, studies of visual stress and pattern-glare have typically associated both symptoms and increased neural activity with gratings at 3 c/deg; less so at higher frequencies. Indeed, PG symptoms at 11-13 c/deg are often used as a control for symptoms at 3 c/deg (Evans & Stevenson, 2008). Therefore, our findings of significant results on our PGI – which treats responses to the high-frequency grating as a control for medium – may be a better correlate of symptoms than responses to high-frequency gratings alone, upon which Fong et al. focus their findings.

**S8 | Trials with hidden stimuli**

Due to health and safety concerns, participants were given the option of turning off the stimulus during a trial. In such cases, the stimulus remained off until the end of the trial but was revealed again thereafter. Unfortunately, due to a programming error, these trials were not automatically marked for removal and we only know the number of times each participant used this option. Three participants hid the stimulus once each so only 3 trials were affected across the whole dataset. Reasoning that the impact of such a small number of epochs was likely to be small we retained these data. Any large artefacts due to eye or muscle movement will have been removed by our other processes so these stimulus free epochs most likely contained small random variations about baseline.

In order to verify that the epochs in question did not contain large fluctuations that drove our subsequent analyses, we recomputed weighted average ERPs and the PGI index at our electrodes of interest for the effects we found on the discomfort and headache factors leaving out all of the data for the three participants who chose to hide stimuli. These new traces are plotted alongside the original weighted ERPs in Figure S4. The revised traces (green and cyan) are very similar to the originals (red and blue) and show larger positive deviations in the PGI. We thus conclude that including the data from the affected epochs will have had little influence on our results (less than removing entire data sets as here) and if anything reduced the effects found.
Figure S4. Weighted ERPs and resultant PGI traces for the full data set (red and blue lines) and after removal of participants 29, 33, and 40 (green and cyan lines). Red and green lines represent participants in the high group for each factor, blue and cyan, low. The factor and electrodes of interest are shown at the top of each column.
S9 | References

Brooks, J. L., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. Psychophysiology, 54(1), 100-113.

Bowman, H., Brooks, J. L., Hajilou, O., Zoumpoulaki, A., & Litvak, V. (2020). Breaking the circularity in circular analyses: Simulations and formal treatment of the flattened average approach. PLoS computational biology, 16(11), e1008286.

Fong, C. Y., Law, W. H. C., Braithwaite, J., & Mazaheri, A. (2020). Differences in early and late pattern-onset visual-evoked potentials between self-reported migraineurs and controls. NeuroImage: Clinical, 25, 102122.

Hillyard S.A., Vogel E.K., and Luck S. J. (1998) Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. Phil. Trans. R. Soc. Lond. B353, 1257–1270.

Kungl, M. T., Bovenschen, I., & Spangler, G. (2017). Early Adverse Caregiving Experiences and Preschoolers' Current Attachment Affect Brain Responses during Facial Familiarity Processing: An ERP Study. Frontiers in Psychology, 8, 2047.

Luck, S. J. (2014). An Introduction to the Event-related Potential Technique. MIT press.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. Psychophysiology, 32(1), 4-18.

Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1993). Electrocortical substrates of visual selective attention. In D. Meyer & S. Kornblum (Eds.), Attention and Performance XIV (pp. 219-243). Cambridge, Massachusetts: MIT Press.

Slotnick, S.D., Klein, S.A., Carney, T., Sutter, E., Dastmalchi, S. (1998) Using multi-stimulus VEP source localization to obtain a retinotopic map of human primary visual cortex, Clinical Neurophysiology, 110, 1793-1800.