Result notes for “Temporal dynamics of auditory bistable perception correlated with a fluctuation of baseline pupil size”

Yuta Suzuki, Hsin-I Liao, Shigeto Furukawa

25 January, 2022

\*Corresponding author: Yuta Suzuki NTT Communication Science Laboratories, NTT Corporation, Atsugi 243-0198, Japan Tel: +81-046-240-3525, E-mail: [yuuta.suzuki.fc@hco.ntt.co.jp](mailto:yuuta.suzuki.fc@hco.ntt.co.jp)

# Results

## Experiment 1

### Behavioral response

Participants reported the number of perceptual alternations from 0 to 5. Since the trial numbers were unbalanced among these responses (**Fig. 1C**), we classified the number of alternation responses into 0, 1, and more than 2 times (hereafter, referred to as 0-, 1-, >1-alt cases, respectively). The average numbers of trials were 26.27 12.4, 35.59 10.18, and 14.68 9.83, respectively ((1.84,38.62) = 15.427, = 0, = 0.424, = 1.4780229^{5}). Post analysis showed that the number of trials for 1-alt cases was significantly larger than for 0- and >1-alt cases ((21) = 2.614, = 0.016, Cohen’s = 0.557, = 3.331; (21) = 6.143, = 0, Cohen’s = 1.31, = 4741.359) and that for 0-alt cases was larger than for >1-alt cases ((21) = 2.703, = 0.013, Cohen’s = 0.576, = 3.921) at alpha level of 0.05/3 corrected by a Bonferroni-Holm method. RTs of each response category were 1.24 0.501 sec, 1.129 0.401 sec, and 1.209 0.864 sec ((1.24,26.09) = 0.307, = 0.633, = 0.014, = 0.161). Although the observation period was jittered from 5-9 second, there was no statistical differences between the number of perceptual alternations and the observation time ((1.55,32.58) = 1.213, = 0.301, = 0.055, = 0.314).

### Baseline pupil size

**Figure 2A** illustrates the grand-averaged baseline pupil changes across participants before the response cue onset, as a function of perceptual alternations number. The one-way repeated measures ANOVA revealed a significant main effect on the number of perceptual alternations ((1.52,31.91) = 11.424, = 0.001, = 0.352, = 1304.695). The post-hoc multiple comparisons showed that the baseline pupil size in the >1-alt case was significantly larger than in the 0- and 1-alt cases ((21) = 3.66, = 0.001, Cohen’s = 0.78, = 25.712; (21) = 3.954, = 0.001, Cohen’s = 0.843, = 47.379, respectively), indicating that the baseline pupil size, prior to counting the number of alternations, was related to a subsequent number of perceptual transitions. The answered number of trials was significantly different among the 0-, 1-, and >1 cases as shown above. Such unbalanced trial numbers can cause biases in the statistical analysis and decreased statistical power. Thus, we performed an alternative analysis to avoid this potential statistical problem. We segregated the trials equally into five bins based on the rank order of baseline pupil size. Results are shown in Fig. 2B. The data were fitted by a simple regression model ( = 0.079 + 0.76, = 0.119, = 5.1447, = 0). Consistent with the previous results, the number of alternations monotonically increased with the baseline pupil size.

## Experiment 2

### Behavioral response

The average number of ‘yes’ and ‘no’ trials (presence and absence of perceptual alternations, respectively) were 36 13.66 and 44.89 15.3, respectively ((1,18) = -1.448, = 0.165, Cohen’s = 0.332). Again, there was no statistical differences between ‘yes’ or ‘no’ trial and the observation time ((1,18) = -0.157, = 0.877, Cohen’s = 0.036). RT in the answer ‘yes’ (1.236 0.44 s) was significantly faster than ‘no’ (1.236 0.44 s)  
((1,18) = 5.724, = 0, Cohen’s = 1.313). This could be because, the participant in the ‘yes’ trial would have been ready to respond as soon as the first occurrence of a perceptual alternation before the response cue was presented, whereas the participant had to wait until the cue to say ‘no’. It is important to note, therefore, that the perceptual load and/or mental effort is expected to be lower in the ‘yes’ trials than in the ‘no’ ones. In Experiment 1, the participant had to keep counting and memorizing the number of alternations throughout the observation period, possibly leading to increasing perceptual load in the trials with increasing number of alternations. Thus, there was a concern that the association between the pupil size and number of perceptual alternations observed in Experiment 1 reflected such a perceptual load, rather than the processes involved in perceptual switching. The current yes/no paradigm serves as a control test to evaluate the confounding effect of this task-related perceptual load.

### Baseline pupil size

Figure 3A shows the grand-averaged time-course of baseline pupil changes parameterized by alternation cases (yes or no). Consistent with Experiment 1, a paired t-test for averaged changes in baseline pupil size from -1000 ms to the response cue onset for each answer (i.e., the presence or absence of perceptual alternation) showed that the baseline pupil size in the presence of a perceptual alternation was significantly larger than in the absence of perceptual alternation ((18) = -2.508, = 0.022, Cohen’s = 0.575, = 2.73).

Following the same analysis procedure as in Experiment 1, we segregated the trials into five bins based on the ranked order of the normalized baseline pupil size. For each participant, we normalized the probability of perceptual alternation by z-scores and averaged them in each pupil size bin (Fig. 3B). The model fitted by a simple regression showed the significance ( = 0.04 + -0.12, = 0.228, = 2.4301, = 0.0167).

## Transient Pupil Dilation/Constriction (PD/PC)

To assess the relationship between perceptual alternations and transient pupil change reported previously (Einhäuser et al., 2008; Grenzebach et al., 2021; Turi et al., 2018), we calculated the rate of PD/PC events (see Methods). Figure 4A shows the occurrence of PD/PC events for each trial across all subjects, over a period of 2 s before the task response to 4 s after it. We averaged the number of PD events over a period of 4 s after the task response (Fig. 4B). To compare by the within-subject design, the participants who were not rejected in both Experiments 1 and 2 were examined in the following analysis. Two-way repeated measures ANOVAs on the averaged PD events with the response content and experiment as within-subject factors revealed that the average number of PD event was significantly larger in alternation trials than in no-alternation trials ((1,18) = 5.973, = 0.025, = 0.249, = 2.186), consistent with the previous studies (Einhäuser et al., 2008; Grenzebach et al., 2021). The number of PD events was larger in Experiment 1 than in 2 ((1,18) = 7.77, = 0.012, = 0.302, = 17.831), which could be explained by the higher task demand in Experiment 1, as the LC-NE system reflects a broad range of cognitive processes. There was no interaction between the response content and experiment ((1,18) = 0.226, = 0.64, = 0.012, = 0.377). Two-way repeated measures ANOVAs on the averaged PC showed that there were no significant main effect and interaction ((1,18) = 1.183, = 0.291, = 0.062, = 0.399); (1,18) = 2.372, = 0.141, = 0.116, = 0.892; (1,18) = 0.004, = 0.948, = 0, = 0.302, respectively). There was a concern that the PD events and observed baseline pupil size were not sufficiently independent measures. For example, if the transient pupil dilation had a sufficiently slow time decay, the apparent baseline diameter would build up with accumulated PD events. To address this concern that the baseline pupil size might be affected by the PD events, we calculated the number of PD events in the baseline pupil size analysis in every 1-s time bin (Supplementary Fig. 2) and the correlation between the baseline pupil size and number of transient PD events (Supplementary Fig. 3). Note that the number of PD events during the early 4-s part of the observation period were summed and their order ranked across trials for each subject. The baseline pupil size after the window for calculating the number of PD events was parameterized by the rank order of the PD events. The results revealed that there was no statistical evidence that increasing the number of PD events can explain the baseline pupil size in both experiments.

## Cross-correlation function

The cross-correlation analysis (Fig. 5) revealed that the baseline pupil size and the number of perceptual alternations had a significant positive correlation over the range of nine trials; -6 to +3 trials with a lag of around 0. This result can be interpreted as indicating that the baseline pupil size at least 35 s (= 5 s x 6 trials + 5 s; see below) before the behavioral response predicts perceptual alternation and the overall correspondence between pupil size and the number of perceptual alternations is sustained for at least 45 s (5 s x 9 trials). It should be recalled that for a given trial (or the lag of 0 in Fig. 5), the window for deriving the baseline pupil size preceded the corresponding response by the duration of the observation period (therefore, an additional 5 s was included to derive the 35-s period), and that the observation period was varied randomly between 5 and 9 s (Fig. 1D, red line). Thus, the 35-s and 45-s durations shown above are conservative estimates, with the assumption that the representative observation period was the minimum of the randomization range.

### Statistical analysis

Mean baseline pupil size was subjected to a one-way repeated-measures analysis of variance (ANOVA) with the number of alternations as the within-subject factor in Experiment 1. A paired t-tests was applied to the mean baseline pupil size between the presence (alternation) and absence (no- alternation) of perceptual alternations in Experiment 2. The level of statistical significance was set to p < 0.05 for all analyses. Pairwise comparisons of the main effects were corrected through multiple comparisons using the Bonferroni-Holm method. Effect sizes were given as partial ; for ANOVA and as Cohen’s d\_z for t-tests **(Cohen, 1988, p.48; Lakens, 2013)**. Greenhouse–Geisser corrections were performed when the results of Mauchly’s sphericity test were significant. To quantify the evidence in the data, we performed Bayesian paired t-tests with a Cauchy prior width of r = 0.707 for effect size using the BayesFactor package (v0.9.12-4.2) **(Morey, 2019)** for the R software (Version 3.6.3) **(R Core Team, 2020)**. Bayesian Factor (BF) estimates the relative weight of the evidence in favor of over as

For baseline pupil size bin analysis, for each participant, all trials were sorted in ascending order based on the baseline pupil size and divided into five equally populated bins. This procedure led to an equal trial number in each pupil size bin. The number of perceptual alternations (Experiment 1) and the z-scored probability of perceptual alternation (Experiment 2) was averaged and fitted by the following two models to assess whether the behavioral variability (Y) can be explained by a monotonic fitting or second-order polynomials.

Model 1 :

Model 2 :

where represents regression coefficients, and represents the baseline pupil response bins. The models were quantified using the Akaike information criterion (AIC), which specifies the evidence of goodness of fit for a model.