Result notes

Yuta Suzuki

3/2/2021

# 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 27.05 13.48, 36.27 10.64, and 15.14 9.65, respectively ((1.9,39.87) = 13.968, = 0, = 0.399, = 5.9323169^{4}). Post analysis showed that the number of trials for 1-alt cases was significantly larger than for 0- and >1-alt cases ( (1, 21) = 2.708, = 0.013, (1, 21) = 5.885, = 0). RTs of each response category were 1.241 0.502 sec, 1.126 0.399 sec, and 1.22 0.855 sec ((1.25,26.21) = 0.359, = 0.602, = 0.017, = 0.165).

### 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 ((1.65,34.68) = 9.79, = 0.001, = 0.318, = 79.727). 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 ((1, 21) = 3.583, = 0.005, Cohen’s = 0.233, = 21.968; (1, 21) = 3.275, = 0.005, Cohen’s = 0.198, = 11.768, 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, = 14.9673, = 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 36.85 13.86 and 45.1 14.99, respectively ((1,19) = -1.379, = 0.184, Cohen’s = 0.572). RT in the answer ‘yes’ (1.22 0.435 s) was significantly faster than ‘no’ (1.22 0.435 s) ((1,19) = 4.608, = 0, Cohen’s = 0.367). 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 ((1, 19) = -2.195, = 0.041, Cohen’s = 0.087, = 1.631).

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 ( = 0.039 + -0.118, = 0.224, = -2.1495, = 0.0338).

## 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 ((1,19) = 0.113, = 0.741, = 0.006, = 0.247), 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 ((1,19) = 0, = 0.998, = 0, = 0.23), 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 ((1,19) = 0.119, = 0.734, = 0.006, = 0.333). Two-way repeated measures ANOVAs on the averaged PC showed that there were no significant main effect and interaction ((1,19) = 1.238, = 0.28, = 0.061, = 0.337); (1,19) = 0.685, = 0.418, = 0.035, = 0.329; (1,19) = 7.128, = 0.015, = 0.273, = 12.547, respectively).

## Cross-correlation function

The baseline pupil size up to six trials ahead of the behavioral response and after the response until three trials was positively correlated to the number of perceptual alternations (Fig. 5A). As the intertrial interval was jittered in 5–7 s and the baseline pupil size was inherently defined as 5–7 s before the response in the 0-lag condition, the result indicates that the baseline pupil size predicted perceptual alternation at least 35 s (= 5 s x 6 trials + 5 s) before the behavioral response and that the overall correspondence between pupil size and perceptual alternation was over a sustained time window of 45 s (5 s x 9 trials) at minimum and of 54 s (6 s x 9 trials) at the averaged interval.

### Statistical analysis

A one-way repeated-measures analysis of variance (ANOVA) was performed using the baseline pupil size and the number of switches in Experiment 1 and presence vs. absence of perceptual switch in Experiment 2 as within-subject factors. The level of statistical significance was set to < 0.05 for all analyses. Pairwise comparisons of the main effects were corrected through multiple comparisons using the modified Bonferroni method. Effect sizes were given as partial ; for ANOVA and as Cohen’s for -tests **(Cohen, 1998, p. 48)**.

To quantify the evidence in the data, we performed Bayesian one-sample t-tests using the BayesFactor package (v0.9.12-4.2) **(Morey, 2019)** for the R software (Version 3.6.3) **(R Core Team, 2020)**. We reported Bayesian Factor (BF) estimating the relative weight of the evidence in favor of over as . Greenhouse–Geisser corrections were performed when the results of Mauchly’s sphericity test were significant.

In the analysis of pupil response bins, we fitted the following two models to assess whether the behavioral variability () can be explained by a second-order polynomials or monotonic fitting.

Model 1 :

Model 2 :

where as regression coefficients and as 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.