Result notes

Yuta Suzuki

3/2/2021

# Results

We measured pupil size using an infrared camera while participants listened to the consecutive auditory tone sequences. As illustrated in **Figure 1A**, the auditory stream patterns elicit alternating percepts of a single galloping or two isolated rhythms. Participants were engaged in counting the number of perceptual switches in Experiment 1 and answering whether they experienced a perceptual switch in Experiment 2 at a certain period (see Method).

## Experiment 1

### Behavioral response

Participants reported the number of perceptual switches every few seconds (see Methods). We classified the number of switches into 0, 1, and more than 2 times(2+); the averaged number of trials were 26.41 12.5, 35.91 10.63, and 14.95 9.7, respectively ((1,2) = 15.041, = 0, = 0.417). Post analysis showed that the number of trials for 1 time switch was significantly larger than 0 and 2+ ( (1, 21) = 2.683, = 0.014, (1, 21) = 5.969, = 0). RT of each response were 1.242 0.502 sec, 1.134 0.423 sec, and 1.211 0.8 sec ((1,2) = 0.335, = 0.717, = 0.016).

### Time course of baseline pupil change

**Figure 2A** illustrates the grand-averaged baseline pupil changes across participants before the task queue onset, as a function of perceptual switches number. A one-way repeated measures ANOVA for averaged changes in baseline pupil size from -1000 ms to the task queue onset for each answer revealed a significant main effect on the number of a perceptual switch ((1,2) = 11.807, = 0, = 0.36) (**Figure 2B**). Following multiple comparisons showed that the baseline pupil size in more than two times of perceptual switch was significantly larger than in less than one-time switch and no-switch ((1, 21) = 4.077, = 0.002, Cohen’s = 1.299, = 27.874; (1, 21) = 3.699, = 0.002, Cohen’s = 0.999, = 61.435, respectively)), indicating that the baseline pupil size, to prior counting the number of switches, was related to a subsequent number of perceptual transition.

### The averaged number of perceptual switches classified by pupil size bins

Since the answered number of trials was significantly different among 0, 1, and 2+ as shown above, we averaged the number of perceptual switches based on baseline pupil size bins which contained an equal number of trials in each participant so as to avoid the influence of the number of trials as shown in **Figure 2B**. The data was fitted by a simple regression model ( = 0.079 + 0.76, = 0.119, = 14.9673, = 0). Consistent with the previous analysis, the number of switches was monotonically increased with the baseline pupil size.

### Correlation heatmap

As a further analysis, we calculated the correlation between the number of switch and several eye-metrics (**Figure 2C**); besides the significant correlation with baseline pupil size ( = 0.501, = 0), besides the significant positive correlation between the number of switch and baseline pupil size ( = , = ). Since the pupil change is mediated by the dilater/sphincter muscle in the iris, this might be interpreted as a trade-off mechanism of the iris musculature or activation of ANS **(Joshi et al., 2016)**.

### Cross-correlation function

Each trial lasted for 5-7 s, segregated by a response cue to ask for behavioral responses. In the cross-correlation analysis, we aligned the timing of behavioral response and pupil size throughout the experimental session (see Method). Cross-correlation functions between the changes in baseline pupil size and the number of perceptual switches were calculated in every single subject. We found that the number of switches was positively correlated with the pupil size up to 5 trials (nearly 25 to 35 s) ahead of the response time (Figure 3A). The maximum cross-correlation coefficient in raw cross-correlation was significantly higher than in shuffled (**Figure 3A**). The maximum cross-correlation coefficient in raw cross-correlation was significantly higher than in shuffled ((1, 21) = 12.097, = 0, Cohen’s = 2.763, = 1.4643872^{8}). Furthermore, the averaged lag at peak correlation was significantly shifted to negative ((1, 21) = -0.493, = 0.314, = 0.249) (**Figure 3C**).

### Liner mixed modeling and ROC curve

In this section, we focused on (a) how long the baseline pupil size effect on the number of perceptual switches at a certain trial lasts, (b) variability within participants, (c) classification accuracy. To address these, we calculated beta weights using a linear mixed-model (LMM) (Figure 3D) regarded the number of perceptual switches as dependent variable, the baseline pupil size as fixed effect and participants as random effect. For each trial, LMM was performed repeatedly changing the baseline pupil size data adopted from -20 to 20 trials corresponding to the answer at a certain trial. In consistent with the cross-correlation analysis, the LMM revealed a significant positive correlation between the number of switch and the baseline pupil size with not larger individual differences at least from -7 to 0 trial (i.e., 30s). Furthermore, the detection performance was quantified by the area under the curve (AUC) using receiver operating characteristic (ROC) curves (Figure 3E). Wilcoxon rank-sum test showed that the AUC from -5 to 5 trial were significantly greater than chance level (i.e., 0.5).

## Experiment 2

In Experiment 1, we reported that the baseline pupil size reflected the frequency of perceptual switch. However, one concern to the result was the effect of perceptual load on baseline pupil size such as counting the number of perceptual switches since baseline pupil size is associated with the performance of working memory task **(Aminihajibashi et al., 2020; Schneider et al., 2016)**. Experiment 2 was conducted to assess the effect of perceptual load on baseline pupil size in order to change the task into a forced-choice task (i.e., yes or no) so that the effect was assumed to be similar or less than Experiment 1.

### Behavioral response

Participants reported whether perceptual switches occurred or not. The averaged number of trials of ‘yes’ and ‘no’ were 35.4 14.4 and 43.8 16.06, respectively ((1,19) = -1.415, = 0.173, Cohen’s = 0.551). RT in the answer of ‘yes’ (1.222 0.433 sec) was significantly faster than ‘no’ (1.065 0.415 sec) ((1,19) = 4.786, = 0, Cohen’s = 0.369), most likely because the participants’ answer was determined in the moment of occurrence of a perceptual switch in ‘yes’ trial. Thus, perceptual load in terms of task resource allocation was expected to be smaller in the trial where a perceptual switch occurs.

### Baseline pupil size

**Figures 4A** and **B** show the grand-averaged time-course of baseline pupil changes and averaged baseline pupil size across participants around the task queue onset. Consistent with Experiment 1, a paired t-test for averaged changes in baseline pupil size from -1000 ms to the task queue onset for each answer (i.e., the presence or absence of perceptual switch) showed that the baseline pupil size in the presence of a perceptual switch was significantly larger than in the absence of perceptual switch ((1, 19) = -2.428, = 0.025, Cohen’s = 0.698, = 2.392).

Following the normalization of the probability of perceptual switch by z-scored for each participant, we averaged the probability of perceptual switch based on baseline pupil size bins as performed in Experiment 1 (**Figure 4C**). The model fitted by a simple regression showed the significance in line with Experiment 1 ( = 0.039 + -0.118, = 0.224, = -2.1495, = 0.0338).

### PD events

To assess the relationship with transient pupil change as reported previously **(Einhäuser et al., 2008; Turi et al., 2018)**, we calculated the rate of dilation/constriction events (see Methods). **Figure 5A** shows the occurrence of dilation/constriction events for each trial across all subjects, over 2 sec before the task response to 5 sec after the response. We averaged the number of dilation events after the task response (**Figure 5B**). Two-way repeated measures ANOVAs for averaged dilation events for each answer and each experiment revealed that the average number of dilation events was significantly larger in switch trials than unswitch trials, consistent with previous study ((1,19) = 7.01, = 0.016, = 0.269). The number of dilation events were marginaly larger in Experiment 1 than 2 ((1,19) = 3.36, = 0.083, = 0.15). Numerous studies showed that the pupil change related to the LC-NE system reflects a broad range of cognitive processes including perceputual load. Thus, the differences between exeriments might be explainiend by the higer perceptual load in the Experiment 1 due to the task difference. There was no interaction between the answer and experiment ((1,19) = 0.238, = 0.631, = 0.012).

### PC events

Main effect on alternation Main effect on experiments Interaction ((1,19) = 1.053, = 0.318, = 0.053; (1,19) = 0.824, = 0.375, = 0.042; (1,19) = 0.024, = 0.88, = 0.001).

### 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 d for -tests. 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.

In Experiment 1, the value of Pearson’s correlation was calculated at a lag between the pupil size and the normalized number of switches using -score in cross-correlation. values for Pearson’s correlation were corrected by FDR for multiple comparisons using the Benjamini and Hochberg method **(Benjamini and Hochberg, 1995)**.

We used a linear mixed-effects modeling (LMM) with participant as a random effect to calculate the correlation between the individuals’ baseline pupil size () and the number of perceptual switch () in Experiment 1 as follwing model, using the lme4 packages **(Bates et al., 2015)**.

where as beta weight and as indivisuals’ total number of trials. The analyzed period was

### Pupil dilation/constriction events analysis

Pupil slope was computed using second order central differences after applying 1 s window of smoothing filter within a trial. Pupil dilation and constriction events were defined as maximum positive and negative values of the slope, separeted by 300 ms (**Joshi et al., 2016; Zhao et al., 2019**). The analysis interval was from 2 s before to 5 s after the task response. For each trial, the number of dilation events were summed and then all averaged between switch and unswitch condition in each participant.