

Cueing the Necker cube: Pupil dilation reflects the viewing-from-above constraint in bistable perception

Fumiaki Sato

Department of Computer Science and Engineering,
Toyohashi University of Technology, Toyohashi, Aichi,
Japan



Bruno Laeng

Department of Psychology, University of Oslo, Oslo,
Norway
RITMO Centre for Interdisciplinary Studies in Rhythm,
Time, and Motion, University of Oslo, Oslo, Norway



Shigeki Nakauchi

Department of Computer Science and Engineering,
Toyohashi University of Technology, Toyohashi, Aichi,
Japan



Tetsuto Minami

Department of Computer Science and Engineering,
Toyohashi University of Technology, Toyohashi, Aichi,
Japan



We hypothesized that a perceptually ambiguous or bistable object (Necker cube) can be more effectively biased to assume a point of view-from-above (VFA) than from below the object by cueing attention. Participants viewed a Necker cube in which one surface was temporarily shaded so as to prime a specific perspective on the cube. Subsequently, the standard (wireframe) Necker cube was viewed for 3 seconds, and participants reported what perspective they had seen initially and whether their perception shifted to the alternative perspective during the brief viewing. Concomitantly, pupil size was monitored with an eye-tracker to obtain an index of cognitive effort. There were two conditions: passive viewing and forced attention to sustain the initially primed perspective. We confirmed the presence of a VFA bias with forced attention, which was accompanied by reduced attentional effort, as indexed by a reduced pupil diameter, compared with the view-from-below. Participants showed no bias during passive viewing. We suggest that the level of intensive attention, when retrieving and maintaining a specific view from memory, is mirrored in the size of the eye pupils and may reflect ecological constraints on visual perception.

Introduction

Our visual system receives two-dimensional information from each image on the retina. From these two-dimensional images, it restores to vision the three-dimensional shape of the object, sometimes resolving ambiguities in the visual input in an automatic and unconscious manner. As a way of resolving perceptual ambiguity, Nakayama and Shimojo (1992) proposed a theoretical framework in which the visual system tends to interpret images and surfaces as seen from a generic point, instead of an accidental vantage point. In other words, the object properties of ambiguous images are estimated based on constraints derived from past visual experiences.

Several visual generic principles have been proposed for disambiguating the perception of objects (e.g., the light-from-above constraint; Ramachandran, 1988). The view-from-above (VFA) constraint or heuristic is a bias in vision so that whenever the input information is ambiguous and consistent with different viewpoints on a same object (Troje & McAdam, 2010), the VFA is preferred in perception over the alternative view-from-below.

Kornmeier and colleagues suggested that observers prefer the VFA due to an asymmetry of perceptual experiences or statistical learning, according to which

Citation: Sato, F., Laeng, B., Nakauchi, S., & Minami, T. (2020). Cueing the Necker cube: Pupil dilation reflects the viewing-from-above constraint in bistable perception. *Journal of Vision*, 20(4):7, 1–14, <https://doi.org/10.1167/jov.20.4.7>.

we more often look down on objects (e.g., artifacts) than look up at them (Kornmeier, Hein, & Bach, 2009). Remarkably, Sundaeswara and Schrater (2008) showed that the VFA preference of Necker cubes is close to 100% for very short presentations and declines for long exposures, but it remains the most likely interpretation. As these authors pointed out, in Bayesian models of perception, the visual system tends to choose the optimal interpretation among alternatives, typically the viewpoint with maximal posterior probability. Hence, a Bayesian account suggests that VFA may be the most frequent experience with many types of common objects, which are usually manipulated at hand level, and therefore examined from a top view. Supposedly, the Necker cube may be spontaneously interpreted as a representation of some kind of smaller-than-the-body box that could be “afforded” by the hands.

Toppino (2003) suggested that the top-down activation or priming of perceptual representations affects a specific interpretation of the cube in the Necker cube. Some studies showed that one can control to some extent the perception of an ambiguous figure by focusing on specific portions of the figure, or by intentionally selecting appropriate focal features by using the Necker cube (Kawabata, Yamagami, & Noaki, 1978) or the so-called “my husband and my father-in-law” drawings (Kawabata & Mori, 1992). Wernery, Atmanspacher, and Kornmeier (2015) showed that the dwell time (i.e., periods of transiently stable percepts), while passively observing the Necker cube for 3 minutes, was longer for the VFA than the one from below. Importantly, Meng and Tong (2004) showed that selective attention can bias reversals of the perception of the Necker cube.

Despite many studies on the VFA bias, most of these are psychophysical experiments in which the introspective reports of participants are the dependent variable. We reason that evidence for a VFA bias can also be revealed independently by using a physiological index. We specifically focused on the method of pupillometry, which has been previously used to investigate bistable perception (Einhäuser, Stout, Koch, & Carter, 2008). A pupillary dilation temporally close to a reversal in perspective of the Necker cube has been interpreted as a physiological signal of a “reset” mode in consciousness (for a review, see Laeng, Sirois, & Gredeback, 2012), driven by noradrenergic activity in the brain. In the present study, we use pupillometry as an index of the intensity of cognitive processing or of attention, as originally suggested by Kahneman (1973), Hess and Polt (1964), Kahneman and Beatty (1966), and Just and Carpenter (1993). Specifically, we hypothesized that if one perspective of the two (above/below) has been more frequent in past encounters with similar objects, then one view-specific memory should be less effortful or more “fluent” than the other.

Several previous studies have investigated voluntary control in bistable stimuli (e.g., Strüber & Stadler, 1999; Toppino, 2003; Van Ee, Van Dam, & Brouwer, 2005; Kornmeier et al., 2009), suggesting that the rate of perceptual switching is modulated by voluntary control, which seems consistent with our present results. For example, eye movements and eye position can affect perceptual switching, including eye movements or saccades (Toppino, 2003; Van Dam & Van Ee, 2006) and eye positions (Einhäuser et al., 2004). In this study, we designed an experimental paradigm that avoids the influence of eye gaze positions by requesting participants to maintain central fixation during stimulus presentation.

Specifically, priming one particular surface of the Necker cube by briefly making it opaque at the start of a trial should be more effective in disambiguating the view of the cube whenever the cued surface is consistent with a VFA than when consistent with a view-from-below (see Figure 1). Moreover, we predicted that when cueing a surface consistent with a view-from-below, reversals to the alternative view should be more frequently reported than when cueing a surface consistent with a VFA. In all of the earlier described cases, the less “fluent” perspective should be indexed by a larger size of the pupil (cf. Yoshimoto, Imai, Kashino, & Takeuchi, 2014) consistent with increased cognitive workload or in intensive attention (Kahneman, 1973).

Finally, we hypothesize that maintaining the perception of a specific view may be more likely to reveal a VFA bias in the forced attention condition than during passive viewing of the Necker cube. We also note that in previous studies of perceptual switching with pupillometry, it has been debated whether perceptual switching causes pupil dilation and whether it may be simply triggered by the motor response (Einhäuser et al., 2008; Hupe, Lamirel, & Lorceau, 2009). In this study, participants’ responses regarding the perspective of perception were collected after the offset of the stimuli; therefore the present paradigm can separate between the occurrence of a perceptual switch and the motor response when reporting it.

Experiment 1

In the first experiment, we ascertained whether the pupil diameter can index reduced effort, as expected if there is a VFA bias, when viewing a bistable (Necker) cube. In each trial, either the top or bottom sides of the Necker cube was filled-in for 3 seconds with an opaque white hue. This was intended to cue attention toward the surface and consequently bias a specific perspective consistent with a front side position of the surface. Subsequently, the surface became transparent, and a classic (wireframe) Necker cube was shown for

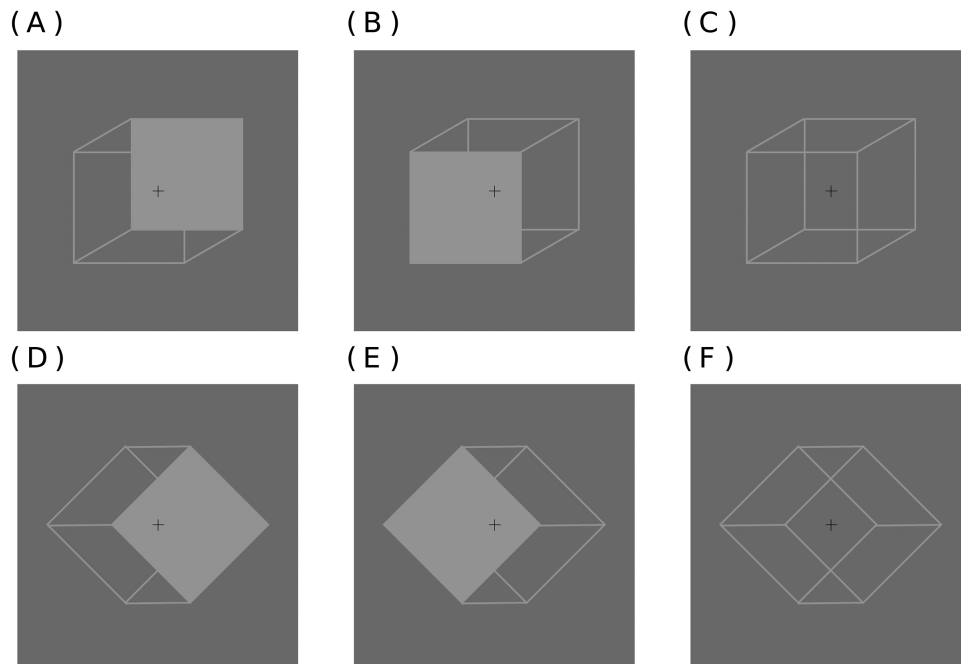


Figure 1 . Stimuli used in Experiments 1 and 2 (top and bottom panels, respectively). Figures (A, B, D, and E) were used for priming. Figures (C and F) shows the standard Necker cube.

3 seconds. In a passive viewing blocked condition, participants simply kept fixation on a cross centered in the middle of the Necker cube. In the forced attention condition, the participants were asked to maintain in their perception the perspective that had been initially cued, while also keeping central fixation (see Figure 2).

We reasoned that maintaining a VFA perspective should require fewer attentional resources than a view-from-below perspective, particularly so when engaging top-down intensive attention in the forced attention condition. Therefore we expected a difference in pupil diameter after cueing either the top or bottom side of the Necker cube, and particularly so in the attention condition.

Materials and methods

Participants

Twenty-eight healthy subjects participated in Experiment 1. Four participants were excluded from pupil analyses because two participants had eye blinks on more than 70% of the trials, one participant misunderstood the instructions of the tasks, and one participant never shifted perception to another perspective, yielding a final study group size of 24 participants (mean age, 23.00 years; *SD*, 1.56; 4 women).

All participants self-reported that they had a normal or corrected-to-normal visual acuity. The experimental procedures received approval from the

Committee for Human Research at the Toyohashi University of Technology. Participants provided written informed consent, and the experiment was conducted in accordance with the guidelines of the committee.

To ensure adequate statistical power = 0.8 ($1 - \beta$; the probability that the test rejects the null hypothesis when a specific alternative hypothesis is true), the sample size was a priori determined by a power analysis based on predicted effect size using G*Power version 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007). Thus the present study's sample size was above the estimated sample size range of $N = 24$ necessary to find large-to-medium effects in a repeated measures analysis of variance (ANOVA), given the error probability of $\alpha = 0.05$ (effect size estimates for each main effect and interaction are given later by partial eta-square (η_p^2) for ANOVA) (Aarts et al., 2015).

Stimuli and apparatus

We used three kinds of images in Experiment 1. Two images were unambiguous figures of cubes generated by shading one surface of the Necker cube drawing (Figures 1A and B). The third image was the standard Necker (wireframe) cube (Figure 1C). The cube occupied 7.7° horizontally and 7.7° vertically in visual angle. The color of the cube edges was light gray (78.8 cd/m^2) on a gray background (60.9 cd/m^2). The fixation cross was black (0.4 cd/m^2) with a $0.9^\circ \times 0.9^\circ$ size in visual angle. These images were created first with Microsoft PowerPoint 2016 (Microsoft Corp., Redmond, WA) and adjusted with MATLAB

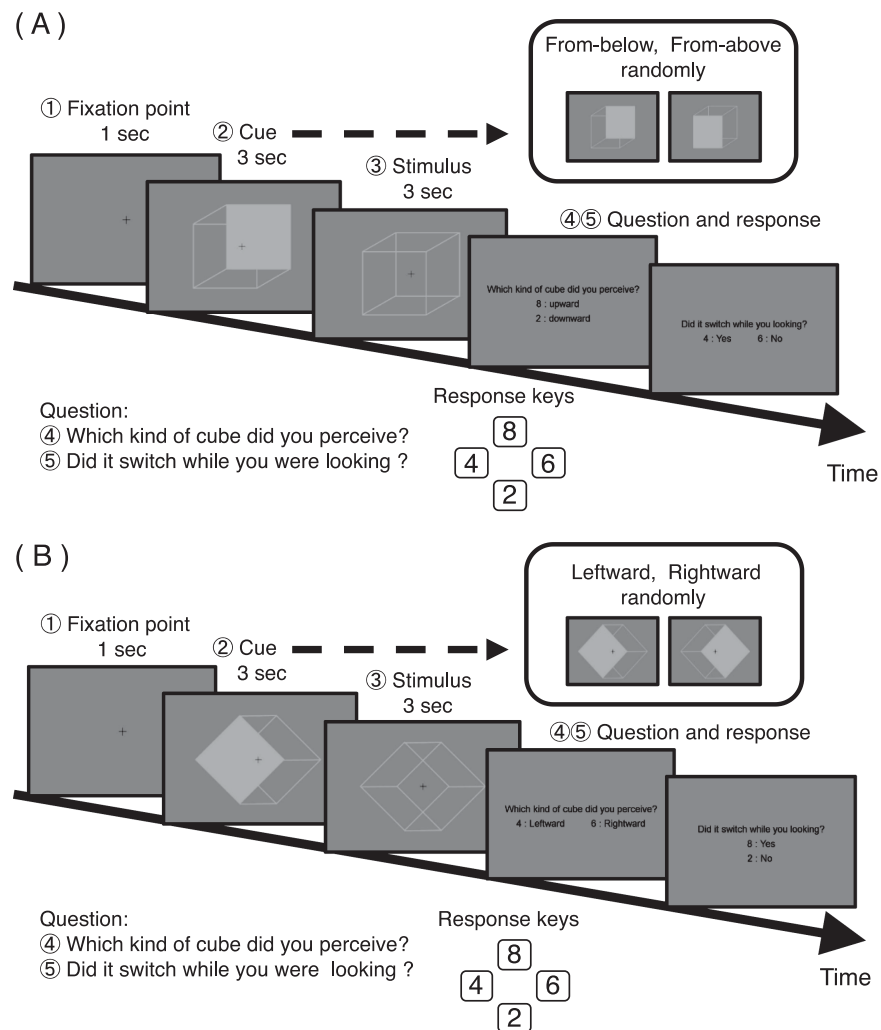


Figure 2 . (A) Timeline of [Experiment 1](#). The arrow shows the sequence of images presented to participants with the corresponding presentation times. Participants answered two questions after stimuli presentation: (1) “Which kind of cube did you perceive?” and (2) “Did it switch while you were looking?” The former question was answered by pressing “8” as “upward” and “2” as “downward” by Two-alternative forced choice (2AFC). “Upward” refers to the appearance of viewing from below, and “downward” refers to viewing-from-below. The inset above illustrates the two possible cues. In the passive viewing condition, participants simply looked at a fixation cross on the Necker cube. In the forced attention condition, the participants were asked to maintain the same perspective as the one initially cued, while also keeping fixated on the central cross. (B) Experimental procedure in [Experiment 2](#). Same as in the above panel except for the rotated Necker cube stimuli. Questionnaire sentences were also the same, but the former was answered by pressing “4” as “leftward” and “6” as “rightward,” and the latter was answered by pressing “8” as “yes” and “2” as “no.”

(MathWorks, Natick, MA) using Psychtoolbox 3.0 ([Brainard, 1997](#)). All stimuli were shown on an LCD display (Display++, Cambridge Research Systems, Kent, United Kingdom) at a refresh rate of 120 Hz.

Procedure

[Figure 2A](#) shows the timeline of one trial in [Experiment 1](#). First, the fixation cross was shown for 1000 ms, and then a cube with either the top or bottom side rendered opaque was presented for 3000 ms as a cue (hereafter called “cue”). Two kinds of images were used randomly as cues: one yielded the perspective of

a cube seen from above and the other the perspective from below. After the cue, the standard Necker cube was shown (by removing the surface shading while leaving the standard wireframe) for 3000 ms, and at its offset the participants reported (by pressing one of two forced-choice keys) which percept they saw initially and whether a reversal to the alternative perspective occurred while viewing the empty cube. The experiment consisted of two conditions (passive viewing and forced attention), each with 40 trials (each cue type \times 20 trials). The order between the two conditions was counterbalanced by inverting it for every other participant. In the passive viewing

condition, participants simply looked at a fixation cross on the Necker cube. In the forced attention condition, the participants were asked to maintain the same perspective as the one initially cued, while also keeping fixated on the central cross. Participants were also instructed to refrain from blinking as much as possible during each trial.

Recording and analysis

Behavioral analysis

From the participants' key-press responses, we calculated the probability that they perceived the same appearance as primed by the cue. The probability of a percept was obtained by dividing the number of trials in which the participant's initial percept was congruent with the cued appearance by all of the trials in each condition (20 trials). Similarly, we calculated the probability that the participant's percept switched while viewing the empty Necker cube. This was obtained by dividing the number of trials in which a participant's percept switched by all the trials in the condition (20 trials). For statistical analysis, a two-way repeated measures ANOVA was performed with perspective (from above, from below) and attention (passive, forced) as the within-subject factors in these probabilities. Pairwise comparisons for main effects in the ANOVA were corrected for multiple comparisons using the Bonferroni method. In addition, to investigate the effect of the cue, the number of trials of all participants who perceived the same appearance of the cue was calculated, and binomial tests were performed in each condition. The level of statistical significance was set to $p < 0.05$ for all analyses.

Pupil recording and analysis

Pupil sizes and eye movements were measured during the task by a noninvasive infrared eye tracker (iViewX RED500, SensoMotoric Instruments, Teltow, Germany) at a sampling rate of 500 Hz. Eye movements were monitored from both eyes. The positions of both eyes were acquired by nine-point calibration at the start of the experiment. For analyses, we averaged the pupil diameters from both eyes. Trials in which the pupil could not be detected were excluded from analysis. Pupil recordings were smoothed using a sliding average (80-ms time window). In the time-course analysis, each trial was normalized by subtracting pupil size at stimulus onset from the baseline pupil size. Baseline pupil size was computed as an average of data collected –500 ms prior to stimulus onset (0 ms). We calculated the time course of the trial's average pupils when the participants perceived either the upward cube (view-from-below) or the downward cube (VFA) in correspondence with the perspective primed by the cue. In the time

course of pupil diameter changes, the significant differences were corrected with false discovery rate for multiple comparisons using the Benjamini and Hochberg method (Benjamini & Hochberg, 1995). Specifically, the average pupil diameters from 220 to 3000 ms after stimulus presentation were calculated, and a repeated measures ANOVA was performed to assess the presence of significant differences in pupil diameter, with perspective (from above, from below) and attention (passive, forced) as the within-subject factors. The reason for excluding data before 220 ms is based on the known latency of the light reflex, which has a minimum of approximately 220 ms (Ellis, 1981). Furthermore, to separate the effects of perceptual perspective and perceptual switching, we extracted and analyzed only those trials in which there occurred no perceptual switching. This additional ANOVA was also performed with perspective and attention as factors. Finally, we calculated the averaged time course of pupil diameter in separate trials in which a perceptual switching occurred or was absent to better reveal how perceptual switching per se affected pupil diameter in each attentional condition. Pairwise comparisons for main effects in the ANOVA were corrected for multiple comparisons using the Bonferroni method, and the level of statistical significance was set at $p < 0.05$ for all analyses as for the behavioral analysis.

Results and discussion

Behavioral results

We computed the participants' key presses indicating their subjective view during perception. Based on the earlier described data, we found that the probability to perceive the Necker cube as having the same perspective as primed by the cue was greater in the forced attention condition than in the passive viewing condition (see Figure 3A). In addition, the probability of a from-below appearance was lower than that from above. The ANOVA revealed a main effect of attention, $F_{(1,23)} = 28.46$, $p < 0.001$, $\eta_p^2 = 0.55$, and perspective, $F_{(1,23)} = 8.91$, $p = 0.007$, $\eta_p^2 = 0.28$. There was no significant interaction between attention \times perspective, $F_{(1,23)} = 0.10$, $p = 0.76$, $\eta_p^2 = 0.004$ (see Figure 3A). As shown by binomial tests, both attention and perspective were significantly different from 0.5 of a chance level ($p < 0.001$).

Next, we also calculated the perceptual switching probabilities. An ANOVA on the switches showed an attention by perspective interaction, $F_{(1,23)} = 12.38$, $p = 0.002$, $\eta_p^2 = 0.35$ (see Figure 3B). To identify more detailed effect of the factors, a post hoc test revealed a simple effect in the forced attention condition, $F_{(1,23)} = 4.88$, $p = 0.04$, $\eta_p^2 = 0.18$, which indicated that the perceptual switching probability toward a from-above

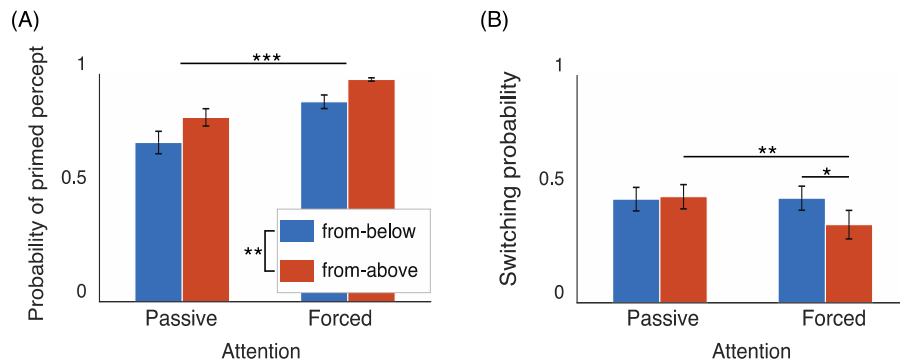


Figure 3 . Bar graphs (A) show grand average probability of the participants perceived the same appearance as the cue. (B) Grand average of switching probability during stimulus presentation. Both results are in [Experiment 1](#). Asterisks on horizontal lines represent a significant difference (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

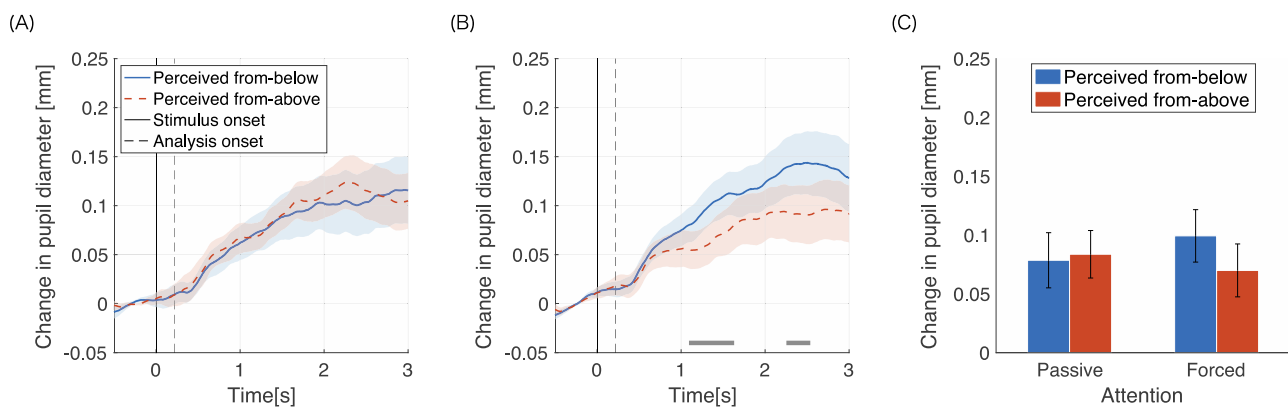


Figure 4 . Pupillometry results classified by perceived perspectives in [Experiment 1](#). (A) Passive viewing condition: grand-averaged time course of pupil changes during the presentation time of 3 seconds. The horizontal axis indicates the time (in seconds) and the vertical axis indicates pupil dilation (in mm) relative to baseline (from -500 to 0 ms). Shaded colors are the standard error of the mean. (B) Forced attention condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. The gray horizontal bars in the middle panel represent significant differences between the perspective condition of view-from-above and from-below trials; p values were corrected for multiple comparisons with an expected False Discovery Rate (FDR) of 0.05. (C) Bars show the mean pupil dilation from 220 to 3000 ms in attention condition and perspective condition, respectively.

perspective was higher than toward a view-from-below. In addition, a post hoc test revealed a simple effect in the from-below perception, indicating that the perceptual switching probability of forced attention condition was lower than in the passive condition, $F_{(1,23)} = 8.55$, $p = 0.008$, $\eta_p^2 = 0.27$. There was no main effect of attention, $F_{(1,23)} = 1.62$, $p = 0.22$, $\eta_p^2 = 0.07$, and perspective, $F_{(1,23)} = 1.16$, $p = 0.29$, $\eta_p^2 = 0.05$.

Pupillometry results

Time course of the pupils when participants perceived either the view-from-below or VFA.: We then analyzed the time course of the pupil diameter to confirm a relationship of pupil size for the perceived perspectives between the attentional conditions. These analyses were conducted for trials in which the perception was congruent with the prior cue because the number of

trials in which perception was incongruent with the prior was too small (i.e., the grand average probability of incongruent perception in all condition and in all participants was 15.8%; see [Figure 3A](#)). [Figures 4A](#) and [B](#) show the grand-averaged time course of changes in pupil diameter when the participants perceived the view-from-below the cube or from above the cube during the stimulus presentation for 3 seconds in passive viewing and forced attention condition, respectively. An ANOVA was carried out to compare the mean pupil dilation from 220 to 3000 ms in the two attention and perspective conditions ([Figure 4C](#)). There was, however, no significant main effect of attention, $F_{(1,23)} = 0.09$, $p = 0.76$, $\eta_p^2 = 0.004$, nor a main effect of perspective, $F_{(1,23)} = 2.47$, $p = 0.13$, $\eta_p^2 = 0.01$. There was also no evidence for an interaction between attention and perspective ($F_{(1,23)} = 2.71$, $p = 0.11$, $\eta_p^2 = 0.10$).

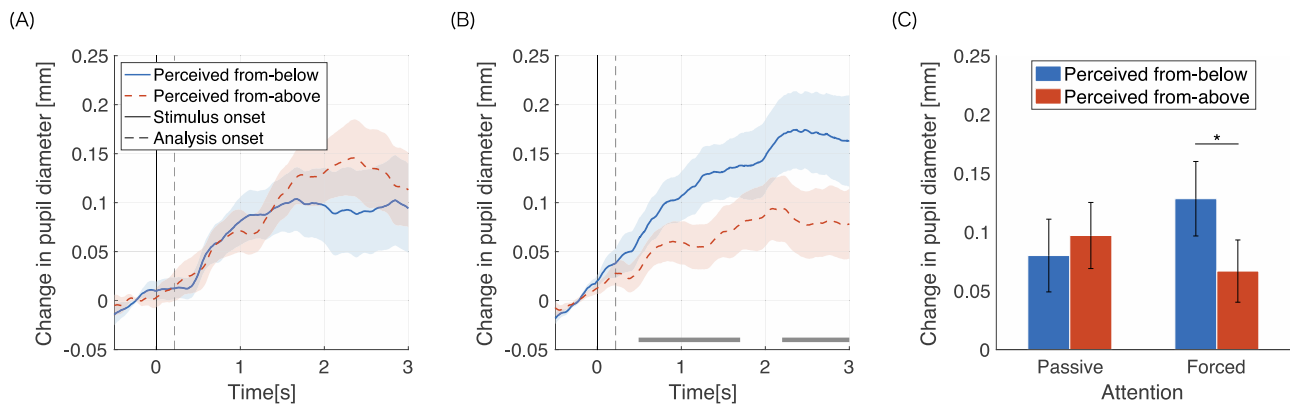


Figure 5 . Pupillometry results classified by perceived perspectives during nonswitch trials in Experiment 1. (A) Passive viewing condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. The horizontal axis indicates the time (in seconds) and the vertical axis indicates pupil dilation (in mm) relative to baseline (from -500 to 0 ms). Shaded colors indicate the standard error of the mean. (B) Forced attention condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. Shaded colors indicate the standard error of the mean. The gray horizontal bars in the middle panel represent significant differences between the perspective condition of view-from-above and from-below trials; p values were corrected for multiple comparisons with an expected False Discovery Rate (FDR) of 0.05 (C) Bars show the mean pupil dilation from 220 to 3000 ms in attention condition and perspective condition, respectively. Asterisk represents a significant difference ($*p < 0.05$). The result shows that there was a significant difference in the amount of change in pupil diameter between perceived from below and from above with forced attention condition, even though there was no button response no perceptual switching.

Trials with no perceptual switching: The previous analysis included only trials with perceptual switching during the stimulus presentations. Thus the previous analysis results included not only the effect of perceptual switching but also the effect of different perceptual appearances after perceptual reversal. In the present analysis, we excluded switch trials so as to examine in detail the effect of a specific view. Two participants for which no perceptual switching occurred under any one of the conditions were rejected from the analysis. Figures 5 (A and B) shows the grand-averaged time course of the changes in pupil diameter when the participants perceived either the upward cube (view-from-below) or the downward cube (VFA) with the condition of nonswitching trials during the stimulus presentation for 3 seconds in passive viewing and forced attention condition, respectively. Importantly, as seen in an ANOVA, there was a significant interaction of attention \times perspective, $F_{(1,21)} = 5.06$, $p = 0.035$, $\eta_p^2 = 0.19$ (see Figure 5C). A post hoc test revealed that the pupil diameter was significantly larger for the perspective from below than from above in the condition of forced attention, $F_{(1,21)} = 7.44$, $p = 0.013$, $\eta_p^2 = 0.26$. There were no significant main effects for either attention, $F_{(1,21)} = 0.19$, $p = 0.66$, $\eta_p^2 = 0.009$, or perspective, $F_{(1,21)} = 1.62$, $p = 0.22$, $\eta_p^2 = 0.07$.

Time course of the pupils during nonswitch versus switch trials: To investigate whether the pupils were dilated due to perceptual switching, we calculated time courses separately for the nonswitch and switch trials. Figure 6 shows the averaged time courses of pupil diameter when perceptual switching occurred or did not during passive

viewing. The t -tests revealed no significant difference in the time sequence data.

Discussion

The present findings confirm that priming one specific view of the Necker cube was effective, and the shading cue preceding the presentation of the wireframe typically evokes the corresponding view. However, as expected, we observed a change in pupil diameter only when we instructed the participants to actively sustain a specific perspective.

Experiment 2

In Experiment 2, we further tested the idea that the ease in sustaining a particular perceptual interpretation of the bistable Necker drawing, and its effect on the pupil diameter, reflects an ecological VFA constraint on internal representations or heuristic. We reasoned that, by rotating the stimuli used in the previous experiment of 90° , we would exclude the presence of any perspective bias because these particular views of the cube appear to be rare. As visible in Figure 1 (bottom panel), these rotated perspectives of the same Necker cube do not seem familiar, especially considering our real-world experience with solid cubes (i.e., a real cube in any of the two possible perspectives in the images would be gravitationally unstable because

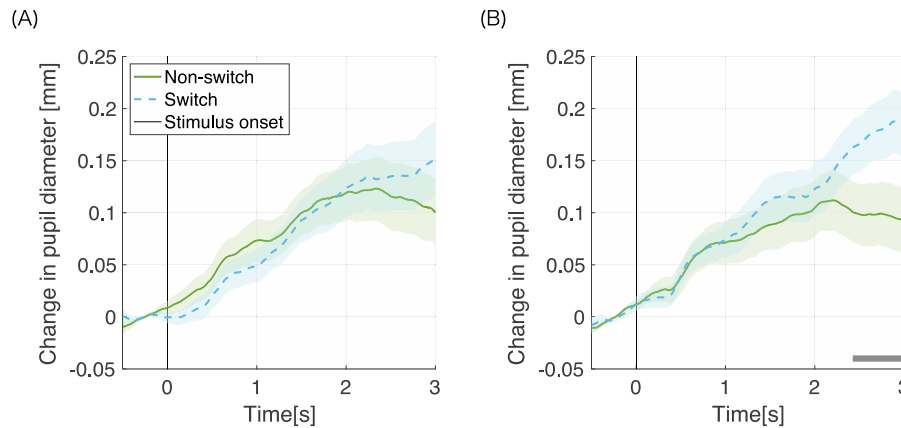


Figure 6 . Pupillometry results separating nonswitch trials versus switch trials in [Experiment 1](#). (A) Passive viewing condition: averaged time courses of pupil diameter separately for perceptual switching trials and nonswitching trials. (B) Forced attention condition: averaged time courses of pupil diameter separately for perceptual switching trials and nonswitching trials. Shaded colors indicate the standard error of the mean. The gray horizontal bar represents a significant difference between nonswitch trials and switch trials; p values were corrected for multiple comparisons with an expected False Discovery Rate (FDR) of 0.05. Please note that 0 seconds represents onset of the ambiguous Necker stimuli.

in both views the cube appears to be poised on the tip of one corner).

Materials and methods

Participants

Twenty-seven healthy subjects participated in [Experiment 2](#). Two participants were excluded from pupil analyses because one showed eye blink in more than 70% of the trials and one participant had difficulty performing the task, yielding a final study group size of twenty-five participants (21 men, 4 women; mean age, 22.76 years; SD, 1.64). All participants had a normal or corrected-to-normal visual acuity. Again, the experimental procedures received the approval of the Committee for Human Research at the Toyohashi University of Technology, and the experiment was strictly conducted in accordance with the approved guidelines of the committee, and all participants provided written informed consent. The study's sample size was based, as explained previously, on an estimated sample size of $N = 24$.

Stimuli and apparatus

The apparatus was the same as in [Experiment 1](#). Stimuli in [Experiment 2](#) were also identical and simply rotated on the frontal plane.

Procedure

The procedure was identical to [Experiment 1](#) ([Figure 2B](#)). Two types of stimuli were used randomly as cue: one was perspective of a rightward ([Figure 1D](#))

cube and the other perspectives of leftward ([Figure 1E](#)). The only difference was that responses of “leftward” and “rightward” perspectives were used instead of “from below” and “from above.”

Recording and analysis

Behavioral analysis

As done previously, we calculated the probability that participants perceived the same appearance as primed by the cue. For statistical analysis, a two-way repeated measures ANOVA was performed with perspective (leftward, rightward) and attention (passive, forced) as within-subject factors. Binomial tests were also performed in each condition.

Pupil analysis

The analyses of the pupils were performed as previously by first calculating the grand-averaged time course of the average pupils when the participants perceived either the leftward or rightward in correspondence to the perspective primed by the cue. We calculated the average pupil diameters from 220 to 3000 ms after stimulus presentation, and an ANOVA was performed with perspective and attention as factors. Moreover, to separate the effects of perceptual perspective and perceptual switching, we analyzed only those trials in which there was no perceptual switching. Finally, we calculated the averaged time course of pupil diameter in trials in which perceptual switching occurred or was absent, so as to clarify how perceptual switching affected the pupil diameter in each attentional condition.

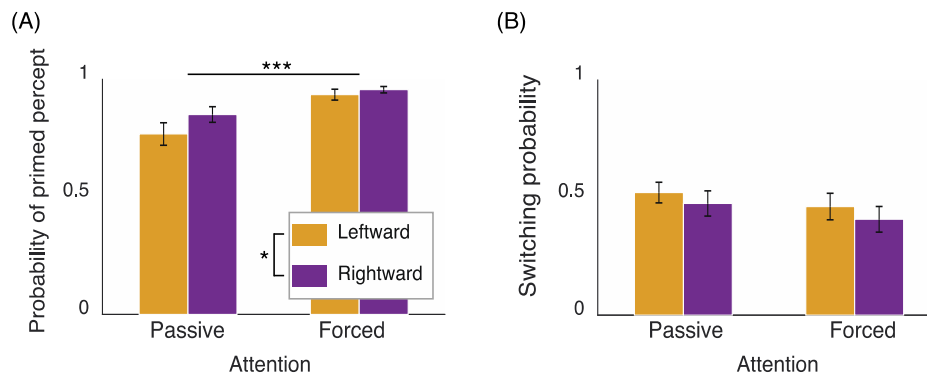


Figure 7 . (A) Grand average probability of the perceptions of the cued perspectives. (B) Grand average of switching probability. Both results are in [Experiment 2](#). Asterisk represents a significant difference (* $p < 0.05$, *** $p < 0.001$).

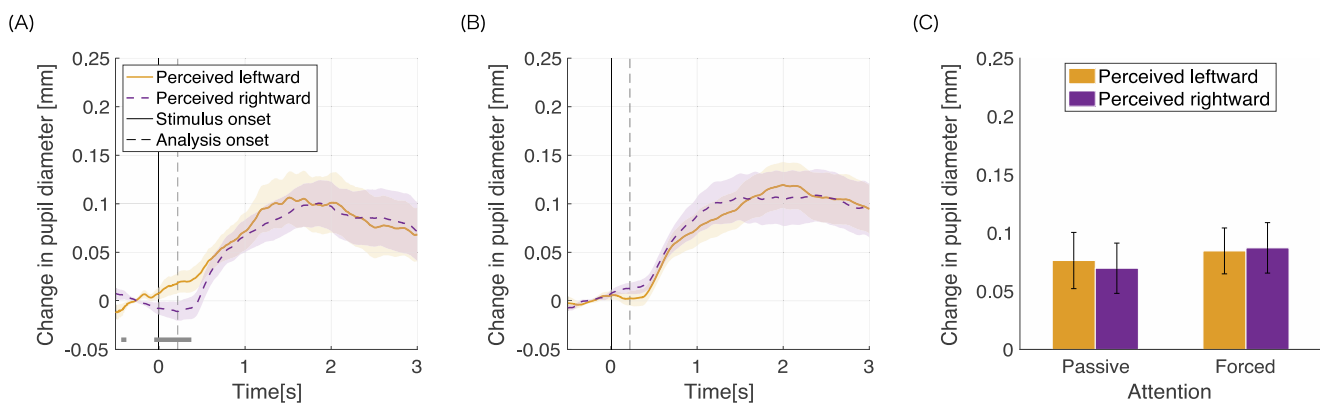


Figure 8 . Pupillometry results classified by perceived perspectives in [Experiment 2](#). (A) Passive viewing condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. The horizontal axis indicates the time (in seconds) and the vertical axis indicates pupil dilation (in mm) relative to baseline (from -500 to 0 ms). Shaded colors indicate the standard error of the mean. The gray horizontal bar represents the significant difference between the perspective condition of leftward and rightward trials; p values were corrected for multiple comparisons with an expected False Discovery Rate (FDR) of 0.05. (B) Forced attention condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. Same conventions as A. (C) Bars show the mean pupil dilations from 220 to 3000 ms in the attention and perspective conditions.

Results and discussion

Behavioral results

The probability to perceive the same appearance of the cube as cued was greater in the forced attention condition than passive viewing condition (see [Figure 7A](#)). As shown by binomial tests, both attention and perspective were significantly different from 0.5 of a chance level ($p < 0.001$). In addition, the probability of leftward perception was lower than rightward. The ANOVA revealed a main effect of attention, $F_{(1,24)} = 17.81$, $p = 0.0003$, $\eta_p^2 = 0.43$, and perspective, $F_{(1,24)} = 7.20$, $p = 0.013$, $\eta_p^2 = 0.23$. There was no significant interaction between attention \times perspective, $F_{(1,24)} = 2.40$, $p = 0.13$, $\eta_p^2 = 0.09$.

[Figure 7B](#) shows the result of perceptual switching probability. There was no significant main effect between either attentional conditions, $F_{(1,24)} = 1.43$, $p = 0.24$, $\eta_p^2 = 0.06$, or perspective, $F_{(1,24)} = 1.97$, $p = 0.17$, $\eta_p^2 = 0.08$, or their interaction, $F_{(1,24)} = 0.03$, $p = 0.87$, $\eta_p^2 = 0.001$.

Time course of pupils when perceiving either the leftward cube or the rightward cube: We analyzed time courses of change in pupil diameter as done for [Experiment 1](#). As expected, an ANOVA on average changes in pupil diameter from 220 to 3000 ms ([Figure 8C](#)) showed no significant difference in any main effect (main effect of attention: $F_{(1,24)} = 1.24$, $p = 0.28$, $\eta_p^2 = 0.05$; main effect of perspective: $F_{(1,24)} = 0.05$, $p = 0.83$, $\eta_p^2 = 0.002$, and the interaction between attention \times

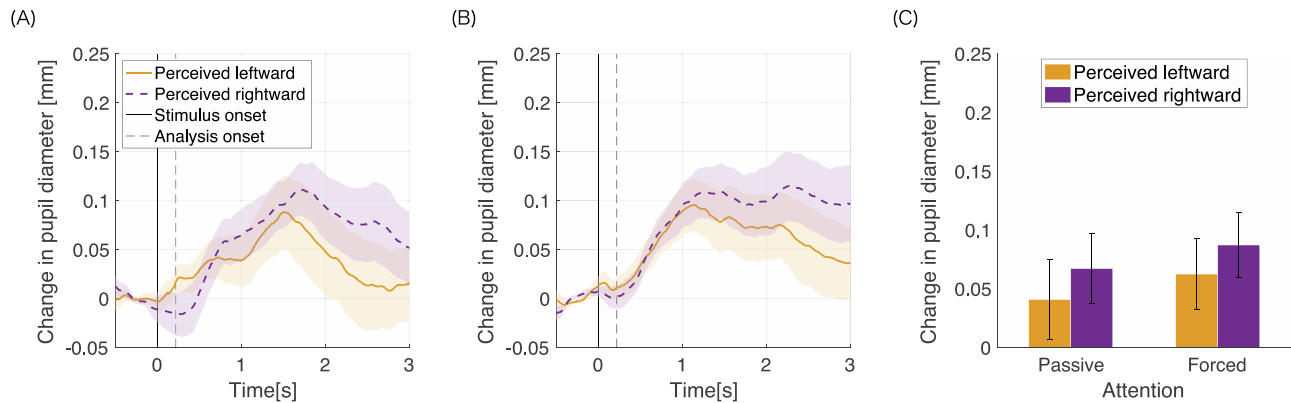


Figure 9 . Pupillometry results classified by perceived perspectives during nonswitch trials in Experiment 2. (A) Passive viewing condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. The horizontal axis indicates the time (in seconds) and the vertical axis indicates pupil dilation (in mm) relative to baseline (from -500 to 0 ms). Shaded colors are the standard error of the mean. (B) Forced attention condition: grand-averaged time course of pupil changes during presentation time for 3 seconds. Same conventions as A. (C) Bars show the mean pupil dilation from 220 to 3000 ms in attention condition and perspective condition.

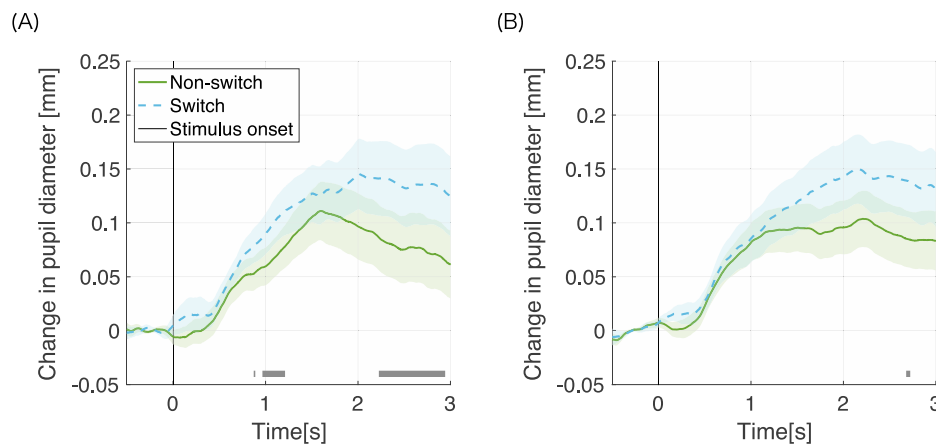


Figure 10 . Pupillometry results separating nonswitch trials versus switch trials in Experiment 2. (A) Passive viewing condition: averaged time courses of pupil diameter separately for perceptual switching trials and nonswitching trials. (B) Forced attention condition: averaged time courses of pupil diameter separately for perceptual switching trials and nonswitching trials. Shaded colors indicate the standard error of the mean. The gray horizontal bar represents a significant difference between nonswitch trials and switch trials; p values were corrected for multiple comparisons with an expected False Discovery Rate (FDR) of 0.05. Please note that 0 seconds represents onset of the ambiguous Necker stimuli.

perspective missed the significant cutoff, $F_{(1,24)} = 0.24$, $p = 0.63$, $\eta_p^2 = 0.01$.

Nonswitch trials: As done earlier, to reveal the effect of perspective, we further analyzed time courses of change in pupil diameter as in Experiment 1 (Figures 9A and B). We excluded five participants for which no perceptual switching occurred under any one of the conditions. An ANOVA showed no significant difference main effect (Figure 9C): attention: $F_{(1,19)} = 0.94$, $p = 0.34$, $\eta_p^2 = 0.047$; perspective: $F_{(1,19)} =$

3.10 , $p = 0.094$, $\eta_p^2 = 0.14$. There was no interaction between attention \times perspective: $F_{(1,19)} = 0.02$, $p = 0.97$, $\eta_p^2 = 0.0001$ (see Figures 9A and B).

Time course of pupils during and switch trials: Finally, we calculated time courses separately between the nonswitch and switch trials, as done in Experiment 1. Figure 10 shows the averaged time courses of pupil diameters when perceptual switching occurred or not in each of attentional condition.

Discussion

As expected, when the Necker cube drawings were rotated to physically unlikely positions, there was no indication of effort related to perspective, which is consistent with the absence of an ecological viewpoint constraint on these particular images. We found, however, confirmation of the effectiveness of forcing attention when sustaining a specific perceptual interpretation of the bistable figure. Interestingly, there was a slight preference for the rightward view of the cube. We surmise that, given that this view would correspond to a clockwise rotation of the VFA cued surface, there may be at work complex interaction between the VFA bias and spontaneous mental rotations to the more natural orientation of the stimuli. We speculate, considering that the right arm is stronger in most people (Peters & Servos, 1989), that most objects are naturally manipulated in a clockwise action when held and turned (e.g., caps and lids). We also note that there is some evidence for a preference in “mentally rotating” abstract shapes in a clockwise direction (e.g., Koriat & Norman, 1985; Liesefeld & Zimmer, 2011).

General discussion

We found that changes in pupil diameter were significantly larger when participants perceived the view-from-below than when they perceived the VFA of the Necker cube and, specifically, when actively attempting to maintain one of these perspectives. Moreover, the probability of maintaining a specific perspective in perception, after disappearance of the cue, was greater during forced attention condition than in passive viewing. In addition, the probability of perceptual switching in the forced condition was lower than in passive condition.

In a second experiment, we showed the same cubes after a 90° rotation, either leftward and rightward, yielding two equally unusual and physically unlikely perspectives (based on a gravitational constraint) of the bistable shape. Given that in this case the alternative views do not differ in terms of an above or below viewpoint, we also expected to find no difference in pupil diameter. Indeed, whereas the views from above and below differed in terms of pupillary response, the views from left or right did not.

Einhäuser and colleagues reported that pupil diameter increases around the time of a perceptual switching during perceptual rivalry (Einhäuser et al., 2008). However, another study reported that the degree of pupil dilation does not predict subsequent stability in perceptual rivalry (Hupe, Lamirel, & Lorenceau, 2008; Hupe, Lamirel, & Lorenceau, 2009). Instead, a small (approximately 5% of change on average) but reliable pupil dilation was observed around the

time of key presses, and 70% of pupil dilation could be accounted for by the motor response (Hupe, J. M. et al., 2009). Another study also showed that a key press can influence both brain activity and pupil diameter (Frässle, Sommer, Jansen, Naber, & Einhäuser, 2014). In the present study, all key-press responses were performed after stimulus offset.

When we separated in the analyses trials with and without a switch, no statistically significant difference was found. Also, it seems likely that the observed differences in pupillary responses in our study reflected a difference in specific perceptual content (e.g., from-below appearance, from-above appearance) instead of effects of motor responses, as also suggested by Kloosterman and colleagues (2015). We also note the interactive effect of attention and perspective remained significant in Experiment 1, even after excluding all trials in which a perceptual switch occurred.

In line with Kahneman’s original account (Kahneman, 1973), we assume that pupil dilation reflects the allocation of attentional resources, and therefore that it provides an index of the level of mental effort exerted in a particular situation. We note that in this study, there was no significant pupil size difference between the two attentional conditions. However, attention affected the pupil diameter as an interactive effect because pupils were significantly dilated only when participants perceived the view-from-below than the VFA during forced attention condition. It seems reasonable to conclude that it is more effortful to actively sustain viewing-from-below than from above. Hence the VFA bias is reflected in the level of mental effort, that is, a reduced dilation of the pupil.

At the physiological level, pupil dilations that are related to cognitive processing are thought to result from an inhibitory effect on the parasympathetic oculomotor complex by release of norepinephrine (NE) from the locus coeruleus (LC: Wilhelm, Wilhelm, & Lüdtke, 1999). The LC sends its noradrenergic projections to virtually all brain regions (except the basal ganglia), with particularly dense projections to areas known to be important in attentional processing, such as the parietal cortex, the pulvinar nucleus of the thalamus, and the superior colliculus (Bouret & Sara, 2005; Foote & Morrison, 1987; Schneider & Kastner, 2009). Single cell recordings in monkeys (Joshi, Li, Kalwani, & Gold, 2016; Rajkowski, 1993) and pharmacologic studies in humans (Koss, 1986; Phillips, Bitsios, Szabadi, & Bradshaw, 2000) have confirmed a physiological link between the activity of the LC-NE system and changes in pupillary diameter, allowing the use of pupillometry to tap task-related changes in attentional states mediated by LC-NE activity (Alnaes et al., 2014; Laeng et al., 2012). Thus our results are consistent with the idea that the modulation of pupil diameter reflects top-down attentional processing and different levels of mental effort.

Another consideration can be based on the Bayesian theory of human perception (Kersten, Mamassian, Yuille, & Yuille, 2004; Knill & Pouget, 2004), in which biases in perception toward specific interpretations, especially of ambiguous stimuli, tends to match the natural statistics of the environment (Girshick, Landy, & Simoncelli, 2011; Barlow, 2001; Jazayeri & Movshon, 2006; Weiss, Simoncelli, & Adelson, 2002; Zhang, Xu, Jiang, & Wang, 2017). Thus in light of this account, the VFA bias during perception of the Necker stimulus should reflect the supposed higher frequency in adopting the from-above perspectives than from-below perspectives in everyday situations. Such a priori perceptual bias can thus be interpreted as a strong influence on long-term perceptual memory of repeated instances of disambiguation of sensory information (Kornmeier, Wörner, Riedel, & Van Elst, 2017). Thus under the present task, a difference in the probability of memory retrieval from long-term memory will reflect a VFA bias and in turn the amount of attentional load that is mirrored in the pupil. Our results seem to support the earlier described interpretation because the VFA bias per se did not affect the pupil unless attention was forced to maintain a specific perspective.

Conclusion

To our knowledge, this is the first study to investigate whether the VFA bias influences effort, as expressed in pupil diameter when maintaining a specific percept. Because VFA may be the most frequent experience with many types of common objects, attending and maintaining this specific perspective requires a lower degree of mental effort.

Keywords: intensive attention, Necker cube, pupillometry, top-down processing, viewing-from-above bias

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Grant no. 17H01807, 17H06292, 17KK0005 and 13904120199937) and the Program for Leading Graduate Schools at Toyohashi University of Technology.

Commercial relationships: none.

Corresponding author: Tetsuto Minami.

Email: minami@tut.jp.

Address: Electronics-Inspired Interdisciplinary Research Institute (EIIRIS), Toyohashi University of Technology, Toyohashi, Aichi, Japan.

References

- Aarts, A. A., Anderson, J. E., Anderson, C. J., Attridge, P. R., Attwood, A., Axt, J., . . . Zuni, K. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), <https://doi.org/10.1126/science.aac4716>.
- Alnaes, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, *14*(4), 1–1, <https://doi.org/10.1167/14.4.1>.
- Barlow, H. (2001). The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences*, *24*(4), 602–607, <https://doi.org/10.1017/S0140525X01000024>.
- Benjamini, Y., & Hochberg, Y. (1995). Benjamini and Y FDR.pdf. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*, 289–300, <https://doi.org/10.2307/2346101>.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436, <https://doi.org/10.1163/156856897X00357>.
- Einhäuser, W., Martin, K. A. C., & König, P. (2004). Are switches in perception of the Necker cube related to eye position? *European Journal of Neuroscience*, *20*(10), 2811–2818, <https://doi.org/10.1111/j.1460-9568.2004.03722.x>.
- Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(5), 1704–1709, <https://doi.org/10.1073/pnas.0707727105>.
- Ellis, C. J. (1981). The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, *65*(11), 754–759, <https://doi.org/10.1136/bjo.65.11.754>.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G * Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.
- Foote, S. L., & Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annual Review of Neuroscience*, *10*(1), 67–95, <https://doi.org/10.1146/annurev.ne.10.030187.000435>.
- Frässle, S., Sommer, J., Jansen, A., Naber, M., & Einhäuser, W. (2014). Binocular rivalry: Frontal activity relates to introspection and action but not to perception. *Journal of Neuroscience*, *34*(5), 1738–1747, <https://doi.org/10.1523/JNEUROSCI.4403-13.2014>.

- Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*, 926.
- Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*, 1190–1192, <https://doi.org/10.1126/science.143.3611.1190>.
- Hupe, J. M., Lamirel, C., & Lorenceau, J. (2008). Pupil dilation does not predict subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, *105*(28), E43–E43, <https://doi.org/10.1073/pnas.0803456105>.
- Hupe, J. M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, *9*(7), 10–10, <https://doi.org/10.1167/9.7.10>.
- Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*(5), 690–696, <https://doi.org/10.1038/nn1691>.
- Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221–234, <https://doi.org/10.1016/j.neuron.2015.11.028>.
- Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, *47*, 310–339, <https://doi.org/10.1037/h0078820>.
- Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585, <https://doi.org/10.1126/science.154.3756.1583>.
- Kawabata, N., & Mori, T. (1992). Disambiguating ambiguous figures by a model of selective attention. *Biological Cybernetics*, *67*(5), 417–425, <https://doi.org/10.1007/BF00200985>.
- Kawabata, N., Yamagami, K., & Noaki, M. (1978). Visual fixation points and depth perception. *Vision Research*, *18*(7), 853–854, [https://doi.org/10.1016/0042-6989\(78\)90127-X](https://doi.org/10.1016/0042-6989(78)90127-X).
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304.
- Kloosterman, N. A., Meindertsma, T., van Loon, A. M., Lamme, V. A. F., Bonneh, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, *41*(8), 1068–1078, <https://doi.org/10.1111/ejn.12859>.
- Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*(12), 712–719, <https://doi.org/10.1016/j.tins.2004.10.007>.
- Koriat, A., & Norman, J. (1985). Mental rotation and visual familiarity. *Perception & Psychophysics*, *37*(5), 429–439, <https://doi.org/10.3758/BF03202874>.
- Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: When bottom-up and top-down coincide. *Brain and Cognition*, *69*(1), 138–147, <https://doi.org/10.1016/j.bandc.2008.06.005>.
- Kornmeier, J., Wörner, R., Riedel, A., & Van Elst, L. T. (2017). A different view on the Necker cube—Differences in multistable perception dynamics between Asperger and non-Asperger observers. *PLoS ONE*, *12*(12), 1–14, <https://doi.org/10.1371/journal.pone.0189197>.
- Koss, M. C. (1986). Pupillary dilation as an index of central nervous system α 2-adrenoceptor activation. *Journal of Pharmacological Methods*, *15*, 1–19, [https://doi.org/10.1016/0160-5402\(86\)90002-1](https://doi.org/10.1016/0160-5402(86)90002-1).
- Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*(1), 18–27, <https://doi.org/10.1177/1745691611427305>.
- Liesefeld, H. R., & Zimmer, H. D. (2011). The advantage of mentally rotating clockwise. *Brain and Cognition*, *75*(2), 101–110, <https://doi.org/10.1016/j.bandc.2010.10.012>.
- Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257* (5075), 1357–1363, <https://doi.org/10.1126/science.1529336>.
- Peters, M., & Servos, P. (1989). Performance of subgroups of left-handers and right-handers. *Canadian Journal of Psychology*, *43*, 341–358, <https://doi.org/10.1037/h0084226>.
- Phillips, M. A., Bitsios, P., Szabadi, E., & Bradshaw, C. M. (2000). Comparison of the antidepressants reboxetine, fluvoxamine and amitriptyline upon spontaneous pupillary fluctuations in healthy human volunteers. *Psychopharmacology*, *149*(1), 72–76, <https://doi.org/10.1007/s002139900334>.
- Rajkowski, J. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Social Neuroscience, Abstract, Washington, DC, 1993*.
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166, <https://doi.org/10.1021/jp904892v>.
- Strüber, D., & Stadler, M. (1999). Differences in top-down influences on the reversal rate of different

- categories of reversible figures. *Perception*, 28(10), 1185–1196, <https://doi.org/10.1068/p2973>.
- Sundareswara, R., & Schrater, P. R. (2008). Perceptual multistability predicted by search model for Bayesian decisions. *Journal of Vision*, 8(5), 12–12, <https://doi.org/10.1167/8.5.12>.
- Toppino, T. C. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics*, 65(8), 1285–1295, <https://doi.org/10.3758/BF03194852>.
- Troje, N. F., & McAdam, M. (2010). The viewing-from-above bias and the silhouette illusion. *I-Perception*, 1(3), 143–148, <https://doi.org/10.1068/i0408>.
- Van Dam, L. C. J., & Van Ee, R. (2006). The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision Research*, 46(6-7), 787–799, <https://doi.org/10.1016/j.visres.2005.10.011>.
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6), 598–604, <https://doi.org/10.1038/mn858>.
- Wernery, J., Atmanspacher, H., Kornmeier, J., Candia, V., Folkers, G., & Wittmann, M. (2015). Temporal processing in bistable perception of the Necker cube. *Perception*, 44(2), 157–168, <https://doi.org/10.1068/p7780>.
- Wilhelm, B., Wilhelm, H., & H., L. (1999). Pupillography: Principles and applications in basic and clinical research. *Pupillography: Principles, methods and applications*, 1–11
- Yoshimoto, S., Imai, H., Kashino, M., & Takeuchi, T. (2014). Pupil response and the subliminal mere exposure effect. *PLoS One*, 9(2), 1–8, <https://doi.org/10.1371/journal.pone.0090670>.
- Zhang, X., Xu, Q., Jiang, Y., & Wang, Y. (2017). The interaction of perceptual biases in bistable perception. *Scientific Reports*, 7, 1–8, <https://doi.org/10.1038/srep42018>.