Behavioral/Cognitive

Pupil Diameter Tracks Statistical Structure in the Environment to Increase Visual Sensitivity

Caspar M. Schwiedrzik[1,2] and Sandrin S. Sudmann[1,2]

[1]Neural Circuits and Cognition Lab, European Neuroscience Institute Göttingen—A Joint Initiative of the University Medical Center Göttingen and the Max Planck Society, 37077 Göttingen, Germany, and [2]Perception and Plasticity Group, German Primate Center, Leibniz Institute for Primate Research, 37077 Göttingen, Germany

Pupil diameter determines how much light hits the retina and, thus, how much information is available for visual processing. This is regulated by a brainstem reflex pathway. Here, we investigate whether this pathway is under the control of internal models about the environment. This would allow adjusting pupil dynamics to environmental statistics to augment information transmission. We present image sequences containing internal temporal structure to humans of either sex and male macaque monkeys. We then measure whether the pupil tracks this temporal structure not only at the rate of luminance variations, but also at the rate of statistics not available from luminance information alone. We find entrainment to environmental statistics in both species. This entrainment directly affects visual processing by increasing sensitivity at the environmentally relevant temporal frequency. Thus, pupil dynamics are matched to the temporal structure of the environment to optimize perception, in line with an active sensing account.

Key words: **active sensing; primate; pupil; statistical learning**

Significance Statement

When light hits the retina, the pupil reflexively constricts. This determines how much light and thus how much information is available for visual processing. We show that the rate at which the pupil constricts and dilates is matched to the temporal structure of our visual environment, although this information is not directly contained in the light variations that usually trigger reflexive pupil constrictions. Adjusting pupil diameter in accordance with environmental regularities optimizes information transmission at ecologically relevant temporal frequencies. We show that this is the case in humans and macaque monkeys, suggesting that the reflex pathways that regulate pupil diameter are under some degree of cognitive control across primate species.

Introduction

Our environment is structured in space and time: visual events unfold over time and present statistical regularities (Billock et al., 2001). Our senses can extract these statistics to form internal models that allow optimizing perception. Such models can also be used to guide sampling of information through motor actions (Friston et al., 2012). This is captured by the notion of "active sensing" (Schroeder et al., 2010), e.g., we use visual information to plan sequences of saccades that target relevant locations in a scene, providing the visual system with bouts of information (e.g., for object recognition). A critical component of active sensing is how much light and therefore how much information hits the retina. This is regulated by the reflexive adjustment of pupil diameter (PD). Here, we investigate whether this adjustment is fully automatic or under the control of flexible internal models involved in active sensing.

When light hits the retina, pupils transiently constrict to limit light influx. This pupillary light response (PLR) is controlled by a parasympathetic brainstem circuit in which retinal luminance information is relayed via the pretectal olivary nucleus to the Edinger–Westphal nucleus, which signals the pupillary sphincter muscle to contract. In addition, a sympathetic pathway adjusts PD to background illumination (McDougal and Gamlin, 2015). By doing so, PD modulates visual cortex activity, which scales with the amount of light passed (Haynes et al., 2004). PD affects the acuity and sensitivity of visual processing (Laughlin, 1992), as follows: smaller pupils sharpen the image and increase depth

Received Jan. 28, 2020; revised Mar. 20, 2020; accepted Apr. 15, 2020.

Author contributions: C.M.S. designed research; S.S.S. performed research; C.M.S. analyzed data; C.M.S. wrote the paper.

This research was supported by Emmy Noether Grant SCHW1683/2-1 from the German Research Foundation to C.M.S. We thank I. Kagan for support; J. Fischer and L. Melloni for comments on this manuscript; L. Burchardt, D. Bertazzi Lazzarini, and R. Brockhausen for technical support; T. Becker, B. Kamp, K. Lampe, and A. Schrod for veterinary care; and A.-L. Beyer for help with data acquisition.

The funders had no role in study design, data collection and interpretation, decision to publish, or preparation of the manuscript. The authors declare no competing financial interests.

Correspondence should be addressed to Caspar M. Schwiedrzik at c.schwiedrzik@eni-g.de.

https://doi.org/10.1523/JNEUROSCI.0216-20.2020

of field, while larger pupils allow more light to hit the retina and thus increase the signal-to-noise ratio and field of view. Thus, visual processing may be optimized by adjusting PD to environmental conditions to maximize information transmission, in line with active sensing (Mathôt and Van der Stigchel, 2015; Ebitz and Moore, 2018). Some environmental statistics are not available from light intensity alone and require more complex computations. Indeed, evidence suggests that the PLR is not a fully automated reflex but under some degree of corticocollicular control (Wang and Munoz, 2015). For example, the PLR is modulated by chromatic isoluminant stimuli, which cannot result from isolated subcortical processes that do not have access to chromaticity (Barbur et al., 1992).

Here, we ask whether PD adjusts to higher-order temporal statistics in the visual input, and whether this has perceptual consequences. Alternating light input causes the pupil to rhythmically dilate and constrict, resulting in temporal entrainment of PD known as "pupil cycling" (Miller and Thompson, 1978). But can the pupil entrain to temporal structure that is not a direct consequence of light variations but needs to be derived from higher-order regularities? By adjusting its temporal dynamics to the environment, PD could act as an adaptive filter that optimizes the signal-to-noise ratio at environmentally relevant frequencies, enhancing sensitivity. This would constitute a sophisticated mechanism of active sensing whereby internal models act on the very onset of visual processing.

We induced temporal statistics by presenting sequences of luminance-equalized images at a fixed temporal rate (2 Hz) while manipulating their order, as follows: we grouped images into pairs, such that the first image in a pair predicted the identity of the second image (Fig. 1*A*). Hence, we could separate luminance-induced pupil cycling at the 2 Hz image rate from modulations of PD at the pair rate, 1 Hz. Because only the statistical structure of the stimulus streams, but not light variations, can drive 1 Hz responses under these conditions, PD modulation at 1 Hz speaks for sensitivity of the responsible pathways to higher-order statistics. To assess whether this reflects a common sensorimotor strategy across species, we tested humans and macaque monkeys. While the brainstem PLR circuit is preserved between these species (Douglas, 2018), pupil dynamics differ between monkeys and humans (Gamlin et al., 1998), and it is unclear whether PD is under the same amount of cortical/cognitive control. Finally, we investigated whether pupil entrainment to higher-order statistics has perceptual consequences, as predicted by active sensing.

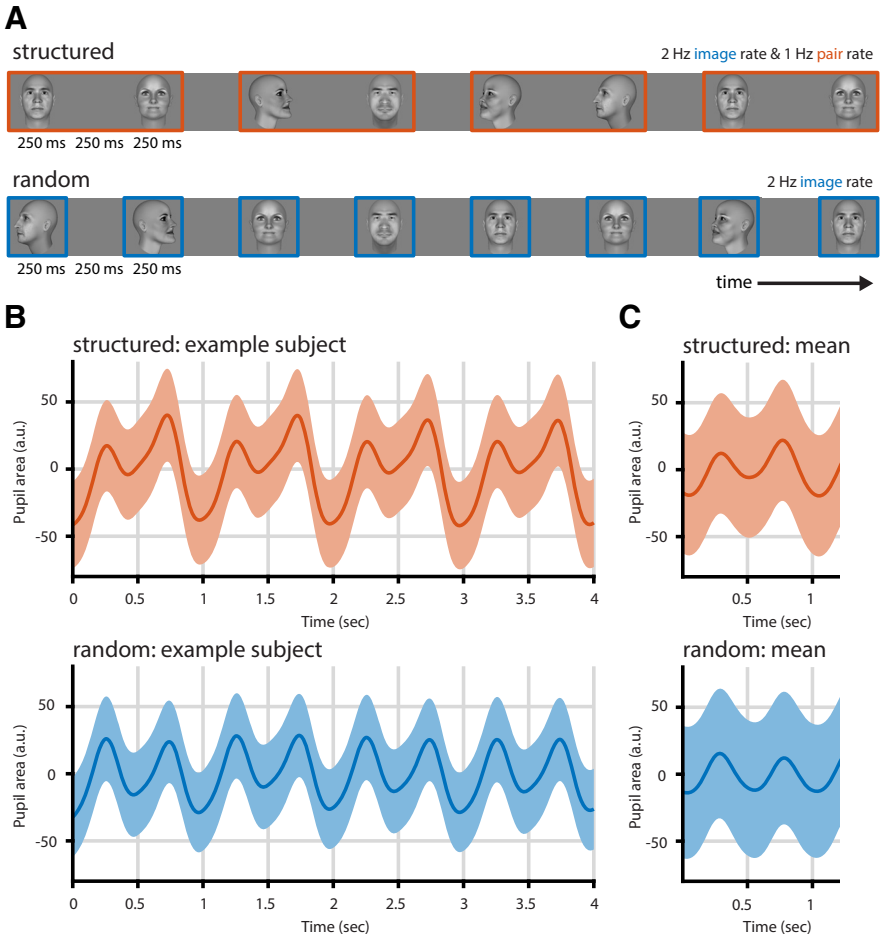

Figure 1. Paradigm and pupil entrainment. ***A***, We presented sequences of faces in structured or random order. In the random condition, faces were shown in random order at 2 Hz, the image rate. In the structured condition, faces were shown at the same rate (2 Hz). To induce statistical structure, images were grouped into pairs, such that one particular image always followed on another particular image. This gives rise to the pair rate at 1 Hz. ***B***, Pupil entrainment was evident at the single-subject level (here: mean over 83 trials). The structured condition (orange) shows a clear modulation at the 2 Hz image rate and slower dynamics, including the pair rate at 1 Hz. The predominant frequency in the random condition (blue) is the image rate (2 Hz). ***C***, Pupil entrainment at 1 and 2 Hz is also clearly evident when averaged across all runs and subjects. In ***B*** and ***C***, gray vertical lines indicate image onsets, shading represents the SEM, and data are lightly detrended.

Materials and Methods

Subjects

A total of 56 healthy human volunteers (27 female; 2 left handed; mean age, 25.46 years; SD, 3.91 years) participated in this study. All subjects had normal or corrected-to-normal vision, reported no history of neurologic or psychiatric disease, and gave written informed consent before participation. No sample size estimate was performed, but sample size was selected based on previous studies. Because all comparisons were within subjects, we used convenience sampling. Thirty-three subjects (20 female; 2 left handed; mean age, 25.45 years; SD, 3.46 years) participated in the main experiment. Three subjects had to be excluded from data analysis because they did not complete the study or failed to follow instructions (final $n = 30$; 19 female; 2 left handed; mean age, 25.07 years; SD, 2.90 years). Twelve subjects participated in a second experiment to assess awareness of the statistical structure ($n = 12$; 2 female; mean age, 26.5 years; SD, 5.09 years). Two subjects were excluded from data analysis because they did not complete the study or because of technical failure (final $n = 10$; 2 female; mean age, 26.2 years; SD, 5.05 years). Eleven subjects (5 female; mean age, 24.36 years; SD, 3.85 years) participated in a third experiment to assess perceptual consequences of PD. One subject had to be excluded from data analysis because of the failure to follow instructions (final $n = 10$; 5 female; mean age, 23.4 years; SD, 2.27 years). All procedures with human subjects were approved

by the Ethics Committee of the University Medical Center Göttingen (protocol no. 29/8/17). Subjects received monetary compensation for their participation.

In addition, two adult male rhesus monkeys (*Macaca mulatta*) participated in the study (at the time of testing: age: 7 years, monkey P; 13 years, monkey C; weight: 7.3 kg, monkey P; 8.4 kg, monkey C). The sample size matched that of earlier studies (Gamlin et al., 1998). Both animals had previously been implanted with cranial head posts under general anesthesia and aseptic conditions, for participation in neurophysiological experiments. The surgical procedures and purpose of these implants were previously described in detail (Dominguez-Vargas et al., 2017; Schwiedrzik and Freiwald, 2017). Animals were extensively trained with positive reinforcement (Prescott and Buchanan-Smith, 2003) to enter into and stay seated in a primate chair, and to have their head position stabilized via the head-post implant. This allows implant cleaning, precise recordings of gaze, and neurophysiological recordings while the animals work on cognitive tasks. Here, we made opportunistic use of these situations to record eye movement and PD data. The experimental procedures were approved by the responsible regional government office [Niedersächsisches Landesamt für Verbraucherschutz und Lebensmittelsicherheit (LAVES)]. The animals were pair housed or group housed in accordance with all applicable German and European regulations. The facility provides the animals with an enriched environment (including a multitude of toys and wooden structures), natural as well as artificial light, and access to outdoor space, exceeding the size requirements of European regulations. The psychological welfare and veterinary welfare of animals were monitored daily by veterinarians, animal facility staff, and scientists.

Stimuli and tasks

We used a set of images depicting human faces from a previous study (Schwiedrzik and Freiwald, 2017). In brief, we generated 36 three-dimensional human faces with a neutral expression and no hair in FaceGen (version 3.5.3; Singular Inversions). Images were converted to black and white, and the luminance histogram was equalized using SHINE (Willenbockel et al., 2010). For this study, we selected 18 unique images from the full set, each showing a different face (Schwiedrzik and Freiwald, 2017, their Supplemental Fig. 1). For humans, images (10×10 degree of visual angle (dva)) were presented foveally on an LCD monitor (ViewPixx EEG; refresh rate, 120 Hz; resolution, 1920×1080 pixels; viewing distance, 68 cm) in a darkened, sound-attenuating booth (Desone Modular Acoustics). The mean luminance of the face stimuli was 41.2 cd/m^2, and 30.75 cd/m^2 for the background. For monkeys, images (40×40 dva) were presented using a projector (Barco F22 WUXGA; refresh rate, 60 Hz; resolution, 1920×1080 pixels; viewing distance, 52 cm) in a darkened training setup. Stimulus delivery and response collection for human subjects were controlled using Presentation (version 19; Neurobehavioral Systems); visual stimulation and reward for monkey subjects were controlled using MWorks (https://mworks.github.io/). The experimenter was not blinded to the experimental conditions.

During the exposure phase, subjects viewed the faces in temporal sequence (Fig. 1). In the random condition, images were presented in random order for 250 ms each, with a 250 ms interstimulus interval (ISI). Human subjects were instructed to fixate on a blue fixation dot. To assure that subjects were paying attention to the images, they performed a 1-back repetition detection task in the main experiment [i.e., they had to report an infrequent (18 of 198/block) immediate repetition of an identical image by means of a button press on a standard keyboard]. Monkeys passively viewed the stimuli but were rewarded with juice or water if they continuously fixated on a red, centrally presented fixation dot for 2–4 s. Effective fixation accuracy was <2 dva [median, 95%; median absolute deviation (MAD), 1.96]. Human subjects were exposed to 1188 images in six blocks of 198 images within a single session. Monkey subjects were exposed to an average of 6000 images per session and completed 38 blocks (monkey P) and 46 blocks (monkey C) of 1200 images in which fixation accuracy was greater than or equal to 85%, respectively. Monkey P completed 19 blocks in the random condition and 19 blocks in the structured condition (see below); monkey C completed 17 blocks in the random condition and 29 blocks in the structured condition.

Stimuli and timing were identical in the structured condition, but, unbeknownst to the subjects, images were now joined into pairs. Pairs were arranged such that one identity–view combination would uniquely predict one other identity–view combination, while assuring that head orientation was fully balanced across pairs (e.g., three different identities at 0° head orientation were paired with three different identities at 0°, 60°, and 300° head orientation, respectively). To induce statistical structure, the sequence of pairs was arranged such that transition probabilities within pairs (i.e., between stimuli) were 100%, while transition probabilities between pairs (i.e., between trials) were at minimum and balanced across pairs. Human subjects were exposed to nine pairs, monkey subjects to three (BK, FO, GP; Schwiedrzik and Freiwald, 2017, their Supplemental Fig. 1). Subjects performed the same tasks as in the random condition. Because reward for monkey subjects was solely delivered on the basis of fixation performance, there was no systematic relationship between the occurrence of a pair and reward.

We performed three additional experiments to investigate the consequences of exposure to statistical structure and pupil entrainment. In experiment 1 [rapid serial visual presentation (RSVP) task], we tested whether human subjects retained the statistical structure they had been exposed to during the structured condition in a subsequent offline test. Specifically, subjects had to detect a target face in a RSVP stream of face images by means of a speeded button press (Turk-Browne et al., 2005). On each trial, we first presented the target image above the central fixation dot. The target was one of the second images from the nine face pairs. The subjects could then initiate the RSVP by a button press. The RSVP consisted of 18 face images presented at fixation and with the same timing as during the exposure phase (250 ms stimulus duration; 250 ms ISI). The target image could not appear as the first or last image in the sequence. As in the structured exposure condition, and again unbeknownst to the subjects, all images in the RSVP were presented as pairs. This served to assess whether the subjects had acquired and retained the statistics of the structured exposure phase. If so, we hypothesized that they could predict the occurrence of the second image in a pair once they saw the corresponding first image, and that this should facilitate face detection (Schwiedrzik and Freiwald, 2017). To test this hypothesis, we created two test conditions that violated the following exposed associations: in the "foil" condition, we replaced the first image in a pair that immediately preceded the target face during the RSVP either with a facial identity that had been shown during exposure but that had been paired with a different target, or with an image that showed the previously paired identity but with a different head orientation; and in the "novel" condition, we replaced the first image in the pair by a novel facial identity (but preserving the head orientation from the exposure phase). If subjects had acquired image-specific predictions about the order and identity of the face pairs during the structured exposure phase, then violating these expectations during the RSVP should lead to lower accuracy and slower reaction times during target detection relative to detecting a target that appears in the known configuration. Subjects completed 63 trials with pairs in the previously exposed configuration (seven per pair), 36 trials in the foil condition, and 18 trials in the novel condition, for a total of 117 trials. We retained a ratio of 1.2 trained over test conditions so as to not overwrite expectations stemming from the preceding structured exposure phase.

In experiment 2 ("card-sorting task"), we assessed whether subjects had awareness of the statistical structure. A new group of subjects ($n = 10$) underwent the same exposure to the random and structured stream as in the main experiment. Subsequently, they were informed that the structured stream had contained pairs and were asked to reproduce these pairs. To this end, subjects performed a card-sorting task in which they were given all 18 images shown during the exposure phase and were asked to sort them into nine pairs. Each image was printed at 13×13 cm. Subjects had 5 min to complete the task.

In experiment 3 ("disk detection task"), we assessed whether pupil entrainment to statistical structure affects visual detection performance in a new group of subjects ($n = 10$). To this end, we again presented random and structured streams of faces. To facilitate rapid learning and pupil entrainment already in the first block, we only used the three pairs we had also shown in the experiments with monkeys. Instead of the 1-back

repetition detection task, subjects now had to detect a small disk (0.22 dva) that was presented for 8 ms in the intervals between the images by means of a button press. The disk appeared at a random foveal location between 0.45 and 3.6 dva from the fixation point. In the structured condition, the disk could appear either between pairs (i.e., 8 ms before the first stimulus in a pair), or within pairs (i.e., 8 ms before the second stimulus in a pair). In the random condition (that was always acquired first), probes were presented at the exact same times as in the structured condition to assure that other factors such as time on task did not systematically affect comparisons. To assess visual sensitivity, we used a weighted up–down staircase procedure (Kaernbach, 1991), as follows: whenever subjects correctly detected the disk, we decreased its luminance by three RGB steps; whenever subjects missed a disk, we increased its luminance by 9 RGB steps. Staircases were run separately for within-pair and between-pair time points and were initialized at 91.65 cd/m^2 (descending) and 55.98 cd/m^2 (ascending), respectively, in alternating blocks. Subjects completed six blocks. Each block contained 97 pairs, 15 discs between pairs, and 15 discs within pairs. Discs could not appear before the fourth pair had been presented.

Pupil and gaze measurements

During all experiments, we continuously acquired pupil and gaze measurements using a high-speed, video-based eye tracker (EyeLink 1000 Plus, SR Research). Data were sampled at 1000 Hz from both eyes.

Data analysis

Pupil preprocessing. All data analyses were conducted in MATLAB (R2017b, MathWorks). We preprocessed pupil area data by first linearly interpolating blinks, otherwise missing values, as well as outliers exceeding 3.5× the MAD over the entire recording per eye per subject. Data were then low-pass filtered at 5 Hz using a one-pass/zero-phase, Kaiser-windowed sinc finite impulse response (FIR) filter (filter order, 1812; transition width, 2.0 Hz; pass band, 0–4.0 Hz; stop band, 6.0–500 Hz; maximal pass band deviation, 0.0010 (0.10%); stop band attenuation, −60 dB; Widmann et al., 2015). Subsequently, data were detrended per block and the mean per block as well as an average over the baseline of 2 s before the beginning of stimulation in each block were subtracted. Data were then averaged between eyes and finally downsampled to 500 Hz (keeping every second sample). We also corrected pupil data for gaze position (Brisson et al., 2013) before filtering, but this did not affect the pattern of results. We thus report uncorrected pupil data.

Pupil spectral power. For analyses of spectral power, we performed a Fourier transform per block using Welch's method (Welch, 1967), as implemented in the MATLAB function *pwelch.m*. We used a Hanning window of 16,866 points (one-third of the block length), 5622 overlapping points (one-third of the window length), and 8000 discrete DFT (discrete Fourier transform) points to obtain a spectral resolution of 0.0625 Hz. Before statistical analyses, blocks in which the preprocessed pupil signal exhibited large remaining artifacts or in which the power spectra were distributed abnormally were excluded (15.8% of human data; 9.5% of monkey data). The remaining power spectra were converted to decibels by taking the decadic logarithm and multiplying by 10.

In humans, the blockwise power spectra were averaged per subject and compared statistically across participants. In monkeys, statistical analyses were performed comparing blocks per animal. We specifically focused on the pair frequency at 1 Hz and the image frequency at 2 Hz. To statistically determine the existence of spectral peaks in the power spectra, we compared spectral power at 1 and 2 Hz, respectively, to the average of the four surrounding frequency bins (two above, two below), by means of a paired *t* test. For visualization, we plotted the respective mean differences and corresponding 95% Bayes-bootstrapped (1000 samples) high-density intervals using the Robust Statistical Toolbox (https://github.com/CPernet/Robust_Statistical_Toolbox) in MATLAB. To assess whether there were statistically significant differences between the random and the structured conditions and whether they were specific to the pair/image frequency, we conducted a repeated-measures ANOVA (rmANOVA) with frequency and condition as within-subjects factors in humans, and a mixed ANOVA with frequency as within-blocks factor and condition as the between-blocks factor in monkeys. In addition, we performed paired *t* tests between conditions at the pair and image frequency, respectively, in humans, and independent samples *t* tests in monkeys.

To assess whether pupil dilation differed after the presentation of the first and the presentation of the second stimulus in a pair, we compared average pupil dilation between 250 and 350 ms after the onset of the first and second stimulus during the random and the structured conditions, respectively, in humans using an rmANOVA. To assess whether pupil constriction differed within and between pairs, we compared the average pupil constriction between 500 and 600 ms after pair onset (within) to the average pupil constriction between 1000 and 1100 ms after pair onset (between) during the random and the structured conditions, respectively, in humans using an rmANOVA.

To compare the modulation of pupil dynamics at 1 Hz between species, we used a two-tailed, Bayesian standardized difference test (https://homepages.abdn.ac.uk/j.crawford/pages/dept/psychom.htm; Crawford et al., 2010) that allows the comparison of individual subjects, in our case individual monkeys, to a group of subjects, in our case the human observers, and estimates the probability that a more extreme score than the one observed in the respective individual subject is found in the group.

Pupil phase consistency. We also assessed whether there was consistent phase locking of the pupil signal to the pairs at 1 Hz. To this end, we cut the continuous data into pseudotrials of 16 s, time locked to the first stimulus in a pair, after the filtering preprocessing step, baseline-corrected each pseudotrial by subtracting an average over the 2 s preblock baseline, and averaged the signal between eyes. We then computed a trial-by-trial Fourier transform using discrete prolate spheroidal sequences (DPSSs) as tapers with the same spectral resolution as for the analyses of spectral power (0.0625 Hz), using the MATLAB toolbox Fieldtrip (version 20170327; http://www.fieldtriptoolbox.org/; Oostenveld et al., 2011). The resulting complex spectra were then used to calculate inter-trial phase coherence (ITC), as follows:

$$\mathrm{ITC} = \left| n^{-1} \sum_{r=1}^{n} e^{ik_{fr}} \right|,$$

where n is the number of trials, and k is the phase angle on trial r at the frequency f. ITC reflects the degree to which the phase angle of an oscillation at any given time relative to a stimulus is consistent across trials. ITC ranges between 0 and 1, with 0 indicating uniformly distributed phase angles, and 1 indicating identical phase angles across trials. Statistical comparisons between conditions were conducted using paired *t* tests in humans, and independent-samples *t* tests in monkeys.

Eye movements. The same preprocessing and analysis steps as for PD were conducted for the horizontal and vertical eye position signals. In addition, we computed saccade rates, time locked to the onset of the first stimulus on a pair. To this end, we detected saccades in the continuous eye signal using the built-in Eyelink algorithm, smoothed the resulting traces with a moving average filter (width 120 ms), and scaled them to the rate per second. We then compared the random and the structured conditions from 250 ms before to 1000 ms after the onset of the first stimulus using paired *t* tests, followed by correction of multiple comparisons using the false discovery rate (FDR) at $q = 0.05$ (Benjamini and Yekutieli, 2001). For spectral analyses of saccade rates, we extracted pseudotrials of 16 s (as for the pupil phase consistency analyses) and computed a Fourier transform using DPSS tapers (spectral resolution, 0.0625 Hz). Statistical tests were the same as for pupil spectral power analyses.

Task performance. To analyze accuracy and reaction times of the 1-back and RSVP tasks (experiment 1), we first excluded trials with outliers in the reaction times per subject using the estimator *Sn* (Rousseeuw and Croux, 1993) at a threshold of 2, as well as trials in which reaction time exceeded 2.5 s. Average accuracy and reaction times per condition (exposed, foil, novel) were then compared by means of planned paired *t* tests. To analyze detection performance in the disk detection task (experiment 3), we averaged reversals per block and condition, and then calculated the Weber fraction as a measure of required luminance

contrast to detect the disk against the uniform background. Weber fractions were compared using an rmANOVA and planned paired t tests. For visualization, we plotted means and corresponding 95% Bayes-bootstrapped (1000 samples) high-density intervals from the RSVP and disk detection tasks, respectively, using the Robust Statistical Toolbox (https://github.com/CPernet/Robust_Statistical_Toolbox) in MATLAB.

Pupil–task correlations. To assess the relationship between pupil entrainment during the exposure phase and accuracy in the subsequent test phase of experiment 1, we correlated accuracy in the RSVP task with the normalized spectral power at the 1 Hz pair frequency in the random and structured conditions, respectively, using Pearson's correlation coefficient. To this end, we first normalized the 1 Hz peak by subtracting the average spectral power of four surrounding frequency bins (two above, two below) within each condition. This estimates the distinctiveness of the peak at 1 Hz relative to the surrounding spectrum. To assess the specificity of the correlation between 1 Hz spectral power in the structured condition and accuracy, we further performed partial correlation analyses, controlling for any correlation between 1 Hz spectral power in the random condition and accuracy, again using Pearson's correlation coefficient. Analyses of (partial) correlations using Spearman's correlation coefficient yielded the same pattern of results.

Effect sizes. For all t tests, we computed Hedges' g and for all ANOVA partial η^2 as measures of effect size using the MATLAB toolbox Measures of Effect Size (version 1.6.1; https://github.com/hhentschke/measures-of-effect-size-toolbox/; Hentschke and Stüttgen, 2011).

Data availability

All data (PD, gaze, and behavior) are publicly available on Figshare (Schwiedrzik and Sudmann, 2019).

Results

To test whether PD is sensitive to temporal environmental statistics, we presented long sequences (≥2 min) of computer-generated, luminance-equalized face stimuli varying in facial identity and head orientation against a gray background (Fig. 1*A*). Each face was shown for 250 ms, followed by a 250 ms gap. Such alternation in light intensity should induce pupil cycling at 2 Hz, the image rate. This frequency is well within the range of pupillary constriction dynamics (Alexandridis and Manner, 1977). In the random condition, the images were presented at a fixed rate but in random order. In the structured condition, we presented the same images at the same rate, but we arranged them such that they were systematically grouped into pairs (i.e., the transitional probabilities between specific images were fixed at 100%, while the transitional probabilities between pairs were minimized). Thus, in addition to the individual images, this pairing gave rise to temporally coherent units at half the image rate. If this statistical structure was extracted, it should be reflected by a modulation of the pupil at 1 Hz, the pair rate. To draw attention to the stimuli, but not explicitly to their temporal structure (Turk-Browne et al., 2005), human subjects were instructed to perform a 1-back task during both conditions in which they had to detect infrequent repetitions of identical faces. Performance in the task was above chance (random condition: mean accuracy, 59.97%; $t_{(29)} = 4.786$; $p < 0.001$; $g = 0.851$; structured condition: mean accuracy, 59.72%; $t_{(29)} = 3.147$; $p = 0.004$; $g = 0.560$) and did not differ between conditions (mean difference, 0.25%; $t_{(29)} = 0.095$; $p = 0.925$; $g = 0.017$).

Main experiment: humans

We first present the human data. Pupil cycling at the 2 Hz image rate was clearly visible in the pupil signal on the single-subject level (Fig. 1*B*). Spectral analyses of the continuous data showed a distinct peak in spectral power at the image rate (2 Hz) in both the structured ($t_{(29)} = 15.018$, $p < 0.001$, $g = 2.314$) and the random ($t_{(29)} = 13.494$, $p < 0.001$, $g = 2.134$) conditions compared with the four surrounding frequency bins (Fig. 2). In addition, we find a strong response at 1 Hz, the pair rate, but only in the structured condition (interaction of condition × frequency: $F_{(1,29)} = 6.114$, $p = 0.019$, $\eta^2 = 0.174$; structured vs random conditions: mean difference, 1.493 dB; $t_{(29)} = 3.451$; $p = 0.002$; $g = 0.449$). In fact, there was no distinct peak at 1 Hz in the random condition compared with the surrounding bins ($t_{(29)} = 0.519$, $p = 0.607$, $g = 0.038$). A complementary analysis of the pupillary phase revealed significantly stronger phase locking at 1 Hz in the structured condition than in the random condition (mean difference, 0.151; $t_{(29)} = 5.326$; $p < 0.001$; $g = 1.447$). Thus, the human pupil can track environmental statistics that are not directly evident in the light pattern that hits the retina.

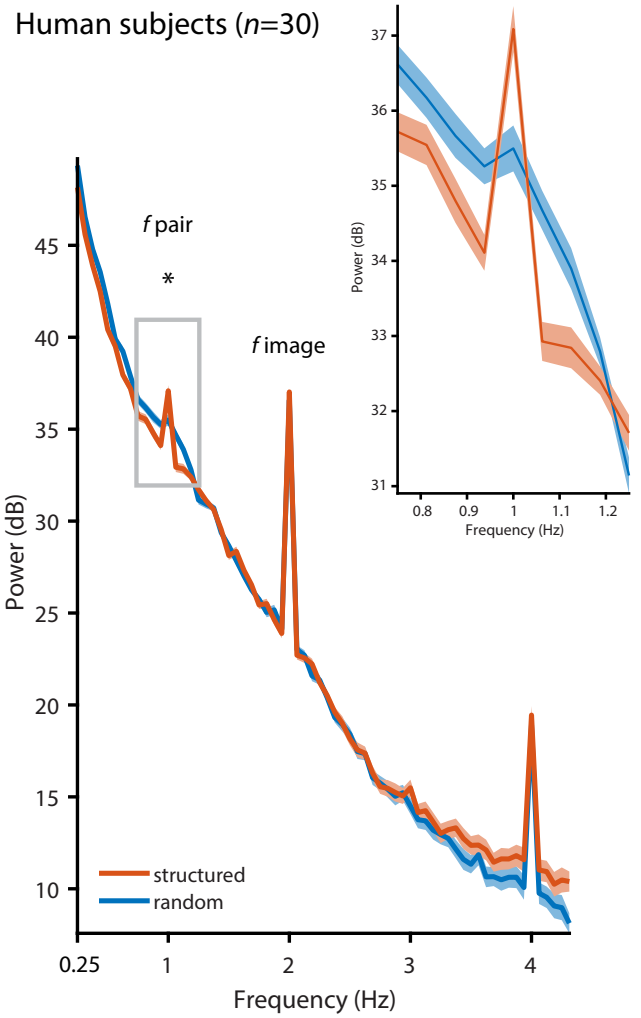

Figure 2. Pupil diameter in humans. Pupil diameter in human subjects ($n = 30$) followed the image rate at 2 Hz in the random condition (blue, $t_{(29)} = 13.494$, $p < 0.001$, $g = 2.140$) and the structured condition (orange, $t_{(29)} = 15.018$, $p < 0.001$, $g = 2.314$; compared with the mean of the four surrounding frequencies), but the pair rate at 1 Hz was only evident in the structured condition (structured vs random: $t_{(29)} = 3.451$, $p = 0.002$, $g = 0.449$). The inset shows a zoomed version of the power spectrum at ~1 Hz. Peaks at 4 Hz likely reflect harmonics. Shading represents the SEM across subjects, corrected for intersubject variability (Morey, 2008). Asterisk indicates statistical significance at $p < 0.05$ for the contrast structured vs. random condition.

Control analyses showed that the 1 Hz peak was not evident in the concurrently recorded eye movement signal in either condition (Fig. 3*A*), and that there were no correlations between 1 Hz spectral power in the pupil and the eye movement signal

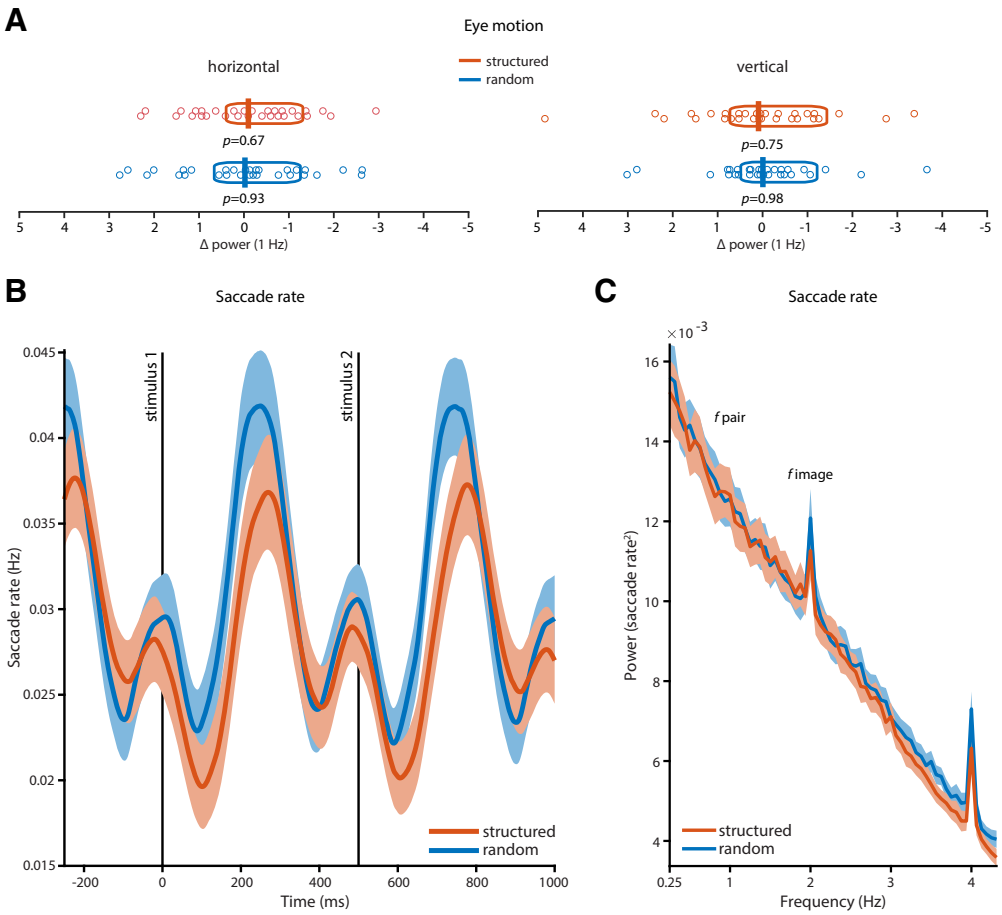

Figure 3. Eye motion in humans. ***A***, There were no statistically significant peaks at 1 Hz in the dynamics of horizontal or vertical eye position in the random or structured condition in human subjects ($n = 30$) relative to the four surrounding frequency bins. Vertical bars indicate the mean; boxes indicate the 95% Bayes-bootstrapped high-density interval; circles indicate the individual subjects' data points; and p values are from paired, two-sided t tests of 1 Hz power against the average of four surrounding frequency bins (all $p > 0.6$). ***B***, Saccade rates time locked to the onset of the first stimulus in a pair showed no statistically significant differences between the structured and the random condition on a time point-by-time point basis (all $p > 0.05$, FDR corrected). ***C***, There was no entrainment of saccade rates at the pair rate (1 Hz); i.e., there was no distinct peak at 1 Hz compared with the surrounding bins in the structured condition ($t_{(29)} = 1.7532$, $p = 0.0901$, $g = 0.0434$) or in the random condition ($t_{(29)} = 0.6698$, $p = 0.5083$, $g = 0.0195$). Power at 1 Hz also did not differ between the structured and random conditions (mean difference, −0.0001; $t_{(29)} = -0.1042$; $p = 0.9178$; $g = -0.0131$). Shading in ***B*** and ***C*** represents the SEM across subjects, corrected for intersubject variability (Morey, 2008).

(horizontal: $r = -0.264$, $p = 0.159$; vertical: $r = 0.117$, $p = 0.539$). We also specifically investigated whether the frequency of saccades (i.e., saccade rates) followed environmental statistics. Saccade rates throughout the experiments were very low (mean rate structured, 0.036 saccades/s; mean rate random, 0.033 saccades/s) and did not differ between conditions (mean difference, 0.002 saccades/s; $t_{(29)} = 1.5230$; $p = 0.1386$; $g = 0.3627$). Residual saccade rates time locked to the onset of the first stimulus in a pair showed a typical pattern of saccades (Reingold and Stampe, 2002; Engbert and Kliegl, 2003) with ~200 ms of saccadic inhibition after stimulus onset, followed by a brief rebound (Fig. 3*B*). When we compared saccade rates between the structured and the random conditions on a time point-by-time point basis, we found no significant differences between conditions (all $p > 0.05$, FDR corrected). We also found no entrainment of saccade rates at the pair rate (1 Hz; Fig. 3*C*): there was no distinct peak at 1 Hz compared with the surrounding bins in the structured condition ($t_{(29)} = 1.7532$, $p = 0.0901$, $g = 0.0434$) or in the random condition ($t_{(29)} = 0.6698$, $p = 0.5083$, $g = 0.0195$), and power at 1 Hz did not differ significantly between conditions (mean difference, −0.0001; $t_{(29)} = -0.1042$; $p = 0.9178$; $g = -0.0131$). Finally, we performed spectral analyses of PD only on saccade-free intervals, which was possible in 19 of 30 subjects (63.3%). As in our original analyses, the analyses of saccade-free data segments show a clear peak at 1 Hz relative to the surrounding bins in the structured condition ($t_{(18)} = 4.6076$, $p = 0.0002$, $g = 0.8179$), and significantly higher spectral power at 1 Hz in the structured condition than in the random condition (mean difference, 1.2497; $t_{(18)} = 2.6560$; $p = 0.0161$; $g = 0.6525$). Together, this rules out that pupil entrainment at this frequency was an artifact arising from other eye movements and blinks that may occur at a similar rate.

To determine which aspects of pupillary movement contribute to the dynamics at 1 Hz in the structured condition, we examined pupil dilations and constrictions, respectively, time locked to the onset of the pairs. Pupil dilation can be elicited by surprising stimuli under constant illumination (Preuschoff et al., 2011). In our paradigm, once the statistics of the structured stream have been acquired, they may render stimuli more or less predictable/surprising. Specifically, the second stimulus in a pair is entirely predictable from the first stimulus. In contrast, the first stimulus in a pair may be surprising as it cannot be predicted from its predecessor. Hence, transient pupil dilations to every first stimulus reflecting surprise could contribute to spectral power at 1 Hz. We find, however, that pupil dilation following the first stimulus was smaller than following the second stimulus

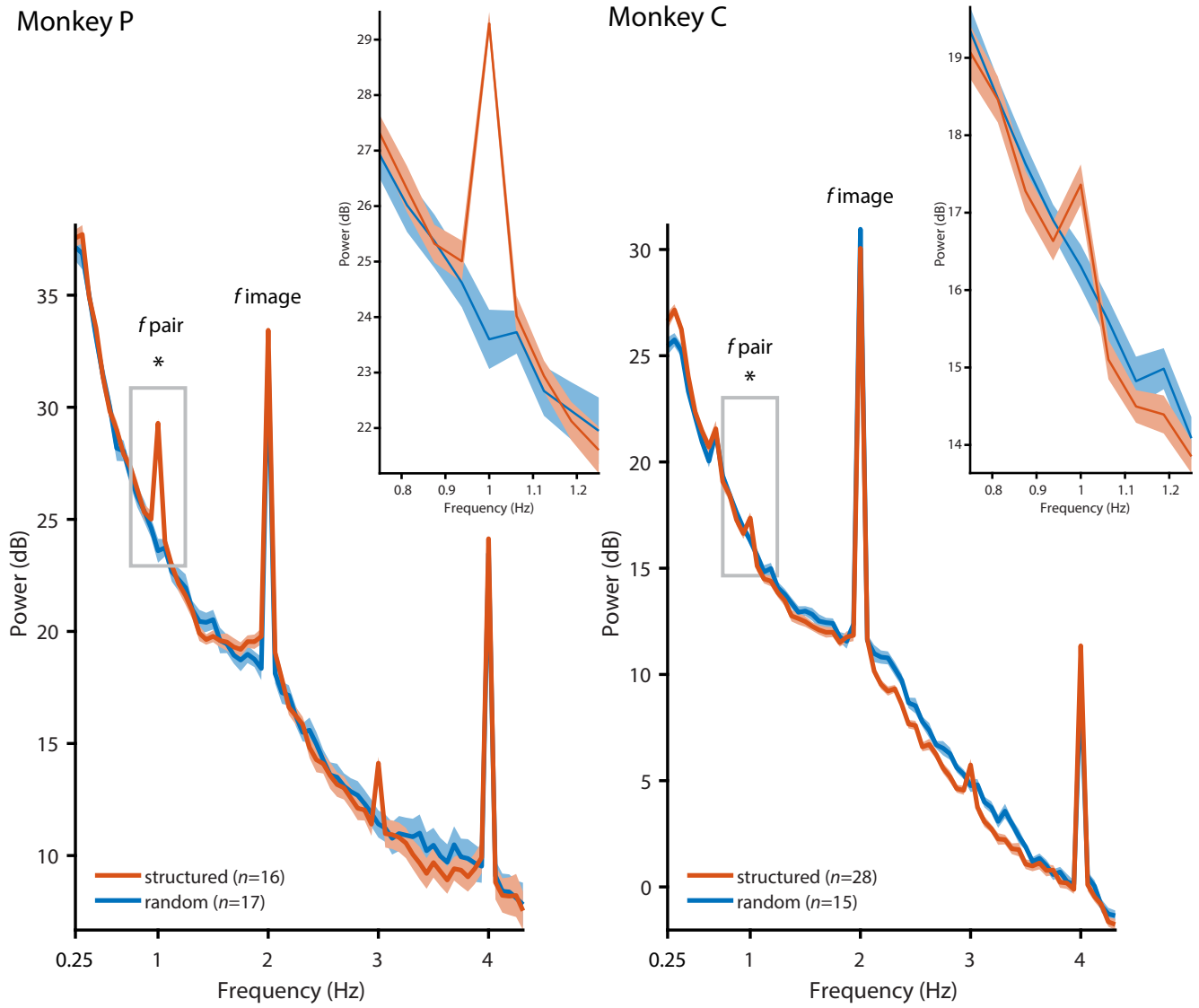

Figure 4. Pupil diameter in monkeys. Pupil diameter in monkeys followed the image rate at 2 Hz in the random (blue) and the structured (orange) conditions, but the pair rate at 1 Hz was only evident in the structured condition (structured vs random: monkey P: $t_{(31)} = 10.37$, $p < 0.001$, $g = 3.524$; monkey C: $t_{(41)} = 2.327$, $p = 0.025$, $g = 0.731$). The inset show a zoomed-in version of the power spectrum at ~1 Hz. Peaks at 3 and 4 Hz likely reflect harmonics of the 1 and 2 Hz response, respectively. Shading represents the SEM across sessions. Asterisk indicates statistical significance at $p < 0.05$ for the contrast structured vs. random condition.

(in the structured condition only: interaction condition × time point: $F_{(1,29)} = 18.526$, $p < 0.001$, $\eta^2 = 0.390$). This suggests that in our paradigm, pupil dilation at 1 Hz reflects statistical structure of the input stream but is not (solely) driven by surprise.

In addition, Figure 1, *B* and *C*, shows how pupil constriction may contribute to the 1 Hz response in the structured condition: pupil constriction between images is suppressed within a pair, while pupil constriction between pairs appears particularly pronounced. Indeed, across subjects, mean pupil constriction was significantly smaller within pairs than between pairs only during the structured condition (interaction condition × time point: $F_{(1,29)} = 14.043$, $p < 0.001$, $\eta^2 = 0.326$). This suggests that the constrictive PLR within pairs is diminished as if not to interrupt information transmission within an environmentally coherent unit, and augmented between pairs, possibly to segment the continuous input stream at the pair rate. Together, this shows that both dilation and constriction dynamics contribute to the 1 Hz response in the structured condition, with the net effect of an overall wider pupil within a pair than between pairs.

Main experiment: monkeys

We ran the same experiments with macaque monkeys, albeit with a reduced number of stimulus pairs (three instead of 9). Instead of performing the 1-back task, monkeys passively viewed the stimulus sequences and were rewarded for maintaining fixation. Like in humans, we find in both monkeys a PLR at 2 Hz in the structured and the random conditions, but a response at 1 Hz only when there is statistical structure (Fig. 4; condition × frequency interaction: monkey P: $F_{(1,31)} = 69.610$, $p < 0.001$, $\eta^2 = 0.692$; monkey C: $F_{(1,41)} = 18.216$, $p = 0.001$, $\eta^2 = 0.308$). The strength of the modulation at 1 Hz, relative to the surrounding frequencies, in both monkeys (mean difference: 4.97 dB in monkey P; 1.25 dB in monkey C) fell well into the range we observed in humans (minimum, 0.15 dB; maximum, 6.32 dB). Bayesian statistics showed that the probability that the difference for human subjects would be more extreme than in our monkey subjects was only 5.12% (monkey P) and 25.51% (monkey C), suggesting that monkeys and humans showed similar pupillary dynamics. As in humans, both monkeys also showed stronger

phase locking at 1 Hz in the structured condition than in the random condition when we analyzed pupillary phase (monkey P: $t_{(31)} = 13.361$, $p < 0.001$, $g = 4.540$; monkey C: $t_{(41)} = 5.582$, $p < 0.001$, $g = 1.753$). There were no significant differences between conditions in other eye movements (Fig. 5). Thus, the cognitive modulation of PD based on environmental statistics that we found in humans is shared with macaques.

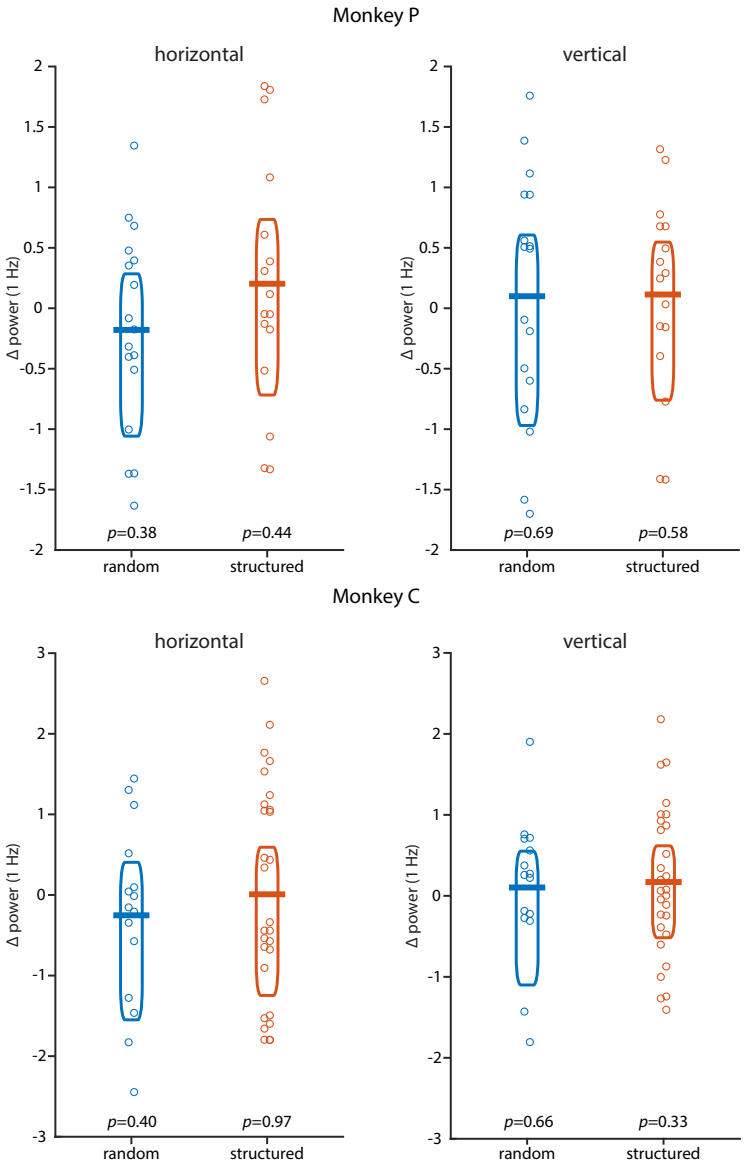

Figure 5. Eye motion in monkeys. There were no statistically significant peaks at 1 Hz in the dynamics of horizontal or vertical eye position in the random or structured condition in monkey P and monkey C relative to the four surrounding frequency bins. Horizontal bars indicate the mean; boxes indicate the 95% Bayes-bootstrapped high-density interval; circles indicate the data points of individual runs; *p* values are from paired, two-sided *t* tests of 1 Hz power against the average of four surrounding frequency bins (all $p > 0.3$).

Experiment 1: RSVP task

To assess whether subjects indeed internalized models of the environment, human subjects also performed an additional task in which the statistics were directly relevant for task performance. Specifically, after watching the random sequences and subsequently the structured sequences, subjects performed a new task in which they had to detect a face stimulus in a RSVP sequence of other faces (Fig. 6*A*). In a majority of the trials, the target face was immediately preceded by the face with which it was paired during the structured exposure phase (the predictor) to induce a priming effect. For comparison, we created two additional test conditions that violated the previously exposed structure: the foil condition, in which we replaced the predictor with a facial identity or head orientation that had been shown during exposure but that had been paired with a different target; and the novel condition, in which we replaced the predictor with a novel facial identity. We hypothesized that if subjects learned the order and identity of the face pairs during the structured exposure phase, then violating learned associations during the RSVP should lead to lower accuracy and slower reaction times during target detection relative to detecting a target that appears in the known configuration.

Indeed, subjects were significantly faster (exposed vs foil: mean difference, −13.692 ms; $t_{(29)} = -4.522$; $p < 0.001$; $g = -0.266$; exposed vs novel: mean difference, −21.344 ms; $t_{(29)} = -5.464$; $p < 0.001$; $g = -0.403$) and more accurate (exposed vs foil: mean difference, 3.17%; $t_{(29)} = 2.928$; $p = 0.007$; $g = 0.376$; exposed vs novel: mean difference, 9.11%; $t_{(29)} = 4.715$; $p < 0.001$; $g = 0.962$) in detecting targets in the exposed condition than in the foil and the novel conditions (Fig. 6*B*), showing that they indeed extracted and retained the statistics of the face sequences from the exposure phase. We then correlated the performance differences in accuracy from the RSVP task with the degree of pupil entrainment at the 1 Hz pair rate (relative to the four surrounding frequency bins; see Materials and Methods) during the preceding exposure phase (Fig. 6*C*). We find that the stronger the pupil entrainment during exposure, the larger the accuracy benefit of the exposed over the test conditions (exposed vs foil: $r = 0.481$, $p = 0.007$; exposed vs novel: $r = 0.402$, $p = 0.027$) in the RSVP task. The same held true when controlling for any correlation between 1 Hz spectral power in the random condition and accuracy (partial correlation; exposed vs foil: $r = 0.450$, $p = 0.014$; exposed vs novel: $r = 0.412$, $p = 0.026$). Hence, pupil entrainment during exposure predicts subsequent detection performance, suggesting that they are based on the same internal models.

Experiment 2: card-sorting task

Only certain internal models (or components thereof) may be available to awareness (Blakemore et al., 2002). To assess whether subjects developed awareness of the statistical structure, a new group of participants ($n = 10$) was exposed to the random and the structured stimulus streams, and then performed a card-sorting task in which they had to reproduce the stimulus pairs they had been exposed to. Performance in the 1-back task during the

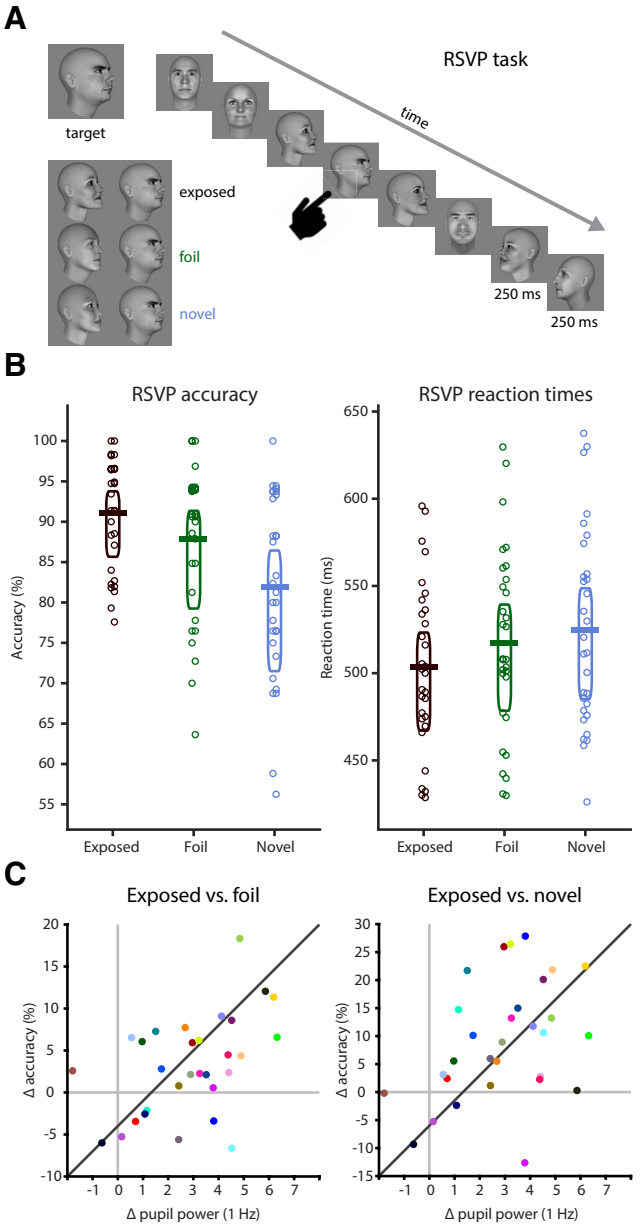

Figure 6. Offline learning test and relation to pupil entrainment. ***A***, On each trial of the RSVP task, subjects had to detect a target face. Targets were embedded into a stream of faces (250 ms on, 250 ms off). The face immediately preceding the target could be the previously exposed predictor ("exposed" condition, black), an identity or a head orientation that had been paired with another face during the exposure phase (foil condition, green), or a completely novel identity (novel condition, blue). ***B***, Human subjects were more accurate (exposed vs foil: $t_{(29)} = 2.928$, $p = 0.007$, $g = 0.376$; exposed vs novel: $t_{(29)} = 4.715$, $p < 0.001$, $g = 0.962$) and faster (exposed vs foil: $t_{(29)} = -4.522$, $p < 0.001$, $g = -0.266$; exposed vs novel: $t_{(29)} = -5.464$, $p < 0.001$, $g = -0.403$) in detecting the target when it was preceded by the previously exposed predictor than in the foil or the novel conditions. Horizontal bars indicate the mean, boxes indicate the 95% Bayes-bootstrapped high-density interval, and circles indicate the individual subjects' data points ($n = 30$). ***C***, Pupil entrainment at the 1 Hz pair rate during the exposure phase (normalized against the mean of the four surrounding frequencies) predicted offline learning effects: the stronger the entrainment at 1 Hz, the larger the accuracy benefit in the exposed over the foil condition ($r = 0.481$, $p = 0.007$) and the novel condition ($r = 0.402$, $p = 0.027$).

exposure phase was similar to that in the main experiment (random condition: mean accuracy, 58.06%; $t_{(9)} = 2.141$; $p = 0.061$; $g = 0.619$; structured condition: mean accuracy, 57.41%; $t_{(9)} = 2.402$; $p = 0.039$; $g = 0.695$; random vs structured condition: mean difference, 0.65%, $t_{(9)} = 0.316$; $p = 0.759$; $g = 0.057$). We also fully replicated the pupil entrainment effects of our main experiment in this smaller sample. Specifically, there was a clear peak at 1 Hz, the pair rate, in the structured condition (relative to the four surrounding frequency bins: $t_{(9)} = 4.868$, $p < 0.001$, $g = 1.207$) but not in the random condition ($t_{(9)} = -1.369$, $p = 0.204$, $g = -0.395$). The difference between the random and the structured conditions was statistically significant (mean difference, 3.419 dB; $t_{(9)} = 4.101$; $p = 0.003$; $g = 1.325$) and specific to 1 Hz (interaction condition × frequency: $F_{(1,9)} = 7.866$, $p = 0.021$, $\eta^2 = 0.466$). After the exposure phase, subjects were informed that the structured stream had contained pairs and were asked to sort the individual pictures into pairs. No subject reproduced more than one pair, and 5 of 10 subjects reproduced no pair. This suggests that pupil entrainment to statistical structure can occur in the absence of an explicit awareness of this structure.

Experiment 3: disk detection task

If PD entrains to environmental statistics to optimize the transmission of visual information about pairs, this may also have direct perceptual consequences. We thus tested in a new group of human subjects ($n = 10$) whether visual detection performance is affected by pupillary entrainment to higher-order temporal statistics. Subjects had to detect a small, briefly flashed disk in their foveal visual field during structured and random sequences, respectively (Fig. 7*A*). During the structured sequence, we presented the disk at times either falling between pairs or within pairs. During the random sequence, the disk was presented at identical times, but in the absence of statistical structure. If pupil entrainment at the pair rate (1 Hz) affected information transmission, then the disk should be detected more readily in the structured sequence within a pair (i.e., when the pupil is relatively more dilated) than at the identical time point in the random condition, where no 1 Hz entrainment takes place (Fig. 1*C*). Indeed, the luminance contrast required to detect the disk was significantly lower in the structured than in the random condition when the disk was presented within a pair (Fig. 7*B*; condition × time point interaction: $F_{(1,9)} = 6.943$, $p = 0.027$, $\eta^2 = 0.435$). Hence, pupil entrainment at temporal frequencies reflecting higher-order environmental statistics directly affects visual sensitivity.

Discussion

We find that the peripheral oculomotor system is involved in actively tracking environmental statistics, in line with an active sensing account. The modulation of PD is conserved between monkeys and humans, whose evolutionary lineage split from ours some 25 million years ago (Kumar and Hedges, 1998). This is not a given, as other modulations of PD show systematic differences between the two species that point to differences in the underlying anatomy and/or neural pathways involved (Gamlin et al., 1998). In both species, the pupil entrains to higher-order structure in the visual input, which leads to an adaptive correlation between visual ecology and response dynamics, as follows: the visuomotor system matches the dilation–constriction dynamics of the pupil to the temporal structure of the environment. This involves an increase in pupil dilation for predictable stimuli, and a reduction of pupil constriction within a pair. The result is an overall wider pupil during the presentation of environmentally coherent units. This maximizes information transmission for downstream visual processing, as evidenced by higher sensitivity to visual information within than between pairs. In

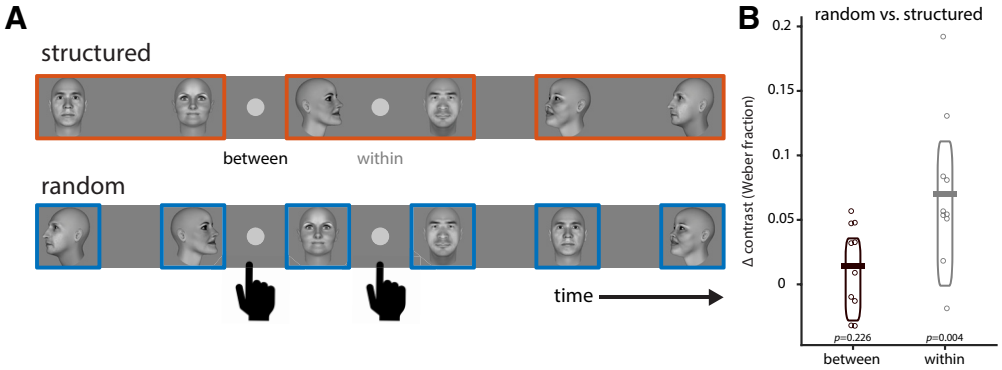

Figure 7. Effects of pupil entrainment to higher-order environmental structure on visual sensitivity. ***A***, Subjects ($n = 10$) had to detect a small, foveally presented disk that could appear either within or between pairs in the structured condition, or at temporally matched time points in the random condition (in the absence of statistical structure and 1 Hz pupil entrainment). The luminance contrast level required to detect the disk was used as a measure of visual sensitivity. ***B***, Luminance contrast required to detect the discs was significantly higher between pairs than within pairs, particularly in the structured condition (condition × time point interaction: $F_{(1,9)} = 6.943$, $p = 0.027$, $\eta^2 = 0.435$). There was no significant difference in luminance contrast between the random and the structured conditions between pairs ($t_{(9)} = 1.298$, $p = 0.226$, $g = 0.306$), but only within pairs ($t_{(9)} = 3.818$, $p = 0.004$, $g = 0.654$), reflecting wider pupil diameter as a result of 1 Hz entrainment in the structured condition. Horizontal bars indicate the mean, boxes indicate the 95% Bayes-bootstrapped high-density interval, and circles indicate the individual subjects' data points ($n = 10$).

addition, we find an augmented PLR between pairs, akin to an active segmentation of the continuous input stream at event boundaries between coherent units. Together, this resembles an adaptive filter in the time domain that optimizes the signal-to-noise ratio at ecologically relevant frequencies.

Our findings imply that pupil-controlling pathways in the brainstem are under control of (or at least influenced by) brain areas that can extract environmental statistics and translate them into useful models for active sampling already at the very first stage of visual interaction with the outside world. Previous research has shown that PD is modulated by other factors than light influx, which likely arise from cortical and/or higher-order thalamic computations [e.g., attention (Daniels et al., 2012; Naber et al., 2013), motion coherence (Barbur et al., 1992), environmental volatility (Nassar et al., 2012; Browning et al., 2015; Muller et al., 2019), complex auditory patterns (Barczak et al., 2018), and even speech (Jin et al., 2018)]. We show here that cognitively induced pupil dynamics directly influence visual sensitivity, which suggests that the temporal modulation is not solely a reflection of the sensitivity of the brain to these parameters but is actively used for information processing.

One possible source of PD modulation is extrastriate visual cortex, where the same paradigm we used here affects neural processing of facial information (Schwiedrzik and Freiwald, 2017) and where lesions affect PD modulation by other visual features like color (Heywood et al., 1998). Statistical structure also entrains brain areas beyond visual cortex, including motor and premotor cortex (Henin et al., 2019). Here, stimulation (e.g., of the frontal eye fields) induces pupil size changes in macaque monkeys (Lehmann and Corneil, 2016; Ebitz and Moore, 2017). Alternatively or additionally, modulation of the pupil may be mediated by the superior colliculus (Wang and Munoz, 2015) or the locus ceruleus (Joshi et al., 2016), whose activity is also known to covary with PD. In this context, it is interesting that pupil cycling is primarily under parasympathetic control and is doubly dissociated from single light pulse-induced pupil constrictions in patients with lesions in the pupil reflex pathways (Martyn and Ewing, 1986; Milton et al., 1988). This suggests that the effects we found here may be dissociable from the often discussed noradrenergic effects on PD (Reimer et al., 2016) and may explain why the modulation of PD during cycling was not easily accounted for by surprise.

Interestingly, PD entrains to higher-order environmental statistics even if subjects are unaware of them, as evidenced by their failure to reproduce the pairs in the card-sorting task. Our results are in line with a recent study showing that PD phasically increases in response to stimuli that violate statistical environmental structure, although subjects are unaware of this structure (Alamia et al., 2019). Thus, evidence from phasic pupil responses as much as our results on pupil cycling support the notion that unconscious processes can guide attention and inference. More so, our results suggest that internal models for active sensing need not be consciously available, at least in the case of an involuntary movement like the adjustment of PD.

Finally, on a technical note, pupil cycling presents itself as an easy-to-obtain, high signal-to-noise behavioral readout of learning models from the environment, possibly permitting the assessment of sensitivity to statistical structure across levels of complexity among individuals, health status, stages of development, or species.

References

Alamia A, VanRullen R, Pasqualotto E, Mouraux A, Zenon A (2019) Pupil-linked arousal responds to unconscious surprisal. J Neurosci 39:5369–5376.

Alexandridis E, Manner M (1977) Folgefrequenz der Pupille bei flimmernden Lichtreizen [Frequency of the pupillary response following flicker stimuli (Author's Transl)]. Albrecht Von Graefes Arch Klin Exp Ophthalmol 202:175–180.

Barbur JL, Harlow AJ, Sahraie A (1992) Pupillary responses to stimulus structure, colour and movement. Ophthalmic Physiol Opt 12:137–141.

Barczak A, O'Connell MN, McGinnis T, Ross D, Mowery T, Falchier A, Lakatos P (2018) Top-down, contextual entrainment of neuronal oscillations in the auditory thalamocortical circuit. Proc Natl Acad Sci U S A 115:E7605–E7614.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188.

Billock VA, de Guzman GC, Kelso JAS (2001) Fractal time and 1/f spectra in dynamic images and human vision. Physica D 148:136–146.

Blakemore SJ, Wolpert DM, Frith CD (2002) Abnormalities in the awareness of action. Trends Cogn Sci 6:237–242.

Brisson J, Mainville M, Mailloux D, Beaulieu C, Serres J, Sirois S (2013) Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. Behav Res Methods 45:1322–1331.

Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015) Anxious individuals have difficulty learning the causal statistics of aversive environments. Nat Neurosci 18:590–596.

Crawford JR, Garthwaite PH, Porter S (2010) Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. Cogn Neuropsychol 27:245–260.

Daniels LB, Nichols DF, Seifert MS, Hock HS (2012) Changes in pupil diameter entrained by cortically initiated changes in attention. Vis Neurosci 29:131–142.

Dominguez-Vargas AU, Schneider L, Wilke M, Kagan I (2017) Electrical microstimulation of the pulvinar biases saccade choices and reaction times in a time-dependent manner. J Neurosci 37:2234–2257.

Douglas RH (2018) The pupillary light responses of animals: a review of their distribution, dynamics, mechanisms and functions. Prog Retin Eye Res 66:17–48.

Ebitz RB, Moore T (2017) Selective modulation of the pupil light reflex by microstimulation of prefrontal cortex. J Neurosci 37:5008–5018.

Ebitz RB, Moore T (2018) Both a gauge and a filter: cognitive modulations of pupil size. Front Neurol 9:1190.

Engbert R, Kliegl R (2003) Microsaccades uncover the orientation of covert attention. Vision Res 43:1035–1045.

Friston K, Adams RA, Perrinet L, Breakspear M (2012) Perceptions as hypotheses: saccades as experiments. Front Psychol 3:151.

Gamlin PD, Zhang H, Harlow A, Barbur JL (1998) Pupil responses to stimulus color, structure and light flux increments in the rhesus monkey. Vision Res 38:3353–3358.

Haynes JD, Lotto RB, Rees G (2004) Responses of human visual cortex to uniform surfaces. Proc Natl Acad Sci U S A 101:4286–4291.

Henin S, Turk-Browne N, Friedman D, Liu A, Dugan P, Flinker A, Doyle W, Devinsky O, Melloni L (2019) Statistical learning shapes neural sequence representations. bioRxiv. Advance online publication. Retrieved May 7, 2020. doi: 10.1101/583856.

Hentschke H, Stüttgen MC (2011) Computation of measures of effect size for neuroscience data sets. Eur J Neurosci 34:1887–1894.

Heywood CA, Nicholas JJ, LeMare C, Cowey A (1998) The effect of lesions to cortical areas V4 or AIT on pupillary responses to chromatic and achromatic stimuli in monkeys. Exp Brain Res 122:475–480.

Jin P, Zou J, Zhou T, Ding N (2018) Eye activity tracks task-relevant structures during speech and auditory sequence perception. Nat Commun 9:5374.

Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89:221–234.

Kaernbach C (1991) Simple adaptive testing with the weighted up-down method. Percept Psychophys 49:227–229.

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917–920.

Laughlin SB (1992) Retinal information capacity and the function of the pupil. Ophthalmic Physiol Opt 12:161–164.

Lehmann SJ, Corneil BD (2016) Transient pupil dilation after subsaccadic microstimulation of primate frontal eye fields. J Neurosci 36:3765–3776.

Martyn CN, Ewing DJ (1986) Pupil cycle time: a simple way of measuring an autonomic reflex. J Neurol Neurosurg Psychiatry 49:771–774.

Mathôt S, Van der Stigchel S (2015) New light on the mind's eye: the pupillary light response as active vision. Curr Dir Psychol Sci 24:374–378.

McDougal DH, Gamlin PD (2015) Autonomic control of the eye. Compr Physiol 5:439–473.

Miller SD, Thompson HS (1978) Edge-light pupil cycle time. Br J Ophthalmol 62:495–500.

Milton JG, Longtin A, Kirkham TH, Francis GS (1988) Irregular pupil cycling as a characteristic abnormality in patients with demyelinative optic neuropathy. Am J Ophthalmol 105:402–407.

Morey RD (2008) Confidence intervals from normalized data: a correction to Cousineau (2005) Tutor Quant Methods Psychol 4:61–64.

Muller TH, Mars RB, Behrens TE, O'Reilly JX (2019) Control of entropy in neural models of environmental state. Elife 8:e39404.

Naber M, Alvarez GA, Nakayama K (2013) Tracking the allocation of attention using human pupillary oscillations. Front Psychol 4:919.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046.

Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput Intell Neurosci 2011:156869.

Prescott MJ, Buchanan-Smith HM (2003) Training nonhuman primates using positive reinforcement techniques. J Appl Anim Welf Sci 6:157–161.

Preuschoff K, 't Hart BM, Einhauser W (2011) Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. Front Neurosci 5:115.

Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun 7:13289.

Reingold EM, Stampe DM (2002) Saccadic inhibition in voluntary and reflexive saccades. J Cogn Neurosci 14:371–388.

Rousseeuw PJ, Croux C (1993) Alternatives to the median absolute deviation. J Am Stat Assoc 88:1273–1283.

Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P (2010) Dynamics of active sensing and perceptual selection. Curr Opin Neurobiol 20:172–176.

Schwiedrzik CM, Freiwald WA (2017) High-level prediction signals in a low-level area of the macaque face-processing hierarchy. Neuron 96:89–97.e4.

Schwiedrzik CM, Sudmann SS (2019) Pupil diameter tracks statistical structure in the environment. Figshare. Advance online publication. Retrieved May 7, 2020. doi: 10.6084/m9.figshare.9970058.v1.

Turk-Browne NB, Jungé J, Scholl BJ (2005) The automaticity of visual statistical learning. J Exp Psychol Gen 134:552–564.

Wang CA, Munoz DP (2015) A circuit for pupil orienting responses: implications for cognitive modulation of pupil size. Curr Opin Neurobiol 33:134–140.

Welch PD (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust 15:70–73.

Widmann A, Schröger E, Maess B (2015) Digital filter design for electrophysiological data–a practical approach. J Neurosci Methods 250:34–46.

Willenbockel V, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW (2010) Controlling low-level image properties: the SHINE toolbox. Behav Res Methods 42:671–684.