JNeurosci
THE JOURNAL OF NEUROSCIENCE

Research Articles: Behavioral/Cognitive

The representational dynamics of sequential perceptual averaging

https://doi.org/10.1523/JNEUROSCI.0628-21.2021

Cite as: J. Neurosci 2021; 10.1523/JNEUROSCI.0628-21.2021

Received: 24 March 2021
Revised: 26 November 2021
Accepted: 3 December 2021

The representational dynamics of sequential perceptual averaging

Jongrok Do (도종록)[a,d], Kang Yong Eo (어강용)[a], Oliver James[b,c], Joonyeol Lee (이준열)[b,c], Yee-Joon Kim (김이준)[a*]

[a] Center for Cognition and Sociality, Institute for Basic Science, Daejeon, Republic of Korea.

[b] Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, Republic of Korea.

[c] Department of Biomedical Engineering, Sungkyunkwan University, Suwon, Republic of Korea.

[d] Department of Biological Science, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea.

Abbreviated Title: Sequential update of mean information in frontocentral area

Number of Pages: 43

Number of Figures: 6

Number of words for Abstract: 216

Number of words for Introduction: 627

Number of words for Discussion: 1722

Conflict of interest: The authors declare no competing financial interests.

Acknowledgment: This research was supported by IBS-R001-D2-a00, IBS-R001-D2-a01, and IBS-R015-D1. We thank Professor Satoru Suzuki for constructive comments on earlier versions of this manuscript.

[*]Proofs and correspondence to:

Yee-Joon Kim

Center for Cognition and Sociality, Institute for Basic Science,
55 Expo-ro, Yuseong-gu, Daejeon, 34126, Republic of Korea.
Telephone:+82 (42) 878-9111
Email address: joon@ibs.re.kr

Abstract

It is clear that humans can extract statistical information from streams of visual input, yet how our brain processes sequential images into the abstract representation of the mean feature value remains poorly explored. Using multivariate pattern analyses of electroencephalgraphy recorded while human observers viewed the sequentially presented ten Gabors of different orientations to estimate their mean orientation at the end, we investigated sequential averaging mechanism by tracking the quality of individual and mean orientation as a function of sequential position. Critically, we varied the sequential variance of Gabor orientations to understand the neural basis of perceptual mean errors occurring during sequential averaging task. We found that the mean-orientation representation emerged at specific delays from each sequential stimulus onset and became increasingly accurate as additional Gabors were viewed. Especially in frontocentral electrodes, the neural representation of mean orientation improved more rapidly and to a greater degree in less volatile environment while individual orientation information was encoded precisely regardless of environmental volatility. The computational analysis of behavioral data also showed that perceptual mean errors arise from the cumulative construction of the mean orientation rather than the low-level encoding of individual stimulus orientation. Thus, our findings provide neural mechanisms to differentially accumulate increasingly abstract feature from a concrete piece of information across the cortical hierarchy depending on environmental volatility.

Significance Statement

The visual system extracts behaviorally relevant summary statistical representation by exploiting statistical regularity of the visual stream over time. However, how the neural representation of the abstract mean feature value develops in a temporally changing environment remains poorly identified. Here, we directly recover the mean orientation information of sequentially delivered Gabor stimuli with different orientations as a function of their positions in time. The mean orientation representation, which is regularly updated, becomes increasingly accurate with increasing sequential position especially in the frontocentral region. Further, perceptual mean errors arise from the cumulative process rather than the low-level stimulus encoding. Overall, our study reveals a role of higher cortical areas in integrating stimulus-specific information into increasingly abstract task-oriented information.

Keywords: EEG, Multivariate pattern analysis, Sequential perceptual averaging, Summary statistical representation

Introduction

Understanding how humans effectively interact with the dynamic and complex sensory environment is of central importance in the behavioral, cognitive, and neural sciences. Growing evidence shows that the perceptual system extracts behaviorally relevant information from complex dynamic sensory signals by summarizing them with their central tendency – the mean – through the exploitation of statistical regularities of sensory data over space (Chong and Treisman, 2005; Greenwood et al., 2009; Alvarez, 2011; de Gardelle and Summerfield, 2011; Whitney and Yamanashi Leib, 2018) or time (Haberman et al., 2009; Albrecht et al., 2012; Piazza et al., 2013; Gorea et al., 2014; Hubert-Wallander and Boynton, 2015). How does the perceptual system compute the mean of sensory features and construct an abstract representation? Although much is known about spatial averaging (for example, via progressively larger receptive fields in the visual system (Dumoulin and Wandell, 2008; Freeman and Simoncelli, 2011)), less is known about temporal averaging (Navajas et al., 2017; McWalter and McDermott, 2018). Particularly, the neural mechanism of sequential perceptual averaging is not well-defined. Although recent neuroimaging studies have shown that the brain simultaneously represents multiple successive images and their changes at each instant (Marti and Dehaene, 2017; King and Wyart, 2019), it is unclear how the abstract representation of the mean feature value develops in this changing visual stream.

Extracting temporally stable information, such as temporal feature averages from dynamic sensory environments, helps optimize behavior by allowing the structure of the environment to be robustly grasped. Integration plays a fundamental role in this process (Navajas et al., 2017; McWalter and McDermott, 2018), but it is unclear whether the mean feature value is updated after each stimulus or multiple stimuli when stimuli are serially delivered. Additionally, it has been established that the precision of the extracted mean is not perfect and declines with increasing feature variability (Dakin, 1999; de Gardelle and

Mamassian, 2015; Haberman et al., 2015; Navajas et al., 2017; McWalter and McDermott, 2018). What are the causes of this error when estimating the mean of a stimulus sequence? Does the error occur at the stage of encoding individual visual images or at the stage of integrating these continuously changing features over time? Does environmental volatility influence the low-level representation of individual visual images, the cumulative construction of the mean feature value, or both? To address these questions, we explored the dynamics of the sequential perceptual averaging using multivariate pattern analyses of electroencephalography (EEG) signals recorded from human observers while they estimated the mean orientation of ten randomly oriented Gabor patches sequentially presented at the fovea. By using an inverted encoding model (IEM) (Brouwer and Heeger, 2009; Garcia et al., 2013; Myers et al., 2015; Foster et al., 2017), we were able to investigate how individual and their mean orientations were represented in multivariate EEG activity during the sequential averaging task. By manipulating the variance of the ten Gabor orientations, we monitored how the neural representations of individual orientations and their means were modulated as a function of environmental volatility. The sequential averaging task encouraged observers to update the mean orientation upon presentation of each Gabor stimulus. The multivariate EEG pattern analyses enabled us to assess the contributions of individual stimuli to perceived mean orientation. By estimating how the individual stimuli were weighted into the mean orientation, we probed the neural mechanisms of estimating the mean orientation.

We found that both individual and mean orientations were represented in the dynamically evolving multivariate EEG activities. Secondly, the representation of the mean orientation emerged at specific delays after each Gabor onset, and its accuracy increased gradually towards the end of the sequence especially in the frontocentral region. For sequences of high orientation variance, however, the weighting of the later sequential stimuli

was decreased, which may account for poor behavioral performance of perceptual mean estimation.

Materials and Methods

Observers

24 human observers (9 females, 15 males) participated in this study. Two participants were excluded from the analysis due to excessive eye movements. All observers had normal or corrected-to-normal visual acuity, gave informed written consent to participate as paid volunteers, and were tested individually in a dark room. The study was approved by the Institutional Review Board of the Korea National Institute for Bioethics Policy.

Stimulus

Visual stimuli were generated and presented using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) along with custom scripts written in MATLAB (Mathworks Inc.). The 19-inch display CRT monitor (ViewSonic PF817) was set to a refresh rate of 100 Hz and a resolution of 800×600 pixels. The CRT monitor gamma tables were adjusted to ensure response linearity and a constant mean luminance of 59 cd/m^2. Participants viewed the stimuli from a distance of 70 cm in a darkened room.

Each trial comprised a sequence of ten randomly oriented Gabor patches, presented centrally for 100 ms, with an inter-stimulus interval of 500 ms to give observers enough time to encode and integrate sensory information across time. The orientation of each Gabor patch was one of 40 possible evenly spaced angles spanning 180°. Each sequence was preceded and followed by a blank period. All Gabor patches had identical parameters (contrast: 50 %,

diameter: 8° visual angle, spatial frequency: 1.25 cycles/degree, Gaussian envelope SD: 2° visual angle), except for their orientation.

Experimental procedure

The observer initiated each trial by pressing the space button. A central fixation cue appeared for 500 ms. During the subsequent six seconds, observers viewed a sequence of ten tilted Gabor patches against a mid-gray background, followed by a blank period of 700 ms (Figure 1a). During this entire period, observers were instructed to maintain fixation on the center of the screen and attempt to withhold eye blinks. After the sequence, a circularly bounded red probe bar appeared in the center of the screen. The observers' task was to estimate the mean orientation of ten Gabor patches by rotating the red probe bar counter-clockwise or clockwise using the left- or right-arrow key and pressing the down-arrow key when the adjusted orientation seemed to match the mean orientation. The probe display remained until the observers responded.

For experiments, there were sixteen trial types: eight mean orientations (11.25°, 33.75°, 56.25°, 78.75°, 101.25°, 123.75°, 146.25°, and 168.75°) and two variances. There were two experimental conditions that differed in the orientation variance of the sequence. For the low-variance condition, the sequence comprised ± 4.5°, ± 9°, ± 13.5°, ± 18°, and ± 22.5°-oriented Gabor patches relative to the mean orientation. For the high-variance condition, the sequence comprised ± 9°, ± 18°, ± 27°, ± 36°, and ± 45°-oriented Gabor patches relative to the mean orientation. Ten orientations of every sequence were randomly shuffled to define the presentation order. Eight possible mean orientations were used for both conditions. The Gabor patch with the mean orientation of the sequence never appeared in the stream of ten Gabor patches. Because this manipulation of the orientation variability alone made the sequential averaging task difficult enough, we did not further manipulate temporal

regularity in the streams of Gabor patches. As previous studies have shown that any perturbation of temporally regular stimulation impedes both perceptual sensitivity and reaction time (Schroeder and Lakatos, 2009; Cravo et al., 2013; Morillon et al., 2016), we used the same periodic stimulation in low- and high-variance conditions.

We tested each observer for 320 trials in 8 blocks of 40 trials each. In each block, a sequence with one of eight possible mean orientations was repeated five times with ten randomly shuffled orientations. Two variance conditions alternated in a block-design manner, and the order of conditions was counterbalanced across observers. We gave the observers breaks within and between blocks as necessary.

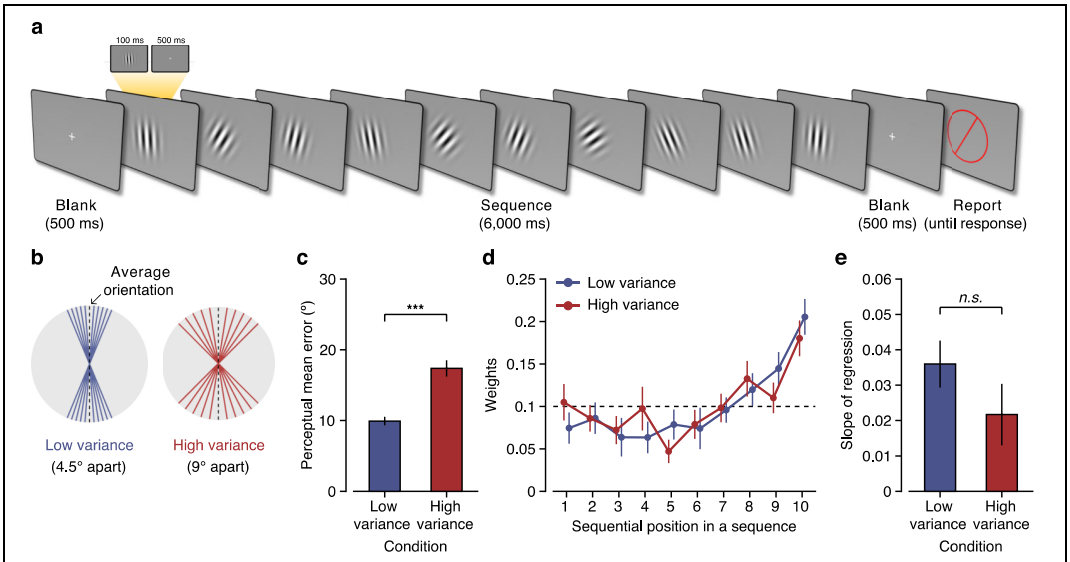

Figure 1. Stimuli and trial sequence, experimental conditions, and behavioral data analysis. **(a)** Trial sequence. After a fixation period, ten randomly oriented Gabor patches were sequentially presented, and the participants were instructed to report the mean orientation of the sequence by adjusting a red probe bar at the end of the sequence. **(b)** Experimental conditions. Each trial had either low (4.5º) or high (9º) inter-stimulus orientation variance of ten sequential Gabor patches. **(c)** Behavioral performance in the low-variance and high-

variance sequences. **(d)** Mean weights (regression coefficients) as a function of sequential position. The X-axis indicates the sequential order of the presented Gabor patches on each trial. The Y-axis indicates the relative influence of each stimulus on the participants' responses in the task. The dashed line indicates the expected weights when all sequential stimuli have the equal amount of influence on participants' responses. **(e)** Linear slope of regression coefficients across sequential positions. **(c-e)** Error bars indicate ±1SEM.

EEG signal acquisition and preprocessing

The EEG data were collected with 128-sensor HydroCel Sensor Nets (Electrical Geodesics, Eugene OR) at a sampling rate of 500 Hz and were band-pass filtered from 2 Hz to 200 Hz. The raw data were then epoched between -150 ms and 6,700 ms relative to the first stimulus onset. We used the FASTER (Nolan et al., 2010) (Fully Automated Statistical Thresholding for EEG Artifact Rejection, http://www.mee.tcd.ie/neuraleng/Research/Faster) package to reject artifacts and interpolation of noisy EEG sensors. Finally, the EEG was re-referenced to the common average of all the sensors. All analyses were performed with 111 EEG channels, excluding 17 channels vulnerable to movement artifacts including electrodes around the ears and on the face. The elimination of these nuisance channels did not change the results of the analyses that used 128 EEG channels.

Behavioral data analysis

First, we performed linear regression analysis to quantify the relative influence of each sequential position on the observers' reported perceptual mean orientation (Juni et al., 2012; Hubert-Wallander and Boynton, 2015),

$$R_j = \sum_{i=1}^{10} w_i x_{ij},$$

where R_j is the observer's reported perceptual mean orientation for trial *j*, x_{ij} is the orientation of the Gabor patch at sequential position *i* and trial *j*, and w_i is the weight for sequential position *i*. The ten relative weight values averaged across observers for both variance conditions are plotted in Figure 1d. We additionally performed linear regression analysis to calculate the regression slope of the relative weight values over ten sequential positions. We used the slope as a proxy of primacy or recency effect (Figure 1e). A negative slope indicates a primacy effect and a positive slope indicates a recency effect.

In order to examine the effect of sequence variance on how the perceptual mean is computed during a sequential averaging task, the observers' behavioral data were fit to the sequential update model (Navajas et al., 2017). This model is based on the assumption that observers keep track of the mean orientation and update it after each stimulus presentation. In this model, observers combine a noisy estimate of the current stimulus with their previous estimate of the mean,

$$\mu_i = (1 - \lambda)\mu_{i-1} + \lambda\theta_i + \gamma\theta_i\xi_i,$$

where μ_i is the estimate of the mean after *i* stimuli ($\mu_0 = 0$), $0 < \lambda < 1$ determines the relative weighting of recent versus more distant stimuli, and θ_i is the actual orientation of the *i*th stimulus in the sequence. ξ_i is sampled from the standard normal distribution and γ is a free parameter that indicates the strength of the noise. For each variance condition, we implemented a constrained nonlinear optimization algorithm to determine the best-fitting parameters λ and γ that minimized the root mean square of the difference between the predicted and reported mean orientations.

Inverted encoding model (IEM)

To reconstruct the orientation information from the spatially distributed pattern of the EEG signals, we used an inverted encoding model (IEM) (Brouwer and Heeger, 2009; Garcia

et al., 2013; Myers et al., 2015; Foster et al., 2017) where each orientation is represented using weights from a linear basis set of population tuning curves. Forty hypothetical channel tuning functions (CTFs) were centered at forty orientations used in trials, evenly spaced from 0° to 180° in steps of 4.5°; each basis function was a half-sinusoidal function raised to the fifth power. The epoched signals were baseline-corrected using the average signal from -150 to -50 ms relative to the onset of the first Gabor patch presented in the sequence of each trial. We focused all of our IEM analyses on the EEG signals above 2 Hz. The main reason was to minimize the effect of the physically driven oscillatory waveform (Steady-State Visual Evoked Potential, SSVEP) at the stimulus presentation rate of 100 ms ON – 500 ms OFF on the representational dynamics, although the periodic stimulation still elicited higher harmonic SSVEP responses in EEG. Also, the current research was aimed at investigating whether the EEG activity pattern dynamics directly represent the stimulus and the mean information during a sequential averaging task rather than confirming the role of the specific frequency band activity in rhythmically modulating the gain of information processing because various frequency band activities are already known to be involved in sensory and cognitive information processing (Maris and Oostenveld, 2007; Busch et al., 2009; Busch and VanRullen, 2010; Landau and Fries, 2012; Fiebelkorn et al., 2013; Landau et al., 2015; Fiebelkorn et al., 2018; Helfrich et al., 2018) regardless of external stimuli being presented at a rate of particular frequency in the range of delta (Schroeder and Lakatos, 2009; Wyart et al., 2012; Cravo et al., 2013) or theta (Hanslmayr et al., 2013; Fiebelkorn et al., 2018; Helfrich et al., 2018; Zhang et al., 2018) band.

Recovering physical orientation during the sequence presentation period

To recover stimulus orientations, the stimulus-evoked activities of all the sequentially presented Gabor patches labeled with their physical orientations were trained and tested in

the leave-one-trial-out (i.e., leave-ten-samples-out) fashion. Throughout the paper, a sample is simply meant by a multivariate EEG activity evoked by an individual oriented Gabor patch. Specifically, ten Gabor patches in each sequence from both variance conditions were independently epoched between -100 ms and 700 ms relative to each Gabor patch onset and labeled with their physical orientations. Out of the 3,200 samples across both variance conditions (80 samples per one of forty orientations), 10 samples from one test trial were tested with the inverted encoding model weight matrix trained on the remaining 3,190 samples. For each sample, the channel tuning function was zero-centered relative to the presented orientation. This procedure was repeated for each time point in the stimulus epoch before moving to the next iteration in the leave-one-trial-out procedure. Zero-centered orientation-selective tuning functions were then separated into each experimental condition (1600 samples per variance condition) and averaged across trials for each condition.

We constructed the inverted encoding model as:

$$B_1 = WC_1$$

where B_1 is the training set (111 sensors x 3,190 samples) and C_1 is the hypothetical channel tuning functions (40 orientations x 3,190 samples). Then, we estimated the weight matrix W (111 sensors × 40 orientations) by multiplying both sides by the pseudoinverse of C_1 as in the ordinary least squares (OLS):

$$\widehat{W} = B_1 C_1^T (C_1 C_1^T)^{-1}$$

We estimated the population orientation response $\hat{C}_2$ (40 orientations × 10 samples) with the estimated weight $\widehat{W}$ and the test set B_2 (111 sensors × 10 samples):

$$\hat{C}_2 = (\widehat{W}^T \widehat{W})^{-1} \widehat{W}^T B_2$$

where $\hat{C}_2$ is the tuning curve of the test set, $\widehat{W}$ is the weight matrix, $\widehat{W}^T$ is its transpose, and $\widehat{W}^{-1}$ is its pseudoinverse. For each time point in the epoch of all training sets (-100–700 ms after each stimulus onset), we applied the estimated weights to the same time point in the test

set, and then zero-centered the output tuning curves $\hat{C}_2$ relative to the labeled physical orientation of the sample. This procedure was repeated for all time points in the epoch (in 10 ms steps, using a sliding window of 40 ms). When we plotted the reconstructed tuning curves before zero-centering, their peak locations at 40 different physical stimulus orientations from 0° to 180° in steps of 4.5° were clearly distinguished from each other (Figure 3b).

To summarize the tuning-curve slope as a function of time, we calculated the linear slope of the zero-centered tuning curve from -90° to 0° at each time point in the epoch (Myers et al., 2015). We averaged the zero-centered tuning curves that were equidistant from 0° (i.e., +4.5° and -4.5°, +22.5°, and -22.5°). For Figures 3c, 3d, 4b, 5a, and 6a, the resulting orientation channel time course was smoothed with a Gaussian kernel (σ = 30 ms). We then fit a linear slope across the orientation channels from -90° to 0°, separately for each time point, variance condition, and observer. Tuning-curve slope was evaluated using one-sample *t*-tests (against 0). In doing so, zero tuning-curve slope corresponded to no orientation selectivity, while higher tuning-curve slope corresponded to greater orientation selectivity. Multiple comparisons across time points were corrected using non-parametric cluster-based permutation testing (Maris and Oostenveld, 2007) (5000 permutations).

Recovering mean orientation during the sequence presentation period

We performed another IEM analysis to determine the mean orientation represented in each of the stimulus-evoked activity patterns of Gabor patches presented in the trial. This analysis had two purposes. The first purpose was to examine how sequential variability influences the accuracy of the mean information represented in the EEG activity patterns during the sequence presentation period. Thus, we checked if the difference in behavioral performance between the two conditions correlated with the difference in the representational quality of the mean information. Another purpose was to examine whether observers used the

sequential update strategy to extract the mean orientation over the sequence. Specifically, we examined whether the neural representation of the mean orientation becomes increasingly more precise as a function of sequential position.

To recover the mean orientation, we trained the same inverted encoding model as described above with 3,190 stimulus-evoked activity patterns epoched between -100 ms and 700 ms relative to each Gabor patch onset, except for 10 stimulus-evoked activity patterns from one test trial. Both training (3,190 samples) and test data (10 samples from 1 trial) were labeled with their mean orientation of the corresponding trials. The population tuning curve was recovered by applying the weight matrix to the left-out samples of the trial. For each test sample, the population tuning curve was zero-centered relative to the mean orientation. This procedure was repeated for each time point in the stimulus epoch before moving to the next iteration in the leave-one-trial-out (i.e., leave-ten-samples-out) fashion.

In order to investigate whether observers could keep track of the mean orientation after each stimulus presentation, we hypothesized that the mean orientation tuning-curve slope gradually increases as a function of sequential position if successive samples of sensory evidence are accumulated across sequential positions. We collapsed data across the low-variance and high-variance conditions and then performed a linear regression analysis on ten tuning-curve slopes across sequential positions (Figures 4-6). This linear regression analysis was repeated at every time point from -100 ms to 700 ms after each Gabor patch onset. These output slopes were used to find the time clusters where regression slopes were significantly higher than zero, using a non-parametric cluster-based permutation testing (Maris and Oostenveld, 2007) (5000 permutations). Only when significant time clusters were found through the repeated linear regression analysis, tuning-curve slopes at each sequential position were averaged within significant time cluster. The averaged ten tuning-curve slopes were then used for testing the hypothesis of the linearly increasing trend of tuning-curve

336 slopes of mean orientation across sequential positions. Specifically, we calculated the linear
337 regression slope of the averaged tuning-curve slopes over sequential positions and compared
338 the steepness between the low-variance and high-variance conditions. It would be worse for
339 the encoding model of mean orientation to include all samples, even those belonging to the
340 early sequential positions, because mean orientation cannot be precisely estimated with only
341 a small portion of sequential stimuli. However, it is difficult to make an assumption about the
342 sequential position where mean orientation starts to be precisely extracted. Therefore, to
343 avoid selection problems and maximize statistical power, we used all samples and labeled
344 them with their mean orientation of a trial. If there are samples where the mean orientation is
345 represented, these data would be more influential when training the tuning curve weight
346 matrix; where the mean information is not represented, marginal influence is exerted on
347 computing the weight matrix.

348 In Figure 7, we split 111 electrodes in three clusters to examine the role of
349 frontoparietal region in sequential perceptual averaging process. In anterior, middle and
350 posterior electrode cluster, we performed the same linear regression analysis to find time
351 points where the linear regression slopes across sequential positions were significantly
352 positive. Once significant time clusters were found, we again performed the same linear trend
353 analysis to test the hypothesis of the linear improvement of mean orientation representation
354 over sequence in each electrode cluster.

355

356 *Cross-temporal generalization of the IEM*

357 When we performed the cross-temporal generalization analysis, we estimated the
358 weight matrix using EEG data at each time point and applied the weight for the estimations
359 of the channel responses across all time points. Specifically, we trained a weight matrix from
360 the training set at time *t* and applied the estimated weight matrix to the test set at time *t'*. This

procedure was repeated so that the weight matrices at every time point had been used to calculate the slope of the population-tuning curves (tuning-curve slope) at every time point, thereby creating a two-dimensional temporal generalization matrix of the population tuning curve slopes (Figure 3e and Figure 4a). All other aspects (e.g., leave-one-trial-out (i.e. leave-ten-samples-out) method) were identical to the inverted encoding model procedure explained above. Multiple comparisons across train-test time point pairs were corrected using non-parametric cluster-based permutation testing (Maris and Oostenveld, 2007) (5000 permutations) to evaluate tuning-curve slope was greater than 0 using one-sample *t*-test.

Results

We analyzed scalp EEG signals from 22 human observers as they performed a sequential averaging task. Observers viewed ten randomly oriented Gabor patches sequentially. The Gabor patch with the mean orientation was not presented to examine the internally generated representations of the mean. Following each sequence, observers were instructed to report the mean orientation by adjusting a red probe bar, preceded by a 700 ms blank period (Figure 1a). The high- and low-variance sequences were presented in separate blocks (Figure 1b; see ***Materials and Methods*** for details).

Modeling of sequential averaging process

The perceptual mean error was larger in the high-variance condition than in the low-variance condition (t_{21} = -8.63, $p < 10^{-7}$, Figure 1c and 1d). To examine the relative influence of individual stimulus orientation on the perceived mean orientation, we first performed linear regression analysis (Juni et al., 2012; Hubert-Wallander and Boynton, 2015) (see ***Materials and Methods***). The behavioral data fit well to this weighted average model in both low-variance (R^2 = 0.94 ± 0.04) and high-variance conditions (R^2 = 0.82 ± 0.01), and the

model showed significantly higher R^2 value in the low- than in the high-variance condition (t_{21} = 6.51, $p < 10^{-5}$). We found that there was a significant recency effect (positive slopes); later stimuli had a greater influence on the perceived mean orientation than earlier ones in both low-variance (t_{21} = 5.43, $p < 10^{-4}$) and high-variance sequences (t_{21} = 2.50, $p < 0.05$). There was no significant difference in the recency effect between the two variance conditions (t_{21} = 1.64, $p > 0.05$, Figure 1e). This result is in contrast to that of a previous study which showed that recent stimuli had a greater influence on the perceptual mean orientation for the low- than the high-variance sequence (Navajas et al., 2017). This contrasting result may be partly due to the difference in sequence lengths; Navajas et al.'s sequence contained 30 stimuli, while ours contained only 10. We speculate that shorter sequences have advantages over longer sequences for remembering the earlier part of the sequence, so that even in the high-variance sequence, observers were able to estimate the mean of the sequence, leading to the same recency effect. Furthermore, low- and high-variance conditions were equivalent to the top two high-variance conditions in Navajas et al.'s study (Navajas et al., 2017), the effect of environmental volatility on recency effect might be similar in our task.

Next, we fitted a variant of a leaky integrator model, called the sequential update model (Navajas et al., 2017), to the behavioral data. Because observers were required to report the mean orientation of each sequence, we assume that they updated their estimate of the mean after each stimulus presentation by combining a noisy estimate of the current stimulus with their previous estimate of the mean (see ***Materials and Methods***).

$$\mu_i = (1 - \lambda)\mu_{i-1} + \lambda\theta_i + \gamma\theta_i\xi_i.$$

The model fitted the behavioral data well in both variance conditions (low-variance sequence: R^2 = 0.93 ± 0.04; high-variance sequence: R^2 = 0.80 ± 0.10), being significantly better for the low-variance than the high-variance condition (t_{21} = 6.81, $p < 10^{-6}$). The model predicted that the leak constant / , the relative weighting of recent vs. more distant stimuli, did not differ

between the two variance conditions (t_{21} = 0.21, p > 0.8. Figure 2a). This result is consistent with the preceding analysis using the weighted average model showing that the recency effect was comparable across both conditions (Figure 1e). In both low- and high-variance conditions, the recency effect was well captured by the leak constant λ that correlated positively with the regression slope of the average weights at ten sequential positions in Figure 1d (r = 0.60, p = 0.003 for low-variance, and r = 0.56, p = 0.007 for high-variance condition; Figure 2c and 2d). We also found that stimuli have a larger amount of multiplicative noise (γ) in high variance than in the low-variance condition (t_{21} = -3.83, p < 0.001. Figure 2b). This indicates that the process of updating mean orientation was noisier in the high-variance condition, leading to accurate representations of mean orientation at the end of the sequence.

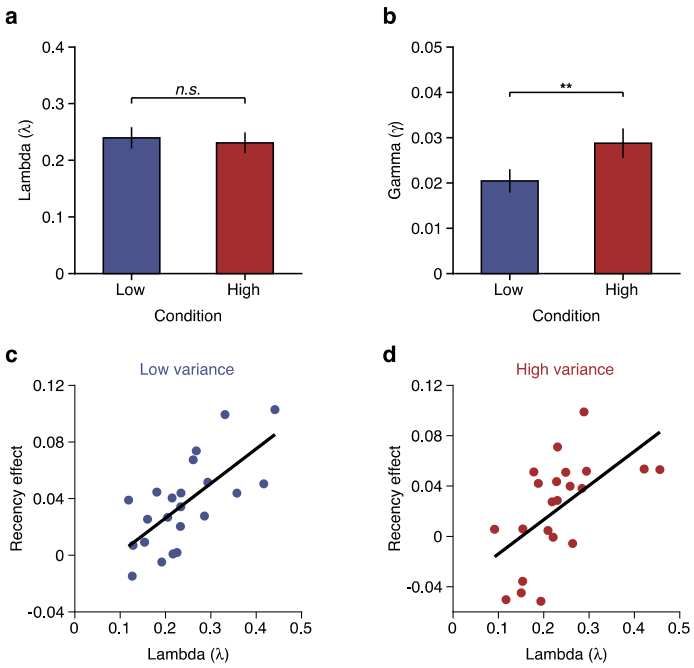

Figure 2. Model fitting of behavioral data. **(a-b)** Two best-fitting free parameters (lambda: λ and gamma: γ) that minimize the sum of squared differences between the output orientation

of the model and the subjectively judged orientation by the observers. **(c-d)** Correlation between the recency effect value and best-fitting parameter λ. The recency effect value is defined as a slope of regression weights (See ***Materials and Methods*** and Figure 1d). Each dot represents each subject.

Probing the neural mechanisms of sequential averaging in the presence of variability

To probe the neural bases of sequential averaging, we used an inverted encoding model (IEM) to recover information about the individual stimulus orientations and the mean orientation from the full EEG signals (Brouwer and Heeger, 2009; Garcia et al., 2013; Myers et al., 2015; Foster et al., 2017) (See ***Materials and Methods***). We characterized the effect of sequential variability on both the neural encoding of individual orientations and the neural integration of orientation by directly visualizing the temporal dynamics of orientation representation. Consistent with the model-fitting of the behavioral data suggesting greater integration noise in the high-variance condition, the IEM analysis confirmed that the process of sequentially integrating individual orientations to update the mean was degraded in the high-variance condition.

Recovering the individual stimulus orientation

First, we investigated whether a more variable sequence caused the individual stimulus orientation to be encoded less precisely in the EEG signals despite the presentation of the same physical stimuli in both sequences. To recover the individual stimulus orientation of the sequence, we applied an inverted encoding model to stimulus-evoked EEG signals. Forty idealized tuning curves, equally spaced between 0° and 180°, were used as basis functions (Figure 3a). Each epoched data was labeled with the presented stimulus orientations and those data were used to train and test the inverted encoding model in the leave-one-

sequence-out cross validation procedure (see ***Materials and Methods***). Consequently, we obtained population tuning curves as a function of time for the presented stimulus orientations for both sequences. The reconstructed population tuning curves averaged from 0 ms to 200 ms post stimulus onset show distinct peaks at forty different stimulus orientations (Figure 3b). When the population tuning curves were zero-centered relative to their presented orientations, they revealed that the stimulus orientations were represented in multivariate EEG activity for almost the entire period after each stimulus onset (Figure 3c; 0 – 510 ms relative to stimulus onset, cluster-corrected $p < 0.001$ for low-variance; 10 – 550 ms relative to stimulus onset, cluster-corrected $p < 0.001$ for high-variance). Orientation-specific coding accuracies measured by the linear slopes of the tuning curves (see ***Materials and Methods***) were not significantly different between the two conditions (Figure 3d; Cluster-based permutation test). We additionally performed a paired t-test on the tuning curve slopes averaged across time period that showed significantly positive tuning curve slopes (from 100 ms to 450 ms relative to stimulus onset). The difference between the two variance conditions was not significant ($t_{21} = 1.75$, $p = 0.094$). These findings show that while individual orientations are encoded in the dynamically changing EEG patterns in both variance conditions, it is not a better encoding of individual orientation that explains better perceptual averaging in the low-variance sequence. (Figure 3e).

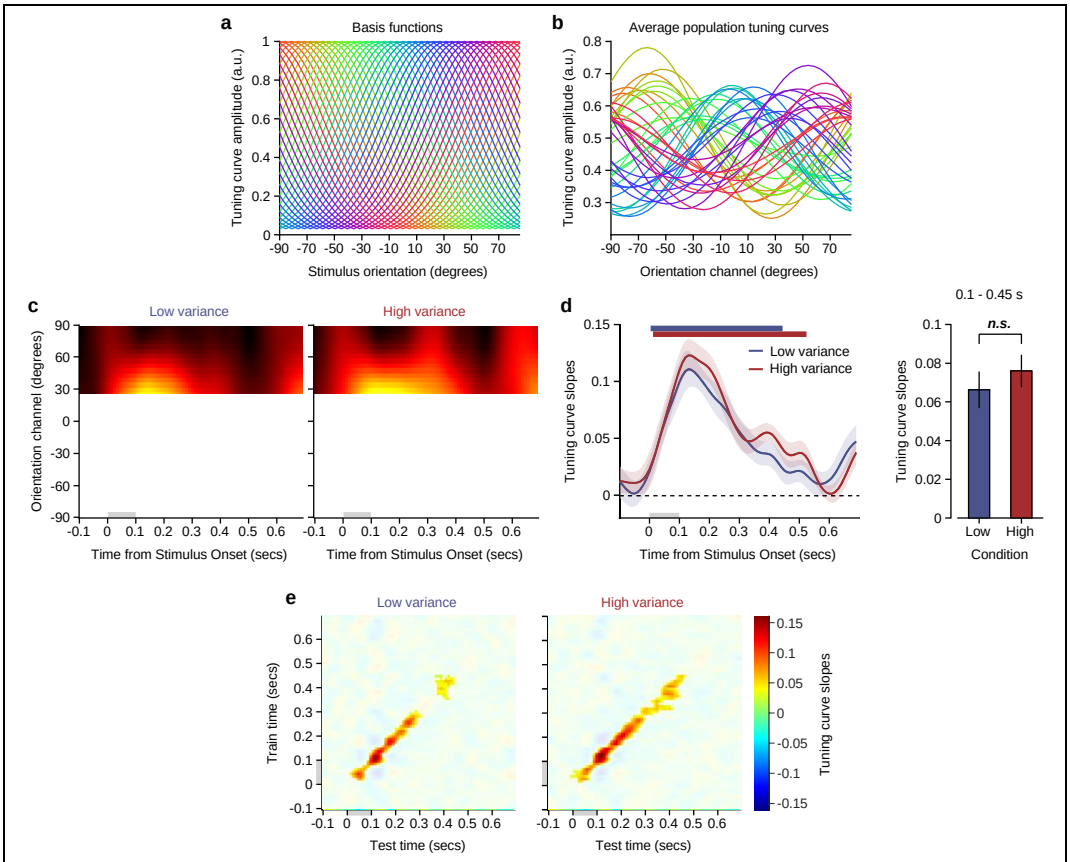

Figure 3. Temporal dynamics of the reconstructed tuning curves of stimulus orientation. The encoding model was trained and tested on stimulus-evoked activities from -100 – 700 ms after stimulus onset, with respect to their physical orientations. **(a)** Hypothetical tuning curve (basis functions). Each color corresponds with one of 40 different orientations. **(b)** Reconstructed stimulus orientation tuning curves (averaged from 100 to 200 ms after stimulus onset). Color coding is same as (a). **(c)** Time-resolved tuning curves. Tuning curves were zero-centered and were averaged across trials separately for both variance conditions (left: low-variance, right: high-variance). **(d)** Tuning-curve slope for both variance conditions. Upper bars indicate the time period when the tuning-curve slope is significantly greater than 0 ($p < 0.05$, based on cluster extent). Shaded areas indicated ±1SEM. On the right, bar graphs indicate tuning curve slopes averaged from 100 ms to 450 ms after stimulus

onset for the comparison between the two variance conditions. Error bars indicate the ±1SEM. **(c-d)** Gray bars represent the time period when the stimulus was presented. **(e)** Cross-temporal generalization of tuning curve slope of stimulus orientation. The tuning curve slope of physical orientation for low- (left) and high- (right) variance conditions was estimated by training weights on one time point in the training data, and applying them to all time points in the test data. The transparency mask highlights the significant clusters where tuning-curve slopes are greater than 0 (one-tailed, $p < 0.05$, based on cluster extent).

Recovering the mean orientation

Next, we investigated whether a more variable sequence caused the mean orientation to be less precisely encoded in the EEG signals. The sequential update modeling of behavioral data predicted that the updated mean representation becomes more accurate towards the end of the sequence. Therefore, we examined if the EEG signals reflected increasingly precise mean orientation representations. Since the integration noise parameter g was significantly larger in the high-variance than in the low-variance condition (Figure 2b), we also investigated if the low-variance condition allowed observers to more precisely update the mean orientation. For this analysis, we assumed that each presentation of a Gabor, the visual evoked response pattern includes the neural representations of both the current stimulus orientation and the updated mean orientation. To isolate the latter, we trained the weight matrix to the mean orientation of the sequence and aligned the recovered orientation tuning curves to the mean orientation of the sequence at every time point after each stimulus onset (see ***Materials and Methods***). Our rationale for labeling sequential samples as their mean orientation was that if the encoding model learns ten samples in each trial as the same mean orientation instead of their own physical orientations, it would discard differences among individual Gabor orientations to only keep the internally integrated orientation that

505 should be close to the mean orientation. This method allowed us to track the internal abstract
506 representation of the mean orientation regardless of the physically presented orientation. It
507 also enabled us to interpret the tuning-curve slope at each sequential position as the distance
508 between the currently represented mean orientation and the actual mean orientation of the
509 whole sequence.

510 We first checked if the neural representation of the mean orientation indeed existed in
511 stimulus-evoked multivariate EEG activity. The mean-orientation-selective tuning curve
512 slopes averaged across ten sequential positions show that the mean orientations were well
513 represented in the dynamically evolving EEG activity pattern for both sequences (Figure 4a).
514 In the low-variance condition, the two off-diagonal significant clusters indicate that the
515 current and the next mean representations encoded in the dynamically changing stimulus-
516 evoked EEG patterns correspond in the earliest period (training time 600 – 690 ms, test time
517 0 – 100 ms, cluster p = 0.007, and the training time 0 – 120 ms, test time 620 – 690 ms,
518 cluster p = 0.003 in the left column of Figure 4a). The diagonal parts in Figure 4a show that
519 the mean orientation information emerges shortly after each Gabor onset (Blue lines in Figure
520 4b: -20 – 420 ms relative to stimulus onset, cluster-corrected $p < 0.001$, and 510 – 690 ms
521 relative to stimulus onset, cluster-corrected p = 0.009 for low-variance; Red lines in Figure
522 4b: 50 – 340 ms relative to stimulus onset, cluster-corrected p = 0.001 for high-variance). The
523 neural representations of the mean orientation were more precise in the low-variance
524 condition than in the high-variance condition (Orange lines in Figure 4b; -70 – 150 ms
525 relative to stimulus onset, cluster-corrected p = 0.005; 230 – 380 ms relative to stimulus onset,
526 cluster-corrected p = 0.015, and 480 – 680 ms relative to stimulus onset, cluster-corrected p =
527 0.029). The first and second columns of Figure 4c show the accuracy of representing the
528 mean orientation as a function of sequential position at every time point after each stimulus
529 onset in the low-variance and high-variance conditions, respectively. We collapsed these data

across both variance conditions and performed a linear regression analysis on ten tuning curve slope values at each time point to track the sequential evolution of the mean orientation represented in the EEG signals. We found two significant clusters of time points at which the mean orientation representation became increasingly precise across the ten sequential positions (140 – 180 ms and 270 – 320 ms relative to stimulus onset, cluster-corrected p = 0.026 and p = 0.029, respectively, for each cluster. These time periods are depicted in Figure 4c using transparency mask highlights). These time points were later than the time points at which the representational accuracy of the presented orientation maximized as shown in Figure 3d. These results are summarized by averaging the mean tuning-curve slopes across time points of the two significant clusters separately for each condition (Figure 4d).

Since a large noise constant g indicates a noisy integration of individual stimulus orientation information, we checked whether the updated mean orientation information became less precise as more stimulus orientations were integrated in a more variable environment. For each sequence, we calculated the linear regression slope of tuning-curve slopes of mean orientation at ten sequential positions. We found a steeper regression slope of tuning-curve slopes of mean orientation across the ten sequential positions in the low-variance condition than in the high-variance condition (t_{21} = 2.88, p = 0.009; Figure 4e). This difference was mainly due to the higher tuning-curve slope values in the later part of the low-variance sequence. The tuning-curve slope value averaged from the sixth to the tenth stimuli was higher in the low-variance condition than in the high-variance condition (t_{21} = 2.30, p = 0.032; Figure 4f), whereas there was no difference in the tuning-curve slope averaged from the first to the fifth stimuli between the two conditions (t_{21} = 0.13, p > 0.8; Figure 4f). For the more variable sequence, the less precise representation of the mean orientation indicated by the shallow regression slope was consistent with the significantly larger noise constant predicted by the sequential update model (Figure 2b), which may lead to a poor behavioral

performance in judging the perceptual mean (Figure 1c). Although the behavioral data in the high-variance condition were well fit by the sequential update model (R^2 = 0.80 ± 0.10, Figure 2), the linear regression analysis of the ten slope values showed a smaller Pearson correlation coefficient in the high-variance condition than in the low-variance condition (r = 0.29 ± 0.06 for low-variance, r = 0.13 ± 0.06 for high variance, and t_{21} = 2.22, p = 0.04). This suggests that the tuning-curve slope of mean orientation as a function of sequential position is less consistent with the linearly increasing trend in the high-variance condition. This is also in line with the fact that the sequential update model fits the behavioral data in the low-variance condition better than in the high-variance condition (t_{21} = 6.81, $p < 10^{-6}$, Figure 2). Thus, we cannot rule out the possibility that the observers might have used different strategies for computing the mean orientation other than the sequential updating in a more variable environment.

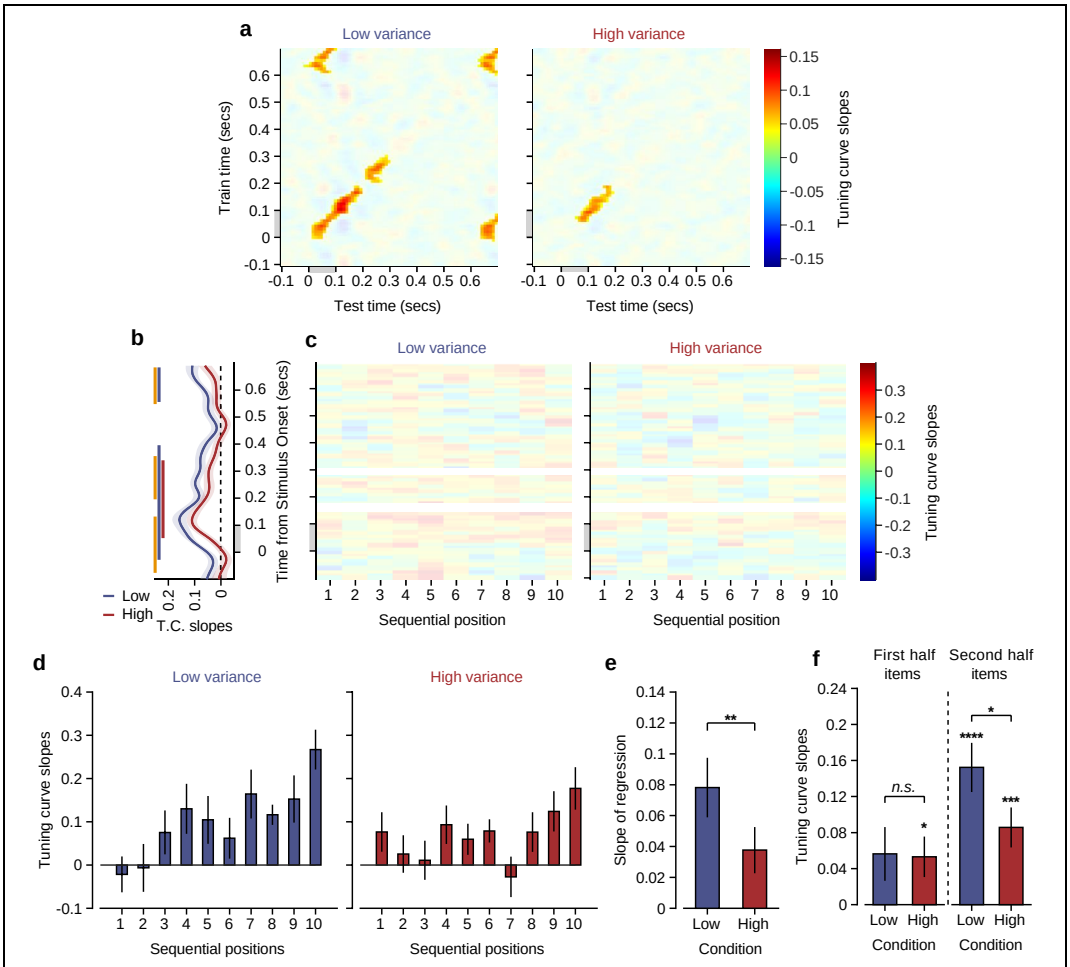

Figure 4. Temporal dynamics of the reconstructed tuning curves of mean orientation. The encoding model was trained and tested on stimulus-evoked activities from -100 – 700 ms after stimulus onset, with respect to their mean orientation of each trial. **(a)** Cross-temporal generalization of tuning curve slope of mean orientation. The tuning curve slope of mean orientation for low- (left) and high- (right) variance conditions was estimated by training weights on one time point in the training data, and applying them to all time points in the test data. The transparency mask highlights the significant clusters where tuning-curve slopes are greater than 0 (one-tailed, $p < 0.05$, based on cluster extent). **(b)** Time-resolved tuning curve slopes with respect to the mean orientation (left). Red and blue bars on the right side

represent the time period when tuning-curve slope is significantly greater than 0 (one-tailed, $p < 0.05$, based on cluster extent; red: high-variance, blue: low-variance). The orange bar indicates the time period when tuning-curve slope is significantly different between the two variance conditions. **(c)** Time-resolved tuning curve slopes with respect to the mean orientation at ten sequential positions for low-variance (left) and high-variance conditions (right). Color represents tuning-curve slope. The transparency mask highlights the significant clusters where linear slopes across sequential positions are positively significant for all trials, regardless of their variance conditions. (See ***Materials and Methods***). Gray bars represent the time period when the stimulus was presented. **(d)** Tuning-curve slope bars as a function of sequential positions averaged across time on the significant clusters in (b). Asterisks in the top center indicate the significance of linear trend across sequential positions. **(e)** The linear regression slope of tuning-curve slopes across sequential positions for the same time period as (d). **(f)** Average tuning-curve slope of the first half stimuli (from the first to fifth stimuli; left) and second half stimuli (from the sixth to tenth stimuli; right) in a sequence for the same period as (d). Error bars in (d-f) indicate ±1SEM. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

When we performed the same linear regression analyses on the accuracy of the representations of individual stimulus orientations recovered from labeling the stimulus-evoked activities as their physical orientations as in Figure 3, we did not find gradual increases in the accuracy of neural representation over sequential positions in either condition (Figure 5). The absence of positive linear trend of the stimulus-coding accuracy suggests that the gradual increase in the mean-coding accuracy is not simply due to the increase in the signal-to-noise ratio with increasing sequential positions. Together, our results indicate that

the positive linear trend of the mean-coding accuracy is due to the sequential updating of the mean orientation information after each stimulus onset.

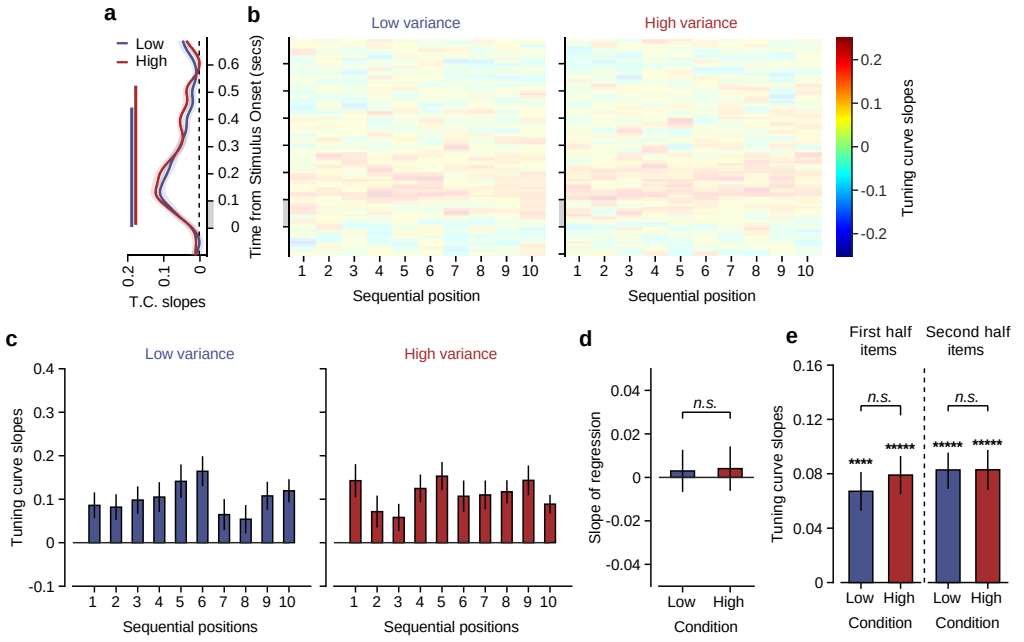

Figure 5. Temporal dynamics of the reconstructed tuning curves of physical orientation. The encoding model was trained and tested on stimulus-evoked activities from -100 – 700 ms after stimulus onset, with respect to their individual physical orientations. **(a)** Time-resolved tuning curve slopes with respect to the physical orientation (left). Red and blue bars on the right side represent the time period when tuning-curve slope is significantly greater than 0 (one-tailed, $p < 0.05$, based on cluster extent; red: high-, blue: low-variance). **(b)** Time-resolved tuning curve slopes with respect to the physical orientation at ten sequential positions for low-variance (left) and high-variance conditions (right). Color represents tuning-curve slope. The transparency mask highlights the significant clusters where linear regression slopes across sequential positions are positively significant for all trials, regardless of their variance conditions. (See ***Materials and Methods***). In this analysis with physical orientation, there was no significant cluster in both variance conditions. Gray bars represent

the time period when stimulus was presented. **(c)** Tuning-curve slope bars as a function of a sequential position averaged across all time points in the significant clusters of Figure 4b. These time points were used because the linearly increasing trend of physical orientation representation was not found in this analysis. **(d)** Linear regression slope of tuning-curve slopes across sequential positions for the same time period as in (c). **(e)** Average tuning-curve slope of the first half stimuli (from the first to fifth stimuli; left) and second half stimuli (from the sixth to tenth stimuli; right) in a sequence for the same period as (c). Error bars in (c-e) indicate ±1SEM. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, ****: $p < 0.0001$.

The role of frontoparietal region in sequential perceptual averaging

Despite the limitation of low spatial resolution of EEG, we investigated where the sequential update was processed in the brain by splitting total electrodes into three electrode clusters (36 anterior, 37 middle, and 37 posterior electrodes). We performed the same linear regression analysis to search for the time points where the regression slope of tuning curve slope values over sequence was significantly positive (see *Materials and Methods* for details). Only when the significant time cluster was found, tuning curve slope values at each sequential position were averaged within the significant time cluster. We then performed the same linear trend analysis on ten averaged tuning-curve slopes to test the hypothesis of the linear improvement of mean orientation representation over sequence in each electrode cluster. The linearly increasing trend of mean-coding accuracy was not identified in the posterior region cluster (Figure 6a), while the neural representation of the mean orientation gradually became more precise as a function of sequential position in the anterior and central regions (Figures 6b & 6c). Specifically in the anterior electrode cluster, the update of the mean orientation occurred every 0.16 s after each sequential stimulus onset as shown in the highlighted time cluster in Figure 6c. This indicates that the mean orientation is updated

regularly in frontal region. On the other hand, the sequential update process occurred at two separate highlighted time clusters around 0.16 s and 0.3 s in the middle electrode cluster (Figure 6b). This suggests that the rate of evidence accumulation is not fixed in parietal region. Note that no bar graphs were plotted in posterior electrode cluster because there was not a single time point where the regression slope of tuning curve slope values over sequence was significantly positive in both variance conditions. These results are in line with previous studies showing that prefrontal and parietal cortex encode task-general information as well as task-specific information (Swaminathan and Freedman, 2012; Ester et al., 2015; Sarma et al., 2015; Kim et al., 2017; Oh et al., 2019).

Caution is necessary regarding the underlying neural sources of the sequential updates because we focused our analyses on the full EEG signals above 2 Hz to minimize the effect of the physically driven SSVEP at the stimulus presentation frequency on the representational dynamics (see ***Materials and Methods*** for details). The analyzed multivariate EEG signals above 2 Hz are still a mixture of exogenous higher harmonic SSVEPs and endogenous oscillatory signals that cannot be completely disentangled from each other. Especially, the strong stimulus-driven SSVEP harmonics appeared in the posterior electrode cluster not in the middle and anterior electrode clusters. In the light of these points, the fact that the linearly increasing trend of mean-coding accuracy was not identified in the posterior electrode cluster strongly suggests that the widespread endogenous rather than exogenous dynamic network activity underlies the sequential averaging process especially in frontoparietal region. The fact that the mean-coding accuracy extracted from the EEG signals including low frequency activity (< 2 Hz) did not linearly increase across sequential positions (data not shown) also suggests that the sequential integration is not due to the high signal-to-noise ratio of SSVEP at physically driven stimulation frequency although this SSVEP component may still play a role in resetting the update time points. Prior neurophysiological studies suggest that large-

scale dynamic network interactions that span multiple brain regions are involved in various perceptual and cognitive processes (Donner et al., 2007; Pesaran et al., 2008; Siegel et al., 2012; Zhang et al., 2018).

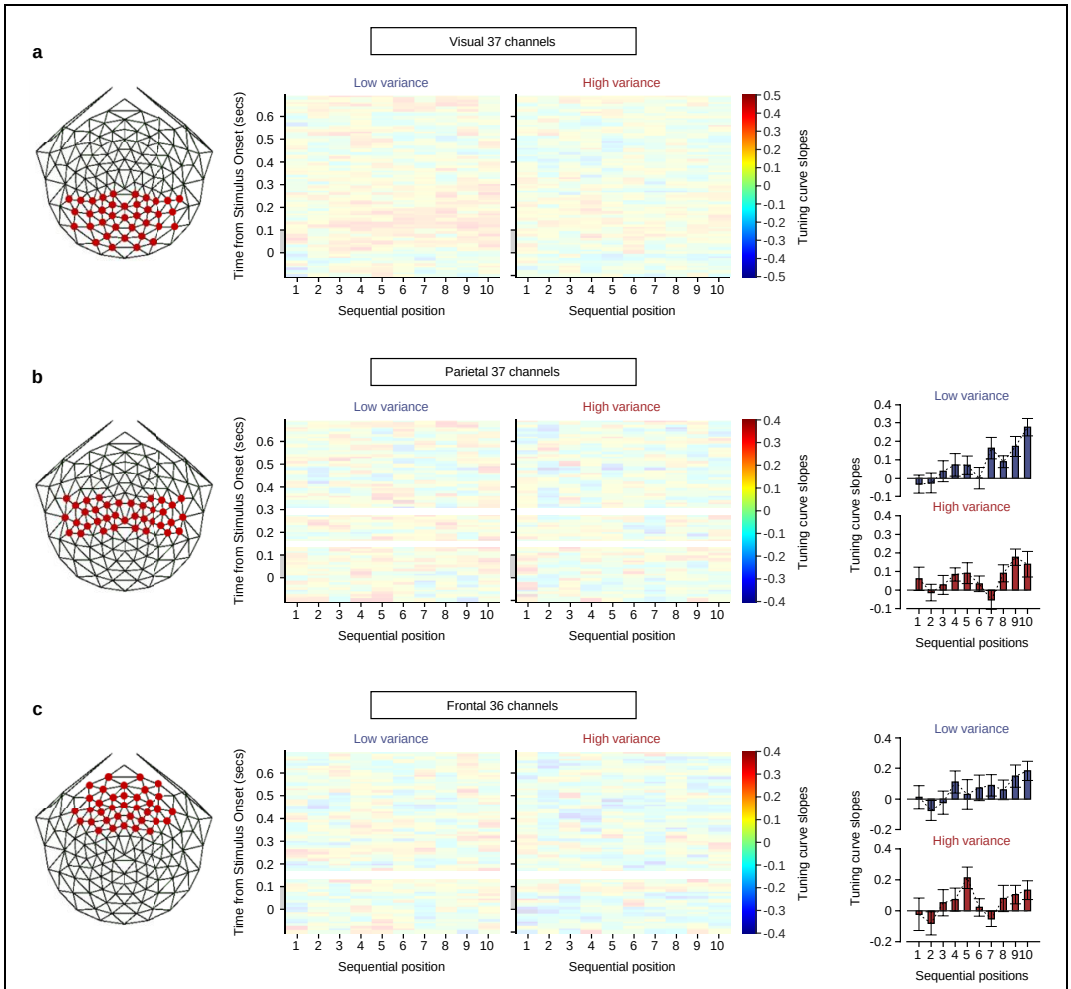

Figure 6. Temporal dynamics of the reconstructed tuning curves of mean orientation for visual, parietal, and frontal electrode clusters indicated by red dots in the 1st columns of (a-c). **(a)** The 2nd and 3rd columns indicate the time-resolved tuning curve slopes with respect to the mean orientation at ten sequential positions in the visual region for low-variance and high-variance conditions, respectively. The transparency mask highlights the time clusters where linear regression slopes across sequential positions are significantly positive for all trials,

regardless of their variance conditions (See ***Materials and Methods***). Note that there is no significant time cluster in the visual region. Gray bars represent the time period when stimulus was presented. **(b)** The 2nd and 3rd columns indicate the time-resolved tuning curve slopes with respect to the mean orientation at ten sequential positions in the parietal region for low-variance and high-variance conditions, respectively. There are two significant time clusters in the parietal region. Both upper and lower graphs of the 4th columns indicate tuning-curve slope bars as a function of a sequential position averaged across all time points in the two significant time clusters, for low-variance and high-variance conditions, respectively. Asterisks in the top center indicate the significance of linear trend across sequential positions. **(c)** The 2nd and 3rd columns indicate the time-resolved tuning curve slopes with respect to the mean orientation at ten sequential positions in the frontal region for low-variance and high-variance conditions, respectively. There is one significant time cluster in the frontal region. Both upper and lower graphs of the 4th columns indicate tuning-curve slope bars as a function of a sequential position averaged across all time points in the significant time cluster, for low-variance and high-variance conditions, respectively. Asterisks in the top center indicate the significance of linear trend across sequential positions. Error bars in (b-c) indicate ±1SEM.

Discussion

To probe the mechanisms of sequential averaging, we combined computational modeling of behavioral data with a multivariate pattern analysis that visualized how the neural representation of the mean orientation developed while viewing a sequence of differently oriented Gabor stimuli. The pattern analysis revealed that the dynamically evolving patterns of the stimulus-evoked EEG activities encoded the mean orientation as well as the stimulus orientation (Figure 3e and 4a). Regardless of whether or not the variance of

the individual stimulus orientations was low or high, the neural representations of individual stimulus orientations were equally precise (Figure 3d) and the tuning-curve slope of the mean orientation increased gradually across the ten sequential positions (Figure 4c and 4d). This linearly increasing trend was steeper in the low- than in the high-variance condition (Figure 4e). These results are consistent with the sequential update model which predicted that multiple stimuli would be sequentially integrated regardless of stimulus variance, but high stimulus variance would make the updating noisier (Figure 2a-b). Together, these findings suggest that environmental volatility mainly influences the noisiness in integrating sequential stimuli and the encoding quality of the mean information towards the end of the sequence, leading to behavioral difference in perceptual mean judgment between the low- and high-variance conditions.

Our results have a number of implications for understanding sequential averaging mechanisms. First, the manner of sequential information processing may depend on the demand of the behavioral task. The present study required observers to integrate a series of briefly presented multiple stimuli, which is likely to make the task more challenging in the high-variance condition than in the low-variance condition. This demanding task under time pressure may not allow observers to spend more time integrating the current stimulus with the previous stimulus, eventually leading to less precise mean information in the high-variance condition. This is in line with previous studies demonstrating the detrimental effects of time pressure on overall decision quality, with general finding that individuals perform significantly worse under time pressure (Payne et al., 1988; Sutter et al., 2003; Kocher and Sutter, 2006; Ahituv et al., 2015). However, when the task is to integrate information on the continuously changing single target stimulus rather than across different successive stimuli as in the current study, observers change the perceptual integration timescale during the perceptual discrimination task (Burr and Santoro, 2001; Kiani et al., 2008). A recent sound

texture perception study revealed an obligatory multi-second averaging process whose duration could not be controlled at will and was much longer for highly variable textures (McWalter and McDermott, 2018). In another perceptual decision-making study, where one of the two concurrently presented streams of luminance changing discs showed a brief increment against the mean luminance of the visual stream, observers employed an integration timescale adapted to the target signal duration (Ossmy et al., 2013). Unlike these studies that revealed the time required for a single target signal integration, the current study revealed the time required for each sequential stimulus to be integrated with the following stimulus. On the other hand, when the task is to categorize or identify the current stimulus instead of integrating the serially presented stimuli, human observers rely on their working memory capacity to remember a few recent stimuli in a volatile environment rather than a sequential update strategy (Summerfield et al., 2011; Fischer and Whitney, 2014; Laquitaine and Gardner, 2018). In this case, the sequential updating may not be efficient because it runs the risk of pooling together stimuli with distinct statistical properties. In contrast, when estimating the mean orientation of the visual stream as accurately as possible, it may be advantageous to keep track of every single stimulus regardless of variability. Thus, previous studies and our results indicate that the task demand adaptively shapes the evidence integration computation even under the same volatile environment. In our case, volatility did not change the integration timescale, but only added uncertainty to the sequential update process shaped by the task goal of extracting the mean across time.

The appearance of the increasingly precise mean information at specific delays after each stimulus onset (Figures 4c & 6b-c) indicated that the integration of successive stimuli occurred regularly. In frontal region, the update of the mean orientation occurred every 0.16 s after each sequential stimulus onset (The highlighted time cluster in Figure 6c). In parietal region (Figure 6b) and all electrodes (Figure 4c), the sequential update process occurred at

two separate highlighted time clusters around 0.16 s and 0.3 s. These results suggest that the sequential update arises with either a shorter latency in the anterior brain area than in the middle brain area, or an equal latency in both brain areas. In any case, the updated mean orientation was encoded in the anterior brain area no later than in the middle brain area, suggesting either an early or simultaneous involvement of the anterior brain area in sequential perceptual averaging process. A recent study on the binary perceptual motion categorization similarly showed that prefrontal cortex leads the decision process when monkey determines whether sequentially presented stimuli belong to the same motion category or not during a delayed match to category task (Zhou et al., 2021). This neurophysiological study has found that neurons in both prefrontal and parietal cortex are involved in categorical encodings of individual stimuli but the prefrontal cortex appears more directly involved than the parietal cortex in transforming categorical encoding into the abstract match/nonmatch decision by integrating the previously presented stimulus with the currently visible stimulus. In the current study, the shorter-latency of regular update in the anterior brain area may imply a flow of sequentially integrated information from the anterior brain area to other brain areas. Since observers employed the same integration timescale in both variance conditions at the expense of the precise integration of highly variable stimuli, our findings provide neural mechanisms to differentially accumulate increasingly abstract feature from a concrete piece of information across the cortical hierarchy depending on environmental volatility.

Our findings also shed light as to whether or not all individual stimuli are encoded during sequential averaging. Previous behavioral studies on sequential averaging indicate that observers use only a subset of stimuli to accomplish mean representation without explicitly encoding every individual stimulus (Corbett and Oriet, 2011; Gorea et al., 2014). However, our finding that spatially distributed EEG activities encoded individual stimulus orientations in both variance conditions with an equivalent precision suggests the robust encoding of

individual stimuli during a sequential averaging task. Thus, one possibility for the difference between our results and other behavioral studies is that even though individual orientations are transiently encoded in the visual system, observers do not form robust memories of individual orientations so that they cannot be reported or recognized later after the sequence. The gradually increasing mean-coding accuracy across sequential positions also suggests that each stimulus is sequentially integrated in both variance conditions (Figure 4c and 4d). This overall trend of approximately linear improvement in the mean representation over sequence is in line with recent neuroimaging and behavioral studies based on the sequential update model framework (Cheadle et al., 2014; Navajas et al., 2017). From the perspective of the sequential update model, the sub-sampling strategy can be regarded as a special case of the weighted whole-set averaging strategy (Juni et al., 2012; Hubert-Wallander and Boynton, 2015), when some of the encoded stimuli are integrated with no weight for computing the mean of the sequence. Nevertheless, it is unlikely that only a subset of the encoded stimuli is integrated to accomplish mean representation because all stimuli in the sequence had positive weights (Figure 1d). In fact, this regression bias toward the mean of the stimulus distribution is one of the most robust empirical regularities in studies of human perceptual judgment across various perceptual domains (Hollingworth, 1910; Stevens and Greenbaum, 1966; Oh et al., 2019; Xiang et al., 2021). Especially in Oh et al.'s work on visual working memory, the neural representation of mean orientation emerges even though there was no task requirement of judging perceptual mean orientation of concurrently presented oriented bars, leading to the participant's biased judgment on target orientation toward the ensemble mean orientation. Additionally, many previous studies reported bias effects of recent stimulation history on perceptual process such as serial dependence (Summerfield et al., 2011; Kiyonaga et al., 2017; Pascucci et al., 2019). Thus, it is highly likely that both the regression effect and the recent history effect make it challenging for observers to perceive and maintain the

801 stimulus attribute as it is. The perceptual distortion of individual stimuli due to these bias
802 effects is another possible factor for contributing to the incorrect or failed report of a single
803 item after the sequence in previous behavioral studies (Corbett and Oriet, 2011; Gorea et al.,
804 2014).

805 Finally, future work on sequential information processing will benefit from
806 characterizing the representational dynamics of the sequential averaging process by
807 systematically manipulating temporal regularity. Such studies could exploit heterochronous
808 streams of events to clarify whether the sequentially updated mean information is reflected in
809 exogenous or endogenous oscillatory signals. This manipulation will allow us to characterize
810 electrophysiological signatures of the sequential averaging mechanism by measuring whether
811 the mean-coding accuracy gradually increases or fluctuates along sequential positions. When
812 observers maintain the sequential update strategy, one can probe whether their integration
813 timescale adaptively changes with temporal regularity and environmental volatility. Further,
814 it is important to examine whether or not the mean information develops automatically over
815 sequence as a function of task. For example, by changing both orientations and spatial
816 frequencies in the stream of Gabor patches and asking observers to report the mean
817 orientation or the mean spatial frequency or one particular item's spatial frequency, one can
818 examine the task-dependency of the sequential averaging process.

819 In summary, stimulus-specific coding is regularly transformed into a brief coding of
820 the integrated information that becomes closer to the sequential mean towards the end of the
821 sequence for the appropriate behavioral response during sequential averaging task. Therefore,
822 the sequential averaging process can be best characterized as perceptual dynamics that
823 swings back and forth between evidence sampling and integration for the perceptual
824 judgment of the mean at the end of the sequence. Finally, the steeper linear trend of the mean
825 tuning-curve slope across sequential positions in the low- than in the high-variance condition

826 suggests that the multivariate activity pattern dynamics underlie the differential sequential
827 averaging process depending on environmental variability.
828
829
830 **References**

831 Ahituv N, Igbaria M, Sella AV (2015) The Effects of Time Pressure and Completeness of
832 Information on Decision Making. J Manag Inf Syst 15:153-172.
833 Albrecht AR, Scholl BJ, Chun MM (2012) Perceptual averaging by eye and ear: computing
834 summary statistics from multimodal stimuli. Atten Percept Psychophys 74:810-815.
835 Alvarez GA (2011) Representing multiple objects as an ensemble enhances visual cognition.
836 Trends Cogn Sci 15:122-131.
837 Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433-436.
838 Brouwer GJ, Heeger DJ (2009) Decoding and reconstructing color from responses in human
839 visual cortex. J Neurosci 29:13992-14003.
840 Burr DC, Santoro L (2001) Temporal integration of optic flow, measured by contrast and
841 coherence thresholds. Vision Res 41:1891-1899.
842 Busch NA, VanRullen R (2010) Spontaneous EEG oscillations reveal periodic sampling of
843 visual attention. Proc Natl Acad Sci U S A 107:16048-16053.
844 Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts
845 visual perception. J Neurosci 29:7869-7876.
846 Cheadle S, Wyart V, Tsetsos K, Myers N, de Gardelle V, Herce Castanon S, Summerfield C
847 (2014) Adaptive gain control during human perceptual choice. Neuron 81:1429-1441.
848 Chong SC, Treisman A (2005) Statistical processing: computing the average size in
849 perceptual groups. Vision Res 45:891-900.

Corbett JE, Oriet C (2011) The whole is indeed more than the sum of its parts: perceptual averaging in the absence of individual item representation. Acta Psychol (Amst) 138:289-301.

Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2013) Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. J Neurosci 33:4002-4010.

Dakin SC (1999) Orientation variance as a quantifier of structure in texture. Spat Vis 12:1-30.

de Gardelle V, Summerfield C (2011) Robust averaging during perceptual judgment. Proc Natl Acad Sci U S A 108:13341-13346.

de Gardelle V, Mamassian P (2015) Weighting mean and variability during confidence judgments. PLoS One 10:e0120870.

Donner TH, Siegel M, Oostenveld R, Fries P, Bauer M, Engel AK (2007) Population activity in the human dorsal pathway predicts the accuracy of visual motion detection. J Neurophysiol 98:345-359.

Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. Neuroimage 39:647-660.

Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87:893-905.

Fiebelkorn IC, Saalmann YB, Kastner S (2013) Rhythmic sampling within and between objects despite sustained attention at a cued location. Curr Biol 23:2553-2558.

Fiebelkorn IC, Pinsk MA, Kastner S (2018) A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. Neuron 99:842-853 e848.

Fischer J, Whitney D (2014) Serial dependence in visual perception. Nat Neurosci 17:738-743.

875 Foster JJ, Bsales EM, Jaffe RJ, Awh E (2017) Alpha-Band Activity Reveals Spontaneous
876 Representations of Spatial Position in Visual Working Memory. Curr Biol 27:3216-
877 3223.e3216.
878 Freeman J, Simoncelli EP (2011) Metamers of the ventral stream. Nat Neurosci 14:1195-
879 1201.
880 Garcia JO, Srinivasan R, Serences JT (2013) Near-Real-Time Feature-Selective Modulations
881 in Human Cortex. 23:515-522.
882 Gorea A, Belkoura S, Solomon JA (2014) Summary statistics for size over space and time.
883 14:22-22.
884 Greenwood JA, Bex PJ, Dakin SC (2009) Positional averaging explains crowding with letter-
885 like stimuli. Proc Natl Acad Sci U S A 106:13130-13135.
886 Haberman J, Harp T, Whitney D (2009) Averaging facial expression over time. J Vis 9:1 1-
887 13.
888 Haberman J, Lee P, Whitney D (2015) Mixed emotions: Sensitivity to facial variance in a
889 crowd of faces. J Vis 15:16.
890 Hanslmayr S, Volberg G, Wimber M, Dalal SS, Greenlee MW (2013) Prestimulus oscillatory
891 phase at 7 Hz gates cortical information flow and visual perception. Curr Biol
892 23:2273-2278.
893 Helfrich RF, Fiebelkorn IC, Szczepanski SM, Lin JJ, Parvizi J, Knight RT, Kastner S (2018)
894 Neural Mechanisms of Sustained Attention Are Rhythmic. Neuron 99:854-865 e855.
895 Hollingworth HL (1910) The Central Tendency of Judgment. The Journal of Philosophy,
896 Psychology and Scientific Methods 7:461.
897 Hubert-Wallander B, Boynton GM (2015) Not all summary statistics are made equal:
898 Evidence from extracting summaries across time. J Vis 15:5.

899 Juni MZ, Gureckis TM, Maloney LT (2012) Effective integration of serially presented
900 stochastic cues. J Vis 12.
901 Kiani R, Hanks TD, Shadlen MN (2008) Bounded Integration in Parietal Cortex Underlies
902 Decisions Even When Viewing Duration Is Dictated by the Environment. J Neurosci
903 28:3017-3029.
904 Kim YJ, Tsai JJ, Ojemann J, Verghese P (2017) Attention to Multiple Objects Facilitates
905 Their Integration in Prefrontal and Parietal Cortex. J Neurosci 37:4942-4953.
906 King J-R, Wyart V (2019) The Human Brain encodes a Chronicle of Visual Events at each
907 Instant of Time. BioRxiv:846576.
908 Kiyonaga A, Scimeca JM, Bliss DP, Whitney D (2017) Serial Dependence across Perception,
909 Attention, and Memory. Trends Cogn Sci 21:493-497.
910 Kocher MG, Sutter M (2006) Time is money—Time pressure, incentives, and the quality of
911 decision-making. J Econ Behav Organ 61:375-392.
912 Landau AN, Fries P (2012) Attention samples stimuli rhythmically. Curr Biol 22:1000-1004.
913 Landau AN, Schreyer HM, van Pelt S, Fries P (2015) Distributed Attention Is Implemented
914 through Theta-Rhythmic Gamma Modulation. Curr Biol 25:2332-2337.
915 Laquitaine S, Gardner JL (2018) A Switching Observer for Human Perceptual Estimation.
916 Neuron 97:462-474 e466.
917 Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. J
918 Neurosci Methods 164:177-190.
919 Marti S, Dehaene S (2017) Discrete and continuous mechanisms of temporal selection in
920 rapid visual streams. Nat Commun 8:1955.
921 McWalter R, McDermott JH (2018) Adaptive and Selective Time Averaging of Auditory
922 Scenes. Curr Biol 28:1405-1418 e1410.

923 Morillon B, Schroeder CE, Wyart V, Arnal LH (2016) Temporal Prediction in lieu of
924 Periodic Stimulation. J Neurosci 36:2342-2347.
925 Myers NE, Rohenkohl G, Wyart V, Woolrich MW, Nobre AC, Stokes MG (2015) Testing
926 sensory evidence against mnemonic templates. Elife 4:e09000.
927 Navajas J, Hindocha C, Foda H, Keramati M, Latham PE, Bahrami B (2017) The
928 idiosyncratic nature of confidence. Nat Hum Behav 1:810-818.
929 Nolan H, Whelan R, Reilly RB (2010) FASTER: Fully Automated Statistical Thresholding
930 for EEG artifact Rejection. J Neurosci Methods 192:152-162.
931 Oh BI, Kim YJ, Kang MS (2019) Ensemble representations reveal distinct neural coding of
932 visual working memory. Nat Commun 10.
933 Ossmy O, Moran R, Pfeffer T, Tsetsos K, Usher M, Donner TH (2013) The timescale of
934 perceptual evidence integration can be adapted to the environment. Curr Biol 23:981-
935 986.
936 Pascucci D, Mancuso G, Santandrea E, Della Libera C, Plomp G, Chelazzi L (2019) Laws of
937 concatenated perception: Vision goes for novelty, decisions for perseverance. PLOS
938 Biology 17:e3000144.
939 Payne JW, Bettman JR, Johnson EJ (1988) Adaptive strategy selection in decision making. J
940 Exp Psychol Learn Mem Cogn 14:534.
941 Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers
942 into movies. Spat Vis 10:437-442.
943 Pesaran B, Nelson MJ, Andersen RA (2008) Free choice activates a decision circuit between
944 frontal and parietal cortex. Nature 453:406-409.
945 Piazza EA, Sweeny TD, Wessel D, Silver MA, Whitney D (2013) Humans use summary
946 statistics to perceive auditory sequences. Psychol Sci 24:1389-1397.

947 Sarma A, Masse NY, Wang X-J, Freedman DJ (2015) Task-specific versus generalized
948 mnemonic representations in parietal and prefrontal cortices. 19:143-149.
949 Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of
950 sensory selection. Trends Neurosci 32:9-18.
951 Siegel M, Donner TH, Engel AK (2012) Spectral fingerprints of large-scale neuronal
952 interactions. Nat Rev Neurosci 13:121-134.
953 Stevens SS, Greenbaum HB (1966) Regression effect in psychophysical judgment. Percept
954 Psychophys 1:439-446.
955 Summerfield C, Behrens TE, Koechlin E (2011) Perceptual classification in a rapidly
956 changing environment. Neuron 71:725-736.
957 Sutter M, Kocher M, Strauß S (2003) Bargaining under time pressure in an experimental
958 ultimatum game. Econ Lett 81:341-347.
959 Swaminathan SK, Freedman DJ (2012) Preferential encoding of visual categories in parietal
960 cortex compared with prefrontal cortex. Nat Neurosci 15:315-320.
961 Whitney D, Yamanashi Leib A (2018) Ensemble Perception. Annu Rev Psychol 69:105-129.
962 Wyart V, de Gardelle V, Scholl J, Summerfield C (2012) Rhythmic fluctuations in evidence
963 accumulation during decision making in the human brain. Neuron 76:847-858.
964 Xiang Y, Graeber T, Enke B, Gershman SJ (2021) Confidence and central tendency in
965 perceptual judgment. Attention, Perception, & Psychophysics.
966 Zhang H, Watrous AJ, Patel A, Jacobs J (2018) Theta and Alpha Oscillations Are Traveling
967 Waves in the Human Neocortex. Neuron 98:1269-1281 e1264.
968 Zhou Y, Rosen MC, Swaminathan SK, Masse NY, Zhu O, Freedman DJ (2021) Distributed
969 functions of prefrontal and parietal cortices during sequential categorical decisions.
970 Elife 10.
971