Moon and Kwon *BMC Biology* (2022) 20:247
https://doi.org/10.1186/s12915-022-01444-7

BMC Biology

RESEARCH ARTICLE Open Access

Attractive and repulsive effects of sensory history concurrently shape visual perception

Jongmin Moon and Oh-Sang Kwon*

Abstract

Background: Sequential effects of environmental stimuli are ubiquitous in most behavioral tasks involving magnitude estimation, memory, decision making, and emotion. The human visual system exploits continuity in the visual environment, which induces two contrasting perceptual phenomena shaping visual perception. Previous work reported that perceptual estimation of a stimulus may be influenced either by attractive serial dependencies or repulsive aftereffects, with a number of experimental variables suggested as factors determining the direction and magnitude of sequential effects. Recent studies have theorized that these two effects concurrently arise in perceptual processing, but empirical evidence that directly supports this hypothesis is lacking, and it remains unclear whether and how attractive and repulsive sequential effects interact in a trial. Here we show that the two effects concurrently modulate estimation behavior in a typical sequence of perceptual tasks.

Results: We first demonstrate that observers' estimation error as a function of both the previous stimulus and response cannot be fully described by either attractive or repulsive bias but is instead well captured by a summation of repulsion from the previous stimulus and attraction toward the previous response. We then reveal that the repulsive bias is centered on the observer's sensory encoding of the previous stimulus, which is again repelled away from its own preceding trial, whereas the attractive bias is centered precisely on the previous response, which is the observer's best prediction about the incoming stimuli.

Conclusions: Our findings provide strong evidence that sensory encoding is shaped by dynamic tuning of the system to the past stimuli, inducing repulsive aftereffects, and followed by inference incorporating the prediction from the past estimation, leading to attractive serial dependence.

Keywords: Vision, Perception, Perceptual bias, Sensory adaptation, Aftereffect, Serial dependence, Bayesian inference, Ideal observer, Encoder-decoder model

Background

Sensory input from the natural environment is highly structured, and humans can learn statistical regularities in sensory input and exploit them to optimize perceptual processing [1–3]. Considerable advances have been made toward understanding whether and how the visual system incorporates knowledge of environmental statistics. For example, knowledge about the distribution of local orientation in natural images can be used not only to produce an efficient representation of the incoming orientation signal but also to optimize the interpretation of that orientation representation [4]. The relevant environmental statistics are not limited to static distributions. Most everyday visual tasks require more than processing static images, as many visual features change continuously over time [5, 6]. For example, orientation signals arising in the natural world tend to change gradually, such that the local orientation at any given moment correlates with the orientation of the previous moment [6]. To process

*Correspondence: oskwon@unist.ac.kr

Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulsan 44919, South Korea

such dynamic sensory inputs efficiently and accurately, the visual system should consider the temporal structure between successive stimuli in the environment [7].

There are two ways for the visual system to exploit knowledge about changing environments. First, temporal structures in sensory input provide an opportunity for the visual system to form an efficient representation of sensory information. The incoming stream of input is partially redundant because the current sensory input tends to resemble the preceding one. Therefore, the visual system can exploit the redundancy to enhance efficiency in neural coding by adaptively adjusting the response properties of sensory neurons according to the changes in input stimuli [8–16]. The perceptual consequence of sensory adaptation is most clearly seen in repulsive aftereffects, a phenomenon in which exposure to a stimulus induces a repulsive bias in the perception of subsequent stimuli [17–21]. Second, predictions derived from recent sensory inputs can be used to interpret the encoded signals. Numerous studies have shown that the visual system utilizes prior knowledge about the sensory environment to infer the state of the world from noisy and incomplete sensory inputs [22–25]. In the context of processing dynamic sensory inputs, the current sensory input is integrated with predictions from the recent past to optimize perceptual estimation [25]. The optimal integration model provides a normative explanation for serial dependence in perceptual behavior [6, 7, 26, 27], a phenomenon in which perceptual estimation of the current stimulus is attracted toward the previous stimuli [28, 29].

Both repulsive aftereffects and attractive serial dependence are well-known phenomena in which the visual system leverages information from the recent past to optimize the processing of incoming visual input. These two opposite effects, however, have been historically described and theorized about in isolation. Only recently have suggestions been made that these two effects may concurrently occur in perceptual processing [30–32]. The implied picture is that one effect is hidden when the other becomes behaviorally observable, but this theoretical consideration lacks direct empirical supports. Most previous studies focused on showing that either, but not both, of the two effects can be signified depending on the specifications of the experimental setup, such as stimulus features [33, 34], time interval between consecutive events [35, 36], and task design [31, 37, 38]. While these studies convergingly showed that the sequential effect varies drastically between attraction and repulsion with only a small change in experimental paradigm, they do not empirically confirm whether and how the two effects with opposite behavioral consequences occur in a single perceptual processing, which is crucial for the investigation of underlying mechanisms [39]. A group of studies showed that perceptual estimates of the current stimulus are attracted to the immediately preceding stimulus and repelled away from stimuli further back in the past (or vice versa) [30, 40–42]. However, even these studies cannot determine whether a single sensory event can induce both attractive and repulsive biases in the immediately subsequent perception.

In this study, we test a key prediction of the current theory about attractive and repulsive sequential effects—that the two effects concurrently occur in perceptual processing [30, 32] but with different determinants [31]. Specifically, we reason that an observer's current estimates would be repelled away from the previous *stimulus*, because the neural population that encodes incoming sensory information naturally adjusts its tuning characteristics according to what was encoded at the previous moment. That is, we consider the previous stimulus as a proxy for sensory measurement made by the observer in the previous trial. Indeed, recent studies showed that neural representations of the current stimulus are repelled away from the previous stimulus, in spite of attractive biases in behavior [32, 43], raising a possibility that repulsive biases might be also hidden in perceptual reports. We also reason that, at the same time, the estimates would be attracted toward the previous *response*, because the observer's prediction that optimizes the perceptual interpretation would be derived from what was perceived by the observer in the previous trial [6, 26]. In the absence of feedback, the observer does not have direct access to the true stimulus in the previous trial, so the final perception in the previous trial is the best estimate the observer has about the previous stimulus. Thus, although it may appear trivial at first glance to determine whether it is a stimulus and/or response that biases the subsequent perception, it has important implications on how and why the visual system processes sensory information according to the recent sensory events.

To test this hypothesis, we aimed to estimate the concurrent influence of the previous stimulus and response on current perception. In a series of trials, we asked subjects to view a field of moving dots and to report the perceived direction of motion (Fig. 1A). The subjects' responses to the direction of motion were systematically attracted toward the direction of motion in the previous trial, consistent with previous studies. Importantly, however, by representing the response error as a function of both the previous stimulus and response, we demonstrated that the estimation responses are repelled from the direction of motion in the previous trial, resembling the repulsive aftereffects, and at the same time attracted toward the reported direction of motion in the previous trial, mirroring the attractive

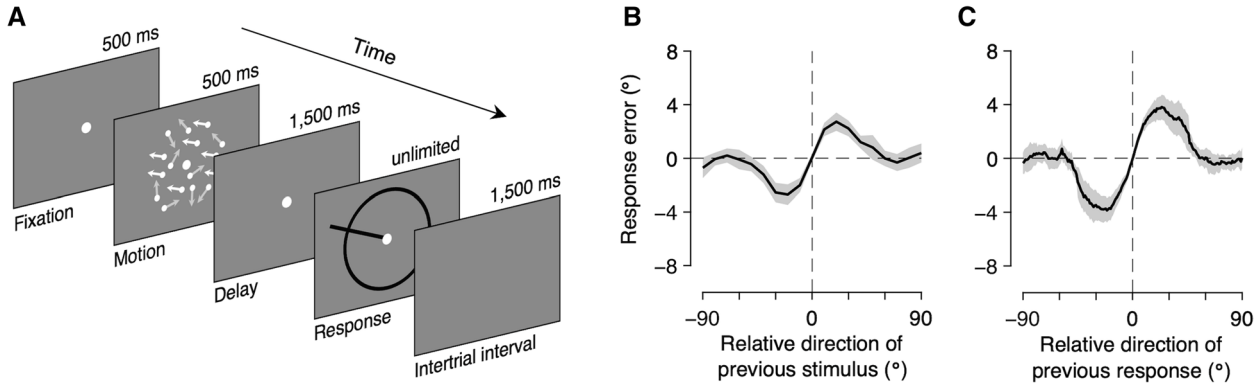

Fig. 1 Task procedure and marginal bias plots. **A** Sequence of events in a trial. On each trial, subjects viewed a random-dot motion stimulus and reported the perceived direction of motion by extending a dark bar from the center point. They were instructed to report the direction of motion of dots by swiping their finger on a touchpad to extend a bar from the center of the display to the direction of their estimate and confirm that report with a click on the touchpad. **B** Bias plot marginalized on the relative direction of the previous stimulus. Response errors are expressed as a function of the relative direction of the previous stimulus (i.e., previous stimulus minus current stimulus). Consistent with earlier studies, estimation responses are systematically biased to the previous stimulus. **C** Bias plot marginalized on the relative direction of the previous response. Response errors are expressed as a function of the relative direction of the previous response (i.e., previous response minus current stimulus). For each subject, bias was computed by averaging the response error over a sliding window (window width: 10°, step size: 1°). Averages with less than ten trials within a subject were excluded before illustration. Again, estimation responses are systematically biased to the previous response. Shaded regions represent 95% confidence intervals

serial dependence. A preliminary version of this work has been presented previously [44].

Results

Processing a sequence of images involves interactions between the current input and past observations, and the repulsive aftereffect and the attractive serial dependence are regarded as behavioral consequences of these interactions [20, 39]. We sought to investigate the determinants of these interactions by separately estimating the respective effects of stimulus and response on subsequent perception. In a typical perceptual estimation task, however, the value of a given stimulus and the subjects' response to it would be strongly correlated, making it challenging to distinguish which of those two is related to the observed bias. To effectively disentangle them, we carefully modulated the stimulus dynamics across subjects, thereby ensuring the collection of a sufficient number of informative trials. Specifically, subjects were divided into four groups of equal size and performed a motion direction estimation task (Fig. 1A) in which the direction of stimulus motion was randomly varied from the direction of motion in the previous trial following a uniform distribution with different ranges (±20°, ±40°, ±80°, or ±180°). This way, we were able to secure a large number of trials in which the previous stimulus and previous response were relatively similar to the current stimulus. Considering that both attractive and repulsive sequential effects are tuned to similarity between successive stimuli [17, 28], our data would be highly informative about the effects of previous stimulus and previous response. In addition, our design also allowed us to directly test the effect of experimentally imposed temporal statistics on subjects' estimation behavior, which will be dealt with in the "Discussion" section.

Attractive sequential effects in motion direction estimation

We first examined whether we successfully replicated previously reported serial dependence. To do this, we adopted the standard approach from earlier studies that focused on the effect of previous stimuli on current perception [28, 29]. First, the response error was computed as the angular difference between the reported and presented motion directions, with positive angles corresponding to clockwise deviations. In doing this, we corrected for the well-known estimation biases toward or away from the cardinal axes [4, 45], which could have been confounded with the sequential effects (Additional file 1: Fig. S1). We then examined the patterns of subjects' response errors as a function of the relative direction of the previous stimulus to the current stimulus (i.e., $\text{Stimulus}_{t-1} - \text{Stimulus}_t$). As expected, we found that the estimation responses were systematically biased to stimuli that were previously viewed (Fig. 1B; see also Additional file 1: Fig. S2). We also examined the response error as a function of the relative direction of the previous response to the current stimulus (i.e., $\text{Response}_{t-1} - \text{Stimulus}_t$) and found that the

responses were again attracted toward responses that were recently made (Fig. 1C). These sequential effects were highly significant (previous stimulus: $t_{(31)} = 8.572$, $P < 0.001$; previous response: $t_{(31)} = 8.317$, $P < 0.001$), and their magnitudes and attraction profiles were consistent with earlier reports [27, 31, 38], following a first derivative of Gaussian (DoG) curve (previous stimulus: amplitude = 2.91° [2.18 3.60], peak location = 20.28° [18.60 22.75]; previous response: amplitude = 4.10° [3.06 5.15], peak location = 22.48° [19.95 25.05]). Note that the observed effect of previous response is not artifacts unrelated to the sequential effect, as our correction procedure has effectively removed the confounding bias function of the current stimulus direction for each subject (Additional file 1: Fig. S3).

However, plotting response errors as a function of either previous stimulus or response direction does not reflect the isolated effect of the previous stimulus or response, since the direction of motion stimulus and the subjects' corresponding response to it are strongly correlated: The circular correlation coefficient between the true and reported direction of motion had a median (and interquartile range) of 0.988 (0.985–0.992) across subjects. Therefore, it would be inappropriate to judge which component of the previous trial truly influenced the response of the current trial based only on the patterns of biases marginalized over the previous stimulus or response (Fig. 1B, C).

Disentangling the effects of previous stimuli and responses

The subjects' estimation behavior in the current trial could have been influenced by the direction of the preceding stimulus, subjects' perceptual estimate of it, or both. To obtain a general insight into which aspects of the previous trial affected subjects' performance and how, we visualized subjects' response error in a two-dimensional map as a function of both previous stimulus and response directions (Fig. 2A; see also Additional file 1: Fig. S4). In this joint bias map, the x-axis represents the relative direction of the previous stimulus, the y-axis represents the relative direction of the previous response, and the color of each pixel represents the subjects' response error. Pixels with warm colors indicate positive errors, and pixels with cool colors indicate negative errors. Thus, both warm-colored pixels with positive labels on the axis and cool-colored pixels with negative labels on the axis

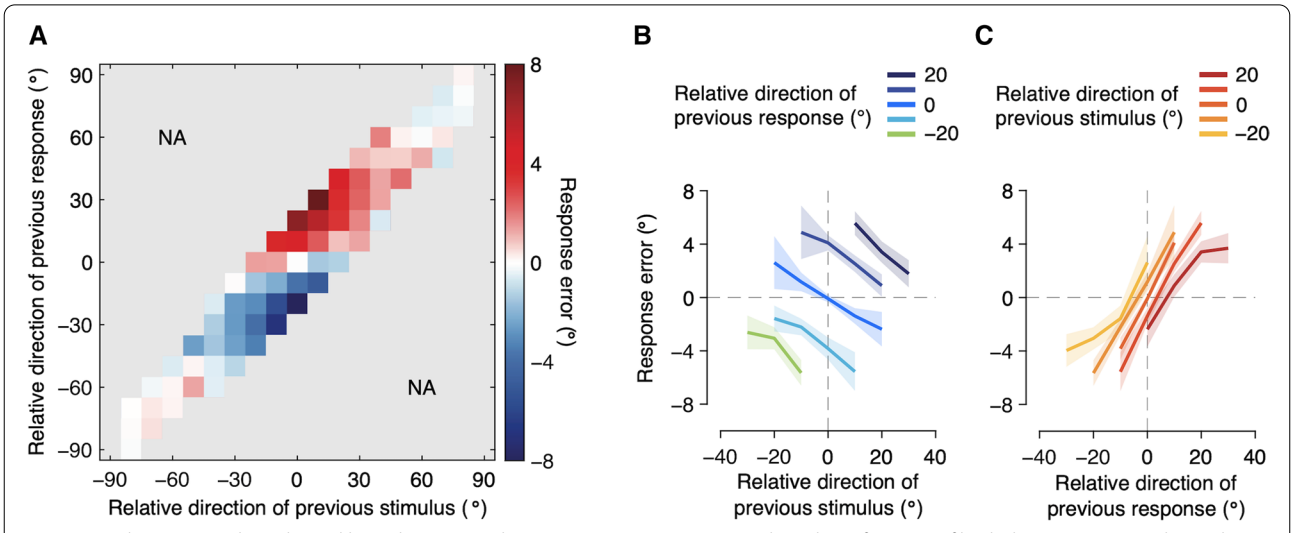

Fig. 2 Joint bias map and conditional bias plots. **A** Joint bias map. Response errors are plotted as a function of both the previous stimulus and response direction relative to the current stimulus. For each subject, we binned the response errors within 10° bins according to the relative direction of the previous stimulus and response. Group means with less than fourteen subjects' data were excluded before illustration. The color of each pixel represents response error on the current trial (warm color: positive error; cool color: negative error). Positive values on the x- and y-axis indicate that the previous stimulus and response direction was more clockwise than the current stimulus, respectively, and positive errors indicate that the estimated direction was more clockwise than the true stimulus direction. Estimation responses on the current trial are systematically repelled away from the stimulus direction of the previous trial, while they are strongly attracted toward the response direction of the previous trial. NA, not available. **B** Bias plot conditioned on relative direction of previous response. Response errors split by relative direction of the previous response were plotted as a function of relative direction of the previous stimulus. Bins with less than five trials within a subject and group means with less than fourteen subjects' data were excluded before illustration. As can be seen across five lines with negative slopes, estimation responses are negatively biased away from the previous stimulus. **C** Bias plot conditioned on relative direction of previous stimulus. Response errors split by relative direction of the previous stimulus were plotted as a function of relative direction of the previous response. Estimation responses are positively biased toward the previous response, as indicated by the positive slopes. Shaded regions represent 95% confidence intervals

represent attractive biases, whereas warm-colored pixels with negative labels on the axis and cool-colored pixels with positive labels on the axis represent repulsive biases.

Visual inspection of the joint bias map reveals that it has a rich structure beyond the traditional view of the sequential effect. Specifically, it becomes immediately apparent from the map that the sequential effect cannot be fully characterized by a simple attractive or repulsive bias and is not solely determined by either the previous stimulus or the previous response. Instead, errors in direction estimates seem to be driven by both attractive and repulsive biases that depend systematically on both the previous stimulus and previous response, relative to the current stimulus. In the following, we further examine the characteristic features of the joint bias map one by one.

Our first and most crucial observation was that the estimation responses were systematically repelled away from the stimulus direction of the previous trial, while they were strongly attracted toward the response direction of the previous trial. Consider, for example, horizontally neighboring pixels labeled as zero on the *y*-axis. Previous responses relative to the current stimulus (i.e., *y*-axis value) are fixed across these pixels, so they would reveal the sole effect of the previous stimulus. These pixels show that the response error decreases as the relative stimulus direction of the previous trial increases, suggesting that the effect of the previous stimulus is repulsive. This indicates that when the random dots in the previous trial moved in a direction more clockwise than the direction of motion of the dots in the present trial, subjects perceived the present random dots as moving in the direction more counterclockwise than their true direction of motion. The repulsive effect of the previous stimulus becomes even more evident when we plot the response error, split by the relative direction of the previous response, as a function of the relative direction of the previous stimulus (Fig. 2B). This would reveal how the estimation responses changed with changes in the previous stimulus, with the effect of the previous response fixed within each line. The bias plot clearly shows negative slopes regardless of previous response directions, which confirms the robust repulsion away from the previous stimulus direction.

In contrast, vertically neighboring pixels that are labeled zero on the x-axis show that the response error increases as the relative response direction of the previous trial increases, suggesting that the effect of the previous response is attractive. This indicates that when subjects perceived and subsequently reported that the random dots in the previous trial had moved in the direction more clockwise than the direction of motion of dots in the present trial, they perceived the present random dots as moving in the direction more clockwise than their true direction of motion. Again, the attractive effect of the previous response becomes even more pronounced by plotting the response error split by the relative direction of the previous stimulus as a function of the relative direction of the previous response (Fig. 2C). The bias plot exhibited positive slopes, confirming the strong attraction toward the previous response direction.

The joint bias map (Fig. 2A) also provides us with an idea about how these two opposite effects interact. Based on the linearly increasing response errors from the lower right to the upper left, we can observe that these two sequential biases are combined in an additive manner. Consequently, the bias was strongest when the directions of the previous stimulus and response were opposite relative to the current stimulus, as revealed by the response errors on the second and fourth quadrants, with their magnitudes being up to 8° on average. In other words, the bias in estimation responses were strongest when the concurrent attractive and repulsive effects happened to work in the same clockwise or counterclockwise direction.

Lastly, we confirmed that it would be inappropriate to quantify the biases with a linear regression that uses the relative direction of the previous stimulus and response as predictors. Both repulsion away from the previous stimulus and attraction toward the previous response were observed when the previous stimulus and response values were similar to the current stimulus, but these patterns were diminished or even slightly reversed when they were very different, as can be seen at the extremities of the joint bias map (Fig. 2A), suggesting that linear regression would not be an appropriate way to quantify the biases.

Additivity of concurrent attractive and repulsive sequential effects

We used a descriptive model to simultaneously quantify the attractive and repulsive biases. Based on existing literature, we assumed that patterns of the attractive and repulsive effects would each resemble the DoG curve and that the two effects would linearly add up to compose the resulting sequential biases in subjects' estimation behavior. We fitted a sum of two independent DoG curves, each representing the bias to or away from the previous stimulus and bias to or away from the previous response, respectively. In the following, we refer to this model as the Stimulus & Response model, as both the previous stimulus and previous response bias the current perception. We compare the Stimulus & Response model to alternative models in which either previous stimulus or previous response biases the current estimates (designated as the Stimulus

model and the Response model, respectively). We first produced the estimation behavior using the best-fitting parameters of each model and plotted joint bias maps to examine the model performance (Fig. 3A–C). While the behaviors of neither Stimulus model (Fig. 3A) nor Response model (Fig. 3B) properly emulated the human data (Fig. 2A), the Stimulus & Response model successfully predicted the pattern of biases observed in the empirical data (Fig. 3C). Models were also quantitatively compared using the Akaike information criterion (AIC) to account for differences in model complexity, with data reported as the mean AIC difference from the best-fitting model followed by a bootstrapped 95% confidence interval in brackets. The Stimulus & Response model outperformed the Stimulus model (AIC difference = 125.7 [79.9 239.1]) and the Response model (AIC difference = 35.2 [24.1 52.3]; Fig. 3D).

We did not constrain the direction of the biases in these models to be attractive or repulsive. Nevertheless, the fitting results of the Stimulus & Response model revealed that the previous stimulus repelled the subsequent responses, while the previous response attracted them (previous stimulus: amplitude = −4.88° [−5.89 −3.90], $t_{(31)} = 14.991$, $P < 0.001$; previous response: 7.95° [6.81 9.01], $t_{(31)} = 27.090$, $P < 0.001$; Fig. 4A; see also Additional file 1: Fig. S5), consistent with our visual inspection of the joint bias map. Note that the absolute magnitudes of both attractive and repulsive biases are considerably larger than those observed in the marginal bias plots (Fig. 1B, C), as well as the typical sequential biases reported in the literature. The attraction toward the previous response was significantly larger in absolute magnitude than the repulsion away from the previous stimulus (difference: 3.11° [1.64 4.35], $t_{(31)} = 6.241$, $P < 0.001$), accounting for the net attractive bias observed in the marginal bias plots of our data (Fig 1B, C) and in earlier studies. In addition, the repulsion away from the previous stimulus was more broadly tuned than attraction toward the previous response (previous stimulus: peak location = 38.45° [34.61 45.65]; previous response: 28.92° [26.75 32.06]; difference: 9.53° [5.35 16.34], $t_{(31)} = 8.799$, $P < 0.001$; Fig. 4B), which explains the small repulsive bias often observed in earlier studies when successive stimuli were markedly different [6, 35, 38, 46, 47]. Overall, our results show that the observed patterns of sequential effects in the joint bias map can be well described by the linear sum of two curves characterizing the attraction toward the previous response and repulsion away from the previous stimulus.

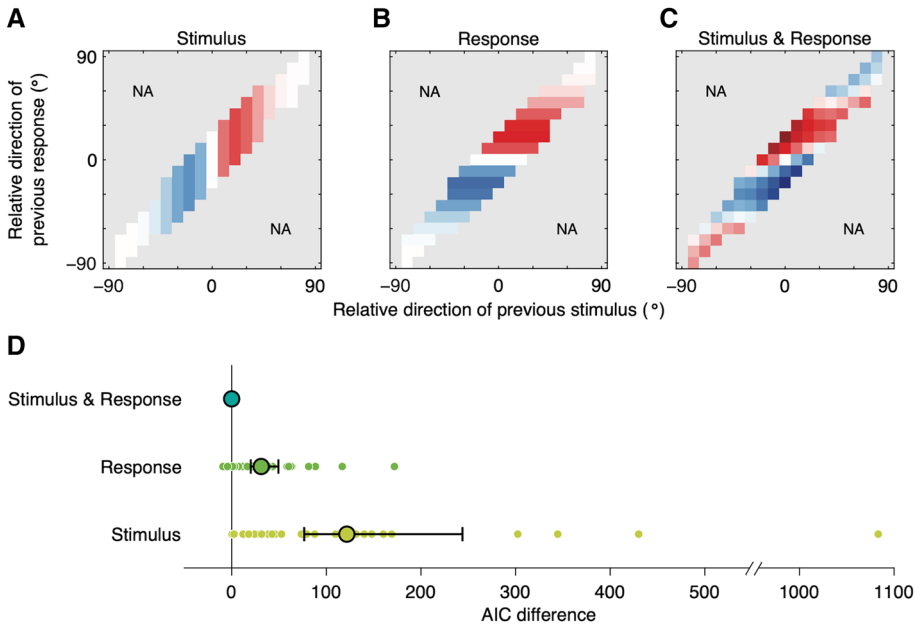

Fig. 3 Model comparison. **A** Prediction from a single DoG function of relative direction of the previous stimulus (Stimulus model). Color conventions are as described in Fig. 2A, and pixels are shown for only those available in Fig. 2A. **B** Prediction from a single DoG function of relative direction of the previous response (Response model). **C** Prediction from the linear sum of two independent DoG functions, each representing bias to (or away from) the previous stimulus and bias to (or away from) the previous response (Stimulus & Response model). Characteristic patterns of the joint bias map are in excellent agreement with empirical data shown in Fig. 2A. **D** Model comparison using Akaike information criterion (AIC). Small dots represent individual subject differences in AIC values between each model and the Stimulus & Response model. Large circles and error bars represent mean and bootstrapped 95% confidence interval. Positive numbers represent a worse fit than the Stimulus & Response model. The Stimulus & Response model fits the best

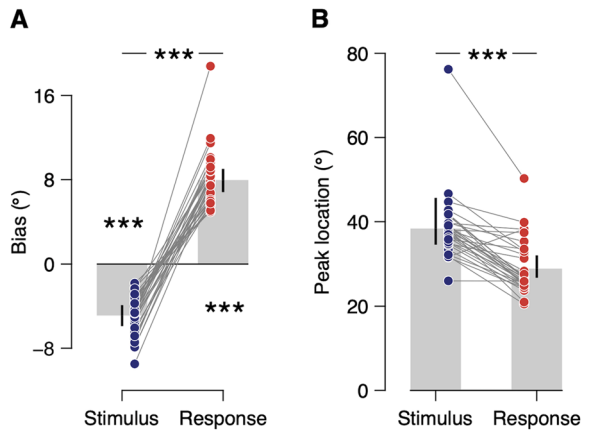

Fig. 4 Quantifying the two opposite biases. **A** Estimated bias magnitudes in the Stimulus & Response model. Bias is negative for the previous stimulus and positive for the previous response, meaning that the estimation responses were repelled away from the previous stimulus and attracted toward the previous response. Both biases are highly significant, with the bias to the previous response being significantly larger in absolute magnitude than the bias away from the previous stimulus. **B** Estimated peak locations of the bias curves in the Stimulus & Response model. Bias away from the previous stimulus is more broadly tuned than the bias to the previous response. Bars represent the posterior estimate of the population mean, error bars represent 95% credible intervals, and dots connected with thin lines represent the parameter estimates for individual subjects. ***$P < 0.001$

Estimating the centers of biases

We constrained the Stimulus & Response model in such a way that the centers of the two bias curves are located at the previous stimulus or response. The rationale was that the previous stimulus is a proxy for the tuning state of the neural population in the encoding stage, which may induce repulsion, and the previous response is a proxy for the prediction from the past estimation, which may induce attraction. However, it is an open question whether the previous stimulus and response are indeed the centers of the biases. More specifically, while we used the previous stimulus as a proxy for the latent sensory measurement in the previous trial, the previous sensory measurement itself could have been repelled away from its own preceding trial, deviating from the previous stimulus. In contrast, the previous response, which we used as a proxy for the final percept in the previous trial, is not likely to systematically deviate from the percept. The percept itself is biased due to the accumulated effects from the past trials; however, the response, we assume, is an unbiased manifestation of the percept with additive noise. Therefore, it is plausible that sensory encoding is affected by not only the previous stimulus but also the stimulus before the previous stimulus, while the prediction from the past estimation is indeed centered at the previous response that reflects the observer's best estimate of the previous stimulus, equivalent to the best predictor for the incoming input in our task.

To estimate the centers of biases, we fitted a new model: a sum of two independent DoG curves, each representing attractive (i.e., $a_{\text{attraction}} > 0$) and repulsive (i.e., $a_{\text{repulsion}} < 0$) biases, respectively, without predetermining the centers of the curves. Specifically, the centers of the attraction and repulsion curves were each parameterized as a combination of the previous stimulus and previous response, $\beta\text{Response}_{t-1} + (1-\beta)\text{Stimulus}_{t-1}$. Note that the centers of the two curves are parametrized identically without any presumption. For each curve, if β is one, the current response is biased by the previous response, and if β is zero, the current response is biased by the previous stimulus. More generally, the beta parameters represent the center of the corresponding bias curves on a normalized scale in which the previous stimulus is zero and the previous response is one. We refer to this model as the Attraction & Repulsion model. We fitted the model to data as for the Stimulus & Response model, but with two βs, each for attraction ($\beta_{\text{attraction}}$) and repulsion ($\beta_{\text{repulsion}}$) curves, as additional free parameters (see Methods). For the attraction curve, we found that $\beta_{\text{attraction}}$ was significantly larger than zero (1.02 [0.89 1.13], $t_{(31)} = 58.149$, $P < 0.001$) and tightly clustered around one (Fig. 5A, orange circle). This provides additional support for the attraction toward the previous response and not toward the previous stimulus (or a combination of the two). We also found that, for the repulsion curve, $\beta_{\text{repulsion}}$ is vastly different from one ($t_{(31)} = 107.887$, $P < 0.001$), suggesting that the current response is not repelled from the previous response. However, the estimated $\beta_{\text{repulsion}}$ was not exactly clustered around zero, either. Instead, $\beta_{\text{repulsion}}$ was significantly negative across subjects (−0.85 [−1.26 −0.53], $t_{(31)} = 51.567$, $P < 0.001$; Fig. 5A, light blue circle), indicating that the center of the repulsion curve was located near the previous stimulus but shifted in the direction opposite to the previous response.

We propose that the systematic deviation of $\beta_{\text{repulsion}}$ from the previous stimulus indicates that the center of repulsion is determined by "sensory measurement" of the previous stimulus rather than the previous stimulus itself. The sensory measurement is defined as the latent value that represents the measured sensory input, which is affected by the repulsive bias in encoding but has not yet been affected by the attractive bias in decoding. The sensory measurement on a given trial can be systematically related to the response error on the trial. Consider a case in which the current response is located clockwise relative to the current stimulus. It is likely that the previous stimulus is also located clockwise relative to the current stimulus because the response is biased, on

average, to the previous stimulus. Consequently, the sensory measurement of the current stimulus would systematically deviate from the current stimulus in a direction away from the current response, as we observed in the analysis (Fig. 5A).

According to our proposal, the magnitude of the attractive bias should have no meaningful relationship with the center of the attractive bias, whereas the magnitude of the repulsive bias should directly affect the previous sensory measurement, and thus the center of the repulsive bias, across individual subjects. To test this prediction, we compared the magnitudes of attractive and repulsive biases for each subject to the shifts in the center of the corresponding biases. As expected, we found no statistically significant relationship between the center and magnitude of the attractive bias ($\rho = 0.149$, $P = 0.702$; Fig. 5B). Crucially, we found that the individual differences in the center of the repulsion curve ($\beta_{\text{repulsion}}$) were correlated with the magnitude of the repulsive bias. Subjects whose repulsion away from the previous stimulus was the strongest tended to have largest shift in the center of the repulsive bias ($\rho = 0.633$, $P = 0.002$; Fig. 5C). These results bear out the qualitative prediction of our hypothesis that the center of the repulsion systematically deviates from the previous stimulus because the previous sensory measurement, which was repelled away from its own preceding trial, is the center of repulsion in the current trial.

Discussion

Our natural environment is abundant with statistical regularities, a typical one being that the environment is stable over time [5, 6]. If it is scorching hot today, no one would expect heavy snow the next day. It is no surprise that an observer utilizes such environmental stability to improve their perceptual performance. However, there is a certain irony in this situation [48], because an ideal observer should not only ensure that sensory neurons adapt to the changes in input stimuli, repelling the current perception, but also perform probabilistic inference based on predictions from past observations, attracting the current perception toward the recent history. These two have been widely suggested as underlying mechanisms of the repulsive aftereffects and attractive serial dependence, but it remains unclear whether and how these two opposite effects interact in the perceptual process [39]. Here, we studied the concurrent attractive and repulsive sequential effects in visual perception to determine how these competing demands are resolved during visual processing. Specifically, we showed that the previous stimulus repels the current response, the previous response attracts the current response, and these two biases add up to shape a unique pattern in estimation behavior. We also directly estimated the centers of the biases and confirmed that whereas the attraction bias is centered precisely at the previous percept, the repulsion bias is centered at the sensory measurement of the previous stimulus. We believe that the concurrent attractive and repulsive sequential effects could be present in many other data in literature (see also ref. [49]), as the overall attraction bias found in our study mirrors several previous reports showing similar attractive biases in numerous perceptual and cognitive domains.

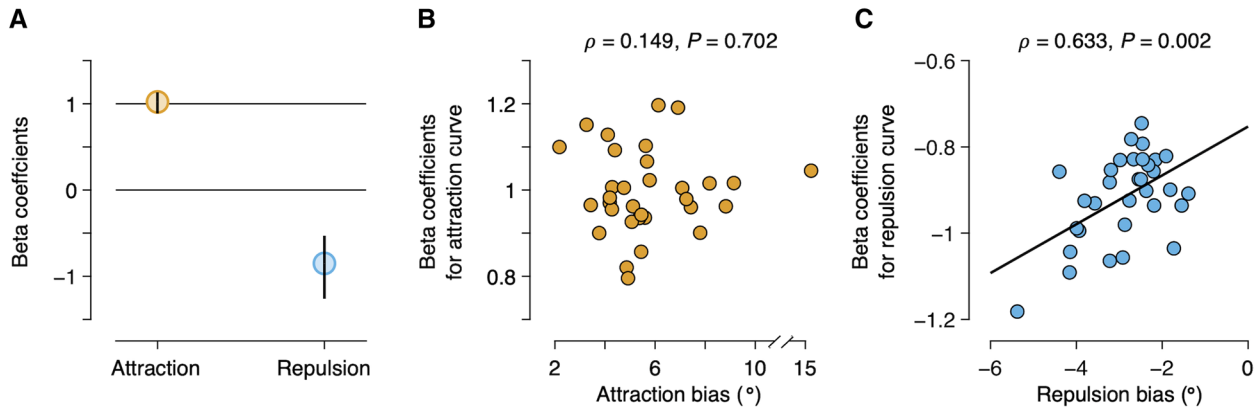

Fig. 5 Estimating the centers of biases. **A** Estimated center of the attraction and repulsion biases in the Attraction & Repulsion model. Circles represent the posterior estimate of the population mean, and error bars represent 95% credible intervals. Beta coefficient zero indicates that the bias is toward (or away from) the previous stimulus, and one indicates that the bias is toward (or away from) the previous response. Center of attraction is located precisely on the previous response, while center of repulsion is located near the previous stimulus but shifted in a direction away from the previous response. **B**, **C** Relationship between the center and magnitude of the bias. Points represent parameter estimates for individual subjects, and line is linear regression

Concurrency and additivity of attractive and repulsive biases

The current results further develop our understanding of sequential effects in three main ways. First, this study dissociated the attractive and repulsive effects with the same task throughout the experiment and without further manipulation in experimental procedures and provided a clear picture of the concurrency of the two opposite biases. Although suggestions have been made on the possibility that the two opposite biases coexist [30–32], their nature and origin remained topics of debate, partly due to a lack of direct empirical support. Notably, if repulsion away from the previous stimulus and attraction toward the previous response occur at the same time, then the bias would be exceedingly strong when the directions of the previous stimulus and response are opposite relative to the current stimulus. This is exactly what we found (Fig. 2A). Second, we empirically demonstrated that the effects of stimulus and response of the preceding trial are largely additive in the perception of the current trial. We fitted the linear sum of two curves representing biases to (or away from) the previous stimulus and previous response, respectively, and found that it could capture the unique pattern in the subjects' sequential estimation (Fig. 3C). Third, we revealed that the repulsive bias, which has been largely considered as determined by the previous stimulus, is actually centered on the sensory measurement of the previous stimulus (Fig. 5), confirming that the repulsive bias occurs during stimulus encoding.

Rapid adaptation to visual stimulus

Our results support the idea that the visual system adapts to changes in briefly presented stimuli, resulting in a repulsive bias away from the previous stimulus. Sensory adaptation and repulsive aftereffects have historically been studied in the context of prolonged stimulus exposure, but increasing evidence also suggests a more rapid form of adaptation. Such adaptation is known to occur rapidly following a stimulus exposure as brief as a few hundred, or even tens, of milliseconds [10, 12, 13, 50, 51] and persist over dozens of seconds despite the presentation of several intervening stimuli [52], inducing robust repulsive biases in subsequent behavioral reports [31, 33–38, 40, 51, 53, 54] over a prolonged timescale [30, 41, 42]. A recent neuroimaging study also found that sensory representations in the visual cortex were significantly and substantially repelled from the previous stimulus even when stimuli were presented only for a second, and successive stimuli were separated by at least ten seconds [32]. Our findings are consistent with all these documentations, which together confirm that the visual system translates input stimuli into sensory representations in an efficient way, inducing a repulsive bias in subsequent perception, even when the stimuli are only briefly presented.

Characteristics of attractive bias

The attractive bias toward the previous response does not necessarily imply that subjects merely reproduced their previous motor commands. Several studies have experimentally precluded the effect of motor replication and observed an attractive bias even when reproducing the previous motor command would not result in such attraction [7, 28, 55–57]. For example, when subjects were asked to make a flipped version of orientation response that is vertically symmetric to the stimulus orientation on every second trial, they still systematically biased their responses toward the previous stimulus, and not toward the previous motor response, suggesting that the bias was introduced before the motor process [55]. Recent neurophysiological studies also denied the effect of motor replication by showing that the biases are already present in neural representations in the visual cortex [58] even without an explicit task [59]. Therefore, care must be taken when interpreting the attractive biases to the previous responses reported in this study. Our proposal is that the previous response, to which the observer's current estimation responses are attracted, is the observer's best estimate of the stimulus in the previous trial. Our sensors can only provide noisy measurements of a given stimulus, so the visual system does not have direct access to the true stimulus value. Instead, the system infers the stimulus value from the sensory measurement and creates a perceptual estimate that sometimes deviates on average from the true stimulus value, as in our experiment. Because it is the final estimate from which the visual system predicts the subsequent stimulus, it is natural for the system to bias its current estimate toward the previous estimate, rather than toward the previous stimulus [6, 26].

Recent studies suggested that not perceptual but later decisional stages bias behavior toward recent perceptual history [31, 38]. This perspective remains a subject of debate, as other studies have shown that attractive serial dependence acts directly on sensory levels [57, 59, 60]. In our study, the final outcome of the perceptual processing (i.e., the percept) is not conceptually equivalent to a perceptual decision. Indeed, the visual system is capable of reconstructing the surrounding visual environment without making implicit or explicit decisions. We rather consider the final percept as what is decoded from the sensory signals encoded by the system. That being said, our study does not exclude the possibility that attractive serial dependence operates on multiple levels including perception [28, 57, 59, 60], memory [35, 61–63], and

decision [31, 38], with additional influence from factors such as attention [28, 46] and confidence [42, 47].

Computational basis of concurrent repulsive and attractive biases

We now consider a generative computational model to confirm that the proposed conceptual framework can be instantiated in a plausible encoding-decoding system. We assume that direction estimates are the outcome of a perceptual encoding-decoding process, beginning with sensory input to a population of direction-selective neurons with independent Poisson spike count variability. The sensory input and uncertainty associated with it are naturally encoded in the population activity of these neurons in the form of the likelihood function [64–67]. The encoding stage of the model was further characterized by the adaptation of these neurons to the direction of the preceding motion. Motivated by a vast amount of neurophysiological evidence [16–21], we assume that the primary effect of sensory adaptation is a reduction in the response gain of neurons selective to the direction of preceding motion (Fig. 6A). The simulation of this encoder model confirms that such changes in the neural coding result in repulsive shifts of the average likelihood function away from the preceding motion direction (Fig. 6B). Due to sensory noise and adaptation, the sensory measurement of the preceding motion direction may differ from the true direction of the preceding motion (Fig. 5). To this end, we also considered an alternative scenario in which the neurons adapt to the sensory measurement (i.e., gain reduction of a neuron according to its spike count in the previous trial), rather than the true stimulus value, and confirmed that the likelihood function is still, on average, repelled away from the preceding motion direction. It is worth noting that adaptation schemes need not be restricted to gain reduction but can be easily generalized to other changes in the tuning mechanism which are known to generate similar repulsive effects [68–70].

Next, we assumed that the model observer uses knowledge about the dynamics of the stimulus to convert its belief about the stimulus value in the previous trial into a prediction about the stimulus value in the subsequent trial [25, 71]. Motivated by previous work [6], we modeled knowledge about the changes in stimuli within a given time interval (i.e., propagation noise) as a mixture of uniform and von Mises distributions with zero mean. The model observer convolved this distribution with posterior belief on the previous stimulus to make predictions about the current stimulus, which is then combined with the new sensory information to produce a posterior belief on the current stimulus based on the past and current sensory information (Fig. 6C). This computation follows a long history of research that models perceptual and sensorimotor behavior as recursive Bayesian inference, such as Kalman filters [7, 72–77]. We simulated the model observer's estimation responses to the sequence of stimuli that our subjects encountered, assuming a squared-error loss function and no motor noise. The results showed that our model observer produced a joint bias map (Fig. 6D) that was close to that of the human subjects (Fig. 2A). The estimation responses of the model observer were repelled away from the previous stimulus because adaptation shifted the likelihood function and attracted toward the previous response because the prediction was made from the observer's belief in the previous trial. The fact that this observer model can approximate the bias behavior of human observers, in line with existing models that share common features [31, 32], implies that human estimation behavior can be understood by adaptive encoding and decoding processes.

Neural substrate of adaptive coding and dynamic Bayesian inference

In this work, perceptual processing is characterized by a mapping from input stimulus to sensory measurement and from sensory measurement to percept. The sensory measurement can be readily identified with activity of direction selective neurons in the primary visual cortex and the middle temporal (MT) area [78]. The stochastic nature of spiking activity in these neurons could provide the means to implicitly encode sensory likelihoods [64, 66, 67], and sensory adaptation has long been observed in the activity of these neurons [17, 68]. For the percept, it would seem reasonable to appeal to activity at, or beyond, MT, as its activity has been strongly linked to the conscious perception of global motion [78–80]. Beyond MT, the lateral intraparietal (LIP) area receives direct inputs from MT, and the participation of its neurons in integration of visual motion signals and decisions about motion direction has been studied quite extensively [81, 82].

How exactly these cortical circuits may realize Bayesian inference are not yet fully identified [83], but mounting evidence points to the probabilistic population codes [66, 84]. It has been shown that task-related firing rates of single neurons before the presentation of sensory information are modulated by prior expectations [85–88], and firing rates after the presentation of sensory information reflect posterior estimate of behaviorally relevant variables [87, 89, 90]. Although these observations have been mostly focused on Bayesian computations with non-dynamic generative models, other studies have also proposed how cortical circuits would support dynamic Bayesian inference with internal propagation noise models, such as Kalman filters

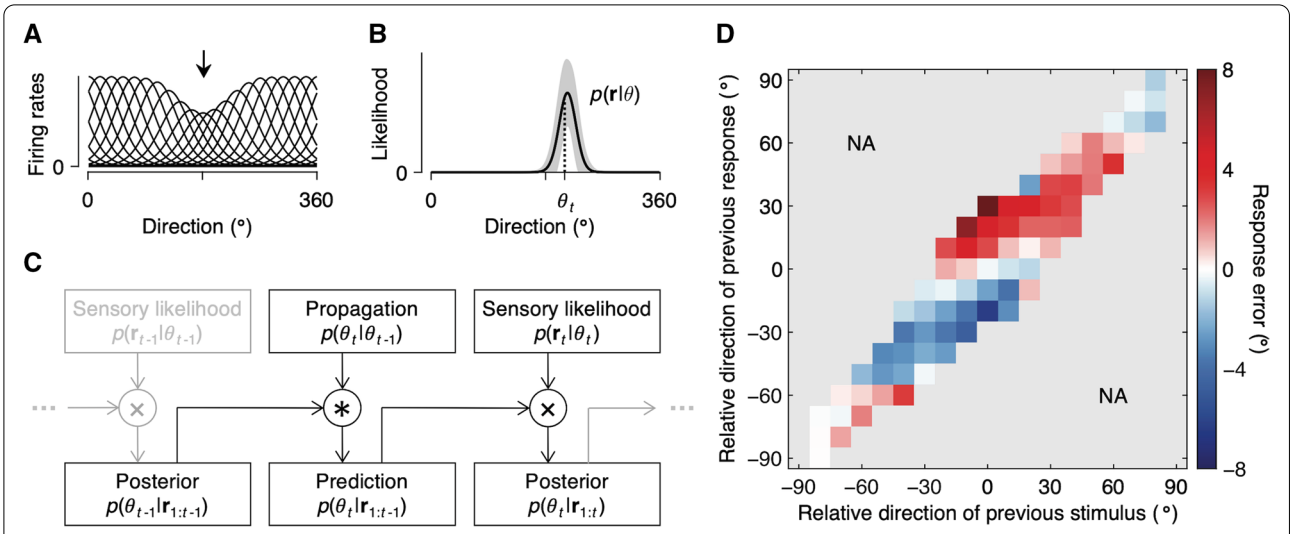

Fig. 6 Generative model. **A** Tuning curves of direction selective neurons after adaptation at 180° (black arrow). Adaptation was modeled as a reduction of the response gain in neurons selective to the adapter direction. **B** Population likelihood. Likelihoods were computed assuming independent Poisson spike count variability and averaged over 1,000 presentations of the stimulus direction θ_t, each following a presentation of the adapter direction as in **A**. Due to adaptation, the likelihoods are on average shifted away from the adapter direction. Shaded regions represent ±1 SD. **C** Recurrent Bayesian inference. The model first convolves posterior belief on the previous stimulus $p(\theta_{t-1}|\mathbf{r}_{1:t-1})$ with the propagation noise distribution $p(\theta_t|\theta_{t-1})$, to predict the current stimulus θ_t based on the past sensory responses $\mathbf{r}_{1:t-1}$. The model then combines this prediction $p(\theta_t|\mathbf{r}_{1:t-1})$ with the new sensory information $p(\mathbf{r}_t|\theta_t)$, to produce posterior belief on the current stimulus $p(\theta_t|\mathbf{r}_{1:t})$. **D** Model simulation. Simulation data were analyzed and plotted as in Fig. 2A. The model successfully generated the characteristic features of the human estimation behavior

[91, 92], and empirically observed such neural implementations in the posterior parietal cortex [93].

An alternative view suggests that a heterogeneous neural population embedding representations of prior knowledge within them facilitates a direct transformation of sensory input into posterior estimates [94–97]. This view is compatible with recent behavioral studies that showed that recent perceptual history constructs priors at a later stage in perceptual processing [60, 98] but interacts directly with early sensory signals, preceding illusory effects driven by spatial context [60]. In line with these observations, while priors generated from recent history are represented at higher-level areas such as the posterior parietal cortex [85], they seem to be fed back to early sensory areas to directly influence cortical representations of incoming sensory information [58].

Timescales of the attractive and repulsive effects

We believe our results can provide a unified framework for interpreting the existing literature. For example, several studies have reported that the magnitude of attractive bias increases with increasing time delay between stimulus presentation and response [35, 38]. These stronger attractive biases may be attributed to more uncertain sensory information corrupted by higher memory noise, while the magnitude of repulsion remains relatively the same, since the time delay took place after the sensory encoding and thus would have not affected the repulsive bias. In the extreme case where there is no time delay between the stimulus and response (and thus minimal memory noise), the net bias becomes repulsive [35] as the magnitude of attraction becomes smaller than that of repulsion. Several studies have shown that estimation responses are attracted to immediately preceding stimuli and are repelled away from stimuli from many trials back [30, 41, 42]. This time, both attraction and repulsion are likely to be affected by the timescale of the trial history. The relative speeds of decay may be each measured by increasing propagation noise with longer timescales [6] and smaller gain reduction of sensory neurons by stimuli from many trials back compared to the immediately preceding one [52], respectively.

Comparison to attraction and repulsion in non-dynamic stimuli

Our proposed model is comparable to the efficient encoding and optimal decoding theory for feature estimation under a static environmental distribution [4]. This theory proposes that the visual system efficiently encodes incoming information, allocating more neural resources to the representation of more probable stimulus values, which typically causes a bias away from the mean of the distribution. The system then decodes the sensory

representations in an optimal manner by combining the representation with prior knowledge about the stimulus distribution, which typically results in a bias toward the mean of the distribution. The relative magnitudes of these opposite effects determine the net bias in perceptual estimates. However, even in that work, there is no behavioral data that show both attractive and repulsive effects at the same time. A notable difference between the model for stationary environments and our proposed model for changing environments is that sensory adaptation for efficient coding, which is supposed to cause repulsive bias, is described as adapting to the preceding stimulus, as evidenced by the joint bias map and conditional bias plot (Fig. 2A, C). Consequently, it becomes possible to separately estimate the attractive bias, which is centered around the previous response, and repulsive bias, which is centered near the previous stimulus (Fig. 4).

Sensitivity to temporal statistics

For the Bayesian observer to be truly optimal, it is necessary for the observer to maintain priors that approximate the true stimulus statistics of the external world [99]. Can we learn novel temporal statistics that might differ from natural statistics and update priors accordingly? In this study, we divided subjects into groups with different propagation noise distributions (range: ±20°, ±40°, ±80°, or ±180°), which allowed us to test the effect of experimentally imposed temporal statistics on estimation behavior. If subjects learn new temporal statistics, then the magnitude of the biases would depend on those temporal statistics, with smaller propagation noise leading to stronger sequential effects. The best-fitting parameters of the Stimulus & Response model showed that there was a significant effect of propagation noise level on the strength of repulsive bias away from the previous stimulus ($P < 0.001$; Kruskal–Wallis test), but no statistical significance was found in the attractive bias to the previous response ($P = 0.125$; Additional file 1: Fig. S6). The fitting results of the Attraction & Repulsion model showed similar patterns ($P < 0.001$ and $P = 0.816$, respectively). The observed difference in repulsive biases could be because the accumulated effect from past trials is stronger when similar stimuli are presented repeatedly. For the non-significant difference in attractive biases, priors specifying the statistical characteristics of the natural environment are often thought to be hard-wired through processes of evolution and development and are difficult to update [100]. In the case of temporal statistics, subjects can adapt their internal models of the temporal statistics to the stimulus set they have encountered but still retain priors on positive temporal correlation within the stimulus set [7]. In addition, it was recently suggested that the visual system might even sidestep the need to learn the temporal statistics by using a mixture of recent stimuli [101]. Future work is required to fully characterize whether and how priors on the temporal statistics of the natural environment can be overridden by experimentally imposed statistics [102].

Conclusions

In summary, we demonstrated that the perceived direction of motion in the current trial depends on both the perceived and presented direction of motion in the previous trial, but in opposite directions. Subjects' estimates of the direction of motion were repelled away from stimuli that had recently been seen, while they were attracted toward responses that had recently been made. We found that the rich pattern of the sequential effects can be well described by the linear sum of two curves, each representing bias away from the sensory measurement of the previous stimulus and bias toward the previous percept. This suggests that the mechanisms underlying the generation of attractive and repulsive sequential biases are largely independent of each other. Our findings suggest that the visual system adaptively encodes and decodes incoming sensory information by referring to the recent history of changing environments to optimize visual processing.

Methods

Subjects

Thirty-two subjects (16 females, aged 18–29 years) participated in the experiment. They were naïve to the purpose of the experiment and had not participated in similar experiments. We required subjects to have normal or corrected-to-normal vision and obtained written informed consent from all subjects prior to their participation. All procedures were approved by the Ulsan National Institute of Science and Technology Institutional Review Board.

Procedure

Figure 1A illustrates the sequence of events in each trial. The subjects viewed the stimuli binocularly from a distance of 137 cm in a dark room, resting their head on a chinrest. Each trial began with the presentation of a fixation point. The subjects were instructed to fixate on the fixation point during the presentation of the motion stimulus. After 0.5 s, the motion stimulus was presented for another 0.5 s, followed by a 1.5-s delay during which only the fixation point was on the screen. After the delay, a circular ring appeared, and the subjects reported the perceived direction of motion by swiping a finger on a touchpad to extend a dark bar from the fixation point in the direction of motion that they had perceived and terminated the trial by clicking on the touchpad. Response

time was not limited, and subjects made a response within 1.10 ± 0.05 s (mean ± SEM across subjects). Trials were separated with a 1.5-s inter-trial interval during which the screen was blank. After each block, the subjects received numerical feedback regarding their mean absolute response error.

In the first trial of each bock, the direction of stimulus motion was randomly chosen from 0 to 360°. After that, the direction of motion in the current trial was solely determined with respect to the direction of motion in the immediately preceding trial. Specifically, subjects were randomly divided into four groups of equal size in which the direction of stimulus motion was randomly varied from the direction of motion in the previous trial following discrete uniform distribution with different ranges (±20°, ±40°, ±80°, or ±180°, in steps of 10°; Additional file 1: Fig. S4D). The number of trials subjects completed in each block was set to be similar across groups (the ±20° group: 5 relative directions × 20 repetitions + 1 first trial = 101 trials; the ±40° group: 9 × 11 + 1 = 100 trials; the ±80° group: 17 × 6 + 1 = 103 trials; the ±180° group: 36 × 3 + 1 = 109 trials). Each session consisted of five blocks, lasting up to 60 min. After one practice session, all subjects went through fifteen blocks over three sessions on consecutive days. As a result, all subjects completed at least 1500 trials in total.

Stimuli

Stimuli were generated using MATLAB and the Psychophysics Toolbox [103] and were displayed by a DLP projector (1920×1080; 120 Hz). All stimuli were presented at the center of a dark gray background of 20 cd/m^2. The fixation point was a white circular point (diameter: 0.4°; luminance: 80 cd/m^2). A gap of 1° between the fixation point and motion stimulus helped the subjects maintain fixation. The motion stimulus was a field of moving dots (diameter: 0.1°; luminance: 80 cd/m^2) contained within a 5-degree circular aperture centered on the fixation point. The dots were plotted in three interleaved sets of equal size. Each set was plotted in one of three successive video frames and was shown for a single frame. Three frames later, randomly chosen 40% of dots from that set moved coherently in a designated direction at a speed of 4°/s; the remainder of the dots were replotted at random locations within the aperture. Dots that moved outside the aperture were placed on the opposite side of the aperture. Together, the three sets produced an average dot density of 48 dots/(deg^{2}s). The presentation of a black circular ring (diameter: 6.6°; width: 0.15°; luminance: 15 cd/m^2) centered on the fixation cued subjects to report their estimate, which they did by swiping their finger on a touchpad to extend and align a black bar (width: 0.15°) to the direction of their estimate and then clicking on the touchpad to confirm.

Data preprocessing

All analyses were performed using MATLAB and the CircStat Toolbox [104]. In the first step of data analysis, we corrected for an individual subject's idiosyncratic estimation bias to or away from the cardinal axes [4, 45], which is unrelated to sequential effects. The bias correction involved inference of a best-fitting function capturing the systematic shape of the estimation biases. Since this bias function is known to take a sinusoidal form [4], earlier studies corrected for the bias by fitting each subject's response errors with a sinusoidal or a polynomial function of stimulus orientation/direction [6, 31, 105]. However, as has been recently shown [45], the patterns of the direction estimation bias in our data varied considerably across subjects: some of our subjects' estimation responses were strongly biased to or away from the cardinal axes, while the others' bias function could not be easily captured by a standard sinusoidal model (Additional file 1: Fig. S1A–C). Capturing these bias functions was necessary because the estimation bias that depends exclusively on the current stimulus direction could confound the observations of bias toward the previous response, no matter how much idiosyncratic the bias functions are.

We compare the performance of three regression models in capturing the systematic shape of the estimation bias: a tenth-degree polynomial regression model, a variant of a sinusoidal model, and a Gaussian process (GP) regression model. The sinusoidal variant that takes into account the asymmetry of biases near the horizontal and vertical axes [106] is given by

$$y = \beta_0 + a\,\mathrm{sgn}(x)\left(\sin\left(4|x| - \sin^{-1}(v)\right) + v\right) \quad (1)$$

where y is the response error, x is the current stimulus direction relative to the vertical axis, β_0 is the general clockwise or counterclockwise bias, a determines the maximum bias, and v determines how much stronger biases to or away from the vertical axis are than biases to or away from the horizontal axis ($-1 \le v \le 1$). To effectively capture idiosyncratic bias functions that do not exhibit strong periodicities (Additional file 1: Fig. S1B, C), we also considered a GP regression [107, 108]. A GP regression model is a nonparametric probabilistic model often used to infer the underlying function from noisy observations without making assumptions about the specific shape of the function (e.g., refs. [109, 110]). Here, the bias function of the current stimulus direction was not restricted to a particular parameterized family. Instead, we searched the function space by assuming that the bias function follows the Gaussian process. Specifically, the response error was modeled as:

$$y = \beta_0 + f(x) \tag{2}$$

where y is the response error, β_0 is the general clockwise or counterclockwise bias, and x is the current stimulus direction. The joint distribution of the latent variables $\mathbf{f} = (f(x_1), \ldots, f(x_N))^{\mathrm{T}}$ across N trials was modeled as an N-dimensional Gaussian distribution:

$$\mathbf{f} \sim \mathrm{Normal}(0, \mathbf{K}) \tag{3}$$

The entries of the covariance matrix $\mathbf{K}$ were determined by the squared exponential kernel function $k(x, x')$, which specifies the covariance of the two latent variables $f(x_i)$ and $f(x_j)$ as:

$$K_{ij} = k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2\sigma_l^2}\right) \tag{4}$$

where σ_f is the signal standard deviation, and σ_l is the characteristic length scale that determines the smoothness of the function. For each subject, these three regression models were each fit to raw data excluding outliers of 2.5 SD from the mean response error. We found that the GP regression performed best, as indicated by Friedman test ($P < 0.001$) and post hoc Tukey test (GP versus sinusoid: $P < 0.001$; GP versus polynomial: $P = 0.001$) on the root mean square of residuals. Based on these results, we used the GP regression to correct for the systematic estimation bias that depends exclusively on the current stimulus direction. We subtracted the bias estimated by the regression model from the response errors (along with the associated responses) and took the residual response errors (and residual responses; Additional file 1: Fig. S1D–F).

To confirm that our correction method successfully removed artifactual sequential effects, we shuffled the order of the trials within each block and then fit the Response model (see below) to each subject's response errors with or without applying the bias correction. This way, any sequential effect of the previous responses was removed, and only artifacts caused by the direction estimation bias alone remained [31]. As expected, we found a small but significant artifactual sequential effect after shuffling and not applying the bias correction (0.67° [0.08 1.28], $t_{(31)} = 3.093$, $P = 0.004$; Additional file 1: Fig. S3A). By contrast, when we applied the bias correction, no significant sequential effect emerged from the shuffling procedure (−0.11° [−0.38 0.15], $t_{(31)} = 1.098$, $P = 0.281$; Additional file 1: Fig. S3B), with significant difference between the artifactual sequential effects in the uncorrected and corrected responses ($t_{(31)} = 4.592$, $P < 0.001$; Additional file 1: Fig. S3C). Having established that our GP regression successfully remove the confounding bias, we used the residual response errors from the GP regression in the remaining analyses. For comparison, we also present results without applying the bias correction in Additional file 1: Fig. S7.

After removing the confounding bias in the raw data, we excluded the first trial of every block, since it does not have any preceding trial to induce sequential effects. We further excluded trials with response error more than 2.5 SD away from the subject's mean response error and the subsequent trials on which subjects could have been affected by those preceding trials with outlying responses. In total, 4.58% of the trials were excluded due to this outlier correction, and the resulting root mean squared response error was 8.15 ± 0.31° (mean ± SEM across subjects).

Descriptive models

The subjects' dependencies on the direction of the previous stimulus and/or previous response in estimating the direction of the current stimulus were quantified by fitting the first derivative of a Gaussian (DoG) curve(s) to their response errors [28]. The DoG curve is given by

$$y = xawc \exp\left(-(wx)^2\right), \tag{5}$$

where y is the response error, a is the amplitude of the curve peaks, w determines the width of the curve, and c is a constant, $\sqrt{2}/e^{-0.5}$. The input to the function, x, was either the relative direction of the previous stimulus (i.e., $\mathrm{Stimulus}_{t-1} - \mathrm{Stimulus}_t$, with index t for trial; Stimulus model) or relative direction of the previous response (i.e., $\mathrm{Response}_{t-1} - \mathrm{Stimulus}_t$; Response model). To effectively capture the characteristic pattern in the estimation data (Fig. 2), we fitted a linear sum of two independent DoG curves, one receiving as input the relative direction of the previous stimulus, and the other receiving the relative direction of the previous response (Stimulus & Response model). To explore the possibility that the centers of biases are not exactly located on the previous stimulus or previous response, we fitted another linear sum of two independent DoG curves, one with a positive amplitude (i.e., $a_{\mathrm{attraction}} > 0$), and the other with a negative amplitude (i.e., $a_{\mathrm{repulsion}} < 0$; Attraction and Repulsion model). Crucially, the inputs to the curves, $x_{\mathrm{attraction}}$ and $x_{\mathrm{repulsion}}$, were each set to be a combination of the previous stimulus and previous response, relative to the current stimulus: i.e., $\beta_{\mathrm{attraction}}\mathrm{Response}_{t-1} + (1 - \beta_{\mathrm{attraction}})\mathrm{Stimulus}_{t-1} - \mathrm{Stimulus}_t$ and $\beta_{\mathrm{repulsion}}\mathrm{Response}_{t-1} + (1 - \beta_{\mathrm{repulsion}})\mathrm{Stimulus}_{t-1} - \mathrm{Stimulus}_t$, respectively. Here, $\beta_{\mathrm{attraction}}$ and $\beta_{\mathrm{repulsion}}$ are free parameters that determine the center of the corresponding bias curves. For example, for the attraction curve, if $\beta_{\mathrm{attraction}}$ is one, $x_{\mathrm{attraction}}$ is reduced to $\mathrm{Response}_{t-1} - \mathrm{Stimulus}_t$, and the current response is attracted toward the

previous response. If $\beta_{\text{attraction}}$ is zero, $x_{\text{attraction}}$ is reduced to $\text{Stimulus}_{t-1} - \text{Stimulus}_t$, and the current response is attracted toward the previous stimulus. Similarly, for the repulsion curve, if $\beta_{\text{repulsion}}$ is one, the current response is repelled away from the previous response, and if $\beta_{\text{repulsion}}$ is zero, the current response is repelled away from the previous stimulus. More generally, the beta parameters can be interpreted as the center of the corresponding bias curves in the stimulus space normalized such that the previous stimulus is zero and the previous response is one.

All parameters were estimated using a hierarchical Bayesian approach that uses the aggregated information from the entire population sample to inform and constrain the parameter estimates for each individual [111]. Specifically, we assumed a hierarchical prior on parameters, in which the parameters for each individual were drawn from independent Gaussian or von Mises distributions characterizing the population distributions of the model parameters. At the trial level, the response error was modeled following a von Mises distribution

$$y_{(i,t)} \sim \text{von Mises}\left(\mu_{(i,t)}, 1/\sigma_{(i)}^2\right), \tag{6}$$

with indices i for individual subject and t for trial (specifically for the description of the hierarchical Bayesian model, we use parenthesis for indices corresponding to specifications of the hierarchical level). The mean of the von Mises distribution, $\mu_{(i,t)}$, was determined by a DoG curve or a linear sum of two DoG curves depending on the model (see above). At the individual level, the model parameters were constrained by population-level parameters. Specifically, all parameters at the individual level were parameterized using Gaussian or von Mises distributions.

$$\begin{aligned} \beta_{(i)} &\sim \text{Normal}\left(\mu_\beta, \sigma_\beta^2\right) \\ a_{(i)} &\sim \text{von Mises}\left(\mu_a, 1/\sigma_a^2\right) \\ l_{(i)} &\sim \text{von Mises}\left(\mu_l, 1/\sigma_l^2\right) \\ \sigma_{(i)} &\sim \text{von Mises}\left(\mu_{l\sigma}, 1/\sigma_\sigma^2\right) \end{aligned} \tag{7}$$

where $l = 1/\sqrt{2}w$ directly represents the peak location of the DoG curve. The noise standard deviation $\sigma_{(i)}$ was further restricted to positive values. Specifically for the Attraction & Repulsion model, the amplitude parameters $a_{\text{attraction}}$ and $a_{\text{repulsion}}$ were further restricted to positive and negative values. At the population level, priors on the mean of the population distributions were set to broad distributions with ranges large enough to cover all practically plausible values.

$$\begin{aligned} \mu_\beta &\sim \text{Normal}\ \left(0.5, 0.5^2\right) \\ \mu_a &\sim \text{Uniform}\ (-\infty, \infty) \\ \mu_l &\sim \text{Normal}\ (5, 90) \\ \mu_\sigma &\sim \text{Normal}\ (0, \infty) \end{aligned} \tag{8}$$

Note that the priors on μ_β favored neither previous stimulus (i.e., $\beta = 0$) nor previous response (i.e., $\beta = 1$) as the center of the bias (i.e., neutral priors). Again, specifically for the Attraction & Repulsion model, the population mean of amplitude parameters, $\mu_{a_{\text{attraction}}}$ and $\mu_{a_{\text{repulsion}}}$, were restricted to positive and negative values, respectively.

$$\begin{aligned} \mu_{a\text{attraction}} &\sim \text{Uniform}(0, \infty) \\ \mu_{a\text{repulsion}} &\sim \text{Uniform}(-\infty, 0) \end{aligned} \tag{9}$$

Priors on the standard deviation of the population distributions were set to gamma distributions

$$\begin{aligned} \sigma_\beta &\sim \text{Gamma}\ \left(s_\beta, r_\beta\right) \\ \sigma_a &\sim \text{Gamma}\ (s_a, r_a) \\ \sigma_l &\sim \text{Gamma}\ (s_l, r_l) \\ \sigma_\sigma &\sim \text{Gamma}\ (s_\sigma, r_\sigma) \end{aligned} \tag{10}$$

with their shape parameter s and rate parameter r are set so that their mode and standard deviation would be approximately a half and twice the standard deviation of the individual-level parameters, respectively, making the hyper-priors vague on the scale of the data [111]. Based on earlier reports, we assumed that the standard deviation of the amplitude, peak location, and noise standard deviation across individual subjects would be 1°, 5°, and 2°, respectively, and parametrized the hyper-priors accordingly. For the beta parameters in the Attraction & Repulsion model, we naively assumed that the standard deviation across individual subjects would be 0.15.

We used a Markov chain Monte Carlo (MCMC) technique, specifically a Metropolis-Hastings algorithm, to compute the posterior probability density of the parameters. Initial values of the individual-level parameters were set to amplitude = 0°, peak location = 30°, and noise standard deviation = 10°. For the Attraction & Repulsion model, in which amplitude parameters $a_{\text{attraction}}$ and $a_{\text{repulsion}}$ were constrained to be positive and negative, respectively, we set them to be ±5°, considering the fitting results of the Stimulus & Response model. Initial values of the beta parameters $\beta_{\text{attraction}}$ and $\beta_{\text{repulsion}}$ were set to 0.5, favoring neither previous stimulus (i.e., $\beta = 0$) nor previous response (i.e., $\beta = 1$) as the center of the bias. Initial values of the population mean were set to be the same with the individual-level parameters, and initial values of the population standard deviation of amplitude, peak location, noise standard deviation, and beta parameters were set to 1°, 5°, 2°, and 0.15, respectively. We used four independent chains in parallel, each with an independent random number generator, and used the first one million iterations for each chain as a burn-in period, thereby minimizing the influence of the initial values of the model parameters. After the burn-in period, the subsequent one million new samples from each chain were used to estimate the posterior probability density

function. We further thinned the samples by selecting every 1000 samples in the chain, resulting in a final set of 4000 samples for each parameter and reducing autocorrelations in the samples to near zero. Convergence of the chains was confirmed by visual inspection of trace plots and Gelman–Rubin tests [112]. All parameters in the models had $\hat{R} < 1.1$, suggesting that all chains successfully converged to the target posterior distribution.

For the statistical significance of the model parameters, we report the mode and 95% credible interval of the posterior distribution of the population-level mean parameters, along with the results of a classical one-sample *t* test on the individual-level parameter estimates after testing for normality using a Kolmogorov–Smirnov test. Similarly, to statistically compare the model parameters, we report the mode and 95% credible interval of the posterior distribution of the difference between population-level mean parameters, along with the results of a paired-sample *t* test on the individual-level parameter estimates.

Model predictions shown in Fig. 3A–C are the point estimate of the response error made with the population-level parameter estimates for each model. We used AIC for model comparison. To report the AIC, we computed the AIC for each model and each individual using the individual-level parameter estimates and then averaged the AIC difference from the best-fitting model (i.e., the one with the highest mean AIC values) across subjects. The confidence interval for the mean AIC difference was computed by bootstrapping using the bias-corrected and accelerated percentile method (10000 samples).

Observer model

We begin with a conventional encoding model for a population of $N = 20$ sensory neurons responding to a stimulus direction θ. We assumed that the number of spikes emitted in a given time interval by the ith neuron is a sample from an independent Poisson process, with the mean rate determined by its tuning curve (von Mises distribution). Given these assumptions, the encoding model is specified as the probability of observing a particular population response $\mathbf{r} = (r_1, \ldots, r_N)^{\mathrm{T}}$ for a given stimulus direction θ:

$$p(\mathbf{r}|\theta) = \prod_{i=1}^{N} \frac{f_i(\theta)^{r_i} e^{-f_i(\theta)}}{r_i!} \tag{11}$$

where $f_i(\theta) = g_0 \exp(\kappa \cos(\theta - \theta_i))$ is the tuning curve of the ith neuron with a response gain $g_0 = 4$, a concentration parameter $\kappa \approx 3.65$ (analogous to a Gaussian function with a 30° standard deviation), and the preferred direction θ_i. Notably, from the perspective of the encoder that generates a noisy sensory response $\mathbf{r}$ in response to an unknown θ, this relationship becomes a function of θ with a fixed $\mathbf{r}$, which is known as the likelihood function.

Sensory coding is not invariant to the temporal context, but is adaptive to the recently encountered stimulus. Based on neurophysiological studies, we assume that the primary effect of adaptation is a change in the response gain g_i of neuron i, such that those neurons most responsive to the adapter reduce their gain the most [16–21]. Specifically, we assume that the amount of gain reduction in the ith neuron is a von Mises function of the difference between the adapter direction and the preferred direction of that neuron [69]

$$g_i = g_0\left(1 - \alpha_{adapt} \exp\left(\kappa_{adapt} \cos\left(\theta_i - \theta_{adapt}\right)\right)\right) \tag{12}$$

where α_{adapt} is an adaptation ratio specifying maximal suppression, and κ_{adapt} is a concentration parameter that determines the spatial extent of response suppression in the direction domain. For the model simulation, the gain reduction due to adaptation was modulated by $\alpha_{adapt} = 0.35$ and $\kappa_{adapt} = 8.2$. The encoding model after adaptation is illustrated in Fig. 6B.

One concern that can be raised here is that sensory neurons should adapt to the sensory measurement rather than the physical stimulus because the neurons would not have direct access to the true stimulus value. To this end, we also tested an alternative mechanism in which the response gain of a neuron was reduced according to the number of spikes emitted by the neuron in response to the adaptor:

$$g_i = g_0 \frac{1}{(1 + r_i)^{\alpha_{adapt}}} \tag{13}$$

where the role of α_{adapt} is qualitatively similar to that in Eq. (12). We confirmed that changing the adaptation scheme from Eq. (12) to Eq. (13) does not change the simulation results: the average likelihood function is shifted away from the adaptor direction.

Next, we proceed with the notion that the natural environment is stable over time, and the model observer incorporates such temporal statistics by performing recursive Bayesian inference (Fig. 6C). Specifically, the model observer knows that in the natural environment, the stimulus variable θ propagates from θ_{t-1} to θ_t in a given time interval, following a probability distribution called the propagation noise distribution. Motivated by previous work [6], the propagation noise was modeled as a weighted average of von Mises and uniform distributions

$$p(\theta_t|\theta_{t-1}) = p_{same} \frac{\exp\left(\kappa_{prop} \cos\left(\theta_t - \theta_{t-1}\right)\right)}{2\pi I_0\left(\kappa_{prop}\right)} + \left(1 - p_{same}\right)\frac{1}{2\pi} \tag{14}$$

where $I_0(\cdot)$ is the modified Bessel function of the order of 0. We set $p_{same} = 0.9$ and $\kappa_{prop} = 14.6$, such that the resulting distribution closely approximates the empirical observations [6]. For an observer who has already inferred the previous stimulus and knows that the stimulus value would change following the propagation noise distribution, it is reasonable to predict the current stimulus by convolving the distribution of knowledge about the previous stimulus, $p(\theta_{t-1}|\,\mathbf{r}_{1:t-1})$, with the distribution of changes that can occur between consecutive time points, $p(\theta_t|\,\theta_{t-1})$, as follows:

$$p(\theta_t|\mathbf{r}_{1:t-1}) = \int p(\theta_t|\theta_{t-1})p(\theta_{t-1}|\mathbf{r}_{1:t-1})d\theta_{t-1} \quad (15)$$

Finally, to infer the (unknown) stimulus value θ_t, the observer combines prediction from the past sensory responses (Eq. (15)) with the current sensory likelihood (Eq. (11)), according to Bayes' rule:

$$p(\theta_t|\mathbf{r}_{1:t}) \propto p(\mathbf{r}_t|\theta_t)p(\theta_t|\mathbf{r}_{1:t-1}) \quad (16)$$

This final distribution characterizes the model observer's posterior belief on the current stimulus, based on all the information available to the model observer at time t, including information obtained from both past and current sensory responses. We assume a squared-error loss function (or L_2 norm), which is equivalent to computing the posterior mean [113].

$$\hat{\theta}_t(\mathbf{r}_{1:t}) = \int \theta_t p(\theta_t|\mathbf{r}_{1:t})d\theta_t \quad (17)$$

The model observer's final estimate of the stimulus direction, $\hat{\theta}_t(\mathbf{r}_{1:t})$, is a function of sensory response $\mathbf{r}$ over all past and current trials, which makes the marginalization over the latent variable $\mathbf{r}$ particularly demanding. Specifically, to compute the distribution of estimates $p\left(\hat{\theta}_t(\mathbf{r}_{1:t})|\theta_{1:t}\right)$ on a given trial, one would have to marginalize over all possible combinations of $\mathbf{r}_{1:t}$ in response to a stimulus sequence $\theta_{1:t}$, which grows exponentially with the number of trials. Therefore, we simulated the estimation behavior of the model observer using parameters fitted by hand to human data and demonstrate that the model can generate the characteristic pattern inherent in human estimation data. Direction estimates of the model observer were obtained using the same sequence of stimuli that our subjects encountered, assuming no motor noise. We analyzed the simulation data as we did with the empirical data, except for the cardinal bias correction. The results are shown in Fig. 6D.

Abbreviations

AIC: Akaike information criterion; DoG: Derivative of Gaussian; GP: Gaussian process; LIP: lateral intraparietal area; MCMC: Markov chain Monte Carlo; MT: middle temporal area; SD: standard deviation; SEM: standard error of the mean.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12915-022-01444-7.

Additional file 1: Figures S1–S7. Figure S1. Idiosyncratic bias as a function of current stimulus direction. **Figure S2.** Marginal bias plot for representative subjects. **Figure S3.** Artifactual sequential effect of previous response with and without the bias correction, after shuffling the order of the trials. **Figure S4.** Individual behavior and distribution of trials. **Figure S5.** Residuals of the Response model and the Stimulus model. **Figure S6.** Minimal group difference in model parameters. **Figure S7.** Results without applying the bias correction.

Acknowledgements

Not applicable.

Authors' contributions

J.M. and O.-S.K. designed the research, performed the research, analyzed the data, and wrote the paper. All authors read and approved the final manuscript.

Funding

This research was supported by the National Research Foundation of Korea (NRF-2020S1A3A2A02097375 to O.-S.K.).

Availability of data and materials

All data generated or analyzed during this study are included in this published article, its supplementary information files, and publicly available repositories (https://osf.io/s3cx2/) [114].

Declarations

Ethics approval and consent to participate

All procedures were approved by the Ulsan National Institute of Science and Technology Institutional Review Board. We obtained written informed consent from all subjects prior to their participation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 3 February 2022 Accepted: 19 October 2022
Published online: 07 November 2022

References

1. Geisler WS. Visual perception and the statistical properties of natural scenes. Annu Rev Psychol. 2008;59:167–92.
2. Schwartz O, Hsu A, Dayan P. Space and time in visual context. Nat Rev Neurosci. 2007;8(7):522–35.
3. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. Annu Rev Neurosci. 2001;24:1193–216.
4. Wei XX, Stocker AA. A Bayesian observer model constrained by efficient coding can explain 'anti-Bayesian' percepts. Nat Neurosci. 2015;18(10):1509–17.
5. Dong DW, Atick JJ. Statistics of natural time-varying images. Network. 1995;6(3):345–58.
6. van Bergen RS, Jehee JF. Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. J Neurosci. 2019;39(41):8164–76.
7. Kwon O-S, Knill DC. The brain uses adaptive internal models of scene statistics for sensorimotor estimation and planning. Proc Natl Acad Sci USA. 2013;110(11):E1064–73.

8. Barlow HB. A theory about the functional role and synaptic mechanism of visual after-effects. In: Blackmore CB, editor. Vision: Coding and efficiency. Cambridge: Cambridge University Press; 1990. p. 363–75.
9. Brenner N, Bialek W, de Ruyter van Steveninck R. Adaptive rescaling maximizes information transmission. Neuron. 2000;26(3):695–702.
10. Fairhall AL, Lewen GD, Bialek W, de Ruyter van Steveninck R. Efficiency and ambiguity in an adaptive neural code. Nature. 2001;412(6849):787–92.
11. Gepshtein S, Lesmes LA, Albright TD. Sensory adaptation as optimal resource allocation. Proc Natl Acad Sci USA. 2013;110(11):4368–73.
12. Gutnisky DA, Dragoi V. Adaptive coding of visual information in neural populations. Nature. 2008;452(7184):220–4.
13. Müller JR, Metha AB, Krauskopf J, Lennie P. Rapid adaptation in visual cortex to the structure of images. Science. 1999;285(5432):1405–8.
14. Sharpee TO, Sugihara H, Kurgansky AV, Rebrik SP, Stryker MP, Miller KD. Adaptive filtering enhances information transmission in visual cortex. Nature. 2006;439(7079):936–42.
15. Wainwright MJ. Visual adaptation as optimal information transmission. Vision Res. 1999;39(23):3960–74.
16. Weber AI, Krishnamurthy K, Fairhall AL. Coding principles in adaptation. Annu Rev Vis Sci. 2019;5:427–49.
17. Clifford CW. Perceptual adaptation: motion parallels orientation. Trends Cogn Sci. 2002;6(3):136–43.
18. Clifford CW, Webster MA, Stanley GB, Stocker AA, Kohn A, Sharpee TO, et al. Visual adaptation: Neural, psychological and computational aspects. Vision Res. 2007;47(25):3125–31.
19. Kohn A. Visual adaptation: physiology, mechanisms, and functional benefits. J Neurophysiol. 2007;97(5):3155–64.
20. Thompson P, Burr D. Visual aftereffects. Curr Biol. 2009;19(1):R11–4.
21. Webster MA. Visual adaptation. Annu Rev Vis Sci. 2015;1:54–67.
22. Kersten D, Mamassian P, Yuille A. Object perception as Bayesian inference. Annu Rev Psychol. 2004;55:271–304.
23. Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 2004;27(12):712–9.
24. Knill DC, Richards W. Perception as Bayesian inference. Cambridge: Cambridge University Press; 1996.
25. Körding KP. Decision theory: what "should" the nervous system do? Science. 2007;318(5850):606–10.
26. Cicchini GM, Anobile G, Burr DC. Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. Proc Natl Acad Sci USA. 2014;111(21):7867–72.
27. Cicchini GM, Mikellidou K, Burr DC. The functional role of serial dependence. Proc R Soc B. 2018;285:20181722.
28. Fischer J, Whitney D. Serial dependence in visual perception. Nat Neurosci. 2014;17(5):738–43.
29. Liberman A, Fischer J, Whitney D. Serial dependence in the perception of faces. Curr Biol. 2014;24(21):2569–74.
30. Fritsche M, Spaak E, de Lange FP. A Bayesian and efficient observer model explains concurrent attractive and repulsive history biases in visual perception. eLife. 2020;9:e55389.
31. Pascucci D, Mancuso G, Santandrea E, Della Libera C, Plomp G, Chelazzi L. Laws of concatenated perception: Vision goes for novelty, Decisions for perseverance. Plos Biol. 2019;17(3):e3000144.
32. Sheehan TC, Serences JT. Attractive serial dependence overcomes repulsive neuronal adaptation. Plos Biol. 2022;20(9):e3001711.
33. Alais D, Leung J, Van der Burg E. Linear summation of repulsive and attractive serial dependencies: orientation and motion dependencies sum in motion perception. J Neurosci. 2017;37(16):4381–90.
34. Taubert J, Alais D, Burr D. Different coding strategies for the perception of stable and changeable facial attributes. Sci Rep. 2016;6:32239.
35. Bliss DP, Sun JJ, D'Esposito M. Serial dependence is absent at the time of perception but increases in visual working memory. Sci Rep. 2017;7:14739.
36. Kanai R, Verstraten FA. Perceptual manifestations of fast neural plasticity: Motion priming, rapid motion aftereffect and perceptual sensitization. Vision Res. 2005;45(25–26):3109–16.
37. Czoschke S, Fischer C, Beitner J, Kaiser J, Bledowski C. Two types of serial dependence in visual working memory. Brit J Psychol. 2019;110(2):256–67.
38. Fritsche M, Mostert P, de Lange FP. Opposite effects of recent history on perception and decision. Curr Biol. 2017;27(4):590–5.
39. Burr DC, Cicchini GM. Vision: efficient adaptive coding. Curr Biol. 2014;24(22):R1096–8.
40. Chopin A, Mamassian P. Predictive properties of visual adaptation. Curr Biol. 2012;22(7):622–6.
41. Gekas N, McDermott KC, Mamassian P. Disambiguating serial effects of multiple timescales. J Vis. 2019;19(6):24.
42. Suárez-Pinilla M, Seth AK, Roseboom W. Serial dependence in the perception of visual variance. J Vis. 2018;18(7):4.
43. Hajonides JE, van Ede F, Stokes MG, Nobre AC, Myers NE. Multiple and dissociable effects of sensory history on working-memory performance. bioRxiv. 2022. https://doi.org/10.1101/2021.10.31.466639.
44. Moon J, Kwon O-S. Additivity of attractive and repulsive sequential effects in motion direction estimation. J Vis. 2019;19(10):295a. https://doi.org/10.1167/19.10.295a.
45. Kosovicheva A, Whitney D. Stable individual signatures in object localization. Curr Biol. 2017;27(14):R700–1.
46. Fritsche M, de Lange FP. The role of feature-based attention in visual serial dependence. J Vis. 2019;19(13):21.
47. Samaha J, Switzky M, Postle BR. Confidence boosts serial dependence in orientation estimation. J Vis. 2019;19(4):25.
48. Press C, Kok P, Yon D. The perceptual prediction paradox. Trends Cogn Sci. 2020;24(1):13–24.
49. Sadil P, Cowell R, Huber DE. The yin-yang of serial dependence effects: every response is both an attraction to the prior response and a repulsion from the prior stimulus. PsyArXiv. 2021. https://doi.org/10.31234/osf.io/f52yz.
50. Dragoi V, Sharma J, Miller EK, Sur M. Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. Nat Neurosci. 2002;5(9):883–91.
51. Glasser DM, Tsui JM, Pack CC, Tadin D. Perceptual and neural consequences of rapid motion adaptation. Proc Natl Acad Sci USA. 2011;108(45):E1080–8.
52. Fritsche M, Solomon SG, de Lange FP. Brief stimuli cast a persistent long-term trace in visual cortex. J Neurosci. 2022;42(10):1999–2010.
53. Aagten-Murphy D, Burr DC. Adaptation to numerosity requires only brief exposures, and is determined by number of events, not exposure duration. J Vis. 2016;16(10):22.
54. Fornaciai M, Park J. Spontaneous repulsive adaptation in the absence of attractive serial dependence. J Vis. 2019;19(5):21.
55. Cicchini GM, Mikellidou K, Burr DC. Serial dependencies act directly on perception. J Vis. 2017;17(14):6.
56. Kowler E. Cognitive expectations, not habits, control anticipatory smooth oculomotor pursuit. Vision Res. 1989;29(9):1049–57.
57. Murai Y, Whitney D. Serial dependence revealed in history-dependent perceptual templates. Curr Biol. 2021;31(14):3185–91.
58. St John-Saaltink E, Kok P, Lau HC, de Lange FP. Serial dependence in perceptual decisions is reflected in activity patterns in primary visual cortex. J Neurosci. 2016;36(23):6186–92.
59. Fornaciai M, Park J. Attractive serial dependence in the absence of an explicit task. Psychol Sci. 2018;29(3):437–46.
60. Cicchini GM, Benedetto A, Burr DC. Perceptual history propagates down to early levels of sensory analysis. Curr Biol. 2021;31(6):1245–50.
61. Barbosa J, Stein H, Martinez RL, Galan-Gadea A, Li S, Dalmau J, et al. Interplay between persistent activity and activity-silent dynamics in the prefrontal cortex underlies serial biases in working memory. Nat Neurosci. 2020;23(8):1016–24.
62. Fischer C, Czoschke S, Peters B, Rahm B, Kaiser J, Bledowski C. Context information supports serial dependence of multiple visual objects across memory episodes. Nat Commun. 2020;11:1932.
63. Kiyonaga A, Scimeca JM, Bliss DP, Whitney D. Serial dependence across perception, attention, and memory. Trends Cogn Sci. 2017;21(7):493–7.
64. Jazayeri M, Movshon JA. Optimal representation of sensory information by neural populations. Nat Neurosci. 2006;9(5):690–6.
65. Li HH, Sprague TC, Yoo A, Ma WJ, Curtis CE. Joint representation of working memory and uncertainty in human cortex. Neuron. 2021;109(22):3699–712.
66. Ma WJ, Beck JM, Latham PE, Pouget A. Bayesian inference with probabilistic population codes. Nat Neurosci. 2006;9(11):1432–8.
67. Walker EY, Cotton RJ, Ma WJ, Tolias AS. A neural basis of probabilistic computation in visual cortex. Nat Neurosci. 2020;23(1):122–9.

68. Kohn A, Movshon JA. Adaptation changes the direction tuning of macaque MT neurons. Nat Neurosci. 2004;7(7):764–72.
69. Seriès P, Stocker AA, Simoncelli EP. Is the homunculus "aware" of sensory adaptation? Neural Comput. 2009;21(12):3271–304.
70. Stocker AA, Simoncelli EP. Sensory adaptation within a Bayesian framework for perception. Adv Neural Inf Process Syst. 2006;18:1289–96.
71. van Bergen RS, Kriegeskorte N. Going in circles is the way forward: the role of recurrence in visual inference. Curr Opin Neurobiol. 2020;65:176–93.
72. Körding KP, Tenenbaum JB, Shadmehr R. The dynamics of memory as a consequence of optimal adaptation to a changing body. Nat Neurosci. 2007;10(6):779–86.
73. Kwon O-S, Tadin D, Knill DC. Unifying account of visual motion and position perception. Proc Natl Acad Sci USA. 2015;112(26):8142–7.
74. Mehta B, Schaal S. Forward models in visuomotor control. J Neurophysiol. 2002;88(2):942–53.
75. Orban de Xivry JJ, Coppe S, Blohm G, Lefèvre P. Kalman filtering naturally accounts for visually guided and predictive smooth pursuit dynamics. J Neurosci. 2013;33(44):17301–13.
76. Saunders JA, Knill DC. Visual feedback control of hand movements. J Neurosci. 2004;24(13):3223–34.
77. Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. Science. 1995;269(5232):1880–2.
78. Pasternak T, Tadin D. Linking neuronal direction selectivity to perceptual decisions about visual motion. Annu Rev Vis Sci. 2020;6:335–62.
79. Block N. Two neural correlates of consciousness. Trends Cogn Sci. 2005;9(2):46–52.
80. Born RT, Bradley DC. Structure and function of visual area MT. Annu Rev Neurosci. 2005;28:157–89.
81. Gold JI, Shadlen MN. The neural basis of decision making. Annu Rev Neurosci. 2007;30:535–74.
82. Huk AC, Katz LN, Yates JL. The role of the lateral intraparietal area in (the study of) decision making. Annu Rev Neurosci. 2017;40:349–72.
83. Sohn H, Narain D. Neural implementations of Bayesian inference. Curr Opin Neurobiol. 2021;70:121–9.
84. Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, et al. Probabilistic population codes for Bayesian decision making. Neuron. 2008;60(6):1142–52.
85. Akrami A, Kopec CD, Diamond ME, Brody CD. Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature. 2018;554(7692):368–72.
86. Basso MA, Wurtz RH. Modulation of neuronal activity by target uncertainty. Nature. 1997;389(6646):66–9.
87. Darlington TR, Beck JM, Lisberger SG. Neural implementation of Bayesian inference in a sensorimotor behavior. Nat Neurosci. 2018;21(10):1442–51.
88. Rao V, DeAngelis GC, Snyder LH. Neural correlates of prior expectations of motion in the lateral intraparietal and middle temporal areas. J Neurosci. 2012;32(29):10063–74.
89. Hanks TD, Mazurek ME, Kiani R, Hopp E, Shadlen MN. Elapsed decision time affects the weighting of prior probability in a perceptual decision task. J Neurosci. 2011;31(17):6339–52.
90. Jazayeri M, Shadlen MN. A neural mechanism for sensing and reproducing a time interval. Curr Biol. 2015;25(20):2599–609.
91. Beck JM, Latham PE, Pouget A. Marginalization in neural circuits with divisive normalization. J Neurosci. 2011;31(43):15310–9.
92. Denève S, Duhamel JR, Pouget A. Optimal sensorimotor integration in recurrent cortical networks: a neural implementation of Kalman filters. J Neurosci. 2007;27(21):5744–56.
93. Funamizu A, Kuhn B, Doya K. Neural substrate of dynamic Bayesian inference in the cerebral cortex. Nat Neurosci. 2016;19(12):1682–9.
94. Ganguli D, Simoncelli EP. Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. Neural Comput. 2014;26(10):2103–34.
95. Narain D, Remington ED, Zeeuw CID, Jazayeri M. A cerebellar mechanism for learning prior distributions of time intervals. Nat Commun. 2018;9:469.
96. Nessler B, Pfeiffer M, Buesing L, Maass W. Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. Plos Comput Biol. 2013;9(4):e1003037.
97. Sohn H, Narain D, Meirhaeghe N, Jazayeri M. Bayesian computation through cortical latent dynamics. Neuron. 2019;103(5):934–47.
98. Kim S, Burr D, Cicchini GM, Alais D. Serial dependence in perception requires conscious awareness. Curr Biol. 2020;30(6):R257–8.
99. Ma WJ. Organizing probabilistic models of perception. Trends Cogn Sci. 2012;16(10):511–8.
100. Seriès P, Seitz A. Learning what to expect (in visual perception). Front Hum Neurosci. 2013;7:668.
101. Kalm K, Norris D. Visual recency bias is explained by a mixture model of internal representations. J Vis. 2018;18(7):1.
102. de Azevedo Neto RM. Commentary: Probabilistic representation in human visual cortex reflects uncertainty in serial decisions. Front Hum Neurosci. 2020;14:580581.
103. Brainard DH. The psychophysics toolbox. Spat Vis. 1997;10:433–6.
104. Berens P. CircStat: a MATLAB toolbox for circular statistics. J Stat Softw. 2009;31(10):1–21.
105. Manassi M, Liberman A, Kosovicheva A, Zhang K, Whitney D. Serial dependence in position occurs at the time of perception. Psychon Bull Rev. 2018;25(6):2245–53.
106. Tomassini A, Morgan MJ, Solomon JA. Orientation uncertainty reduces perceived obliquity. Vision Res. 2010;50(5):541–7.
107. Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. New York: Springer; 2006.
108. Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge: The MIT Press; 2006.
109. Acerbi L, Wolpert DM, Vijayakumar S. Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. Plos Comput Biol. 2012;8(11):e1002771.
110. Zhang R, Kwon O-S, Tadin D. Illusory movement of stationary stimuli in the visual periphery: evidence for a strong centrifugal prior in motion processing. J Neurosci. 2013;33(10):4415–23.
111. Kruschke JK. Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan. London: Elsevier; 2014.
112. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci. 1992;7:457–72.
113. Jazayeri M, Shadlen MN. Temporal context calibrates interval timing. Nat Neurosci. 2010;13(8):1020–6.
114. Moon J, Kwon O-S. Attractive and repulsive effects of sensory history concurrently shape visual perception. OSF; 2022. Available from: https://osf.io/s3cx2/.

Publisher's Note