UC Riverside
UC Riverside Previously Published Works

Title

Expectations developed over multiple timescales facilitate visual search performance

Permalink

Journal

ISSN

Authors

Gekas, Nikos
Seitz, Aaron R
Seriès, Peggy

Publication Date

DOI

Expectations developed over multiple timescales facilitate visual search performance

Nikos Gekas

IANC, School of Informatics, University of Edinburgh, Edinburgh, UK

Aaron R. Seitz

UC Riverside, Department of Psychology, Riverside, CA, USA

Peggy Seriès

IANC, School of Informatics, University of Edinburgh, Edinburgh, UK

Our perception of the world is strongly influenced by our expectations, and a question of key importance is how the visual system develops and updates its expectations through interaction with the environment. We used a visual search task to investigate how expectations of different timescales (from the last few trials to hours to long-term statistics of natural scenes) interact to alter perception. We presented human observers with low-contrast white dots at 12 possible locations equally spaced on a circle, and we asked them to simultaneously identify the presence and location of the dots while manipulating their expectations by presenting stimuli at some locations more frequently than others. Our findings suggest that there are strong acuity differences between absolute target locations (e.g., horizontal vs. vertical) and preexisting long-term biases influencing observers' detection and localization performance, respectively. On top of these, subjects quickly learned about the stimulus distribution, which improved their detection performance but caused increased false alarms at the most frequently presented stimulus locations. Recent exposure to a stimulus resulted in significantly improved detection performance and significantly more false alarms but only at locations at which it was more probable that a stimulus would be presented. Our results can be modeled and understood within a Bayesian framework in terms of a near-optimal integration of sensory evidence with rapidly learned statistical priors, which are skewed toward the very recent history of trials and may help understanding the time scale of developing expectations at the neural level.

Introduction

There is a plethora of evidence that perception is strongly influenced by expectations. Particularly in situations of high uncertainty, we rely not only on the information we can gather at the present moment but also on our knowledge of the world and our previous experience in it. Expectations can be formed automatically and continuously, on shorter or longer timescales, and have universal impact or apply only in specific situations. Based upon how expectations generalize across time and environment, they can be divided into two major categories; "structural" and "contextual" (Seriès & Seitz, 2013). Structural expectations are developed over long time frames based on implicit learning of the statistics of the natural environment, or they can be innate (e.g., the expectation that "light comes from above;" Adams, Graf, & Ernst, 2004). Structural expectations apply equally to already experienced situations and novel ones. In contrast, contextual expectations modulate perception in isolated temporal or spatial situations. Contextual expectations can be manipulated explicitly or implicitly over short time frames through sensory cues, specific instructions, or the context in which a stimulus is shown (e.g., Haijiang, Saunders, Stone, & Backus, 2006; Kok, Brouwer, van Gerven, & de Lange, 2013; Sterzer, Frith, & Petrovic, 2008).

At the same time, a growing body of work suggests that visual perception can be thought of as a continuous process of Bayesian inference (Fiser, Berkes, Orbán, & Lengyel, 2010). Under that framework, expectations correspond to the "prior" probability and combine with the observed "likelihood" to form the "posterior" probability of a hypothesis to be

true. In visual perception, the hypothesis could correspond to the presence or a feature of a visual stimulus. The higher the uncertainty of the observed visual information is, the stronger the influence of the prior on the posterior and on the final interpretation of the visual information. In previous work (Chalk, Seitz, & Seriès, 2010; Gekas, Chalk, Seitz, & Seriès, 2013), we found that human observers quickly, automatically, and implicitly developed expectations based on the statistical distribution of visual motion stimuli. These expectations induced biases on the perceived motion direction of presented stimuli but also induced false alarms (so-called "hallucinations" of motion) in the absence of a stimulus. This behavior was well explained by models that assumed observers acted as Bayesian observers, using a prior distribution that approximated the stimulus statistics and suggests that Bayesian models are a parsimonious way to describe how expectations of the environment modulate perception.

Expectations can be developed over long time frames (hours, days, or years). For example, Stocker and Simoncelli (2006) provided evidence toward the hypothesis that the visual system expects objects to be static or move slowly, and this prior expectation can explain perceptual phenomena, such as the aperture problem and why speed perception can differ between high- and low-contrast stimuli (Stone & Thompson, 1992). Although this slow-speed prior is thought to develop over the lifetime, recent research shows that experience within an hour-long experimental session and across days of exposure with quickly moving stimuli can alter this prior toward an expectation of more quickly moving speeds (Sotiropoulos, Seitz, & Seriès, 2011). Expectations can also be developed over very short time frames (seconds or minutes). When we search for a target with a particular feature (shape, orientation, color, etc.), it is easier to detect or discriminate that target or one of its features if we have seen it or interacted with it in the immediate past. This effect is formalized as perceptual priming and suggests that an implicit memory system strongly influences how visual attention is allocated after exposure to a stimulus (Kristjánsson & Campana, 2010). Perceptual priming in visual search has been studied for a variety of features, including orientation (Olivers & Meeter, 2006), motion direction (Kristjánsson, 2009), shape (Fecteau, 2007), and color (Maljkovic & Nakayama, 1994). Many studies have also shown that repetition of trials with a target in the same location can improve search performance significantly (e.g., Geng & Behrmann, 2005; Maljkovic & Nakayama, 1996; Miller, 1988), and this improvement can be very location-specific. For example, Le Dantec and Seitz (2012) showed that repeatedly performing a visual search task to find a subtly different line orientation led to long-lasting performance improvements in a large number of

independent locations that incompletely transferred to neighboring locations as close as 1.5° of visual angle. Although these studies have shown that the statistical predictability of a target's location due to repetition can facilitate performance, a study by Druker and Anderson (2010) found that the statistical properties of a target's location could influence the observer's performance even outside of priming effects. In order to dissociate the effect of a high-probability location from a simple location repetition, they used continuous probability distributions that included a very large number of possible locations in contrast to typical visual search experiments that use a limited number of possible locations. They found that subjects learned the distribution of the stimulus implicitly, and their performance was improved more than what would account to just priming effects given the distance to recently presented targets. Together, these results suggest that people are continuously integrating the statistics of the environment and using this information to update their expectations of future experiences.

Visual search provides a useful framework in which to investigate the formation of expectations. Although mainly signal detection approaches have been used in accounting for visual search phenomena (Verghese, 2001), studies have also used a Bayesian framework to successfully model human visual search behavior and investigate how it compares to that of a Bayesian optimal observer (Eckstein, Abbey, Pham, & Shimozaki, 2004; Eckstein, Peterson, Pham, & Droll, 2009; Elazary & Itti, 2010; Ma, Navalpakkam, Beck, Van Den Berg, & Pouget, 2011). Droll, Abbey, and Eckstein (2009) compared the performance of learning the statistics of cue validity by human observers in a visual search task to that of an ideal Bayesian observer. The authors found that human observers were able to learn the statistics in a single experimental session, but learning was slower compared to that of an ideal observer even using supervised feedback. Recently, Vincent (2011) investigated whether better performance in a visual search task is achieved by combining visual evidence and prior beliefs in a Bayesian optimal way. Observers' prior expectations were manipulated in two experiments via peripheral cuing and via explicit information about the stimulus spatial probabilities. It was found that observers improved their detection rates by optimally combining slightly biased priors with sensory evidence irrespective of how expectations were manipulated. Interestingly, counterpredictive peripheral cues (i.e., cues that indicated a location was less likely than average to contain a target) increased choice reaction times whereas counterpredictive spatial probabilities slightly decreased reaction times. This suggested that counterpredictive cues guided observers' attention involuntarily and unavoidably to the less probable cued locations.

These results suggest a link between perceptual priming and expectation formation and that these may both be parsimoniously described within a Bayesian model. One possibility is that perceptual priming influences how attention shifts to repeated stimulus features or locations. Sigurdardottir, Kristjánsson, and Driver (2008) showed that priming improved detection performance for a target but did not facilitate acuity judgments for the same target. This suggests that priming might influence the speed of attentional shifts rather than stimulus sensitivity directly. Also, in an experiment in which eye movements were analyzed, Becker (2008) found that the accuracy and time course of the first saccade within a trial was modulated by priming effects. Observers would saccade faster and more accurately when the same target was repeated than when it changed between trials. At the same time, many studies have manipulated contextual expectations so as to direct attention to particular locations or features. For example, Posner (1980) developed a task in which a cue explicitly predicts the location of a subsequent target with a certain probability. Subjects were found to process the target stimuli faster and more accurately on correctly cued trials than on incorrectly cued trials, and the difference has been shown to increase with cue validity (e.g., Downing, 1988). In that sense, priming may be considered as a form of contextual expectation in a very short timescale. This outstanding issue naturally extends to the question of what is the distinction between attention and expectations. The focus of spatial attention has been successfully compared to a spotlight (Posner, 1980) as well as to a zoom lens (Eriksen & Yeh, 1985). A Bayesian account for attention was described by Eckstein, Drescher, and Shimozaki (2006), in a study in which subjects looked at pictures with targets at expected or unexpected locations or targets completely absent. A differential weighting Bayesian model was consistent with subjects' pattern of first saccades toward probable locations in target-absent images. Both attention and expectations are thought to be controlled by similar cognitive processes (Corbetta & Shulman, 2002), and on a behavioral level, both can have superficially identical effects on performance. However, the exact mechanisms that produce these effects remain unclear as is the exact nature of the interaction between attention and expectations (Summerfield & Egner, 2009).

In this study, we addressed these issues by investigating the form in which perceptual priming acts in a statistical learning experimental paradigm and how it interacts with expectations formed over longer timescales. We do this through a novel visual search paradigm (although with some similarities to the task in Droll et al., 2009) in which we presented human observers with brief displays of low-contrast stimuli and asked them to report the presence of a stimulus (yes/no task) as well as the exact location of the stimulus (localization task) at the same time. We manipulated their expectations by presenting stimuli in some locations more frequently than others. This task provides an interesting glimpse into the impact of priors in perceptual judgments as we are able to track a number of separate but likely related types of errors: *Mislocalizations* are errors in detection with which the location of an item reported is incorrect, *false alarms* are errors in detection with which stimuli are detected when none were present, and *positional errors* refer to small but systematic biases in localization estimates within the neighborhood of a target location. We hypothesized that subjects would learn the stimulus distribution and use that information to improve their performance in the task and, consistent with prior work (Chalk et al., 2010; Gekas et al., 2013), to observe biases in subjects' localization performance toward the more frequently presented locations (*positional errors*) and increased false detections in the absence of stimulus (*false alarms* and *mislocalizations*), matching the probability distribution of the actual stimuli. In addition to these effects, we investigated the form of interaction between the rapidly learned expectations of the stimulus distribution and priming from very recent stimulus presentations. Our hypothesis was that subjects would integrate both sources of information, which would facilitate their detection performance in the more probable stimulus locations but also induce more false alarms in the process.

Methods

Subjects and stimuli

Twenty-eight naive subjects (17 of them female; 19–33 years of age) with normal or corrected vision were recruited from the University of California, Riverside. All subjects gave informed written consent in accordance with the University of California, Riverside Human Research Review Board and the Declaration of Helsinki and received course credit for their participation.

The stimuli consisted of one, two, or three white dots (0.5° in diameter) at 12 possible locations equally spaced on a circle at 4° of visual angle from the center of the screen. They were generated using the Matlab programming language with the psychophysics toolbox (Brainard, 1997; Pelli, 1997) and displayed on a CRT monitor with a resolution of 1400 × 1050 at 100 Hz. Subjects viewed the display in a darkened room at a viewing distance of 70 cm. A chin rest was used to maintain a constant head location and viewing

distance. The display luminance was calibrated and linearized with a Cambridge Research Systems Color-imeter, and the background luminance was set to 5.05 cd/m^2.

Procedure

At the beginning of each trial, a central white cross was presented as a fixation point (Figure 1A). Then, the stimulus was presented for 100 ms. The display cleared, and subjects were presented with a central circle and a cursor, which they could move freely with a mouse. Subjects were instructed to click inside the circle to finish the trial if no stimulus was perceived. On the other hand, they were instructed to move the cursor outside the circle if they had perceived one or more stimuli. Then, a white dot appeared at the same eccentricity as the stimulus and moved in conjunction with the mouse cursor. Its function was to help subjects make an accurate localization of where they had perceived a stimulus. Subjects were instructed to move the cursor to the location at which they detected a stimulus and click to validate their decision. At the same time, a bar extended from the center of the screen to the point of the cursor. Subjects used the bar to indicate their confidence level of the stimulus being present. The longer the length of the bar the more confident they were of their choice. After clicking, a small blue dot and a blue bar remained on screen. Subjects were free to report as many stimuli as they wanted. To finish the trial, they had to return the cursor inside the central circle and click. No immediate feedback was given after each trial. However, block feedback was given every 50 trials in the form of detection performance in 10% steps (e.g., "Your performance rate was between 60% and 70%") along with a motivational message.

Design

The experiment consisted of two 1-hr sessions (conducted on successive days) of 900 trials each. The stimuli were presented in three different contrast levels: In 60% of trials, contrast was determined using a 2/1 staircase on detection performance (staircase contrast); in 10% of trials, contrast was high (1.05 cd/m^2 above the background luminance), and stimuli were easily visible (high contrast); and in 30% of trials, there was no stimulus presented at all (zero contrast). High-contrast trials were used as a metric of subjects' confidence and localization performance. Because stimuli were easily detected in those trials, subjects should, on average, be very confident of their choice and also be fairly accurate in the localization of the
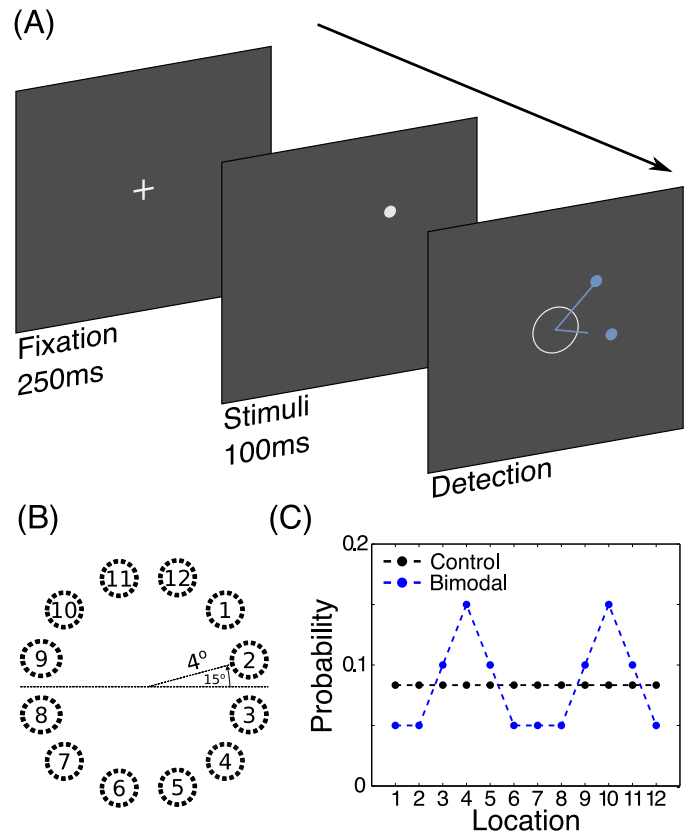

Figure 1. Experimental procedure. (A) Subjects were presented with a fixation point followed by the stimulus for a brief 100 ms. After the screen was cleared, subjects were presented with a circle and a cursor, which they could freely move. If they had not perceived a stimulus, they were instructed to click inside and finish the trial. If they had perceived a stimulus, they were instructed to move the cursor outside of the circle, and a dot similar to the stimulus would appear to allow them to indicate the exact location of the target. Simultaneously, they could extend a bar away from the circle to indicate their confidence level of seeing a stimulus at that location. (B) There were 12 possible stimulus locations at 4° of visual angle, equally spaced on a circle, 15°, 45°, and 75° away from the horizontal cardinal. (C) Probability distributions of presented stimulus locations for the control (black dots) and bimodal (blue dots) groups of subjects. In the control group, all locations were equally presented, and in the bimodal group, four locations were two times more likely to be presented, and two locations were three times more likely to be presented.

stimulus. Thus, performance in the high-contrast trials allowed us to calculate a baseline behavior for each subject regarding his or her confidence and localization error when reporting a stimulus and compare it to behavior in the staircase and zero-contrast trials. In the staircase and high-contrast levels, up to three stimuli could be presented in the same trial; one stimulus was presented in 73.33% of trials, two stimuli in 20%, and three stimuli in 6.67%. We used multiple stimuli in

some trials in order to encourage more false alarms during the experiment. If subjects were uncertain about the exact number of stimuli presented in each trial, they would be more likely to report a false alarm even in trials in which they had already reported a presented stimulus.

Stimuli could appear at 12 possible locations (Figure 1B). The probability that a stimulus would appear at a given location depended upon the subject's group. In the *control* group ($n = 12$), stimuli were equally likely to appear at all locations (Figure 1C). In the *bimodal* group ($n = 16$), stimuli were more likely to appear at locations in quadrants II and IV than at locations in quadrants I and III; in particular, stimuli were twice as likely to appear at locations 15° (locations 3 and 9) and 75° (locations 5 and 11) away from the horizontal cardinal and three times more likely at locations 45° away from the horizontal cardinal (locations 4 and 10). This grid was rotated for each subject such that the location on the circle for location 1 was located at one of four possible orientations (45°, 135°, 225°, 315°). As the bimodal distribution is symmetrical, this created two possible subgroups of frequent versus nonfrequent locations that were counterbalanced between subjects (eight subjects in each subgroup) in order to cancel out any existing biases.

Data analysis

The first 150 trials from each session were excluded from the analysis in order for the staircase to reach stable levels. No significant differences were observed in subjects' behavior across the two experimental sessions (Supplementary Figure 1). Thus, we combined data across both sessions. High-contrast trials were excluded from the analysis.

As discussed in the experimental procedure, subjects were free to make an exact localization of a stimulus they reported. In order to count correctly detected stimuli, we divided all possible angles in 12 30° bins. The center of each bin was the exact angle of each presented location. We then assumed that subjects correctly identified the location of a stimulus, classified as a *correct detection,* when they localized inside the respective 30° bin of each location; otherwise, the response was classified as a *mislocalization*. During our analysis, it became clear that subjects' performance was significantly affected by the absolute location of a stimulus and, in particular, by the distance from the horizontal cardinal. In order to simplify our analysis and be able to show any underlying effects of the stimulus distribution, we divided the presented locations into three categories based on the distance from the horizontal cardinal. Thus, locations were divided into *horizontal* (15° away from the cardinal; absolute

locations 2, 3, 8, and 9 in Figure 1B), *intermediate* (45° away; locations 1, 4, 7, and 10), and *vertical* locations (75° away; locations 5, 6, 11, and 12). In the Results, we will present subjects' performance over the different location categories in a within-subjects analysis.

Subjects would sometimes report that they had perceived one or more stimuli in trials in which no stimulus was presented (zero-contrast trials). They did so in approximately 2% of trials for the control and the bimodal groups. However, that frequency was not large enough to allow for a within-subjects analysis. Instead, we resampled each subject's data with replacement for all 12 locations and then aggregated all false alarms from all subjects. We repeated the process 100,000 times and calculated 95% confidence intervals. As we discussed in the Introduction, we will refer to these responses as *false alarms*. The aforementioned *mislocalizations* were more frequent than *false alarms* consistently across all subjects (in approximately 10% of trials for the control and 8% for the bimodal group), allowing us to do an analysis similar to the detection performance (within subjects).

Results

Stimulus distribution effects on detection performance

First, we investigated whether the stimulus distribution had an effect on subjects' detection performance and the probability that subjects would make a mislocalization (e.g., reporting the wrong stimulus location when a stimulus was present). As discussed previously, we separated stimuli locations into three categories by their distance from the horizontal cardinal: horizontal, intermediate, and vertical (Figure 2). Further, we separated the locations of stimuli presented to the bimodal group by the probability of a stimulus to appear at a location into *frequent* and *nonfrequent* locations. Figure 2A shows the probability distribution of each location category for the control group and for the two conditions of the bimodal group. It was twice as likely for a stimulus to appear at a frequent horizontal or vertical location than at a nonfrequent similar location, and three times more likely for a frequent intermediate location than a nonfrequent intermediate location. If there is a strong effect of the stimulus distribution on performance, we expect to find a significant difference between the detection performance of frequent and nonfrequent locations as well as significantly more false alarms at frequent than nonfrequent locations.

Looking at the fraction of correctly detected stimuli (i.e., the fraction of correct detections over total stimuli
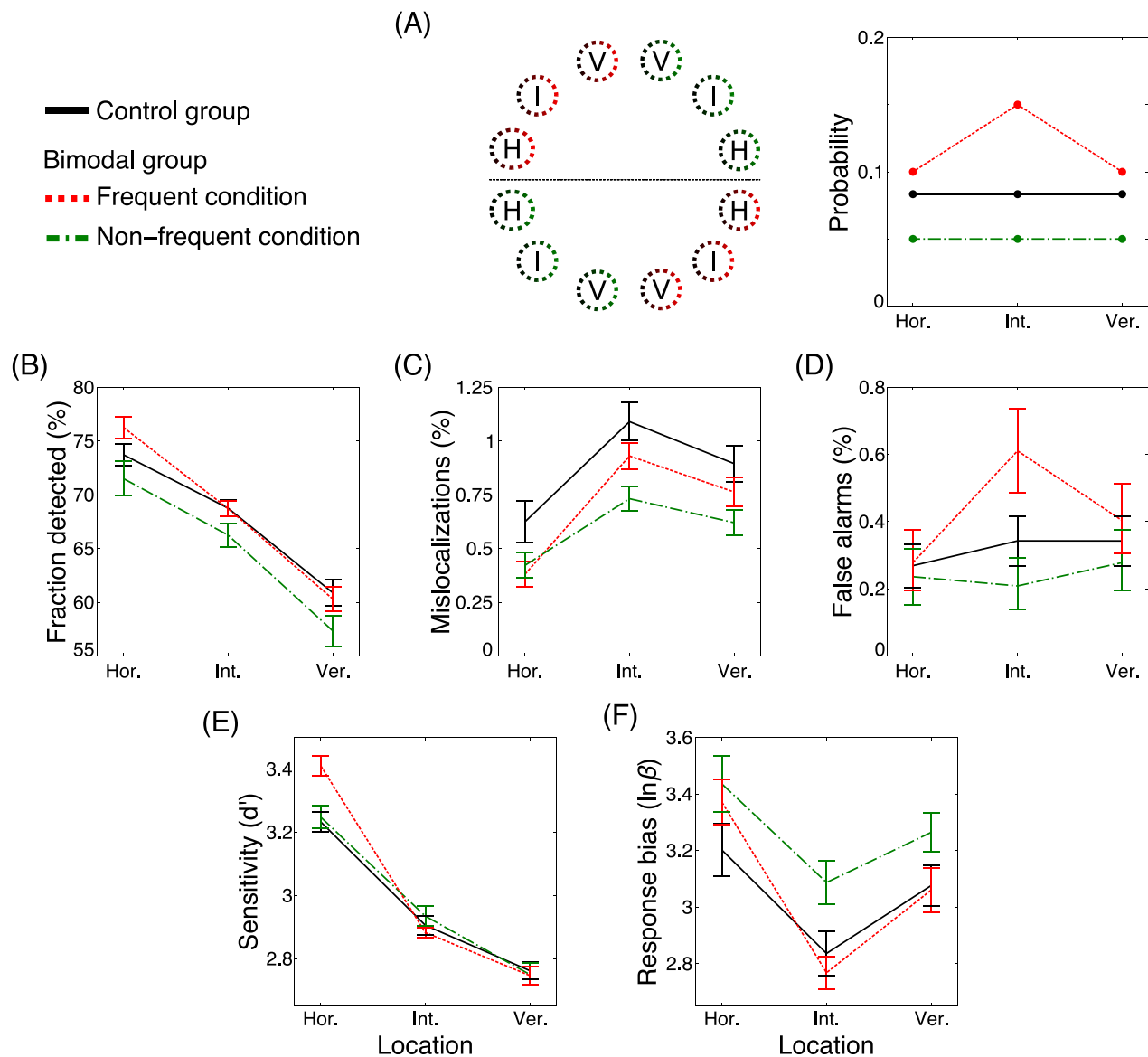
Figure 2. Effect of stimulus distribution on detection performance. (A) (Left) Dividing stimulus locations by their distance from the horizontal cardinal. (Right) Probability distributions of stimulus locations divided by their distance from the horizontal cardinal for the control group (black solid line), and for the frequent (red dashed line) and nonfrequent (green dashed-dotted line) conditions of the bimodal group. (B) The fractions of correctly detected stimuli are plotted against presented stimulus location. (C) Relative frequencies of subjects' mislocalizations and (D) false alarms are plotted against stimulus location. (E) Subjects' mean sensitivity and (F) response bias are plotted against stimulus location. Results are averaged over all subjects and error bars show within-subject standard error, except for (D) false alarms in which results are summed over all subjects and error bars show 95% confidence intervals.

presentations; Figure 2B), there is a significant effect of absolute stimulus location on performance. Subjects were significantly better at detecting a stimulus at locations that were closer to the horizontal cardinal and increasingly worse at locations away from that cardinal ($p < 0.0001$, one-way within-subjects ANOVA for control, frequent, and nonfrequent). Regarding the effect of the stimulus distribution, detection rates were consistently higher for stimuli presented at the frequent locations than at the nonfrequent locations, there was a significant effect of a location being frequent on the

detection performance ($p = 0.007$, three-way within-subjects ANOVA), and there was no significant interaction between location frequency and distance from the horizontal cardinal ($p = 0.69$). It is not surprising that the average detection performance is very similar between the control group and the frequent condition of the bimodal group because its upper bound is set dynamically by the 2/1 staircase procedure on correct detection. Overall, subjects in the bimodal group appeared to have learned the stimulus distribution, and that facilitated their performance at the

frequently presented locations but at the cost of reduced performance at the nonfrequently presented locations.

Looking at subjects' mislocalizations, there was a significant effect of absolute stimulus location on their relative frequencies ($p = 0.02$, $p = 0.0001$, and $p = 0.01$, one-way within-subjects ANOVA for control, frequent, and nonfrequent, respectively; Figure 2C). The frequency of mislocalizations was calculated out of the total number of trials in which a stimulus was presented at a different location to the mislocalization. The total number of trials in which a stimulus was presented at a frequent location was larger than at a nonfrequent location, so it was important to calculate relative frequencies and not just absolute frequencies. In contrast to detection rates, subjects were significantly more likely to report a stimulus in the intermediate or vertical locations than in the horizontal locations. There was no significant effect of a location's condition (frequent or nonfrequent) on the relative frequency of mislocalizations ($p = 0.13$, three-way within subjects ANOVA) and no significant interaction between a location's condition and distance from the horizontal cardinal ($p = 0.2$). Overall, it was more likely, but nonsignificantly so ($p = 0.18$, two-way between-subjects ANOVA), for subjects of the control group to make an incorrect response than for subjects of the bimodal group.

We also found that subjects reported *false alarms* in no-stimulus trials (Figure 2D) and were consistently more likely to report a stimulus at a frequent location than at a nonfrequent one, but the difference was significant only at intermediate locations ($p = 0.0005$). This observation is consistent with prior work in which false alarms were shown to be consistent with perceptual hallucinations of motion (Chalk et al., 2010; Gekas et al., 2013; Seitz, Nanez, Holloway, Koyama, & Watanabe, 2005). The average frequency that subjects of the bimodal group would report false alarms was almost identical to subjects of the control group for horizontal locations. However, it was more likely for subjects of the bimodal group to report a false alarm at a frequent location than subjects of the control group, nonsignificantly for vertical locations and significantly for intermediate locations ($p = 0.01$). This suggests that the stimulus distribution had a significant effect not only between locations (frequent vs. nonfrequent intermediate locations) but induced significantly more overall false alarms in the most frequently presented locations (frequent intermediate). As in our previous work, the close similarity between subjects' false alarm distributions and the distribution of the stimulus suggests that false alarms might directly reflect the prior beliefs of subjects in the task. In summary, the results suggest that the stimulus distribution facilitated detection performance but also increased subjects' false alarms at the frequent stimulus locations.

Finally, we computed subjects' sensitivity (Figure 2E) and response bias (Figure 2F). The hit rate at each location was defined as the number of correct detections divided by the total number of trials in which a stimulus was presented at that location, and the false positive rate was defined as the sum of mislocalizations and false alarms divided by the total number of trials in which a stimulus was not presented at that location. Because the number of false alarms was low, we used a loglinear approach before calculating the hit and false positive rates, which involves adding 0.5 to both the number of hits and false positives and adding 1 to both numbers of trials (Stanislaw & Todorov, 1999). For sensitivity, we computed d' and for response bias the natural logarithm of β. Not surprisingly, sensitivity significantly decreased for locations away from the horizontal cardinal ($p = 0.008$ and $p < 0.0001$ for subjects of the control and bimodal groups, respectively, one-way within-subjects ANOVA). Looking at the subjects of the bimodal group, there was no significant effect of condition on sensitivity ($p = 0.44$, two-way within-subjects ANOVA) but significant interaction of location × condition ($p = 0.002$) due to the large sensitivity difference at the frequent horizontal locations. The opposite was found regarding subjects' response biases. There was a significant effect of location and condition on response bias ($p = 0.0001$ and $p = 0.0125$, respectively) but no significant interaction of location × condition ($p = 0.43$). The difference between conditions was smallest at horizontal locations and largest at the intermediate locations. Interestingly, subjects of the control group were more likely to report a stimulus at the intermediate locations, but there was no significant effect of location of response bias ($p = 0.35$, one-way within-subjects ANOVA). Overall, the signal detection theory analysis highlights the significant sensitivity differences between locations closer and further away from the horizontal cardinal and the effect of the stimulus distribution on subjects' bias to report the presence of a stimulus.

Stimulus distribution effects on localization performance

Absolute stimulus location and location frequency had a significant effect on detection performance. We next looked at whether these properties had an effect on subjects' positional estimates of presented stimulus. We expected that if there were a strong effect, subjects' *positional errors* in reporting the target locations, on average, would be biased toward the frequently presented locations. Figure 3A shows the averaged
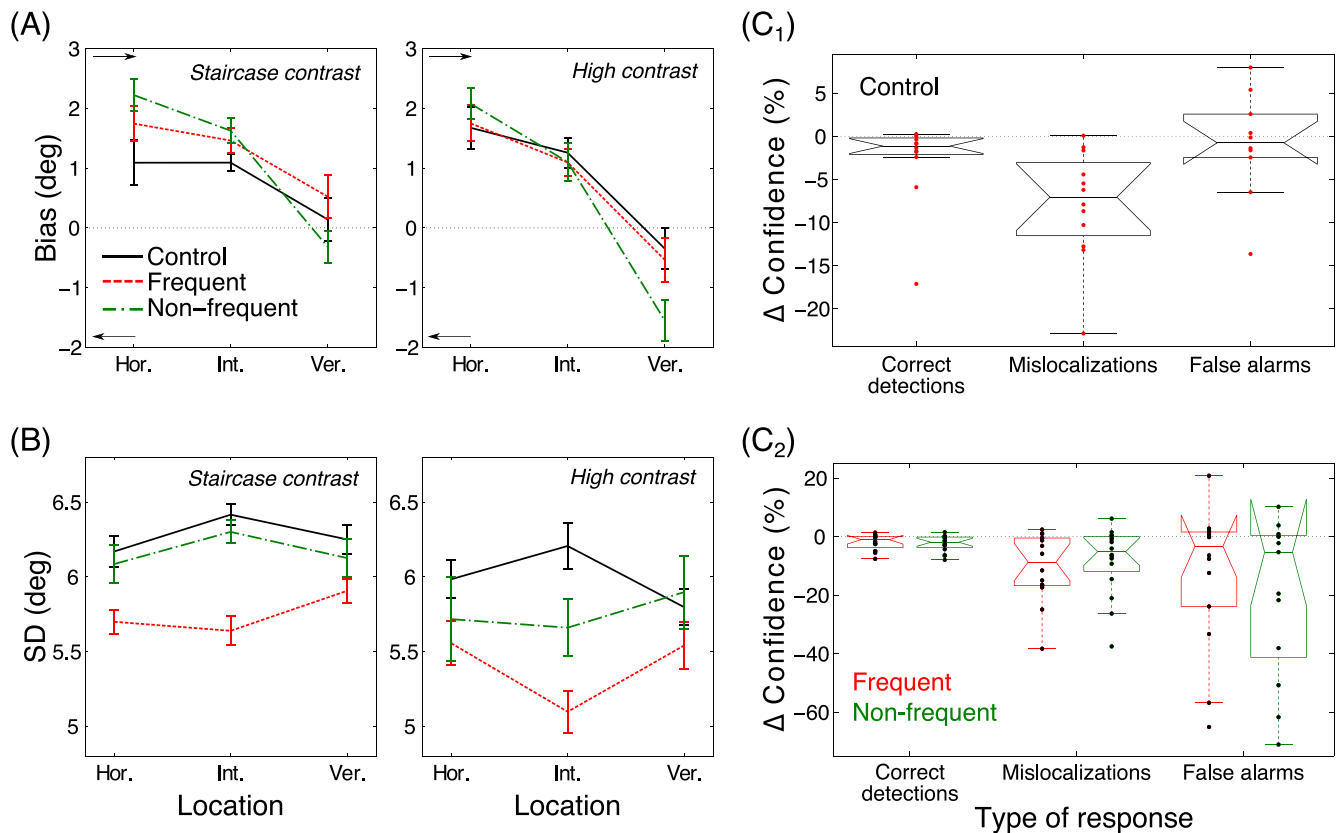
Figure 3. Effect of stimulus distribution on localization performance and confidence. (A) Subjects' mean positional error biases are plotted against stimulus location for (left) staircase-contrast stimuli and (right) high-contrast stimuli. A positive bias indicates localizations away from the horizontal cardinal (0° angle), and a negative bias indicates localizations toward the horizontal cardinal. (B) Standard deviations in subjects' position estimate distributions are plotted against stimulus location for (left) staircase-contrast stimuli and (right) high-contrast stimuli. Results are averaged over all subjects, and error bars show within-subject standard error. (C) Box plots, along with individual subject data, of the differences in confidence level reported by subjects of the (C_1) control group and of the (C_2) bimodal group (divided into frequent and nonfrequent conditions) between high-contrast trials and correct detections of a stimulus (left), mislocalizations (center), and false alarms (right). Dots indicate values for each subject. Each box shows the interquartile range, the horizontal line within the box shows the median, and the notches show 95% confidence intervals on the median.

positional error biases for staircase-contrast and high-contrast trials. A positive bias indicates systematic positional errors away from the horizontal cardinal and a negative bias systematic positional errors toward the horizontal cardinal. Subjects of both the control and bimodal groups appear to be biased away from the horizontal cardinal at horizontal and intermediate locations both in staircase and high-contrast stimuli. In vertical locations, they were slightly biased or unbiased in staircase-contrast stimuli but negatively biased in high-contrast stimuli. There was no significant difference between subjects' positional error biases at frequent and nonfrequent locations in staircase or high-contrast trials ($p = 0.85$ and $p = 0.51$, respectively), but there was a significant effect of absolute position on positional error biases ($p < 0.0001$). The similarity of biases for both low- (staircase) and high-contrast stimuli at frequent and nonfrequent locations suggest

that these biases are preexisting and largely unaffected by presented contrast or stimulus distribution.

However, we still expect position estimates in high-contrast trials to be more consistent across locations and subjects in comparison to staircase-contrast trials. Figure 3B plots the averaged standard deviations of position estimates for staircase-contrast and high-contrast trials. Indeed, standard deviations in high-contrast trials were significantly smaller than in staircase-contrast trials across all data ($p = 0.041$, $p = 0.011$, $p = 0.016$, three-way within-subjects ANOVA for control, frequent, and nonfrequent, respectively). There is also a strong effect of the stimulus distribution; standard deviations of position estimates at frequent locations were significantly smaller than at nonfrequent locations in staircase-contrast trials ($p = 0.008$, three-way within-subjects ANOVA), and the smallest deviations were exhibited at the most frequently presented

Gekas, Seitz, & Seriès

intermediate locations. However, across all data, there was no significant effect of absolute stimulus location on the standard deviation. These results suggest that the stimulus distribution had no effect on the direction of positional errors but had a significant effect on the consistency of position estimates.

Confidence levels over correct detections, mislocalizations, and false alarms

Subjects were more likely to detect a stimulus and more consistent in reporting the actual location of that stimulus when it was presented at a frequent location. We then asked whether subjects were also more confident when they detected a stimulus at a frequent location. We used the high-contrast trials as a benchmark for confidence levels reported by subjects when they had detected a stimulus. Figure 3C shows box plots (along with data points for each subject) of the differences in confidence level reported by subjects of the control group between high-contrast trials and trials in which they detected a stimulus (left), reported a mislocalization (center), and a false alarm (right). The median lines are significantly smaller than 0 for correct detections and mislocalizations (with 95% confidence) but not for false alarms. Moreover, confidence levels were significantly higher in false alarms than in mislocalizations. Interestingly, some subjects were more confident in their false alarms than their average confidence in successful high-contrast trials at the same stimulus location. It is worthwhile to remind the reader here that subjects did not receive any immediate feedback for their reports during the task. Thus, one could argue that subjects could have reported a stimulus as a response strategy when uncertain about the presence of a stimulus, and they would never be directly penalized for using such a strategy. However, we saw that false alarms were rare (Figure 2D). This, in conjunction with the high, on average, confidence, suggests that subjects may have felt quite certain when they chose to report a false alarm and provides some additional evidence toward the argument that subjects may have actually "perceived" these stimuli.

A similar behavior was exhibited by subjects of the bimodal group. Subjects' reported confidence for false alarms across locations and for correct detections at frequent locations were not significantly different from confidence for high-contrast correct detections. Subjects' median confidence in frequent locations was higher than in nonfrequent locations for correct detections and false alarms, but neither effect was significant. In contrast to subjects of the control group, we saw a much larger variation in the bimodal group with some subjects being very uncertain about their mislocalizations and false alarms. Overall, there was no strong effect of the stimulus distribution on subjects' confidence at detecting a stimulus.

Presented stimulus proximity influences subjects' mislocalizations

As seen in Figures 2 and 3, mislocalizations were more frequent than false alarms but with largely reduced confidence on average. That hints at a qualitative difference between mislocalizations and false alarms. We investigated whether we could distinguish between mislocalizations that represented genuine false alarms and those that represented just extreme errors in position estimates. Figure 4A shows the proportion of mislocalizations as a function of the distance from the nearest presented stimulus in the same trial. The distance ranges from 1 to 6. In the vast majority of mislocalizations across all data (≈80%), a stimulus was presented at a nearby location in the same trial (distance = 1). It is more likely that such responses correspond to large errors (localizations outside the 30° window that defines a correct detection) rather than to false alarms unrelated to the stimulus. If so, the presented stimulus would be scored as undetected. Indeed, only a small fraction of these stimuli were scored as detected (Figure 4B). This might explain the significantly reduced reported confidence of mislocalizations (Figure 3C). These data are consistent with the idea that mislocalizations are aptly termed and reflect cases in which stimuli were detected but the location of the stimuli were not well encoded or recalled.

Another way to visualize subjects' behavior in these mislocalizations is to plot a histogram of subjects' position estimates. Figure 4C shows histograms of subjects' (of the bimodal group) position estimates when a stimulus was presented at a nearby location (distance = 1, gray bars) and at locations further away (distance > 1, blue bars). Position estimates are grouped into 3° bins. Because the stimulus distribution is symmetrical (Figure 1C), it is possible to fold the 12 presented locations around the horizontal cardinal into six. For example, we can combine the results from the two frequent horizontal locations (locations 3 & 9 in Figure 1) and so on. In Figure 4C, 0° angle indicates the horizontal cardinal, positive angles indicate frequent locations, and negative angles indicate nonfrequent locations. When mislocalizations were at a distance of 1 from a stimulus, most position estimates were made at the boundaries between locations (vertical dotted black lines), and very few happened close to the stimulus locations (red and green vertical dashed lines). When mislocalizations were at a distance larger than 1 from an actual stimulus, position estimates were less frequent but more evenly distributed across the visual space. These larger mislocalizations are most consistent with
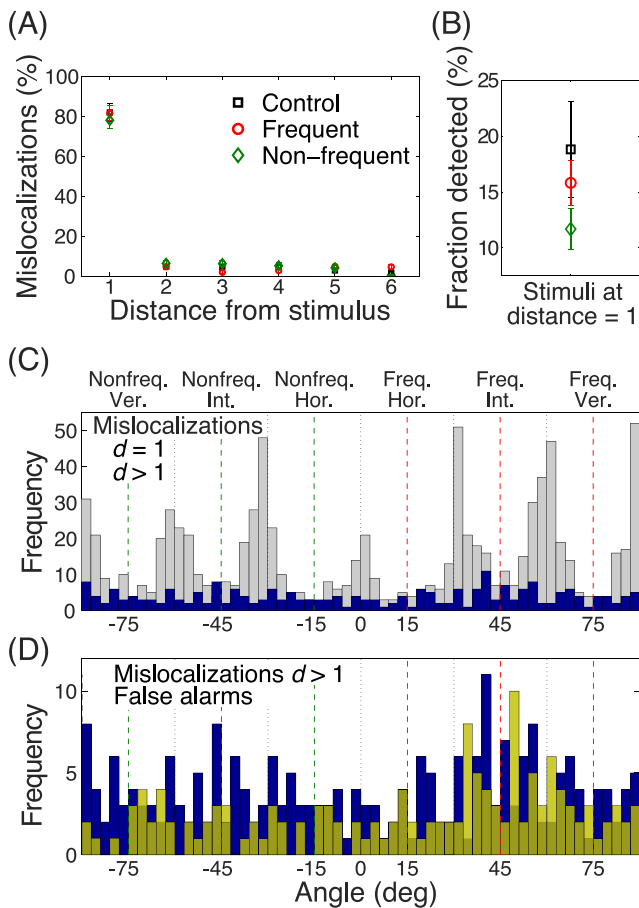
Figure 4. Effect of stimulus proximity on mislocalizations. (A) Proportions of subjects' mislocalizations as a function of the distance from the nearest presented stimulus in the same trial for the control group (black squares) and for the frequent (red circles) and nonfrequent (green diamonds) conditions of the bimodal group. (B) The fraction of correctly detected stimuli that were presented at a nearby location (distance = 1) to a mislocalization in the same trial. Results are averaged over all subjects, and error bars show within-subject standard error. (C) Histograms of bimodal group's position estimates when they made a mislocalization. Gray bars indicate mislocalizations with a distance of 1 from the nearest stimulus, and blue bars indicate a distance larger than 1. Vertical red and green dashed lines indicate frequent and nonfrequent locations, respectively, and vertical dotted black lines indicate the boundaries between locations. The 0° angle indicates the horizontal cardinal. (D) Histograms of bimodal group's position estimates when they made a mislocalization with a distance larger than 1 from the nearest stimulus (blue), and when they reported a false alarm (yellow).

false alarms than misjudgments of the stimulus position because it would have to be an error of at least 45°. However, we cannot say with complete certainty whether these responses represent very large errors or genuine false alarms. Overall, subjects' mislocalizations appear to be significantly affected by the proximity to a

presented stimulus and so are very difficult to correctly classify as errors in position estimates or false alarms. For that reason, we do not include mislocalizations in our following analysis regarding recency effects and our modeling of subjects' behavior in the task.

Figure 4D plots a zoomed-in version of the behavior in mislocalizations when distance > 1 along with subjects' position estimates of false alarms. For both sets of data, most position estimates were made around the area +45° away from the horizontal cardinal where the most frequent intermediate stimuli were presented. This effect was more pronounced for false alarms. We calculated the probability density function for the false alarms data using a kernel comprised of a normal distribution with an automatically computed optimal sigma (Silverman, 1986) adapted for circular data. The probability density function matches the stimulus distribution; subjects were 2.4 times more likely to report a false alarm at +45° (frequent intermediate location) than at −45° (nonfrequent intermediate), 1.4 times at +15° (frequent horizontal) than at −15° (nonfrequent horizontal), and 1.5 times at +75° (frequent vertical) than at −75° (nonfrequent vertical). This result agrees with our previous findings regarding false alarms matching the stimulus distribution (Chalk et al., 2010; Gekas et al., 2013) and, along with the high, on average, confidence levels reported on false alarms, suggests that subjects were certain of their reports of the presence and location of these false alarms.

Effect of stimulus presentation at the same location *n* trials back

So far, we showed that the stimulus distribution had a strong effect on subjects' detection performance, positional errors, and false alarms. Finally, we investigated the effect of location priming on subjects' behavior and whether there was an interaction with the stimulus distribution. Figure 5A shows subjects' detection performance as a function of whether a stimulus was presented at the same location *n* trials back for the control group and the frequent and nonfrequent conditions of the bimodal group. In this analysis, we did not divide locations based on their distance to the horizontal cardinal in order to have as many data points as possible. For the control group and the frequent condition of the bimodal group, there is a strong effect on detection performance for stimuli presented at the same location in the preceding trial, but the effect steadily weakens when exposure is further in the past. Overall, there was a significant effect of recent exposure to the same stimulus location for both control group and frequent condition ($p = 0.009$ and $p = 0.005$, respectively, one-way within-subjects ANOVA). However, we see an
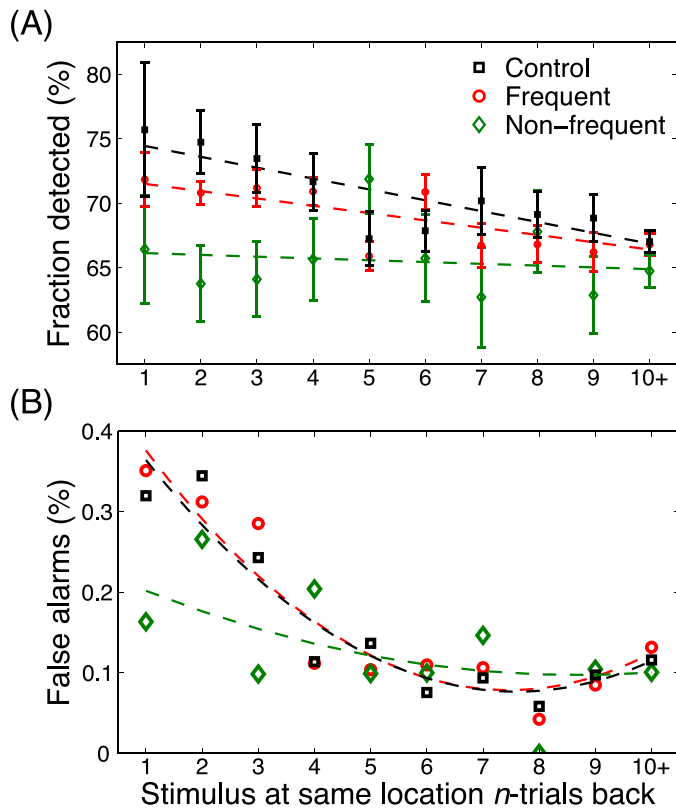
(A)

(B)

Figure 5. Effect of stimulus presentation at the same location n trials back. (A) The fraction of correctly detected stimuli as a function of whether a stimulus was presented at the same location n trials back for the control group (black square) and for the frequent (red circles) and nonfrequent (green diamonds) conditions of the bimodal group. Results are averaged over all subjects, and error bars show within-subject standard error. Dashed lines indicate the best-fitting linear functions. (B) Relative frequencies of subjects' false alarms in the absence of stimulus as a function of whether a stimulus was presented at the same location n trials back. Dashed curves indicate the best-fitting quadratic functions. Response probabilities are calculated out of the total number of trials in which subjects could make an incorrect response.

important difference in subjects' performance between the frequent and nonfrequent conditions. There was no significant effect of recent stimulus exposure to detection rates for nonfrequent locations ($p = 0.73$), and there was a significant effect of a location being frequent on the detection performance ($p = 0.0325$, three-way within-subjects ANOVA). With the exception of when $n = 5$ (i.e., a stimulus was presented at the same location in the fifth preceding trial), detection rates were consistently higher at frequent locations than at nonfrequent, and there was only a marginal improvement for the non-frequent locations even when a stimulus was presented at the same location in the preceding trial ($n = 1$). A possible explanation would be that the same stimulus being presented in two consequent trials is very rare for

the nonfrequent locations (Supplementary Figure 2, green bars) and that subjects might "learn" that a stimulus repeat at a nonfrequent location was highly unlikely. However, a stimulus repeat was equally unlikely for subjects of the control group (Supplementary Figure 2, black bars), but there was still a very strong effect of a stimulus repeat on detection performance for that group. Even when a stimulus was not presented at the same location in the last nine trials ($n = 10+$), which accounts for the majority of trials in each session, detection rates were higher at the frequent than at the nonfrequent locations although nonsignificantly ($p = 0.18$, signed rank test).

We next looked at the effect of location priming on false alarms. Figure 5B shows the relative frequencies that subjects would report a false alarm as a function of whether a stimulus was presented at the same location n trials back. Again, exposure to a stimulus in the most recent three trials had a strong effect only for the control group and the frequent condition of the bimodal group but not for the nonfrequent condition. It was only marginally more likely than average for subjects to report a false alarm at a nonfrequent location even after a stimulus was presented at the same location in the preceding trial. Stimulus exposure further in the past seemed to have very little effect for the control group or for either of the two conditions. The difference can be seen also by fitting the data to a quadratic function (dashed curves). It is interesting to note that the fits for the control and frequent data are very similar even though they represent the behavior of two different groups of subjects. In summary, location priming had a significant effect on subjects' detection performance and probability of reporting false alarms for all subjects regardless of the stimulus distribution. However, when the stimulus distribution was bimodal, this effect only extended to locations at which it was more likely that a stimulus would be presented.

Computational model

In previous work (Chalk et al., 2010; Gekas et al., 2013), we described subjects' performance in a motion estimation and detection task by using models that assumed subjects used a Bayesian strategy in which they combined a learned prior of the stimulus statistics with their sensory evidence in a probabilistic way. These models were shown to outperform models that assumed subjects developed response strategies unrelated to perceptual changes. Moreover, they successfully fit the experimental data and predicted subjects' behavior in trials in which no stimulus was presented but subjects reported a stimulus. Here, we describe a simple Bayesian model of the experimental task and

Bayesian model
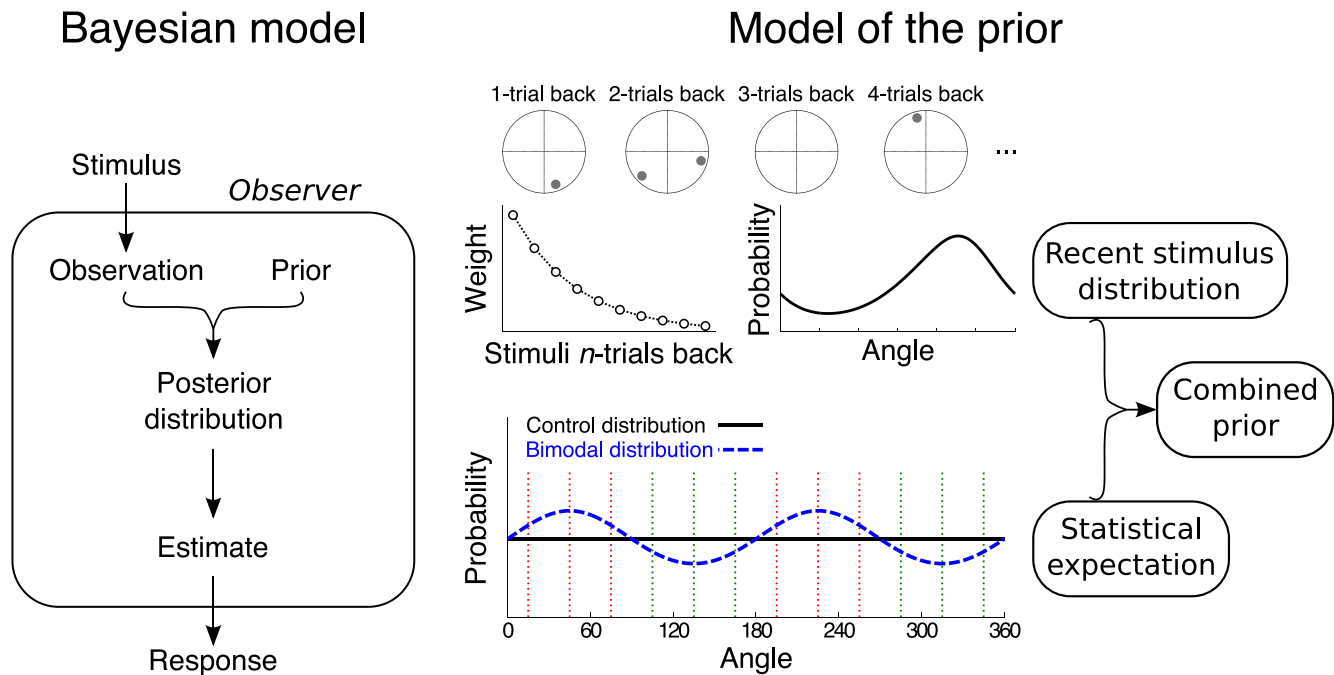
Model of the prior

Figure 6. (Bayesian model). The sensory evidence based on a noisy observation of the stimulus is combined with the prior to form the posterior distribution. A perceptual estimate is made by taking the mean of the posterior, and a response is made of the presence and location of the stimulus. (Model of the prior distribution). The *recent stimulus distribution* is a weighted sum of the stimuli the observer detected *n* trials back. The *statistical expectation* is an approximation of the true stimulus distribution implicitly learned after hundreds of trials. The prior distribution before each new trial is constructed by combining the recent stimulus distribution with the statistical expectation. When the expectation is uniform (e.g., for the control group), the prior distribution is just the recent stimulus distribution.

implement it in order to investigate the form of the prior distribution that would predict subjects' behavior in the experiment. In particular, we were interested in understanding why there were strong recency effects at frequently presented locations but weaker effects at nonfrequently presented locations. The model replicates the behavior of a suboptimal Bayesian observer performing the experimental task. The observer combines a learned prior of the stimulus statistics with her sensory evidence in a probabilistic manner. Trial-by-trial variability is driven by noise in the sensory likelihood (which, in turn, generates false alarms in the absence of stimulus), and recency effects are driven by a dynamically changing prior of recent stimulus history. We did not explicitly fit the model to the data; however, we did a systematic exploration of parameter space in order to find values under which the model approximated subjects' average performance in the task.

According to the model (Figure 6, *Bayesian model*), in each trial, the observer computes the posterior probability of a stimulus' presence at each of the 12 possible stimulus locations given the sensory input at all locations (*x*). The posterior probability $post_i(s|x_i)$ of a stimulus being present (*s*) at the ith location is the combination of the likelihood $likelihood_i(x_i|s)$ of the input given stimulus presence at the ith location with

the prior probability $prior_i(s)$ of a stimulus being present at the ith location, using Bayes' rule:

$$post_i(s|x_i) \propto likelihood_i(x_i|s) \cdot prior_i(s) \qquad (1)$$

In each trial, up to three stimuli *j* combine linearly to generate the sensory input received by the observer at each location (x_i). The probability of observing a stimulus at the ith location is calculated according to

$$likelihood_i(x_i|s) = c \cdot \sum_j V(x_j, \kappa_j) + \gamma_i, \qquad (2)$$

where *c* is the stimulus contrast, $V(x_j, \kappa_j)$ is a von Mises (circular normal) distribution centered on the stimulus location x_j and with width $1/\kappa_j$, and γ_i is a Gaussian noise variable. The term *c* has as a multiplicative effect on the likelihood of a stimulus being observed at a location and can take values ranging from 0, when no stimulus is presented, to 1, when a stimulus is presented with high contrast. We assume that the width $1/\kappa_j$ varies with the stimulus location so that it is narrower at locations closer to the horizontal cardinal and wider at locations further away. The variance of the noise term γ does not vary with absolute stimulus location.

The observer then makes perceptual estimates by comparing the posterior at each location with the posterior distribution $post_i(n|x_i)$ of the stimulus being

absent (n) given the observation. This is calculated similarly so that $post_i(n|x_i) \propto likelihood_i(x_i|n) \cdot prior_i(n)$. If the ratio of $post_i(s|x_i) / post_i(n|x_i)$ is greater than 1, the observer reports that a stimulus was present at the location, otherwise that it was absent. For simplicity, in the model, the observer detects a stimulus at the ith location when the posterior $post_i(s|x_i)$ is larger than a threshold level α; otherwise, the stimulus is not detected. The model follows the same staircase procedure in regard to the stimulus' contrast as in the experiment. If the observer successfully detects all presented stimuli in a trial, the stimulus' contrast is decreased whereas if the observer fails to detect any of the stimuli the contrast is increased. However, if the observer detects at least one of the stimuli (if more than one was presented) but not all, the contrast remains the same. The starting value for c is 0.5 and changes in steps of 0.005. When c is equal to 0, only random noise affects the observer's likelihood. This allows the model to generate false alarms in trials in which no stimulus is shown. For each of the possible 12 locations, if the posterior is larger than the threshold α, the observer reports a false alarm at that location.

We ran simulations of 16 "observers" presented with the same set of stimuli as each of the 16 experimental subjects of the bimodal group. The result for each observer was obtained after 1,000 simulations, and we averaged the results over all observers (16,000 simulations in total). The model requires five free parameters: the widths of the likelihood for each location in relation to its distance from the horizontal cardinal ($\kappa_{horizontal}$, $\kappa_{intermediate}$, $\kappa_{vertical}$), threshold α, and the variance of the Gaussian noise σ_{noise}. We adjusted the values of the free parameters to approximate the average detection performance of the subjects in the experiment (in regard to the performance gap between horizontal and vertical locations) as well as their frequency of false alarms in the no stimulus trials. We did not fit the free parameters to the experimental data; instead we used values that provided a good qualitative fit with subjects' average performance. We leave the fitting of individual subjects' performances to future work.

Using the computational model, we can compare the effect of different priors on the observer's behavior. We consider three different types of priors: a *recent stimulus distribution*, a *statistical expectation* of the stimulus distribution, and a combination of the two (Figure 6, *Model of the prior*). The recent stimulus distribution is a weighted sum of the stimuli j n trials back defined as

$$recent\_stimulus\_distribution = U + \sum_{j,n} w_n \cdot V(x_j, \kappa_j), \quad (3)$$

where U is a uniform prior over each location, and w_n is the weight given to stimuli of the n trial defined as $w_n = w_1 \cdot \exp[-\lambda(n-1)]$, with w_1 the weight of the stimuli one trial back and λ the rate of the weight's decrease over

time. Importantly, only stimuli detected by the observer are considered when calculating the recent stimulus distribution. For the simulations of the model, we defined $n = 10$, so that the observer has a memory of the last 10 trials.

The statistical expectation is defined as an approximation of the true stimulus distribution presented to the bimodal group formalized as the sum of two circular normal distributions centered on the most frequently presented locations (45° and 225°). The widths of the distributions can vary in order to manipulate the degree of the effect of the prior. Finally, the combined prior distribution is obtained by multiplying the recent stimulus distribution with the expectation. The two distributions are combined equally.

The different models of the prior make distinct predictions regarding recency effects for frequent and nonfrequent locations. Figure 7 shows three successive example trials of the task and the prior distribution before the presentation of a new stimulus for each trial. In Trial 1, a stimulus is presented at the same location as in the preceding trial at a nonfrequent location. The recent stimulus distribution (black solid line) is strongly biased toward that location (vertical green dotted line). However, because the location is nonfrequent, the combined prior (dashed orange line) is biased toward the nearby frequent locations and not as much on the presented primed nonfrequent location. Thus, under equal noise levels, an observer utilizing only the recent stimulus distribution as a prior is more likely to correctly detect and report the stimulus than an observer utilizing the combined prior. In Trial 3, a stimulus is presented at the same location as in the preceding trial but now at a frequent location. The recent stimulus distribution is biased toward that location (vertical red dotted line) but not very strongly because of the recent presentation of stimuli at other locations. However, the combined prior is strongly biased toward the location. Thus, an observer utilizing the combined prior is more likely now to correctly detect and report the stimulus than an observer utilizing just the recent stimulus distribution.

The combined model of the prior offers a parsimonious explanation for the recency effects on subjects' detection performance at frequent locations and also the weaker evidence of such effects at the nonfrequent locations. Additionally, it can provide an explanation for subjects' false alarms. In a trial in which no stimulus is presented, the posterior collapses to the prior distributions. If we imagine that no stimulus was presented in Trial 1 and the observer reports a false alarm, the observer is more likely to report it at a nearby frequent location and not at the primed nonfrequent location. In contrast, in Trial 3, the
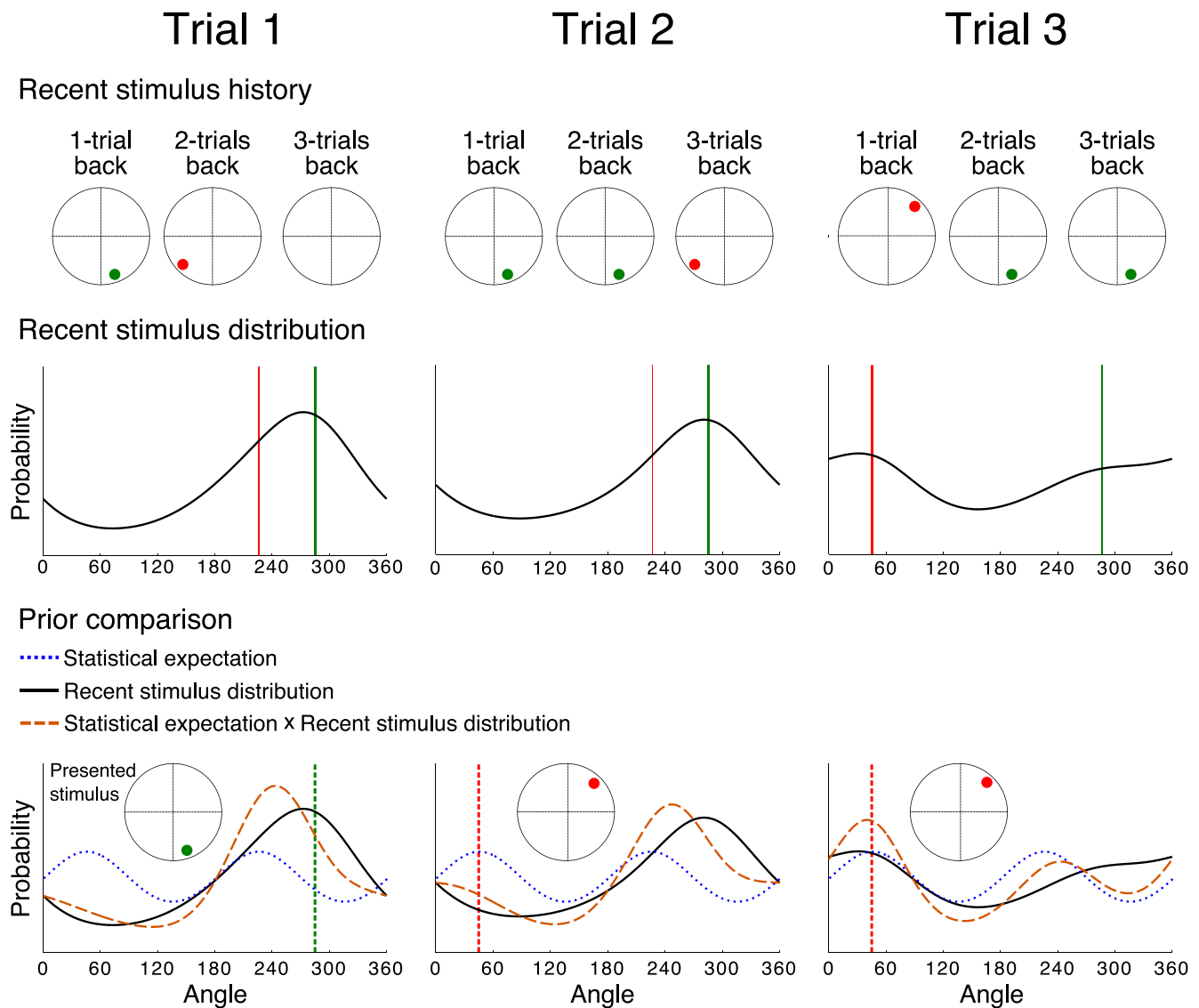
Figure 7. Example run of three successive trials. For each trial, we show the recent stimulus history and compare the different prior distributions before the new stimulus is presented. The recent stimulus history shows the stimuli presented in the last three trials, which are used to calculate the recent stimulus distribution (black solid line). The vertical solid lines indicate the exact recent stimuli angle of presentation, and the width of the line indicates the weight of the stimulus on the prior. The recent stimulus distribution is combined with the Bimodal expectation (blue dashed line) to form the combined prior distribution (orange dashed-dotted line). The vertical dashed colored lines indicate the exact angle of presentation. Red stimuli indicate frequent locations and green stimuli nonfrequent locations. When a frequent location is primed (Trial 3), the combined prior distribution is strongly skewed toward that location. When a nonfrequent location is primed (Trial 1), the combined prior is less affected as the peak of the distribution is still closer to the nearby frequent locations.

observer is even more likely to report a false alarm at the primed frequent location.

All simulations of the model use the same values for $\kappa_{horizontal}$, $\kappa_{intermediate}$, $\kappa_{vertical}$, threshold α, and noise variance σ_{noise}. Supplementary Figure 3 shows the behavior of the model using a flat prior: $prior_i (s) = b \cdot (1/12)$ for each location, where b is the fraction of trials in which a stimulus is presented ($b = 0.7$). As can be expected, in this case, there are no performance differences between frequent and nonfrequent locations

and no recency effects. We ran simulations using the three different priors. We set the parameters of the priors so that the overall number of false alarms was very similar between the three simulations and the experimental data. Figure 8 shows the results of the simulations along with the experimental data for detection performance and false alarms. We calculated the root mean square error (RMSE) between the linear and quadratic fits of the simulations and the experimental data. When the prior is limited to the statistical
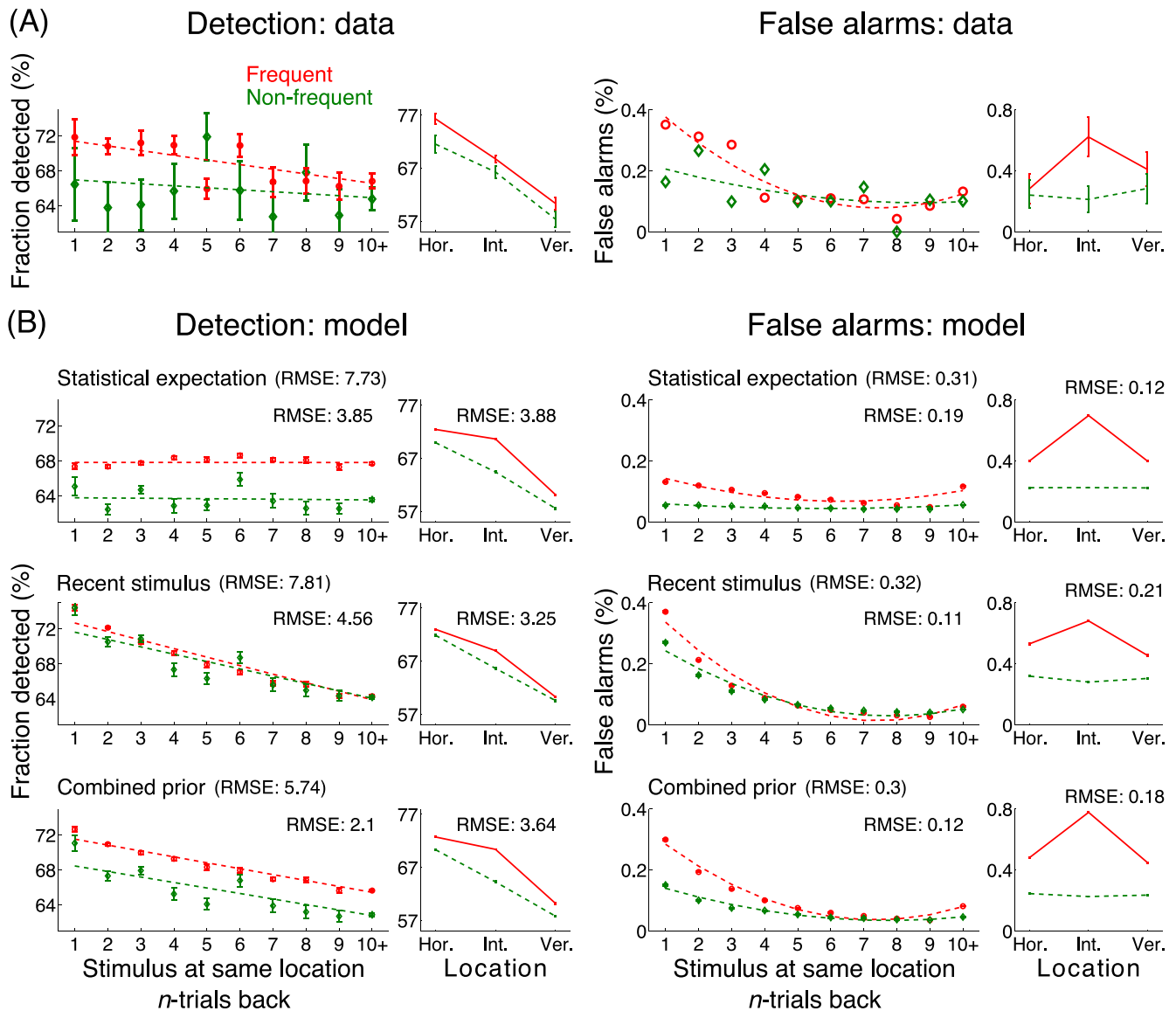
Figure 8. Comparison between experimental data and model. (A) Experimental data. (Left) The fraction of correctly detected stimuli as a function of whether a stimulus was presented at the same location *n* trials back and plotted against the presented location. Dashed lines indicate the best-fitting linear functions. (Right) Relative frequencies of subjects' false alarms as a function of whether a stimulus was presented at the same location *n* trials back and plotted against the presented location. Dashed curves indicate the best-fitting quadratic functions. (B) Simulations of 16 observers presented with the same stimuli as the experimental subjects using three distinct priors: "statistical expectation" of the stimulus distribution, "recent stimulus distribution," and the combined distribution of the former two. Insets show the RMSE of each simulation with the experimental data. For detection, we calculated the percentage of correctly detected stimuli over the total number of stimulus presentations and for false alarms the percentage of false alarms over the total number of no-stimulus trials. The simulations of the combined prior more successfully matched the experimental data than the simulations of the other prior distributions.

expectation, the model does not reproduce any recency effects, resulting in an overall RMSE of 7.73 in the percentage of stimuli correctly detected over the total number of stimulus presentations and 0.31 in the percentage of false alarms over the total number of no-stimulus trials. On the other hand, when the prior is limited to the recent stimulus distribution, the model predicts strong recency effects but for both conditions

resulting in a larger overall error in both detection (7.81%) and false alarms (0.32%) in comparison to the statistical expectation. The simulations using the combined prior are the closest to the experimental results with the smallest overall error (5.74% and 0.3%). Although recency effects are observed in both conditions, they are stronger at frequent locations than at nonfrequent locations. Further, detection performance

is different between the two conditions even when a stimulus was presented at the same location 10 or more trials in the past.

The Bayesian model presented here can successfully, albeit simplistically, describe our experimental results. However, it is important to note that the model does not correspond to an optimal Bayesian model of the task. For example, it does not take into account the statistics of the number of stimuli presented in each trial or the motor noise in the localization task. A more complete model would have to take into account the aforementioned issues as well as the mislocalizations of presented stimuli, how multiple stimuli presented in the same trial combine (e.g., a nonlinear integration of stimuli information), and the preexisting biases we observed in subjects' localization performance at different absolute stimulus locations. The implementation of such a model could help explain certain deviations between model and data, such as the prediction of a large detection performance gap between frequent and nonfrequent intermediate locations, which was not observed experimentally. That model will be the focus of future work.

Discussion

Both perceptual priming and statistically driven expectations have been shown to have a strong influence on visual perception. In the current study, we investigated their interaction in a visual search task. Our results showed that both priming and expectations had a significant effect on visual perception by facilitating detection performance and by inducing more false alarms in the absence of stimulus. However, recency effects were subdued or even nonexistent at locations at which it was less likely that a stimulus would be presented. We also found that subjects' detection and localization performance were significantly affected by absolute stimulus location and that statistically driven expectations had a strong effect on subjects' localization consistency and the probability distribution of false alarms.

Cardinal effects on performance

Visual search performance has been shown to vary across the visual field even at equal eccentricities. A horizontal–vertical anisotropy in which performance is better on the horizontal than the vertical meridian is well documented (Carrasco, Evert, Chang, & Katz, 1995; Rijsdijk, Kroon, & Van der Wildt, 1980) as is a vertical asymmetry in which performance is better in the lower than the upper visual field (Edgar & Smith,

1990; Rubin, Nakayama, & Shapley, 1996). The horizontal–vertical anisotropy has also been shown to lead to more saccades to the upper and lower visual fields during visual search (Najemnik & Geisler, 2008). Physiological studies in human and nonhuman primates have found that along the vertical meridian of the retina there are lower densities of ganglion cells (Perry & Cowey, 1985) and cones (Curcio, Sloan, Packer, Hendrickson, & Kalina, 1987) than along the horizontal meridian, and similar asymmetries have been found in the lateral geniculate nucleus (Connolly & Van Essen, 1984) and V1 (Tootell, Switkes, Silverman, & Hamilton, 1988) of macaque monkeys.

Carrasco, Talgar, and Cameron (2001) investigated whether covert attention affects these performance asymmetries in discrimination, detection, and localization tasks and found that attentional manipulations did not have an effect on performance asymmetries. Our results agree with these findings. Subjects showed significantly better detection performance at locations closer to the horizontal cardinal and increasingly worse away from it. This performance gap was unaffected by the stimulus distribution; even though detection performance was, on average, better at frequent locations, performance at frequent vertical locations was still worse than performance at nonfrequent intermediate locations and so on. This suggest that subjects preexisting horizontal–vertical anisotropy, which may be considered a structural expectation, provides a strong constraint on subjects' performance that was minimally impacted by our relatively brief intervention.

Although there was a strong effect of the absolute stimulus location on detection performance, the same was not observed for accuracy in localization performance. Subjects' localizations at horizontal locations were not significantly more accurate than localizations at other locations. However, subjects exhibited systematic positional error biases away from the horizontal cardinal and toward locations between the intermediate and vertical locations (between 45° and 75° away from the horizontal cardinal), and these biases appear to be unaffected by the stimulus distribution. We have observed similar biases toward oblique (45° away from the cardinals) locations in a previous statistical learning experiment (unpublished) in which low-contrast coherent motion stimuli were shown at multiple motion directions, which were not restricted to a part of the visual field but encompassed the whole circular annulus. We found that subjects' estimates of the presented stimuli motion directions were strongly biased toward the oblique directions and that this bias seemed to mask the possible influence of the stimulus distribution on estimation behavior.

These findings relate to the ongoing discussion on whether structural expectations match the statistics of the environment and whether they are continuously

updated over the observer's lifetime. Recent studies have successfully managed to measure observers' biases of visual stimuli and compare it to the environment's statistics. For example, Girshick, Landy, and Simoncelli (2011) investigated subjects' performance on comparing different orientations of uncertain stimuli and found that it was strongly biased toward the cardinal axes. These biases were shown to match the distribution of local orientations in a data set of photographs. So, if these biases are learned over very long-term exposure, can they be quickly updated in an experimental task? In Sotiropoulos et al. (2011), we showed that the structural prior on slow speeds of moving stimuli is not fixed and can change through experimental training. Interestingly, this change occurred inside the experimental session but also carried over incompletely between different sessions. Our experimental findings in the current task suggest that there are structural expectations that affect subjects' perception of the stimulus location, which were largely unaffected by exposure to the stimulus distribution during our task. The exact nature of these expectations is unknown, and further work will be needed to identify their origin. Further, it would be interesting to investigate whether they can be modulated by a longer-lasting perceptual learning experiment. An alternate explanation for this behavior unrelated to perceptual priors could be that subjects used the physical structure of the monitor as a reference with which they could improve their localization performance. Although we cannot rule out this possibility, we suspect it unlikely given that the distance of stimulus presentation to the monitor edges was large (12.5° visual angle to the top and 21.6° to the sides of the monitor).

Developing expectations of different timescales

A growing body of work shows that expectations can be quickly developed in experimental settings (e.g., Adams et al., 2004). In previous work (Chalk et al., 2010; Gekas et al., 2013), we found that, after a few minutes of presenting low-contrast coherent moving stimuli to subjects, they perceived new stimuli as moving in directions closer to the most frequently presented directions than they actually were. Additionally, we found that subjects were more consistent in their estimations at the most frequently presented directions, that they were better at detecting stimuli that were moving in these directions, and that they were more likely to report motion in these directions in trials in which no stimulus was presented but they reported seeing a stimulus. In the current study, we found that the stimulus distribution had the same effects on subjects' behavior (with the exception of positional

error biases) inside of a similar time frame (around 5 to 8 min of stimulus presentation).

Interestingly, we found that perceptual priming had similar effects as the manipulated stimulus distribution on the behavior of subjects of the control group for which the stimulus distribution was uniform. In our task, priming significantly facilitated detection performance and induced significantly more false alarms. Notably, all forms of priming are not the same, and we do not suggest that all forms of priming involve the same underlying mechanisms; priming in visual search has been shown for many different forms of stimulus characteristics from easy pop-out search tasks to more difficult conjunction search tasks, and importantly, it has been shown that the level of priming can be affected by the stimuli used to test the priming effects (e.g., McBride, Leonards, & Gilchrist, 2009). In our study, priming seems to act as a form of very short-term expectation that changes dynamically over time. When the average statistics of the stimuli were uniform, statistical regularities over a few recent trials induced identical results as statistical regularities that would need much longer exposure to be learned by subjects. This suggests that the updating process of expectations works continuously from the very short timescale of a few trials to the medium timescale of an experimental session and the long timescales of structural expectations formed over an observer's lifetime.

However, statistical regularities over the last few trials still affected subject's performance even when the average statistics of the stimuli were bimodal. Thus, expectations of different timescales appear to interact depending on the properties of the task and the environment. In our results, priming, as a form of short-term expectation, interacts with the longer-term expectation formed by the group of subjects that were presented with the bimodal stimulus distribution, which, in turn, interacts with the longer-term expectation of a horizontal–vertical anisotropy. Although the role of expectations and their effect on perception are increasingly being studied, the interactions of different types of expectations have earned less attention from the scientific community. A synergistic effect between spatial and temporal expectations was observed by Doherty, Rao, Mesulam, and Nobre (2005) in an EEG experiment, and this effect was recently shown to enhance visual discrimination (Rohenkohl, Gould, Pessoa, & Nobre, 2014); temporal expectations significantly increased the effectiveness of spatial expectations, but they did not facilitate performance at locations that were unattended. Likewise, Kingstone (1992) found a synergistic interaction between location and form expectations. However, response times were lower when a cued form appeared at an uncued location or when an uncued form appeared at a cued location. Our findings are in agreement with these

studies. If we consider priming as a form of short-term expectation, it has a positive interaction with the longer-term expectation of the stimulus distribution when the primed location is a more probable location but a neutral interaction when the primed location is a less probable location. It is also worthwhile to note that priming does not have a negative interaction in that it does not divert attention from a more probable location toward a less probable location. In contrast to counterpredictive external cues in Vincent (2011), priming did not reflexively draw attention to the primed nonfrequent location. Otherwise, we would expect to see increased detection performance and more false alarms at primed nonfrequent locations in comparison to nonprimed nonfrequent locations.

Mechanisms of developing expectations

The timescale over which expectations are developed and/or updated is not yet clearly understood. In sensorimotor tasks, for example, it has been shown that observers combine prior knowledge obtained over short timescales and uncertain sensory information in a near-optimal way (Körding & Wolpert, 2004; Tassinari, Hudson, & Landy, 2006). However, some studies have suggested that this process might be suboptimal (Eckstein et al., 2004; Raviv, Ahissar, & Loewenstein, 2012). Raviv et al. (2012) showed that subjects exhibited biases in a two-tones discrimination task that matched the stimulus distribution. These biases were strongly skewed toward the most recent trials and deviated from biases of an optimal Bayesian observer. They suggested that subjects did not learn a close approximation of the true stimulus distribution and that their behavior could be better described by an "implicit memory" model in which the representation of past stimuli is a continuously updating single scalar. An interpretation of this finding is that subjects assumed the statistics of the stimuli in the experiment were highly volatile and only the very recent stimulus history was informative. However, the task was purposely brief, consisting of tens of trials, so it is possible that subjects could not form a complete picture of the prior in that time frame. In our study, on the other hand, subjects had abundant exposure to the statistics of the stimuli, and we found that subjects learned the (static) bimodal statistics of the stimuli while also being affected by the recent stimulus history for the frequent locations. Compared to the study of Raviv et al., this suggests that which aspects of the statistics of the stimuli are learned and used could depend on specific task properties (e.g., time frame of exposure, complexity, initial instructions, etc). Complex recency effects can arise even in simple two-alternative forced choice (2AFC) tasks. For example, in a speeded 2AFC task, Jones, Curran, Mozer, and Wilder (2013) define as a "first-degree recency effect" the reduction in response time due to a physical match between a current stimulus and past stimuli in recent trials. They also define as a "second-degree recency effect" the reduction in response time in a repetition trial (i.e., the current trial matches the previous one) when recent trials were repetitions and in an alternation trial (i.e., the current trial mismatches the previous one) when recent trials were alternations. They proposed two simple learning mechanisms that can explain these recency effects: learning the base rate, which is the proportion of trials in which each stimulus occurs, and learning the repetition rate, which is the proportion of trials that repeat the previous trial. However, it is the interaction of these two mechanisms that can explain specific phenomena, such as the alternation advantage that was not predicted by previous models of sequential effects (Wilder, Jones, & Mozer, 2010; Yu & Cohen, 2008). More work will be needed to identify how these learning processes differ in different task situations and whether such differences could be explained in terms of the system trying to optimize its task performance.

A number of questions regarding the underlying mechanisms of expectations remain open. For example, how expectations developed over short (minutes) or medium (hours) time frames persist over time and, eventually, become structural expectations. Contextual expectations can persist for long periods (e.g., Sotiropoulos et al., 2011) or even transfer to different tasks (e.g., Turk-Browne & Scholl, 2009), so the same mechanisms that are responsible for the formation of these short-term expectations should, to some extent, be used for the formation of long-term structural expectations. Another important question is how expectations are encoded in the properties of populations of sensory neurons. Although Bayesian models provide us with mechanisms to successfully describe behavioral performance, they usually fail to be predictive at the neural level (Colombo & Seriès, 2012; O'Reilly, Jbabdi, & Behrens, 2012). Unfortunately, it is still unknown how probability distributions are neurally implemented, and it is generally difficult to propose experimental setups that would distinguish between different models of neurally plausible probabilistic inference. Nonetheless, our findings and modeling work suggest some of the constraints that such models should adhere to. We show that a prior of the very recent stimulus history is constantly updated and interacts directly with priors formed further in the past in a synergistic way. Moreover, we show that priors can dynamically change over very short timescales (seconds to a few minutes) whereas the formation of a longer-term prior requires at least 5 to 10 min of stimulus exposure. Similar timescales were reported by Chopin and Mamassian (2012) in a visual adaptation task.

They showed that visual adaptation could lead to negative correlation of the current percept with visual events presented recently (up to 3 min) and a positive correlation with a reference window of stimuli further into the past (5 to 10 min). This result seems to contradict our findings at first glance. However, we should note that the negative correlation arises after repeated presentation of the same stimulus. It was unlikely that we would observe negative correlations in our experiment as the same stimulus was never presented at the same location for more than two consecutive trials.

In the current study, we described a Bayesian model in which the prior is a combination of a continuously updating distribution of the recently presented stimuli and an expectation of the average stimuli statistics developed over hundreds of trials. We believe this model of the prior distribution offers a parsimonious explanation for our experimental results. However, the model is unable to predict whether the final prior is indeed a combination of two separate processes of different timescales or one single process extended over time. For example, we could implement a prior that is a sum of all presented stimuli (going back as far as the start of the experimental session) in which the importance of the last few presented stimuli is overvalued. The behavioral effects of such a prior would be identical to the effects of the combined prior we described. How could we distinguish between the two alternatives? More physiology and imaging studies investigating the neural loci of expectations could help us answer that question.

Finally, the current study replicates our previous findings that statistically driven expectations can induce increased false alarms in the absence of stimulus and that the probability distribution of these false alarms matches the distribution of the presented stimuli. Fiser et al. (2010) argued the interesting notion that, in the absence of sensory inputs, the prior distribution might be reflected on the spontaneous activity of neurons. This notion accounts for the observed similarity between spontaneous activity and evoked activity. For example, Berkes, Orbán, Lengyel, and Fiser (2011) found that the spontaneous activity of awake ferrets in the primary visual cortex at different stages of development is similar to the averaged evoked activity and that this similarity increased with age and was specific to responses evoked by natural scenes. Moreover, it has been found that spontaneous activity is sufficient to evoke firing in some cells without sensory input (Tsodyks, Kenet, Grinvald, & Arieli, 1999). We believe the link between spontaneous activity and the prior distribution is a promising direction for future research as are alternative approaches, such as the top-down modulation of sensory signals or the shift in the selectivity of neurons. However, more theoretical and experimental work is needed to answer the outstanding questions regarding the potential neurobiological mechanisms of priors.

In conclusion, our results show that human observers are able to probabilistically combine their noisy observations with a learned expectation of likely stimulus locations. Furthermore, learned expectations over a large number of trials are combined with recent exposure to a stimulus, which facilitates correct detection of stimuli but at the cost of increased false detections (mislocalizations and false alarms). Our work suggests that prior expectations may develop simultaneously over different timescales, potentially through multiple mechanisms, and interact synergistically depending on the demands of the behavioral task. These findings may help in the effort of understanding how probabilistic inference could be implemented in the cortex.

Keywords: expectation, perceptual priming, psychophysics, Bayesian

Acknowledgments

References

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057–1058.

Becker, S. I. (2008). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research, 48*(5), 664–684.

Berkes, P., Orbán, G., Lengyel, M., & Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of

an optimal internal model of the environment. *Science, 331*(6013), 83–87.

Brainard, D. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433–436.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics, 57*(8), 1241–1261.

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial Vision, 15*(1), 61–75.

Chalk, M., Seitz, A., & Seriès, P. (2010). Rapidly learned stimulus expectations alter perception of motion. *Journal of Vision, 10*(8):2, 1–18, doi:10.1167/10.8.2. [PubMed] [Article]

Chopin, A., & Mamassian, P. (2012). Predictive properties of visual adaptation. *Current Biology, 22*(7), 622–626.

Colombo, M., & Seriès, P. (2012). Bayes in the brain-on Bayesian modelling in neuroscience. *The British Journal for the Philosophy of Science, 63*(3), 697–723.

Connolly, M., & Van Essen, D. (1984). The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology, 226*(4), 544–564.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201–215.

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science, 236*(4801), 579–582.

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *The Journal of Neuroscience, 25*(36), 8259–8266.

Downing, C. J. (1988). Expectancy and visual-spatial attention: Effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance, 14*(2), 188.

Droll, J. A., Abbey, C. K., & Eckstein, M. P. (2009). Learning cue validity through performance feedback. *Journal of Vision, 9*(2):18, 1–22, doi:10.1167/9.2.18. [PubMed] [Article]

Druker, M., & Anderson, B. (2010). Spatial probability aids visual stimulus discrimination. *Frontiers in Human Neuroscience, 4*, 63.

Eckstein, M. P., Abbey, C. K., Pham, B. T., &

Shimozaki, S. S. (2004). Perceptual learning through optimization of attentional weighting: Human versus optimal Bayesian learner. *Journal of Vision, 4*(12):3, 1006–1019, doi:10.1167/4.12.3. [PubMed] [Article]

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science, 17*(11), 973–980.

Eckstein, M. P., Peterson, M. F., Pham, B. T., & Droll, J. A. (2009). Statistical decision theory to relate neurons to behavior in the study of covert visual attention. *Vision Research, 49*(10), 1097–1128.

Edgar, G. K., & Smith, A. T. (1990). Hemifield differences in perceived spatial frequency. *Perception, 19*(6), 759–766.

Elazary, L., & Itti, L. (2010). A Bayesian model for efficient visual search and recognition. *Vision Research, 50*(14), 1338–1352.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*(5), 583.

Fecteau, J. H. (2007). Priming of pop-out depends upon the current goals of observers. *Journal of Vision, 7*(6):1, 1–11, doi:10.1167/7.6.1. [PubMed] [Article]

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: From behavior to neural representations. *Trends in Cognitive Sciences, 14*(3), 119–130.

Gekas, N., Chalk, M., Seitz, A. R., & Seriès, P. (2013). Complexity and specificity of experimentally induced expectations in motion perception. *Journal of Vision, 13*(4):8, 1–18, doi:10.1167/13.4.8. [PubMed] [Article]

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics, 67*(7), 1252–1268.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14*(7), 926–932.

Haijiang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences, USA, 103*(2), 483–488.

Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review, 120*(3), 628.

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology, 44*(1), 69–104.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *The Journal of Neuroscience, 33*(41), 16275–16284.

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427*(6971), 244–247.

Kristjánsson, Á. (2009). Independent and additive repetition priming of motion direction and color in visual search. *Psychological Research, 73*(2), 158–166.

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics, 72*(1), 5–18.

Le Dantec, C. C., & Seitz, A. R. (2012). High resolution, high capacity, spatial specificity in perceptual learning. *Frontiers in Psychology, 3*, 222.

Ma, W. J., Navalpakkam, V., Beck, J. M., Van Den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience, 14*(6), 783–790.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672.

Maljkovic, V., & Nakayama, K. (1996). Priming of pop-out: II. The role of location. *Perception & Psychophysics, 58*(7), 977–991.

McBride, J., Leonards, U., & Gilchrist, I. D. (2009). Flexible target representations underlie repetition priming in visual search. *Visual Cognition, 17*(5), 655–678.

Miller, J. (1988). Components of the location probability effect in visual search tasks. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 453.

Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision, 8*(3):4, 1–14, doi; 10.1167/8.3.4. [PubMed] [Article]

Olivers, C. N., & Meeter, M. (2006). On the dissociation between compound and present/absent tasks in visual search: Intertrial priming is ambiguity driven. *Visual Cognition, 13*(1), 1–28.

O'Reilly, J. X., Jbabdi, S., & Behrens, T. E. (2012). How can a Bayesian approach inform neuroscience? *European Journal of Neuroscience, 35*(7), 1169–1179.

Pelli, D. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Perry, V. H., & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: Implications for central magnification factors. *Vision Research, 25*(12), 1795–1810.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Raviv, O., Ahissar, M., & Loewenstein, Y. (2012). How recent history affects perception: The normative approach and its heuristic approximation. *PLoS Computational Biology, 8*(10), e1002731.

Rijsdijk, J. P., Kroon, J. N., & Van der Wildt, G. J. (1980). Contrast sensitivity as a function of location on the retina. *Vision Research, 20*(3), 235–241.

Rohenkohl, G., Gould, I. C., Pessoa, J., & Nobre, A. C. (2014). Combining spatial and temporal expectations to improve visual perception. *Journal of Vision, 14*(4):8, 1–13, doi:10.1167/14.4.8. [PubMed] [Article]

Rubin, N., Nakayama, K., & Shapley, R. (1996). Enhanced perception of illusory contours in the lower versus upper visual hemifields. *Science, 271*(5249), 651–653.

Seitz, A. R., Nanez, J. E., Holloway, S. R., Koyama, S., & Watanabe, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences, USA, 102*(25), 9080–9085.

Seriès, P., & Seitz, A. R. (2013). Learning what to expect (in visual perception). *Frontiers in Human Neuroscience, 7*, 668.

Sigurdardottir, H. M., Kristjánsson, Á., & Driver, J. (2008). Repetition streaks increase perceptual sensitivity in visual search of brief displays. *Visual Cognition, 16*(5), 643–658.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Sotiropoulos, G., Seitz, A. R., & Seriès, P. (2011). Changing expectations about speed alters perceived motion direction. *Current Biology, 21*(21), R883–R884.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137–149.

Sterzer, P., Frith, C., & Petrovic, P. (2008). Believing is seeing: Expectations alter visual awareness. *Current Biology, 18*(16), R697–R698.

Stocker, A., & Simoncelli, E. (2006). Noise character-

istics and prior expectations in human visual speed perception. *Nature Neuroscience, 9*(4), 578–585.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research, 32*(8), 1535–1549.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13*(9), 403–409.

Tassinari, H., Hudson, T. E., & Landy, M. S. (2006). Combining priors and noisy visual cues in a rapid pointing task. *The Journal of Neuroscience, 26*(40), 10154–10163.

Tootell, R. B., Switkes, E., Silverman, M. S., & Hamilton, S. L. (1988). Functional anatomy of macaque striate cortex. II. Retinotopic organization. *The Journal of Neuroscience, 8*(5), 1531–1568.

Tsodyks, M., Kenet, T., Grinvald, A., & Arieli, A. (1999). Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science, 286*(5446), 1943–1946.

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 195.

Verghese, P. (2001). Visual search and attention: A signal detection theory approach. *Neuron, 31*(4), 523–535.

Vincent, B. (2011). Covert visual search: Prior beliefs are optimally combined with sensory evidence. *Journal of Vision, 11*(13):25, 1–15, doi:10.1167/11.13.25. [PubMed] [Article].

Wilder, M., Jones, M., & Mozer, M. C. (2010). Sequential effects reflect parallel learning of multiple environmental regularities. In *Advances in neural information processing systems, 22* (p. 2053–2061). La Jolla, CA: NIPS Foundation.

Yu, A., & Cohen, J. (2008). Sequential effects: Superstition or rational behavior? In *Advances in neural information processing systems, 21* (pp. 1873–1880). La Jolla, CA: NIPS Foundation.