Near-optimal integration of orientation information across saccades

Elad Ganmor

Center for Neural Science, New York University, New York, NY, USA

Michael S. Landy

Center for Neural Science, New York University, New York, NY, USA
Department of Psychology, New York University, New York, NY, USA

Eero P. Simoncelli

Center for Neural Science, New York University, New York, NY, USA
Courant Institute of Mathematical Sciences, New York University, New York, NY, USA
Department of Psychology, New York University, New York, NY, USA
Howard Hughes Medical Institute, New York University, New York, NY, USA

We perceive a stable environment despite the fact that visual information is essentially acquired in a sequence of snapshots separated by saccadic eye movements. The resolution of these snapshots varies—high in the fovea and lower in the periphery—and thus the formation of a stable percept presumably relies on the fusion of information acquired at different resolutions. To test if, and to what extent, foveal and peripheral information are integrated, we examined human orientation-discrimination performance across saccadic eye movements. We found that humans perform best when an oriented target is visible both before (peripherally) and after a saccade (foveally), suggesting that humans integrate the two views. Integration relied on eye movements, as we found no evidence of integration when the target was artificially moved during stationary viewing. Perturbation analysis revealed that humans combine the two views using a weighted sum, with weights assigned based on the relative precision of foveal and peripheral representations, as predicted by ideal observer models. However, our subjects displayed a systematic overweighting of the fovea, relative to the ideal observer, indicating that human integration across saccades is slightly suboptimal.

Introduction

The human retina is highly nonuniform. In the fovea, cones are packed tightly together, but in the periphery they are increasingly sparse, separated by a sea of rods. As a result, the quality of photopic visual input varies substantially across the retina. To obtain a more spatially uniform representation of a scene, humans move their eyes several times per second, each time acquiring high-resolution information at the center of gaze, and lower-resolution information elsewhere.

What happens to the low-resolution information obtained from a peripheral location moments before a saccade to that location? Does the visual system simply discard it, replacing it with the more accurate foveal information acquired after the saccade, or does it somehow combine these qualitatively different sources? On one hand, it would seem sensible to use all available information, since this can only improve visual capabilities. Many studies of sensory cue integration demonstrate that the human visual system is quite good at integrating information from different sources (Ernst & Banks, 2002; Young, Landy, & Maloney, 1993). On the other hand, retaining the information obtained in the periphery and correctly combining it with foveal

information obtained after the saccade is likely to be costly, in terms of memory storage, wiring, and computation (Bullmore & Sporns, 2012; Laughlin, 2001). Given the large discrepancy in the precision of representation in fovea and periphery, saccadic integration might not be worth the cost.

Several studies have demonstrated that humans are able to combine peripheral and foveal information. For example, reading speed is increased when at least part of the upcoming word in a text is visible in the periphery, even though the words cannot actually be deciphered in the periphery (McConkie & Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982). However, other studies found that integration across eye movements is incomplete (Bridgeman, Van der Heijden, & Velichkovsky, 1994; Irwin, Zacks, & Brown, 1990), a fact most directly evident in the phenomenon of "change blindness" (Bridgeman, Hendry, & Stark, 1975; Grimes, 1996). Though these studies provide compelling evidence that not all information obtained in the periphery is discarded, the amount and type of information retained (e.g., the set of visual features or attributes) is not known.

Here, we develop an experimental task and analysis to examine whether humans integrate orientation information across saccadic eye movements. Specifically, observers made judgments about the orientation of a localized target that was seen both before and after a saccade, and their performance on this task was used to determine whether and how they combined the two pieces of information in forming their percept. To maximize the potential benefits of trans-saccadic integration, we altered the contrast during the saccade, artificially matching the amount of information provided by the foveal and peripheral views. Using analysis paradigms derived from the cue-integration literature, we find that humans do indeed integrate orientation information across saccadic eye movements. When compared to a statistically optimal observer, subjects were close to optimal, but exhibited a systematic tendency to rely more on the foveal than the peripheral information. Finally, we examined the role of active saccades in the integration process by repeating the experiment with "replayed" artificial saccades, and found no evidence of integration in the absence of eye movements. An early version of these results was presented in abstract form (Ganmor, Landy, & Simoncelli, 2015).

Methods

Subjects

Thirty-seven subjects participated in the experiments (11 in Experiment 1, four in Experiment 2, four in

Experiment 3, and 18 in Experiment 4). All subjects were naïve as to the purposes of the experiments and were paid for participation (except one author, EG). Subjects were given written instructions before the outset of the experiment. Experimental procedures were approved by the human subjects committee of New York University.

Apparatus

Eye movements were monitored using an Eyelink 1000 (SR Research) eye tracker, operating at 500Hz. The eye tracker was calibrated at the beginning of each experimental session, and drift correction was applied every 50 trials. Stimuli were presented on a Multisync FP214SB (NEC) CRT monitor at a distance of 43 cm and a refresh rate of 85 Hz. Software was written in Matlab using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007).

Procedure

Experiment 1

The basic structure of each trial was as follows (Figure 1A): (a) A fixation point appeared at the center of the screen. (b) After subjects achieved fixation (gaze within 1° of fixation dot), a Gabor patch appeared randomly to the left or right of fixation (0.5 cycle/°, Gaussian envelope with SD 1° and eccentricity 5°). Subjects were required to maintain fixation for 100 ms. (c) The fixation point jumped to the peripheral target. This served as a cue for the subject to saccade to the target. Subjects were required to initiate the saccade within 400 ms. (d) After the saccade was detected (once the gaze crossed a threshold of 1° horizontally in the direction of the target), the target remained on the display for 200 ms. (e) Subjects reported by key press whether the target was oriented clockwise or counter-clockwise of vertical. The intertrial interval was chosen randomly from the interval 0.5–1.0 s.

If subjects failed to maintain fixation or failed to initiate a saccade within 400 ms, the trial was aborted and the subject was notified of the reason. In addition, trials in which the postsaccadic eye position was more than 2° away from the center of the target or in which a saccade was detected less than 100 ms after the saccade cue were discarded in the analysis (Figure 1B through C).

There were three trial types. In "Both" trials the target was visible both before and after the saccade. In periphery-only trials the contrast of the target was set to zero once a saccade was detected. Thus, the target was visible only in the periphery. In fovea-only trials
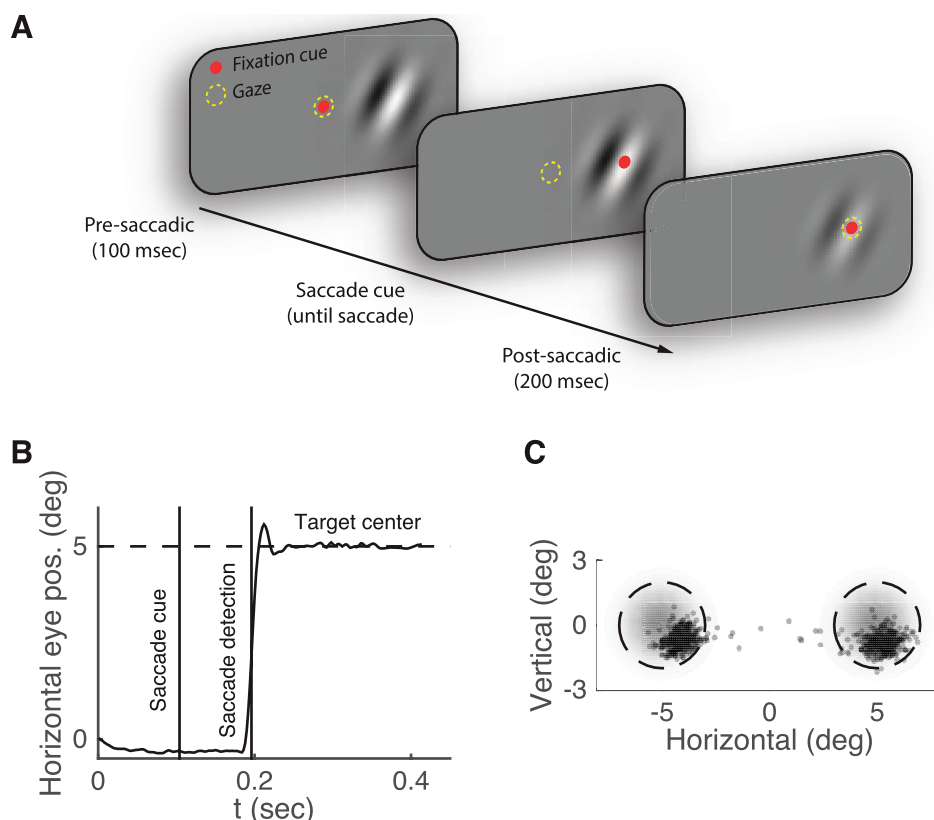
Figure 1. 2-AFC Orientation-discrimination task. (A) The time course of a "both" trial in Experiment 1. Subjects fixated a spot at the center of the screen after which a Gabor patch appeared in the periphery (fixation cue: red dot; subject gaze: dashed yellow circle). After 100 ms the fixation dot jumped to the target, cueing the subject to initiate a saccade to the target. Once the saccade was detected, the contrast of the target changed (to equate foveal and peripheral sensitivity). The target remained visible for 200 ms after saccade detection. Subjects reported whether the target was oriented clockwise or counterclockwise relative to vertical. (B) Horizontal gaze trajectory on a single example trial. Screen center was at 0° and the target was at 5° (i.e., a saccade-right trial). Vertical lines indicate the saccade cue and the moment of saccade detection, at which point the contrast of the target changed. (C) Saccade end points for a typical subject. Each dot is the final position of the subject's gaze on a single trial. Shaded gray regions represent the Gaussian envelope of the target Gabor. Dashed circles mark the 2° error criterion. Trials that fell outside this region were not included in the analysis.

the target was displayed only after a saccade was detected. Thus, the target was only visible foveally.

Each experiment began with 100 practice trials chosen randomly from all three trial types. During practice, subjects were given feedback as to whether their responses were correct. Next, foveal and peripheral contrast thresholds corresponding to 70% correct performance were determined. Subjects performed 150 single-cue trials (periphery-only and fovea-only, randomly assigned). Targets were oriented ±2° (positive orientation refers to clockwise relative to vertical), and contrast was controlled by two interleaved one-up/two-down staircase procedures per trial type (Levitt & Treisman, 1969) with contrast ranging from 3.1% to 8.8%. In addition, on 10% of the trials (randomly assigned) contrast was chosen at random from this range to guarantee an occasional easy trial.

In the main experiment subjects performed all three trial types with no feedback for 250–1,000 trials, depending on the subject's pace. Orientation was selected from the set {±4, ±3, ±2, ±1, 0} using two interleaved staircase procedures per trial type (one-up/two-down, and two-up/one-down, initiated at each end of the sequence), for a total of six staircases. In addition, on 10% of the trials (randomly assigned) orientation was chosen at random from the set to guarantee an occasional easy trial. Trial types were assigned at random. Target contrast changed during the saccade. The contrast values for peripheral and foveal contrast were the 70%-correct values (for ±2°) estimated previously. These saccade-contingent contrast changes were not noticed by observers. The session was terminated after 55 min.

Following each session, subjects filled out a questionnaire about their experience. When asked if they noticed anything unusual during the experiment, none

reported noticing changes in the target during the saccade. Some subjects were asked more explicitly, but did not report noticing the change.

Experiment 2

A second perturbation experiment was similar to Experiment 1. In addition to the three trial types mentioned above, there were four perturbation conditions (seven conditions altogether, trial types assigned at random) in which the presaccadic target orientation was rotated by $\pm 3.2°$ or $\pm 1.6°$ relative to the foveal target. Orientation changes occurred along with the change of contrast after saccade detection and were not noticed by observers.

Perturbation experiments were carried out over two sessions conducted on consecutive days. The procedure for the first session was identical to Experiment 1, with the addition of the four perturbation conditions to the main experiment. The second session included 50 practice trials, and did not include contrast-threshold trials. Rather, the contrast values from the first session were used. Subjects performed between 600 and 2,000 trials over the two sessions, depending on their pace.

Experiment 3

In a third "saccade-away" experiment, the target appeared initially at the center of the screen and the saccade moved the target to the periphery. The procedure was identical to Experiment 2 except that the foveal view was perturbed relative to the periphery.

Experiment 4

The goal of the fourth experiment was to qualitatively reproduce the sequence of images on the retina in Experiment 1, but without eye movements. Thus subjects were asked to maintain fixation at the center of the screen throughout each trial. Targets initially appeared in the periphery (except in the fovea-only condition), and then jumped to the fovea (except in the periphery-only condition). The target was not visible during the jump from periphery to fovea, simulating saccadic suppression. Subjects were told to judge the orientation, and that the peripheral and foveal stimuli in each trial had identical orientation. Simulated saccade initiation delays and saccade durations were sampled from Gaussian distributions with mean and variance estimated from Experiment 1's eye-tracking data pooled over subjects (simulated saccade initiation delay $= 124 \pm 34$ ms, simulated saccade duration $= 60 \pm 12$ ms). To simulate the errors in saccade end-points, we added normally distributed position noise to the foveal view of the

target with mean and covariance estimated from Experiment 1's eye-tracking data. The error in the horizontal direction was $-0.4° \pm 0.47°$ (hypometric on average), and $-0.05° \pm 0.27°$ in the vertical direction (slightly below target on average), the covariance was $0.004°^2$.

Ideal observer model

We assume that the foveal and peripheral estimates of orientation are independent and normally distributed given the target orientation, θ. More explicitly we assume that $\hat{\theta}_{\text{Fov}} = \theta + \eta_{\text{Fov}}$, where $\eta_{\text{Fov}} \sim N(\mu_{\text{Fov}}, \sigma^2_{\text{Fov}})$ is Gaussian estimation noise characterized by a perceptual bias μ_{Fov} and variance σ^2_{Fov} (similarly for $\hat{\theta}_{\text{Peri}}$). Thus, the probability distribution over target orientation, θ, given the foveal estimate $\hat{\theta}_{\text{Fov}}$ and the peripheral estimate $\hat{\theta}_{\text{Peri}}$ can be written as

$$P(\theta|\hat{\theta}_{\text{Fov}}, \hat{\theta}_{\text{Peri}}) = \frac{P(\hat{\theta}_{\text{Fov}}|\theta)P(\hat{\theta}_{\text{Peri}}|\theta)P(\theta)}{P(\hat{\theta}_{\text{Fov}}, \hat{\theta}_{\text{Peri}})}$$

where $P(\theta)$ is the prior over orientation. If the prior is flat (although note that the results readily generalize to a Gaussian prior), then the posterior is Gaussian with mean $\mu = \frac{\hat{\theta}_{\text{Fov}}\sigma^2_{\text{Peri}} + \hat{\theta}_{\text{Peri}}\sigma^2_{\text{Fov}}}{\sigma^2_{\text{Fov}} + \sigma^2_{\text{Peri}}}$, and variance $\sigma^2 = \left(\frac{1}{\sigma^2_{\text{Fov}}} + \frac{1}{\sigma^2_{\text{Peri}}}\right)^{-1}$ (Oruç, Maloney, & Landy, 2003). We assume that an ideal observer's estimate of orientation, $\hat{\theta}^*_{\text{Both}}$, is the mode (which in the Gaussian case is equal to the mean) of the posterior (Stocker & Simoncelli, 2006), also known as the maximum a posteriori (MAP) estimate.

Note that $\hat{\theta}^*_{\text{Both}}$ is a weighted sum of the individual estimates:

$$\hat{\theta}^*_{\text{Both}} = w^*_{\text{Fov}}\hat{\theta}_{\text{Fov}} + w^*_{\text{Peri}}\hat{\theta}_{\text{Peri}}$$

where $w^*_{\text{Fov}} = \frac{1}{\sigma^2_{\text{Fov}}}\left(\frac{1}{\sigma^2_{\text{Fov}}} + \frac{1}{\sigma^2_{\text{Peri}}}\right)^{-1}$ and $w^*_{\text{Peri}} = \frac{1}{\sigma^2_{\text{Peri}}}\left(\frac{1}{\sigma^2_{\text{Fov}}} + \frac{1}{\sigma^2_{\text{Peri}}}\right)^{-1}$ (Landy, Maloney, Johnston, & Young, 1995).

The psychometric curve represents the probability of a "clockwise" response given a target orientation of θ:

$$P(\hat{\theta} > 0|\theta) = \int_0^{\frac{\pi}{2}} P(\hat{\theta}|\theta) \, d\hat{\theta}.$$

Given our assumption of normally distributed estimates, $P(\hat{\theta} > 0|\theta)$ should follow a Gaussian cumulative distribution function. For the ideal observer, the mean and variance are $\mu^*_{\text{Both}} = E[\hat{\theta}^*_{\text{Both}}] = w^*_{\text{Fov}}(\theta + \mu_{\text{Fov}}) + w^*_{\text{Peri}}(\theta + \mu_{\text{Peri}})$ and $\sigma^{*2}_{\text{Both}} = \text{Var}[\hat{\theta}^*_{\text{Both}}] = \left(\frac{1}{\sigma^2_{\text{Fov}}} + \frac{1}{\sigma^2_{\text{Peri}}}\right)^{-1}$. Notice that the
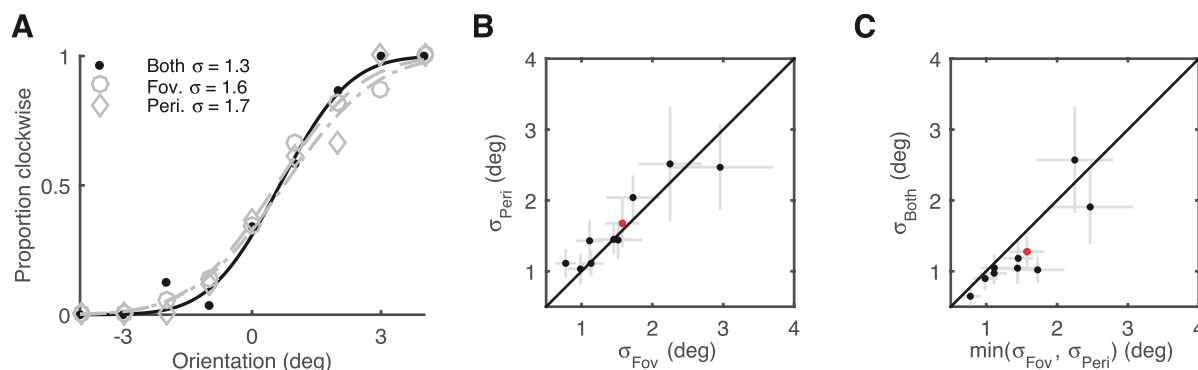
Figure 2. Integration of orientation information across saccades. (A) Psychometric curves for an example subject. The proportion of the subject's "clockwise" responses is plotted as a function of target orientation in the three conditions of Experiment 1. Dots: data; solid lines: Gaussian CDF fits. (B) Orientation-discrimination performance in fovea and periphery was well matched for all subjects. The *SD* of the CDF fit to the periphery-only condition is plotted as a function of the corresponding fovea-only *SD* for each subject. (C) Orientation-discrimination performance in the "both" condition as a function of the best single-cue performance (i.e., smallest *SD*) for each subject. Almost all subjects exhibit a smaller *SD* (i.e., better discrimination) in the "both" condition. Gray error bars: ±1 *SE*. Red dots: example subject in panel A.

maximal reduction in variance (relative to the smaller of σ^2_{Fov} and σ^2_{Peri}) occurs when $\sigma^2_{\text{Fov}} = \sigma^2_{\text{Peri}}$ in which case $\sigma^{*2}_{\text{Both}} = \frac{\sigma^2_{\text{Fov}}}{2}$. On the other hand, if $\sigma^2_{\text{Fov}} \ll \sigma^2_{\text{Peri}}$ then $\sigma^{*2}_{\text{Both}}$ will not be much smaller than σ^2_{Fov}.

Analysis

Gaussian cumulative distribution functions (CDFs) were fit to data by maximizing likelihood. Perceptual bias was defined as the negative of the mean of the CDF fit (the 50% point of the fit curve). Uncertainty was defined as the *SD* of the CDF fit. Subjects who performed the task poorly (average *SD* > 2.5° or *SD* > 3° for any single condition) were not included in the final analysis. Eleven subjects were dropped from the analysis due to the aforementioned criterion: one from Experiment 1, and 10 from Experiment 4. Note that Experiment 4 was substantially harder for subjects to perform, and thus, a higher proportion of subjects had to be dropped to keep discrimination performance on par with previous experiments. Nevertheless, we found that the conclusions do not change qualitatively if these subjects are included in the analysis.

Integration weights were estimated as the slope of a straight line fit to a subject's perceptual bias in the four perturbation conditions and the "both" condition (Young et al., 1993). In Experiment 2 the periphery was perturbed relative to the fovea and in Experiment 3 the fovea was perturbed relative to the periphery. Peripheral weights in Experiment 3 were estimated as $w_{\text{Peri}} = 1 - w_{\text{Fov}}$, where w_{Fov} was the inferred foveal weight from the straight-line fit.

Standard errors of the mean for the above quantities were estimated using bootstrap resampling of the experimental trials (100 repeats).

Results

Humans integrate information across saccades

We first performed an experiment to determine whether humans combine presaccadic and postsaccadic information when estimating a target's orientation. The experimental task is built on the natural behavior of foveation, in which a presaccadic peripheral view of a target is followed by a postsaccadic foveal view. Subjects were asked to judge whether the target appeared to be oriented clockwise or counterclockwise relative to vertical. We tested whether the pre- and postsaccadic views were treated as independent orientation cues and integrated, yielding an improvement in performance compared to performance measured for each view alone. Specifically, if subjects do integrate orientation information across eye movements, the slope of the psychometric function should be steeper when the target is viewed both pre- and postsaccade, compared to the slope obtained under either pre- or postsaccade presentations alone (Figure 2A). Orientation discrimination accuracy was quantified by the standard deviation of a Gaussian cumulative distribution function fit to the psychometric curve.

There is typically a large discrepancy in orientation-discrimination accuracy between fovea and periphery.
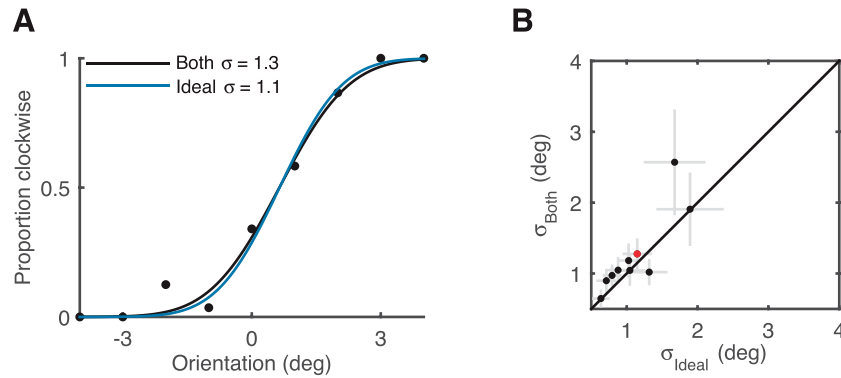
Figure 3. Humans integrate orientation information suboptimally. (A) Psychometric curve of an example subject in the "both" condition (black dots and black fit curve; same data as Figure 2A) compared to the ideal observer that optimally integrates information from the fovea and periphery (blue curve). (B) *SD* in the "both" condition plotted as a function of the *SD* of an ideal integrator. For most subjects the *SD* in the "both" condition was slightly greater than ideal, meaning worse orientation discrimination. Gray error bars: ±1 *SE*. Red dot: example subject in Figure 2A.

In such a situation, the expected effect of cue integration is quite small and hard to distinguish from sole use of the more accurate cue (Rock & Victor, 1964). The maximum performance improvement arising from optimal cue integration is obtained when the accuracies associated with the two cues are equal (Ernst & Banks, 2002). Thus, we equalized the information acquired in the foveal and peripheral views, by adjusting the contrasts of the foveal and peripheral stimuli to levels that produced comparable single-cue performance. For trials in which both stimuli were presented, the contrast was altered during the saccade; subjects were unaware of this change. For most subjects we were able to match performance at the fovea and periphery quite well (Figure 2B).

A comparison of performance in the "both" condition to the best single-cue performance is shown in Figure 2C. Almost all subjects displayed better orientation discrimination in the "both" condition, compared to either single-cue condition ($n = 10$, $p = 0.02$, one-sided paired Wilcoxon signed-rank test). These results clearly demonstrate that humans integrate orientation information across saccadic eye movements.

Given that humans integrate orientation information across eye movements, we next tested whether integration is optimal. Assuming independent Gaussian measurement noise at the fovea and periphery, the variance of a statistically ideal integrator is $\sigma^{*2} = \left(\frac{1}{\sigma_{\text{Fov}}^2} + \frac{1}{\sigma_{\text{Peri}}^2} \right)^{-1}$. In the case of equal foveal and peripheral variance, this results in a factor of two reduction in variance for the combination. Figure 3 compares performance in the "both" condition to predicted performance by the ideal integrator based on estimated uncertainty in the two single-cue conditions.

Integration was generally slightly worse than ideal ($p = 0.04$ one-sided paired Wilcoxon signed-rank test). Thus, whereas humans do integrate orientation information across eye movements, they do so suboptimally.

Humans overweight the fovea

To better understand why humans integrate orientation information across eye movements suboptimally, we sought a more precise characterization of the integration rule. Under independent Gaussian noise assumptions, a statistically optimal estimator produces values that are a linear combination of the estimates obtained under each view alone, $\hat{\theta}_{\text{Both}}^* = w_{\text{Fov}}^* \hat{\theta}_{\text{Fov}} + w_{\text{Peri}}^* \hat{\theta}_{\text{Peri}}$, with weights inversely proportional to the variance of the associated cue,

$$w_{\text{Peri}}^* = \frac{1}{\sigma_{\text{Peri}}^2} \left(\frac{1}{\sigma_{\text{Fov}}^2} + \frac{1}{\sigma_{\text{Peri}}^2} \right)^{-1},$$

and similarly for w_{Fov}^* (Oruç et al., 2003; Young et al., 1993). To determine the rule used by human observers, we measured the effect of perturbing the target orientation during the saccade (as with the contrast changes, these saccade-triggered orientation changes were not noticed by the subjects). Analysis of responses under these conditions allows us to determine whether the human combination rule is indeed linear, and if so, the values of the weights that are used (Young et al., 1993).

Figure 4A shows the data along with psychometric functions fit to the data for one subject in the five perturbation conditions (including no perturbation, identical to the "both" condition of Experiment 1). As the orientation of the peripheral stimulus was per-
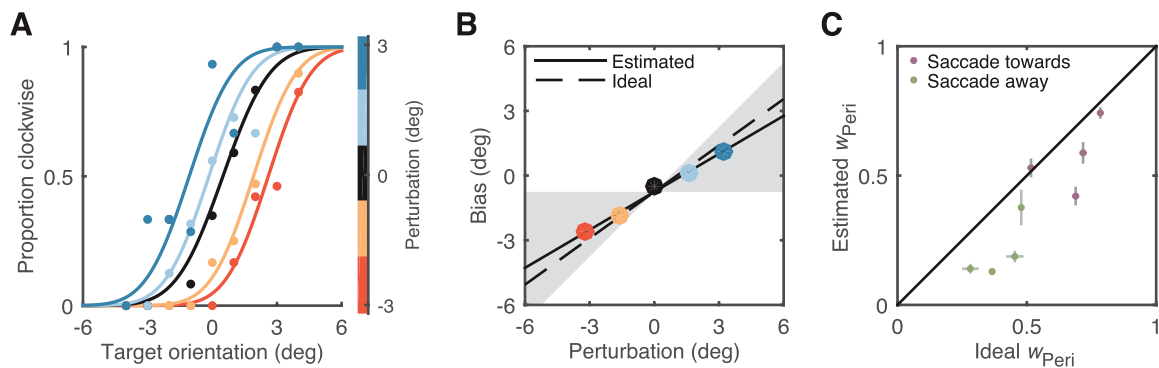
Figure 4. Perturbation analysis demonstrates that humans overweight the fovea. (A) Psychometric curves of an example subject as the orientation of the peripheral stimulus was perturbed relative to the foveal stimulus. The proportion of "clockwise" responses is plotted as a function of the orientation of the foveal stimulus. Dots: experimental data; solid lines: Gaussian CDFs fit separately to the data for each perturbation condition. Color indicates the amount of peripheral perturbation. Perturbing the peripheral stimulus induces a systematic bias of perceived orientation, as seen in the horizontal shift in the fitted psychometric curves. (B) Perceptual bias (amount of shift in psychometric curves in panel A) plotted as a function of the perturbation. Solid line: regression fit to the data. The slope of this line provides an estimate of the weight of the peripheral cue (w_{Peri}). Dashed line: prediction of the ideal-observer model, which predicts a higher peripheral weight. Diagonal and flat boundaries of the gray region indicate performance of an observer that utilizes only peripheral or foveal information, respectively. (C) Inferred peripheral weight (slope of line in panel B) compared to that of an ideal observer, for all subjects. Purple dots: Experiment 2. Green dots: Experiment 3. Error bars: ± 1 SE. Almost all subjects assign a smaller weight to the periphery than the ideal.

turbed clockwise or counterclockwise relative to the foveal stimulus, the observer was more or less likely to respond "clockwise," and the psychometric function shifts to the left or right, respectively.

Figure 4B shows the perceptual bias (horizontal shift of the fitted psychometric curve) as a function of the perturbation amplitude for this subject. These data are well fit by a straight line, consistent with a linear integration rule, and the slope of this line provides an estimate of the weight of the perturbed cue (Young et al., 1993). This slope lies between zero (corresponding to an observer that relies only on the foveal view) and one (an observer that relies only on the peripheral view). The data lie on a line that is slightly shallower than that of an ideal observer, suggesting that this subject weighted the foveal information more heavily than an ideal observer. Three of the four observers weighted the fovea more heavily than the corresponding ideal observer (Figure 4C).

This result could either indicate that observers overweight the fovea in general, or they simply gave higher weight to the most recently viewed stimulus (in Experiments 1 and 2, in the "both" condition the foveal stimulus was always the last one viewed). To distinguish these two alternatives, we ran an additional perturbation experiment (Experiment 3) in which the stimulus was initially viewed foveally at fixation and then was viewed peripherally after a saccade *away* from the stimulus. The results of Experiment 3 (Figure 4C, green symbols) confirm that humans integrate orien-

tation information across saccades, but give more weight to the foveal view than predicted by the ideal observer ($n =$ eight subjects combined across Experiment 2 and 3; $p = 0.02$ two-sided paired Wilcoxon signed-rank test).

Note that the ideal weights in Experiments 2–3 range widely (from 0.3 to 0.8), indicating that foveal and peripheral performance were not perfectly matched across these observers and conditions. The fact that ideal weights are highly correlated with the estimated weights implies that subjects assign cue weights based on relative cue reliabilities, as has been found previously (Landy & Kojima, 2001; Young et al., 1993).

No information integration in the absence of eye movements

Our results clearly demonstrate that humans integrate orientation information across saccadic eye movements. But is the subject-initiated saccade a necessary component of the integration process, or does the visual system have a more general capability for integrating across spatially and temporally distinct stimulus patterns?

To discriminate between these two hypotheses, we conducted an experiment (Experiment 4) in which subjects maintained fixation throughout the trial while image sequences were played on the screen such that
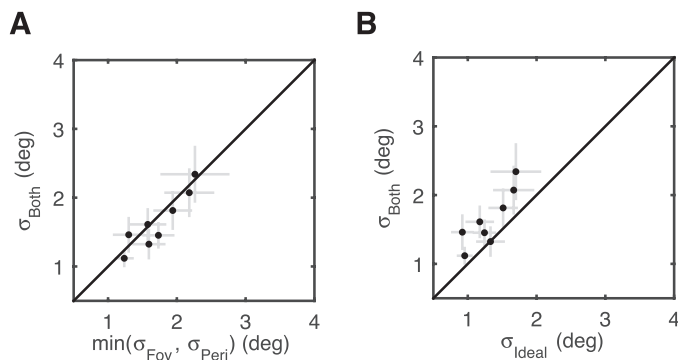
Figure 5. Lack of information integration in the absence of saccades. (A) Orientation-discrimination performance in the "both" condition as a function of the best single-cue performance (i.e., smallest *SD*) for each subject in Experiment 4 (no eye movements). Performance is not significantly different, providing no evidence for integration. Gray error bars: ± 1 *SE*. (B) Performance in the "both" condition plotted as a function of the performance of an ideal integrator. For most subjects the *SD* in the "both" condition was greater than ideal. Gray error bars: ± 1 *SE*.

the spatiotemporal sequence of retinal images corresponded to typical sequences in Experiment 1 (see Methods).

Figure 5A demonstrates that subjects' performance in the "both" condition was not significantly better than the best individual cue condition ($n = 8$; $p = 0.13$ one-sided paired Wilcoxon signed-rank test), and is thus also worse than ideal (Figure 5B; $p = 10^{-4}$ one-sided paired Wilcoxon signed-rank test).

Taken together with the results presented in Figure 2, we conclude that the active process of saccadic eye movements is a necessary ingredient for the integration of information derived from spatio-temporally distinct retinal inputs.

This result can be interpreted in the context of causal inference (Ernst, 2007; Körding et al., 2007): When subjects move their eyes, the two views are perceived as one object and thus should be integrated; whereas in the absence of eye movements, the two views are not likely to arise from one underlying cause and should be kept separate.

Discussion

We have examined whether humans integrate orientation information across saccadic eye movements. Our results show that human orientation-discrimination performance is improved when both presaccadic peripheral and postsaccadic foveal views of a small target are available, relative to perfor-

mance based on either view alone, indicating that humans integrate the two views. We also found that this integration was dependent on eye movement: When subjects maintained fixation and the target jumped between retinal locations, there was no evidence of integration. Additional measurements and analysis indicate that humans combine trans-saccadic information linearly (i.e., as a weighted sum), with weights that depend on the relative precision of the fovea and periphery in this particular task. Finally, we find that the weights used by our subjects are slightly suboptimal relative to an ideal observer, revealing a systematic overweighting of the fovea.

Saccadic integration (or lack thereof) in different perceptual tasks

Trans-saccadic perception has been studied for decades, yet the literature is divided as to whether saccadic integration takes place. Whereas some studies report little to no integration (Bridgeman et al., 1975, 1994; Irwin et al., 1990), others demonstrate better performance in the presence of peripheral information (Henderson & Anes, 1994; McConkie & Rayner, 1975; Pollatsek, Rayner, & Collins, 1984). And although these observations of enhanced performance provide evidence for utilization of peripheral information, it is hard to deduce from these results which visual information is being used.

Recent studies demonstrate that the peripheral view can bias perception of shape (Demeyer, De Graef, Wagemans, & Verfaillie, 2009, 2010b; Herwig, Weiß, & Schneider, 2015), color (Wijdenes, Marshall, & Bays, 2015), and spatial frequency (Herwig & Schneider, 2014). Furthermore, some studies suggest that integration is statistically optimal since sensory noise affects integration weights (Vaziri, Diedrichsen, & Shadmehr, 2006; Wijdenes et al., 2015). Yet, these studies do not quantitatively examine whether the altered weights match the statistically optimal weights. Thus, they do not provide an estimate of how much information is integrated, and how much is discarded. By using models of cue integration, we were able to compare human performance to an ideal observer, placing human performance on a scale from no integration to ideal integration.

It seems surprising that subjects were unaware of orientation perturbations, even though these perturbations did bias their percept of the target. Furthermore, the perturbations were significantly above most subjects' expected discrimination threshold, given their performance in the fovea-only and periphery-only conditions. Whereas it is possible that subjects would be able to report the perturbation if explicitly asked to

do so, there is a growing body of evidence demonstrating that even when humans cannot report saccade contingent changes to a scene, the presaccadic information is not entirely lost. In particular, several studies show that changing the appearance of a target (Demeyer, De Graef, Wagemans, & Verfaillie, 2010a; Tas, Moore, & Hollingworth, 2012), or introducing a postsaccadic blank (Deubel, Schneider, & Bridgeman, 1996), can greatly improve subjects' ability to detect saccade-contingent changes.

In cue integration tasks, it has generally been found that separate cues are integrated only when they are not too discrepant (Hedges, Stocker, & Simoncelli, 2011; Knill, 2007; Landy et al., 1995), but once integration has occurred the observer is no longer aware of the discrepancy. We can adopt a behavioral interpretation of our results: Observers assumed the perturbed pre- and postsaccadic views arose from the same object (Experiments 2 and 3), leading to integration of the two views, whereas in the artificial-saccade case they were assumed to come from distinct objects (Experiment 4), resulting in no integration. This is consistent with previous reports that suggest spatio-temporal integration is contingent on saccades (Cox, Meier, Oertelt, & DiCarlo, 2005; Herwig & Schneider, 2014). In any case, consistency of object identity provides a normative restatement of the problem, but does not offer additional constraints or predictive power; one must still determine the conditions under which the visual system decides that pre- and postsaccadic views arise from the same object.

One of the most intriguing outcomes of the current study is that while humans do integrate orientation information, they do so suboptimally. Humans have been shown to integrate cues optimally under many circumstances (Ernst & Banks, 2002; Landy & Kojima, 2001), so it is not obvious why they would fail to do so in this setting. This is particularly perplexing given the frequency of saccades (several times each second on average). One possibility is that the visual system is only able to partially adapt the integration weights in response to the artificial contrast change that we induced during the saccades. In this case, one would expect a systematic bias of these weights toward their typical values, which would be expected to favor the fovea (see figure S3 of Zaidel, Goin-Kochel, & Angelaki, 2015).

Saccadic integration in natural vision

Our results are unequivocal and consistent across observers; nevertheless, we cannot be certain that they generalize beyond the conditions of our experiments. In particular, we made several choices regarding stimuli,

whose impact on the outcome can, and should, be explored in depth.

First, our measure of integration arises from an orientation-discrimination task. Orientation is a well-studied stimulus attribute, fundamental for both perception and computational vision, and is associated with the most prominent response selectivity of primary visual cortex (Hubel & Wiesel, 1959). However, the substantial differences in orientation-discrimination capabilities in fovea and periphery (Paradiso & Carney, 1988) meant that to enable a substantial effect of integration, we had to equalize performance at the two locations using an artificial trans-saccadic contrast change. This begs the question: Do humans integrate orientation information across saccades under natural viewing conditions? Although we cannot answer this question with the current data, the fact that humans weight the peripheral and foveal inputs based on their relative accuracies suggests that the neural apparatus necessary to perform the integration is in place (Ma, Beck, Latham, & Pouget, 2006). It would be worthwhile to explore integration in visual tasks for which performance at the periphery is comparable to (or better than) the fovea, such as temporal-frequency discrimination (Waugh & Hess, 1994) or texture segmentation (Kehrer, 1987).

Second, we used stimuli that occupied only a small region of visual space and were presented in isolation. This made the task easy for our subjects, but if integration takes place under natural conditions, it should presumably occur across the entire visual field and in the presence of many different visual objects. It would thus be interesting to examine whether (a) The spatial extent of the stimuli affects integration; (b) Integration takes place between pairs of *peripheral* locations; and (c) Integration takes place in the presence of distractors. Masking and crowding effects, especially prominent in the periphery, are known to substantially interfere with discrimination performance, but their effects on trans-saccadic integration have not been characterized and need not be the same.

The mechanism underlying saccadic integration

A physiological mechanism that has been proposed to be involved in saccadic integration is that of receptive field remapping (Duhamel, Colby, & Goldberg, 1992), in which receptive fields shift just before a saccade to their future (postsaccadic) spatial location, providing a temporal window in which pre- and postsaccadic views could be integrated. Evidence for this mechanism has been found in multiple brain regions (for a review, see Sommer & Wurtz, 2008). However, a recent study challenged previous findings,

suggesting that, at least in prefrontal cortex, receptive fields contract toward the saccadic target rather than predictively moving in the direction of the saccadic path (Zirnsak, Steinmetz, Noudoost, Xu, & Moore, 2014). Our results are agnostic with regards to these findings, and do not provide evidence for either mechanism (predictive or contractive remapping).

Nonetheless, our results do place certain constraints on the underlying physiological implementation responsible for trans-saccadic integration of orientation information. A natural candidate for representation of local orientation is area V1, the earliest stage of the visual hierarchy in which cells exhibit orientation selectivity. Could trans-saccadic integration take place within V1? Our subjects combined two views of a target separated by 5°, center-to-center, in retinal coordinates. This corresponds to roughly 3.5–4 V1 receptive field widths (in macaque) at 5° eccentricity and even more at the fovea (see Freeman & Simoncelli, 2011, and references therein). It seems implausible therefore that integration is carried out locally in V1, given the limited extent of lateral connectivity (Angelucci et al., 2002; Cavanaugh, Bair, & Movshon, 2002), even when taking into account the target size and inaccuracies in estimates of receptive field width. Hence, our results suggest that integration most likely takes place in higher visual areas, where receptive fields are larger. This might involve lateral connectivity and feedback to V1 (Olshausen, Anderson, & Van Essen, 1993), or the fusion of information may be directly achieved by extrastriate neurons whose larger receptive fields encompass both views of the target.

Our visual experience provides the illusion of continuity and stability, but is effectively assembled from a sequence of discrete views. Although the internal representation that supports this illusion is surely both incomplete and abstract in form, it would seem necessary that it rely heavily on the fusion of information acquired across saccades. The results presented here provide direct quantitative evidence of such integration, and provide a potential paradigm for its further investigation.

Keywords: eye movements, cue integration, psychophysics, saccadic integration

References

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupé, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience*, *22*(19), 8633–8646.

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

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research*, *15*(6), 719–722.

Bridgeman, B., Van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*(02), 247–258, doi.org/10.1017/S0140525X00034361.

Bullmore, E., & Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, *13*(5), 336–349, doi.org/10.1038/nrn3214.

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, *88*(5), 2530–2546, doi.org/10.1152/jn.00692.2001.

Cox, D. D., Meier, P., Oertelt, N., & DiCarlo, J. J. (2005). "Breaking" position-invariant object recognition. *Nature Neuroscience*, *8*(9), 1145–1147, doi.org/10.1038/nn1519.

Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2009). Transsaccadic identification of highly similar artificial shapes. *Journal of Vision*, *9*(4):28, 1–14, doi:10.1167/9.4.28. [PubMed] [Article]

Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2010a). Object form discontinuity facilitates displacement discrimination across saccades. *Journal of Vision*, *10*(6):17, 1–14, doi:10.1167/10.6.17. [PubMed] [Article]

Demeyer, M., De Graef, P., Wagemans, J., & Verfaillie, K. (2010b). Parametric integration of visual form

across saccades. *Vision Research*, *50*(13), 1225–1234, doi.org/10.1016/j.visres.2010.04.008.

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research*, *36*(7), 985–996, doi.org/10.1016/0042-6989(95)00203-0.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*(5040), 90–92.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5):7, 1–14, doi:10.1167/7.5.7. [PubMed] [Article]

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1201, doi.org/10.1038/nn.2889.

Ganmor, E., Landy, M., & Simoncelli, E. (2015). Near-optimal integration of orientation information across saccadic eye movements. *Journal of Vision*, *15*(12): 1306, doi:10.1167/15.12.1306. [Abstract]

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In A. Elizabeth (Ed.), *Perception* (Vol. 5). New York, NY: Oxford University Press. Retrieved from http://psycnet.apa.org/psycinfo/1996-97608-003

Hedges, J. H., Stocker, A. A., & Simoncelli, E. P. (2011). Optimal inference explains the perceptual coherence of visual motion stimuli. *Journal of Vision*, *11*(6):14, 1–16, doi:10.1167/11.6.14. [PubMed] [Article]

Henderson, J. M., & Anes, M. D. (1994). Roles of object-file review and type priming in visual identification within and across eye fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(4), 826–839, doi.org/10.1037/0096-1523.20.4.826.

Herwig, A., & Schneider, W. X. (2014). Predicting object features across saccades: Evidence from object recognition and visual search. *Journal of Experimental Psychology: General*, *143*(5), 1903–1922, doi.org/10.1037/a0036781.

Herwig, A., Weiß, K., & Schneider, W. X. (2015). When circles become triangular: How transsaccadic predictions shape the perception of shape. *Annals of the New York Academy of Sciences*, *1339*(1), 97–105, doi.org/10.1111/nyas.12672.

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, *148*(3), 574–591.

Irwin, D. E., Zacks, J. L., & Brown, J. S. (1990). Visual memory and the perception of a stable visual environment. *Perception & Psychophysics*, *47*(1), 35–46, doi.org/10.3758/BF03208162.

Kehrer, L. (1987). Perceptual segregation and retinal position. *Spatial Vision*, *2*(4), 247–261, doi.org/10.1163/156856887X00204.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.

Knill, D. C. (2007). Robust cue integration: A Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *Journal of Vision*, *7*(7):5, 1–24, doi:10.1167/7.7.5. [PubMed] [Article]

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9), e943, doi.org/10.1371/journal.pone.0000943.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A*, *18*(9), 2307–2320, doi.org/10.1364/JOSAA.18.002307.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. J. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*(3), 389–412, doi.org/10.1016/0042-6989(94)00176-M.

Laughlin, S. B. (2001). Energy as a constraint on the coding and processing of sensory information. *Current Opinion in Neurobiology*, *11*(4), 475–480, doi.org/10.1016/S0959-4388(00)00237-3.

Levitt, H., & Treisman, M. (1969). Control charts for sequential testing. *Psychometrika*, *34*(4), 509–518, doi.org/10.1007/BF02290604.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*, 1432–1438.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*(6), 578–586, doi.org/10.3758/BF03203972.

Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *The Journal of Neuroscience*, *13*(11), 4700–4719.

Oruç, İ., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly

correlated error. *Vision Research*, *43*(23), 2451–2468, doi.org/10.1016/S0042-6989(03)00435-8.

Paradiso, M. A., & Carney, T. (1988). Orientation discrimination as a function of stimulus eccentricity and size: Nasal/temporal retinal asymmetry. *Vision Research*, *28*(8), 867–874, doi.org/10.1016/0042-6989(88)90096-X.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology*: *General*, *113*(3), 426–442.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, *31*(6), 537–550, doi.org/10.3758/BF03204186.

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the two senses. *Science*, *143*(3606), 594–596, doi.org/10.1126/science.143.3606.594.

Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, *31*(1), 317–338, doi.org/10.1146/annurev.neuro.31.060407.125627.

Stocker, A. A., & Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, *9*(4), 578–585, doi.org/10.1038/nn1669.

Tas, A. C., Moore, C. M., & Hollingworth, A. (2012). An object-mediated updating account of insensitivity to transsaccadic change. *Journal of Vision*, *12*(11):18, 1–13, doi:10.1167/12.11.18. [PubMed] [Article]

Vaziri, S., Diedrichsen, J., & Shadmehr, R. (2006). Why does the brain predict sensory consequences of oculomotor commands? Optimal integration of the predicted and the actual sensory feedback. *The Journal of Neuroscience*, *26*(16), 4188–4197, doi.org/10.1523/JNEUROSCI.4747-05.2006.

Waugh, S. J., & Hess, R. F. (1994). Suprathreshold temporal-frequency discrimination in the fovea and the periphery. *Journal of the Optical Society of America A*, *11*(4), 1199, doi.org/10.1364/JOSAA.11.001199.

Wijdenes, L. O., Marshall, L., & Bays, P. M. (2015). Evidence for optimal integration of visual feature representations across saccades. *The Journal of Neuroscience*, *35*(28), 10146–10153, doi.org/10.1523/JNEUROSCI.1040-15.2015.

Young, M. J., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*(18), 2685–2696, doi.org/10.1016/0042-6989(93)90228-O.

Zaidel, A., Goin-Kochel, R. P., & Angelaki, D. E. (2015). Self-motion perception in autism is compromised by visual noise but integrated optimally across multiple senses. *Proceedings of the National Academy of Sciences*, *USA, 112*(20), 6461–6466, doi.org/10.1073/pnas.1506582112.

Zirnsak, M., Steinmetz, N. A., Noudoost, B., Xu, K. Z., & Moore, T. (2014). Visual space is compressed in prefrontal cortex before eye movements. *Nature*, *507*(7493), 504–507, doi.org/10.1038/nature13149.