Domain Specificity of Oculomotor Learning after Changes in Sensory Processing

Yuliy Tsank and Miguel P. Eckstein

Department of Psychological and Brain Sciences, University of California, Santa Barbara, California 93106-9660

Humans visually process the world with varying spatial resolution and can program their eye movements optimally to maximize information acquisition for a variety of everyday tasks. Diseases such as macular degeneration can change visual sensory processing, introducing central vision loss (a scotoma). However, humans can learn to direct a new preferred retinal location to regions of interest for simple visual tasks. Whether such learned compensatory saccades are optimal and generalize to more complex tasks, which require integrating information across a large area of the visual field, is not well understood. Here, we explore the possible effects of central vision loss on the optimal saccades during a face identification task, using a gaze-contingent simulated scotoma. We show that a new foveated ideal observer with a central scotoma correctly predicts that the human optimal point of fixation to identify faces shifts from just below the eyes to one that is at the tip of the nose and another at the top of the forehead. However, even after 5000 trials, humans of both sexes surprisingly do not change their initial fixations to adapt to the new optimal fixation points to faces. In contrast, saccades do change for tasks such as object following and to a lesser extent during search. Our findings argue against a central brain motor-compensatory mechanism that generalizes across tasks. They instead suggest task specificity in the learning of oculomotor plans in response to changes in front-end sensory processing and the possibility of separate domain-specific representations of learned oculomotor plans in the brain.

Key words: foveation; optimality; saccades; scotoma

Significance Statement

The mechanism by which humans adapt eye movements in response to central vision loss is still not well understood and carries importance for gaining a fundamental understanding of brain plasticity. We show that although humans adapt their eye movements for simpler tasks such as object following and search, these adaptations do not generalize to more complex tasks such as face identification. We provide the first computational model to predict where humans with central vision loss should direct their eye movements in face identification tasks, which could become a critical tool in making patient-specific recommendations. Based on these results, we suggest a novel theory for oculomotor learning: a distributed representation of learned eye-movement plans represented in domain-specific areas of the brain.

Introduction

Brain plasticity is an important property of the human nervous system and is critical for visual, cognitive, and motor development as well as for recovery from disease. Eye movements, which point the high-acuity foveola during exploration of the environment, are critical for accomplishing evolutionary important tasks ranging from visual search to face identification. The brain programs eye movements by taking into account the foveated properties of the visual system in conjunction with the distribution of

task-relevant information in the environment (Legge et al., 1997, 2002; Najemnik and Geisler, 2005, 2009; Peterson and Eckstein, 2012; Paulun et al., 2015) to maximize the acquisition of information during basic perceptual tasks (optimal or near-optimal oculomotor strategies; Legge et al., 1997, 2002; Najemnik and Geisler, 2005, 2009; Peterson and Eckstein, 2012). In addition, the brain can adapt eye movements for changes in the properties of the visual system due to disease. One such case is the development of central vision loss, which can be caused by a number of diseases (e.g., macular degeneration), and results in the fovea becoming partially or fully unavailable (scotoma) for visual processing (Sunnness et al., 1995). Individuals with such impairments undergo a long adjustment period in which a new preferred retinal locus (PRL) is developed and used in place of the fovea as an area with a higher quality of processing in the damaged retina (Walsh and Liu, 2014). Recent studies using a gaze-contingent simulated scotoma technique (Kwon et al., 2013; Walsh and Liu,

Received April 26, 2017; revised Aug. 28, 2017; accepted Sept. 26, 2017.

Author contributions: Y.T. and M.P.E. designed research; Y.T. performed research; Y.T. analyzed data; Y.T. and M.P.E. wrote the paper.

The authors declare no competing financial interests.

Correspondence should be addressed to Yuliy Tsank, Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA 93106-9660. E-mail: yuliy.tsank@psych.ucsb.edu.

DOI:10.1523/JNEUROSCI.1208-17.2017

Copyright © 2017 the authors 0270-6474/17/3711469-16\$15.00/0

2014) have shown that humans generalize compensatory eye movements across visual tasks. Such a finding suggests that a central mechanism in a visuo-motor area (Kaku et al., 2009) might mediate the learning of compensatory eye movements, which can then be applied to all perceptual tasks. However, the only tasks evaluated were those for which the scotoma completely covered the visual target. For those tasks, the compensatory eye movements were rather unambiguous: to uncover the target. The ability of the brain to generalize compensatory eye movements to more complex tasks for which the scotoma only partially covers the stimulus has not been thoroughly studied partly because there has not been an appropriate method to determine what the optimal compensatory eye movements would be.

One such task, which requires integration of features along larger spatial regions that include a scotoma and intact areas of the visual field, is face identification. Humans perform this task many times a day beginning in infancy and exhibit high levels of accuracy with even just a single saccade (Hsiao and Cottrell, 2008; Or et al., 2015). Although some studies have shown that a scotoma will influence the overall distribution of fixations for patients with central vision loss (Seiple et al., 2013), little is known about the functional contribution of these compensatory eye movements and whether they develop with time courses similar to ones for simpler visual tasks.

Here we develop a Bayesian ideal observer that incorporates properties of the foveated visual system and the central scotoma [scotoma-foveated ideal observer (S-FIO)] to make predictions about optimal points of fixation to faces for visual systems with central vision loss. We use a gaze-contingent simulated scotoma to ask whether learned compensatory eye movements from simple tasks (object following and visual search; Kwon et al., 2013) generalize to more complex face identification tasks. If the compensation for a central scotoma involves an area such as the brainstem (Kaku et al., 2009), which is able to learn a compensatory motor command and then apply it to all stimuli and tasks, then we should observe robust generalization and similar learning time courses of the compensatory eye movements across different perceptual tasks. On the other hand, no generalization across tasks might suggest that the learning of eye movements is task specific and may involve distinct domain-specific brain areas.

Materials and Methods

Human psychophysics studies

Participants

The first experiment was completed with a group of 16 undergraduate students of either sex, who participated in the study for course credit. Informed consent was obtained from all subjects and guidelines provided by the institutional review board of the University of California, Santa Barbara, were followed. Seven undergraduate students of either sex participated in the second experiment. Six undergraduate students of either sex participated in the third experiment.

Apparatus and materials

MATLAB Psychtoolbox (RRID:SCR_002881) and Eyelinktoolbox software were used to run the eyetracker from a display computer as well as present visual stimuli on the display screen. The display used was a Barco MDRC 1119 Monitor set to a 1280×1024 pixel resolution and was located 76.5 cm away from the observer's eyes. The display was linearly calibrated with a minimum luminance of 0.05 cd/m^2 and a maximum luminance of 126 cd/m^2 .

Eye tracking

The left eye of each participant was tracked using an SR Research Eyelink 1000 Tower Mount Eye Tracker (RRID:SCR_009602) sampling at 250 Hz. A nine-point calibration and validation were run before each 125-

trial session, with a mean error of not $>0.5^\circ$ of visual angle. Saccades were classified as events in which eye velocity was $>35^\circ/\text{s}$ and eye acceleration exceeded $9500^\circ/\text{s}^2$. The recommended thresholds by SR Research for cognitive research are an eye velocity of 30° and an eye acceleration of $8000^\circ/\text{s}^2$. The minor increase of the velocity and acceleration thresholds in our parameter settings allowed us to better control the number of "broken fixations" during the initial fixation stage at the beginning of every trial before the presentation of the stimulus.

Stimuli

Experiment 1. In this experiment, observers only completed a 1 in 10 face identification task. Ten face images were taken with constant diffuse lighting, distance, and camera settings. A Canon digital camera was used. The digital pixel value was a nonlinear saturating function of luminance (a standard Canon log-cine transfer function). The images were normalized by scaling and cropping, such that the center of the eyes was 200 pixels below the top of the image and the chin was 10 pixels above the bottom of the image. The faces were luminance-mean normalized to 63 cd/m^2 and shown to participants at a contrast of 0.14 (rms contrast of 0.152; and 0.3, rms contrast of 0.2) for the practice blocks), where part of that contrast variation came from added Gaussian white noise with a SD of 7 cd/m^2 (corresponding to a noise rms contrast of 0.11). In scotoma-present blocks, the pixel noise was added to the face stimulus before the simulated scotoma. Participants viewed the face stimuli 46 cm away from the display, resulting in a square stimulus (face and mask) that subtended 18° ($\sim 15^\circ$ for the part of the face that is not covered with the mask). The large size of the faces, more typical of conversational distance, was chosen (1) to allow measurements of larger variations of perceptual performance with fixation position (for small faces, perceptual performance is less sensitive to fixation position within the face); (2) to allow more precise measurements of fixation positions relative to facial features. In addition, the large faces (10° width, 15° height) have been shown to be the face size that optimizes face identification (Yang et al., 2014).

Experiment 2. In this experiment, observers completed a 1 in 10 face identification task and an object-following (and search) task in alternating blocks. For the face identification task, the same stimuli were used as in Experiment 1. For the object-following and search tasks, a set of high-contrast 24 bit RGB color images of 49 indoor scenes (1024×768 pixels; $37.88^\circ \times 28.42^\circ$) and 140 objects ($\sim 50 \times 30$ pixels; $1.85^\circ \times 1.11^\circ$) were used. The images were obtained from Kwon et al. (2013) and were used in that study as stimuli for an object-following and search task, part of which we reproduced in this experiment. Permission was obtained to use the images.

Experiment 3. Everything was the same as in Experiment 1, except the face stimuli that were shown on the screen were halved in size to $\sim 7.5^\circ$ for the part of the face that is not covered with the mask. Importantly, the size of the scotoma stayed the same and thus covered the entire face.

Procedure

Experiment 1. Observers performed a (1 in 10) face identification task with 64 total blocks consisting of several different conditions. Each block consisted of 125 trials. The conditions varied depending on the use of high or low contrast (0.3 vs 0.14), the presence or absence of the scotoma, stimulus presentation time, and whether subjects were allowed to freely make saccades (free-viewing) or were forced to fixate specific locations (forced fixation). Participants were first run through a series of familiarization/practice blocks during which they could get familiarized with the faces and the simulated scotoma as well as get accustomed to the long and short presentation times. Then, experimental blocks were run and were set up in a way that a scotoma-present free-viewing condition was placed before and after a forced-fixation condition. This was done to observe a possible adaptation due to an increase in performance at the new optimal spot. See Table 1 for a summary of the order of the blocks and the conditions that they belonged to. The participants were given instructions to maximize their accuracy in the task by trying to choose the correct face on each trial. They were also told the order of the conditions when they got to each one and received instructions on the differences between them. All 16 observers completed the free viewing blocks before the forced-fixation conditions. A subset of eight of the observers continued to complete additional blocks of free viewing after the forced-fixation conditions.

Table 1. Experiment design

	Blocks	Scotoma presence	Viewing condition	Timing
Practice blocks	1–4	Absent	Free viewing, high contrast	1500 ms
	5–12	Absent	Free viewing	Alternating 1500 ms, 350 ms
	13–28	Present	Free viewing	Alternating 1500 ms, 350 ms
Experimental blocks	29–36	Present	Free viewing	Alternating 1500 ms, 350 ms
	37–44	Absent	Free viewing	Alternating 1500 ms, 350 ms
	45–56	Present	Forced fixation	350 ms
	57–64	Present	Free viewing	Alternating 1500 ms, 350 ms

Experiment 2. A new set of seven observers of either sex performed alternating blocks (44 total) of an object-following/search task and a face identification task with different conditions. All blocks of the face identification task were free viewing and consisted of 125 trials. The object-following/search task blocks consisted of 30 trials each. The first eight blocks (four from each task) were practice blocks. They also served as control blocks in which participants performed the task without a simulated scotoma. The first two practice blocks of the face identification task (blocks 2 and 4) had a 1500 ms stimulus presentation time and a contrast of 0.3. The rest of the face identification blocks had a 350 ms stimulus presentation time and a contrast of 0.14.

Experiment 3. A new set of seven observers of either sex performed the same (1 in 10) face identification task as in Experiment 1, except with faces that are half the size used in Experiment 1. Data from only six of the seven observers was used, however, because the seventh observer had significantly lower performance in the task compared with the other observers. Also, in this experiment, only a free-viewing condition was tested, resulting in a total of 10 blocks. The first two blocks were in scotoma-absent conditions used as practice with a high contrast (0.3; the same as in Experiment 1). Blocks 3 and 4 were in scotoma-absent conditions with regular contrast. Blocks 5–10 were in scotoma-present conditions.

Experimental conditions

Experiment 1. During free-viewing blocks, participants started a trial by pressing the space bar while fixating a cross ($0.74^\circ \times 0.74^\circ$) in one of eight randomly chosen locations located on average 19° from the center of the stimulus. During forced-fixation blocks, the cross was located in one of four locations that corresponded to the forehead, eyes, nose, and mouth every 5.07° downward, respectively. The cross was displayed for a random period of time between 500 and 1500 ms to prevent anticipatory eye movements. If participants moved their eyes $>1^\circ$ from the center of the fixation cross before the stimulus was displayed or while the stimulus was present during the forced-fixation condition, the trial would abort and restart with a new stimulus. The stimulus was then shown for either 1500 ms in the long condition or 350 ms in the short condition. During scotoma-present blocks, a simulated (8° in diameter) circular scotoma, which was centered on the fovea, obstructed the subjects' view of the face and moved with each new saccade. Figure 1 shows the simulated scotoma at three different fixation points. The scotoma consists of a local luminance average of the pixels from the face image behind it, which then gradually transitions into the face image as it gets further from the center of the fixation point. The size of the scotoma was based on the size used in previous simulated-scotoma studies (Kwon et al., 2013; Walsh and Liu, 2014). We used a slightly smaller scotoma base (8° vs 10°) to account for the thickness of the blurred edges. See the S-FIO model description in Materials and Methods above for details on the construction of the simulated scotoma. At the end of each trial, participants had unlimited time to select with the mouse 1 of 10 faces that were displayed on the screen without noise. As soon as a decision was made, feedback was given by outlining the correct face. Figure 2 shows a time line of a single trial.

Experiment 2. The face identification blocks were conducted in the same way as the free-viewing condition in Experiment 1. Trials from the object-following/search task blocks each consisted of four parts in suc-

cession (Fig. 3). This experimental setup was used in a previous study by Kwon et al. (2013).

Part 1: object following with gray background. Participants started a trial by pressing the space bar while fixating a central fixation cross. As in the face task, the trial was aborted if fixation was not maintained for a pre-determined period of time. A randomly chosen object was then presented in place of the fixation cross and was covered with a simulated central scotoma. The scotoma was simulated and behaved in the same way as for the face task, except that a local luminance average of the pixels from the image of the object with a darker gray background (19.69 cd/m^2) was used so that the scotoma would be visible on the lighter gray background (34.45 cd/m^2), which is seen in the first and second squares of the time line in Figure 3. The object then moved to random positions on the screen three times, but only after being uncovered by the scotoma for 1.5 s after each move. This was done to ensure eccentric fixation. A threshold of 4° from the center of the scotoma to the center of the object was chosen to define whether the object was uncovered by the scotoma or not. After the object had been uncovered for 1.5 s at its last position, the trial moved on to the second part.

Part 2: scotoma centering. Here, participants centered their gaze in the middle of the screen so that the simulated scotoma remained inside a square black box (11.7° width) for 1.5 s (Fig. 3, second square in the time line). This was done in order for the participant to start a search task with their gaze centered.

Part 3: search task. Participants performed a yes/no search task with 50% target presence where they had unlimited time to view a cluttered randomly chosen scene (Fig. 3, third square in the time line) with five objects and to determine whether one of them was the object that they followed in the first object-following part of the trial. A simulated scotoma was created in the same way as in previous experiments. On target-present trials, one of the objects was the same as in the first object-following part, while the other four objects were randomly chosen. On target-absent trials, all five objects were randomly chosen excluding the one that was present in the first object-following part. A keyboard response was given with an "up arrow" for target-present trials and with a "down arrow" for target-absent trials. After the response, feedback was immediately given with a green "correct" or a red "incorrect" text centered on the screen for 1000 ms.

Part 4: object following with cluttered background. Participants performed the same object-following task as in part 1 except that the background was a cluttered scene and a different random object was used.

Experiment 3. Everything was the same as in the free-viewing part of Experiment 1, except that only a 1500 ms presentation time was used throughout this experiment.

Statistical analysis

Experiment 1. Paired *t* tests were used to compare average fixation locations and performance between free eye-movement conditions with and without a scotoma, as well as performance between fixation locations in the forced-fixation condition. An *F* test was used to find the effect of the interaction of stimulus duration with scotoma presence on performance in the task. The Pearson correlation was also computed for vertical and horizontal fixation positions between conditions. All comparisons were within subjects. The false discovery rate (FDR) method (RRID:SCR_009473; Benjamini and Hochberg, 1995) was used for multiple-comparison correction on the *p* values. The tests for fixation position and performance were separated into two groups. However, the difference in fixation position between the scotoma-absent and scotoma-present condition remained insignificant even without the correction. See the Experiment 1 in Results for full details.

Experiment 2. The bivariate contour ellipse area (Crossland et al., 2004) was used to quantify the use of peripheral retinal locations in the faces task and object-following task. Paired *t* tests were used to compare differences in this measure across tasks within subjects. Paired *t* tests were also used to compare distances of average fixations from the preferred point of fixation in each of the three tasks (face identification, object following, and search) as well as to separately compare three more metrics of a change in eye movement strategy specifically in the search task.

Gaze Contingent Scotoma

Figure 1. Three different fixations are shown at different parts of the face. A gaze-contingent scotoma is centered on each fixation position and moves with each saccade. The images are shown in high contrast here, the same contrast that was used in the first four practice blocks.

Figure 2. Trial time line. In the free-viewing condition, observers made saccades to the centrally presented face from a fixation cross in one of eight randomly chosen locations. Stimulus presentation times of 1500 and 350 ms were used in alternating blocks for the free-viewing condition. In the forced-fixation condition, observers fixated one of four locations (corresponding to the top of the forehead, the eyes, the nose, and the chin). A stimulus presentation time of 350 ms was used for the forced-fixation condition. Here the stimulus is shown with a scotoma centered on the face, corresponding to a central fixation location around the nose region. At the end of each trial, participants had unlimited time to select with the mouse 1 of 10 faces that were displayed on the screen without noise. As soon as a decision was made, feedback was given by outlining the correct face.

The FDR method (Benjamini and Hochberg, 1995) was used for multiple-comparison correction.

One main result for this article is a lack of a significant effect in the first fixation position between the scotoma-absent and scotoma-present conditions for a face identification task. This result remained insignificant even without the multiple-comparison condition (FDR) correction.

Experiment 3. Paired *t* tests were used to compare average fixation locations and performance between free eye-movement conditions with and without a scotoma. The Pearson correlation was also computed for vertical and horizontal fixation positions between conditions. All comparisons were within subjects. See Experiment 3 in Results for full details.

S-FIO model

A spatially variant contrast sensitivity function (SVCSE) was used to model the degradation of the quality of information obtained in the

periphery of a foveated visual system (Peterson and Eckstein, 2012), as follows:

$$\text{SVCSE}(f, r, \theta) = c_0 f^{a_0} \exp(-b_0 f - d_0(\theta) r^{n_0} f), \quad (1.1)$$

where f is spatial frequency in cycles per degree of visual angle. The terms a_0 , b_0 , and c_0 were chosen constants set to 1.2, 0.3, and 0.625, respectively, to set the maximum contrast at 1 and the peak at 4 cycles/° of visual angle at fixation. The polar coordinates r and θ specify the distance in visual angle and direction from fixation. d_0 specifies the eccentricity factor as a function of direction, which represents how quickly information is degraded in the periphery. n_0 specifies the steep eccentricity roll-off factor. In the model simulations, different parameters are used for directions, d_0 : for the vertical up, du ; for the vertical down, dd ; and for horizontal, dh . The parameters du , dd , dh , and n_0 are fit with the FIO model to match

Figure 3. Trial time line for the object-following/search task trials. A trial was divided into four parts. In the first part, a small object (seen here in the center right of the lowest panel) moved to four different locations on the screen while participants tried to get an adequate look at the object in order to do a search task with the same object later in the trial. In the second part, participants centered the scotoma on the screen in order to provide the same starting point for the search task on every trial. The trial did not continue to the search task until the scotoma was centered within the square for 1.5 s. In the third part, participants did a yes/no search task for the object from the first part of the trial. In the fourth part, a different small object moved to four different locations with a cluttered background.

Figure 4. Human performance (proportion correct) in a forced-fixation face identification experiment is used to fit the FIO model. Here participants are forced to fixate on one of five positions down the vertical midline of the face that cover regions from the top of the forehead to the chin. The FIO performance curve is then fit to these points using the AIC as a distance measure. All of the parameters from this fit, except the internal noise parameter, are used to run the S-FIO model. The internal noise parameter shifts the performance curve for the FIO and S-FIO up or down but does not change the qualitative results (i.e., the relative levels of performance for the different positions down the midline of the face).

human performance (proportion correct) as a function of fixation position (five different fixations down the vertical midline of the face) of a face identification task with the same stimuli, viewing distance, and monitor calibration as in the main experiments of the present article. The data were obtained from a set of 20 different observers who did not participate in the main experiments of the study. Each observer participated in 110 trials per fixation. The values used for parameters du , dd , dh , and n_0 , respectively, are $3E-6$, $1.5E-6$, $5E-5$, and 5.2 . The Akaike information criterion (AIC; Akaike, 1974), which takes into account the variance for each data point, is used as a distance measure. See Figure 4 for the fit. The same parameters are used for the S-FIO, except for the internal noise parameter, which shifts the performance curve downward or upward but does not significantly alter the shape of the curve and the relative rank order of accuracies across fixation points. The circular plots between a

and b in Figure 5 show examples of 2D contrast sensitivity functions at two different locations with respect to the fixation position. Contrast sensitivity functions that correspond to the center of fixation preserve the higher spatial frequencies (Fig. 5, higher contrast in red in the plots), while contrast sensitivity functions that are far from the fixation position act as low-pass filters and mostly leave the low spatial frequencies (Fig. 5, low contrast in blue in the plots).

Here, we run a face identification task with a set of 10 front-view face images that are normalized for the position of the eyes and chin as well as for contrast (see the Stimuli subsection of Human psychophysics studies above for details). For a given fixation point, k , all of the face images $\{f_1, \dots, f_{10}\}$ are updated with a simulated scotoma centered at that fixation point resulting in a set of fixation-dependent templates, $\{s_1, \dots, s_{10}\}$. The simulated scotoma is created by first low-pass filtering f_i (where i

S-FIO Flow Chart

Figure 5. A summary of the process of the computations in the S-FIO for two fixations. The top panels show a fixation point that is below the eyes, which is optimal without a scotoma, but suboptimal in the presence of a scotoma. The bottom panels show a fixation that is at the tip of the nose, which is one of two optimal points in the presence of a scotoma. **A–C.** The filtering operation for a noiseless template. **A.** A face image is conceptually divided into bins that correspond to specific CSFs as a function of retinal eccentricity. Contrast sensitivity functions that correspond to the center of fixation preserve the higher spatial frequencies (seen as a higher contrast in red in the CSF plots), while contrast sensitivity functions that are far from the fixation position act as low-pass filters and mostly leave the low spatial frequencies (seen as a low-contrast blue in the CSF plots). **B.** The image is transformed into the frequency domain, filtered separately by each possible CSF (here only two are shown), and then transformed back into the spatial domain, resulting in a set of differently filtered images corresponding to each bin. **C.** Corresponding bins are then extracted from the filtered images and input into a composite image that simulates foveation. The procedures in A–C are then repeated for each of the rest of the noiseless face images, as well as for the noisy input to the model on a particular trial. A set of response variables are then calculated, from which a set of likelihoods is found of each face given the noisy image input. **D.** A decision of which face was shown is made by taking the maximum likelihood. Across many trials, a set of proportion correct (PC) values is found, one for each fixation point, and then combined into a heatmap. iFFT, Inverse FFT.

represents 1 of 10 faces) with a Gaussian kernel with a side length of 4° of visual angle and an SD of 3.7° , resulting in a blurred image, $\mathbf{b}_i$. An element-by-element weighted sum at each pixel of the original image, $\mathbf{f}_i$, and the blurred image, $\mathbf{b}_i$, is then computed. The weights are taken from a weights matrix, $\mathbf{w}_k$, which is found by placing a circular area, 8° in diameter, centered at fixation point k , with values of 1 that is surrounded by values of 0 and then filtering it with a square Gaussian kernel, sized 8° with an SD of 1° . The weights are then used in the following way to create $\mathbf{s}_{i,k}$:

$$\mathbf{s}_{i,k} = \mathbf{f}_i \circ \mathbf{w}_k + \mathbf{b}_i \circ (1 - \mathbf{w}_k). \quad (1.2)$$

This results in a set of templates that consists of the combination of the original set of face images with a scotoma, centered at fixation point k (Fig. 5b, example of a face image with a superimposed scotoma at two different fixation positions). The scotoma itself consists of a local average of the original images and has blurred edges. In the experiment, the scotoma is created in the same way and with the same parameters, except MATLAB Psychtoolbox automatically displays a weighted sum of images based on transparency values that act as weights. See Figure 1 for an example of what the scotoma looks like in Psychtoolbox.

The only difference between the FIO and the S-FIO is that in the S-FIO, the templates (as well as the signals created from the templates) contain a fixation-dependent scotoma inserted into the face image rather than a fixation-independent face image. On each trial of the simulation, the same contrast and additive white noise that was used for humans is then added to a chosen template, i , before being linearly filtered with the SVCSF and corrupted with additional internal white noise to become the input data, $\mathbf{g}_k$, to the ideal observer, as follows:

$$\mathbf{g}_k = \mathbf{E}_k (\mathbf{s}_{i,k} + \mathbf{n}_{\text{ex}}) + \mathbf{n}_{\text{in}}, \quad (1.3)$$

where $\mathbf{n}_{\text{ex}}$ is the external Gaussian white noise, $\mathbf{n}_{\text{in}}$ is the internal Gaussian white noise, and $\mathbf{E}_k$ is the linear operator that simulates the fixation-dependent foveation of the input. $\mathbf{E}_k$ describes a set of filtering operations, followed by extraction and recombining parts of the filtered noisy templates in the following way: here, for ease of notation we will describe $\mathbf{E}_k$ as it acts on a random noise-free template, $\mathbf{s}_{i,k}$, rather than a signal with noise present. However, the computations are the same for a chosen template with added noise. Each combination of eccentricity (r) and direction (θ) from fixation defines its own CSF. The complete set of CSFs can be described in one equation, which we refer to as the SVCSF, where r and θ remain variables. Due to computational constraints, each image is divided into small bins with a single CSF assigned to each bin. The template, $\mathbf{s}_{i,k}$, is separately filtered 480 times (30 eccentricities and 16 directions) corresponding to the different CSF functions (and bins) to produce a set of noisy filtered signals. Figure 5a shows an example of two fixations (one at the eyes, and another at the nose) where each face image is conceptually divided into the 480 bins that correspond to different CSF functions relative to the fixation position. Each signal is filtered by taking its fast Fourier transform (FFT), multiplying it on an element by element basis with the corresponding contrast sensitivity function, CSF_b , and then transforming it back into the spatial domain using the inverse FFT, resulting in a noisy filtered image, $\mathbf{s}_{i,k,b}$, where b represents the spatial parameters that correspond to bin b (Fig. 5b,c, examples of filtered images corresponding to two different CSFs are shown between panels):

$$\mathbf{s}_{i,k,b} = \text{IFFT}(\text{FFT}(\mathbf{s}_{i,k} \circ \text{CSF}_b)). \quad (1.4)$$

A composite foveated image, $\tilde{\mathbf{s}}_{i,k}$, is then formed by extracting the regions of each $\mathbf{s}_{i,k,b}$ image for the corresponding angle and eccentricity and placing them into $\tilde{\mathbf{s}}_{i,k}$ (Fig. 5c).

Due to the foveation procedure using the $\mathbf{E}_k$ operator, spatial correlations are formed on the additive white noise field. In general, it is optimal (maximizes decision accuracy) to use templates that undo the spatial correlation through a process known as prewhitening (Barrett et al., 1993; Burgess, 1994; Eckstein et al., 2000). When using the prewhitening process, correlations are usually corrected by applying various transformations or incorporated into the templates. However, when modeling humans, a common model used is one in which the observer uses templates that match the filtering operations of the visual system (Burgess, 1994; Zhang et al., 2004; Peterson and Eckstein, 2012). This modeling approach is known as non-prewhitening with an eye filter and uses templates that match each possible signal with the filtering by the human visual system. Given the template responses, $\{\mathbf{r}_{1,f,k}, \dots, \mathbf{r}_{10,f,k}\}$, we can find the covariance matrix relating them. This is done by first filtering both the original face template, $\mathbf{s}_{i,k}$, and the input signal plus noise, $\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}}$, and then taking the dot product between them as follows:

$$\begin{aligned} \mathbf{r}_{f,k} &= \{\mathbf{r}_{1,f,k}, \dots, \mathbf{r}_{10,f,k}\} \\ &= \{(\mathbf{E}_k \mathbf{s}_{1,k})^T (\mathbf{E}_k (\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}}) + \mathbf{n}_{\text{in}}), \dots, (\mathbf{E}_k \mathbf{s}_{10,k})^T (\mathbf{E}_k (\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}}) + \mathbf{n}_{\text{in}})\} \\ &= \{\mathbf{E}_k \mathbf{s}_{1,k}, \dots, \mathbf{E}_k \mathbf{s}_{10,k}\}^T (\mathbf{E}_k (\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}}) + \mathbf{n}_{\text{in}}). \end{aligned} \quad (1.5)$$

Using the Bayes rule, the S-FIO finds a set of posterior probabilities, one for each hypothesis that face f was shown, H_f , given a set of responses $\mathbf{r}_{f,k}$. The maximum posterior probability is then chosen as follows:

$$\text{decision} = \underset{f}{\operatorname{argmax}} (P(H_f | \mathbf{r}_{f,k})). \quad (1.6)$$

Since there is an equal chance of each face being shown, the prior probabilities, $P(H_f)$, are uniform and the calculation of the posterior, $P(H_f | \mathbf{r}_{f,k})$, is reduced to finding the likelihood, $P(\mathbf{r}_{f,k} | H_f)$, of the set of responses, given the presence of each face, f , and the observer's fixation at spatial location, k , as follows:

$$P(H_f | \mathbf{r}_{f,k}) = \frac{P(\mathbf{r}_{f,k} | H_f) P(H_f)}{P(\mathbf{r}_{f,k})} \propto P(\mathbf{r}_{f,k} | H_f) = l_{f,k}. \quad (1.7)$$

Calculating the likelihood requires knowing the statistical distribution (means, variances, and covariances) of the template response ($\mathbf{r}_{f,k}$). Noting that $\mathbf{E}_k$ is a linear operator, we will write the distribution of the response, $\mathbf{r}_{i,f,k}$, of the template, i , given face, f , using simpler notation. We will denote a single-filtered template, $\mathbf{s}_{i,k}$, as $\hat{\mathbf{s}}_{i,k}$ and a double-filtered template as $\check{\mathbf{s}}_{i,k}$:

$$\begin{aligned} \mathbf{r}_{i,f,k} &= (\mathbf{E}_k \mathbf{s}_{i,k})^T (\mathbf{E}_k (\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}})) \\ &= (\mathbf{E}_k^2 \mathbf{s}_{i,k})^T (\mathbf{s}_{f,k} + \mathbf{n}_{\text{ex}}) + (\mathbf{E}_k \mathbf{s}_{f,k})^T \mathbf{n}_{\text{in}} \\ &= \check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k} + \check{\mathbf{s}}_{i,k}^T \mathbf{n}_{\text{ex}} + \hat{\mathbf{s}}_{f,k}^T \mathbf{n}_{\text{in}}. \end{aligned} \quad (1.8)$$

By using zero mean white noise, we are able to write a simple one-term expression for the mean of the $\mathbf{r}_{i,f,k}$ distribution, as follows:

$$\begin{aligned} \boldsymbol{\mu}_{i,f,k} &= E[\mathbf{r}_{i,f,k}] \\ &= E[\check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k} + \check{\mathbf{s}}_{i,k}^T \mathbf{n}_{\text{ex}} + \hat{\mathbf{s}}_{f,k}^T \mathbf{n}_{\text{in}}] \\ &= E[\check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k}] + E[\check{\mathbf{s}}_{i,k}^T \mathbf{n}_{\text{ex}}] + E[\hat{\mathbf{s}}_{f,k}^T \mathbf{n}_{\text{in}}] = \check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k}, \end{aligned} \quad (1.9)$$

where $E[\bullet]$ is the expectation operator. The mean of $\mathbf{r}_{f,k}$ when face f is chosen is then the vector $\boldsymbol{\mu}_{f,k} = \{\boldsymbol{\mu}_{1,f,k}, \dots, \boldsymbol{\mu}_{10,f,k}\}$. We are now able to find the covariance between each set of i th and j th responses that is independent of the presented face f as follows:

$$\begin{aligned} \sum_{i,j,k} &= \operatorname{cov}(\mathbf{r}_{i,f,k}, \mathbf{r}_{j,f,k}) \\ &= E[(\mathbf{r}_{i,f,k} - E[\mathbf{r}_{i,f,k}])(\mathbf{r}_{j,f,k} - E[\mathbf{r}_{j,f,k}])] \\ &= E[(\check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k} + \check{\mathbf{s}}_{i,k}^T \mathbf{n}_{\text{ex}} + \hat{\mathbf{s}}_{f,k}^T \mathbf{n}_{\text{in}} - \check{\mathbf{s}}_{i,k}^T \mathbf{s}_{f,k}) \\ &\quad (\check{\mathbf{s}}_{j,k}^T \mathbf{s}_{f,k} + \check{\mathbf{s}}_{j,k}^T \mathbf{n}_{\text{ex}} + \hat{\mathbf{s}}_{f,k}^T \mathbf{n}_{\text{in}} - \check{\mathbf{s}}_{j,k}^T \mathbf{s}_{f,k})] \\ &= \check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} E[\mathbf{n}_{\text{ex}}^2] + (\check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} + \check{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k}) E[\mathbf{n}_{\text{ex}}^T \mathbf{n}_{\text{in}}] + \hat{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k} E[\mathbf{n}_{\text{in}}^2]. \end{aligned} \quad (1.10)$$

Using the property of the expectation of independent random variables, $E[XY] = E[X]E[Y]$, and the fact that $E[\mathbf{n}_{\text{ex}}]$ and $E[\mathbf{n}_{\text{in}}]$ are both equal to zero, we reduce the middle expression in Equation 1.10 as follows:

$$(\check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} + \check{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k}) E[\mathbf{n}_{\text{ex}}^T \mathbf{n}_{\text{in}}] = 0. \quad (1.11)$$

Then, using the property $\operatorname{Var}(X) = E[X^2] - (E[X])^2$, we reduce the first term in Equation 1.10, as follows:

$$\begin{aligned} \check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} E[\mathbf{n}_{\text{ex}}^2] &= \check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} (\operatorname{Var}(\mathbf{n}_{\text{ex}}) + (E[\mathbf{n}_{\text{ex}}])^2) \\ &= \sigma_{\text{ex}}^2 \check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k}, \end{aligned} \quad (1.12)$$

where σ_{ex}^2 is the variance chosen for the external noise. Similarly, the third term in Equation 1.10 is reduced as follows:

$$\begin{aligned} \hat{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k} E[\mathbf{n}_{\text{in}}^2] &= \hat{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k} (\operatorname{Var}(\mathbf{n}_{\text{in}}) + (E[\mathbf{n}_{\text{in}}])^2) \\ &= \sigma_{\text{in}}^2 \hat{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k}. \end{aligned} \quad (1.13)$$

This results in a simple expression for the covariance that consists of double-filtered templates, single-filtered templates, and the variances for the external and internal noise as follows:

$$\sum_{i,j,k} = \sigma_{\text{ex}}^2 \check{\mathbf{s}}_{i,k}^T \check{\mathbf{s}}_{j,k} + \sigma_{\text{in}}^2 \hat{\mathbf{s}}_{i,k}^T \hat{\mathbf{s}}_{j,k}. \quad (1.14)$$

Due to the use of Gaussian external noise, the response vector $\mathbf{r}_{f,k}$ for a trial, given that a face f comes from a multivariate normal distribution for which we know the mean and covariance matrix, as follows:

$$\mathbf{r}_{f,k} \sim \operatorname{MVN}(\boldsymbol{\mu}_{f,k}, \sum_k). \quad (1.15)$$

Knowing $\mathbf{r}_{f,k}$, $\boldsymbol{\mu}_{f,k}$, and $\sum_k$ allows us to find the likelihood $l_{f,k}$ of the responses using the multivariate Gaussian probability density function, as follows:

$$l_{f,k} = \exp\left(-\frac{1}{2}(\mathbf{r}_{f,k} - \boldsymbol{\mu}_{f,k})^T \sum_k^{-1}(\mathbf{r}_{f,k} - \boldsymbol{\mu}_{f,k})\right). \quad (1.16)$$

The normalizing factor is left out because it is just a scaling factor that is unnecessary for comparing the likelihoods.

Here, we run a simulation with 100,000 trials. Due to computational constraints, we only run the simulation for fixations corresponding to every 10th pixel, which results in a 50×50 performance map. This map is then resized using bilinear interpolation to a 500×500 pixel performance map to match the size of the face images (Fig. 5d).

Bayesian ideal observer

In addition to running an S-FIO and FIO, we also run a standard ideal observer model, which uses image information to achieve the highest possible performance and does not simulate the foveation of the visual system. In contrast with Equation 1.3, the data, $\mathbf{g}$, is now the sum of a random (1 of 10) face template, $\mathbf{s}_i$, and external noise, $\mathbf{n}_{\text{ex}}$, as follows:

$$\mathbf{g} = \mathbf{s}_i + \mathbf{n}_{\text{ex}}. \quad (2.1)$$

The ideal observer does not have any sources of suboptimality such as internal noise or filtering operations on the face template, $\mathbf{s}_i$, that models

foveation or a scotoma. Using the Bayes rule, the ideal observer finds a set of posterior probabilities, one for each hypothesis, H_f , that face, f , was shown, given the image data, $\mathbf{g}$. The maximum posterior probability is then chosen as follows:

$$\text{decision} = \underset{f}{\operatorname{argmax}}(P(H_f|\mathbf{g})). \quad (2.2)$$

Since there is an equal chance of each face being shown, the prior probabilities, $P(H_f)$, are uniform, and the calculation of the posterior, $P(H_f|\mathbf{g})$, is reduced to finding the likelihood, $P(\mathbf{g}|H_f)$, of the image data, $\mathbf{g}$, given the presence of each face, f , as follows:

$$P(H_f|\mathbf{g}) = \frac{P(\mathbf{g}|H_f)P(H_f)}{P(\mathbf{g})} \propto P(\mathbf{g}|H_f) = l_f. \quad (2.3)$$

On each trial, independent Gaussian noise, σ_{ex} , is added to each pixel, $s_{i,p}$, of a random face template, $\mathbf{s}_f$, where p indexes 1 to n (500^2) pixels in the 500×500 image, resulting in a noisy image, $\mathbf{g}$. At the pixel level, the likelihood, $l_{f,p}$, of an individual pixel, g_p , of the data coming from pixel, $s_{f,p}$, in face template, f , is as follows:

$$l_{f,p} = \frac{1}{\sqrt{2\pi\sigma_{\text{ex}}^2}} \exp\left(-\frac{(g_p - s_{f,p})^2}{2\sigma_{\text{ex}}^2}\right). \quad (2.4)$$

As a result of the statistical independence of the image noise, the likelihood, l_f , of the data, $\mathbf{g}$, given the presence of the f th face can be written as a product of the likelihoods of individual pixels, which reduces to a simpler expression involving the original signal template, $\mathbf{s}_f$, the image data, $\mathbf{g}$, and the external noise SD, σ_{ex} , as follows:

$$\begin{aligned} P(\mathbf{g}|H_f) &= P(g_1, \dots, g_n|H_f) = \prod_p P(g_p|H_f) \\ &= \prod_p \frac{1}{\sqrt{2\pi\sigma_{\text{ex}}^2}} \exp\left(-\frac{(g_p - s_{f,p})^2}{2\sigma_{\text{ex}}^2}\right) \propto \prod_p \exp\left(-\frac{g_p^2 + 2g_p s_{f,p} - s_{f,p}^2}{2\sigma_{\text{ex}}^2}\right) \\ l_f &\propto \exp\left(-\frac{\mathbf{g}^T \mathbf{g} + 2\mathbf{g}^T \mathbf{s}_f - \mathbf{s}_f^T \mathbf{s}_f}{2\sigma_{\text{ex}}^2}\right) \propto \exp\left(\frac{2\mathbf{g}^T \mathbf{s}_f - \mathbf{s}_f^T \mathbf{s}_f}{2\sigma_{\text{ex}}^2}\right). \end{aligned} \quad (2.5)$$

Efficiency calculation

It is often useful to assess how well a human performs relative to the upper bound of performance by calculating the absolute efficiency. This is calculated by taking the squared contrast thresholds of humans relative to the ideal observer to achieve a given performance (e.g., experimental accuracy achieved by the human observer; Barlow, 1980; Burgess et al., 1981; Tjan et al., 1995; Pelli et al., 2006).

$$\text{Efficiency} = \frac{C_{\text{ideal}}^2}{C_{\text{human}}^2}. \quad (2.6)$$

For example, if comparing the efficiency between a human observer and an ideal observer, an ideal observer requires much lower signal contrast, C_{ideal} , than humans, C_{human} , to match the experimentally measured human performance.

In addition, the relative efficiency of the human observer relative to the FIO (Peterson and Eckstein, 2014) and S-FIO is similarly defined, as follows:

$$\text{Efficiency} = \frac{C_{\text{FIO}}^2}{C_{\text{human}}^2}, \quad (2.7)$$

where C_{FIO} is replaced by $C_{\text{S-FIO}}$ for doing the calculation relative to the S-FIO.

Results

S-FIO predicts new optimal point of fixation

We first evaluated the theoretical predictions for an optimal point of fixation for an S-FIO model, which took into account the foveated nature of the visual system and the presence of the cen-

tral scotoma. The model incorporated a spatially variant filtering of the visual input and integrated the information in face images optimally to compute posterior probabilities and make trial-to-trial decisions about the identity of the face presented (Fig. 5; Materials and Methods for mathematical details). The parameters for the eccentricity-dependent contrast sensitivity function of the S-FIO were obtained by fitting the standard FIO model (no scotoma) to an independent dataset of 20 different observers participating in a forced-fixation experiment with no simulated scotoma (see Materials and Methods; Fig. 4). Figure 6a shows the S-FIO predictions for identification accuracy as a function of point of fixation. For comparison, we also show the model predictions for the FIO without the scotoma. The internal noise, which was modeled as an additive random Gaussian variable added to the decision variables, was adjusted so that the peak performances are comparable for the FIO and S-FIO in Figure 6a. The results show that the presence of the central scotoma alters the original theoretical optimal point of fixation (just below the eyes), which now leads to the lowest performance of the four experimentally evaluated points of fixation. The fixation points with the highest S-FIO predicted accuracy were at the forehead and nose. The S-FIO makes a strong prediction of how human identification performance and eye movements might change with the simulated central scotoma and generalizes across a large range of internal noise values.

Experiment 1: human perceptual performance and eye movements with simulated central scotoma

A group of 16 observers participated in a face identification task with conditions that evaluated whether the simulated scotoma altered the human initial preferred point of fixation as well as identification performance as a function of fixation position. These conditions were counterbalanced through 64 blocks (125 trials/block) to avoid confounding learning effects and to determine consistency in preferred points of fixation. See Figure 2 for a task time line.

Scotoma-absent, free-viewing

In the first condition of the experimental blocks, participants viewed faces without a scotoma and were able to make free eye movements. Two different stimulus presentation times were used in separate blocks. We used a short presentation time of 350 ms to assess a preferred fixation location for a single saccade. We then used a longer presentation time of 1500 ms, which allowed for several saccades, to determine whether the expected stimulus presentation time would change the initial fixation strategy. We found no effect of presentation time on first fixation position ($F_{(1,42)} = 0.023$, $p = 0.879$). However, there was an increase in performance with the longer viewing duration since more information was gathered in additional saccades (Fig. 6b). On their first saccade, all of the observers tended to fixate a region slightly below the eyes, which is consistent with the previous results from Peterson and Eckstein (2012), showing that this behavior is observed in $\sim 90\%$ of subjects. There was also no observed significant difference between the mean fixation position across the scotoma-absent blocks and the theoretical optimal point of fixation for the FIO ($t_{(15)} = 0.262$, $p = 0.797$, two-tailed), with an average magnitude difference of 0.08° .

Scotoma-present, free-viewing before forced-fixation study

In the second condition, the same experimental setup was used except that a gaze-contingent simulated circular scotoma 8° in diameter was centered on the fovea during the stimulus presentation pe-

Figure 6. *a*, Comparison between the 2D performance (proportion correct) maps of an FIO and an S-FIO. The S-FIO shows two new optimal points of fixation (one at the tip of the nose and the other at the top of the forehead) and decreased performance between the eyes and nose, which is the original optimal point for the FIO. Observers' average fixation positions from free-viewing tasks with and without a scotoma, respectively, are overlaid to compare model predictions to observers' behavior. *b*, A comparison of average performance during scotoma-absent and scotoma-present conditions with two different viewing durations is shown. There is an overall decrease in performance for the shorter stimulus duration and a larger performance difference between the scotoma-absent and scotoma-present conditions when compared with the longer stimulus duration. *c*, The progression of average vertical fixation locations for the face identification task across blocks is shown. The scotoma is introduced at block 13, where the shaded region starts. The average fixation position when the scotoma is present remains far from the theoretical optimal points found from the S-FIO, shown as green lines. Instead, the fixations for all conditions remain close to the theoretical optimal point found from the FIO, shown in yellow. *d*, The high correlation between average vertical fixation locations in degrees (deg) below the eyes during scotoma-absent vs scotoma-present blocks is shown. *e*, Individual initial fixations for each observer are shown along with a block average (red) for the last block of the scotoma-absent condition and last block of the scotoma-present condition. The fixations for each observer represent a unimodal distribution rather than a bimodal one in both the scotoma-absent and scotoma-present conditions. Here, the pixel average of the 10 faces that were used in the experiment is used as the face in the background. *f*, The progression of average distance away from observers' preferred point of fixation for the free-viewing task is shown. Obs, Observer.

riod. The scotoma consisted of a local average of the face image behind it and contained gradually fading blurred edges (for further details, see the description of the S-FIO model in Materials and Methods).

Effects on perceptual performance. The presence of the scotoma led to a significant decrease in performance in both the short 350 ms duration ($t_{(15)} = 7.124$, $p = 3.48 \times 10^{-6}$, one-tailed) and the long 1500 ms duration ($t_{(15)} = 6.84$, $p = 5.56 \times 10^{-6}$, one-tailed). There was also a significant interaction of stimulus duration with scotoma presence ($F_{(1,42)} = 15.37$, $p = 3.2 \times 10^{-4}$). Figure 6*b* shows that the absolute difference was smaller in the long duration (5% vs 9% difference in proportion correct). This may be explained by the fact that there is enough time to gather information from multiple saccades, which helps to decrease the detrimental effects of the scotoma. Even with a single fixation to a spot that is now suboptimal, the performance in this task is far above the chance performance of 10%. This may be explained by the fact that participants had lots of training with the same 10 faces and are able to use less prominent features of the face, like the mouth and chin.

A confusion matrix averaged across observers was calculated for each presentation time and showed that, although there was some variability in performance for different faces, even the lowest-scoring face (37% correct for one of the faces in the scotoma-present, 350 ms condition) was a far above chance performance. The confusion matrices for individual subjects were also analyzed with statistical comparisons (using a multinomial likelihood ratio test; Williams, 1976). For each participant, the confusion matrix for the scotoma-absent conditions was compared with the scotoma-present condition for each stimulus presentation time (1.5 and 0.35 s) separately. The diagonal terms of each confusion matrix were taken out, and the remaining entries were renormalized so that the comparison would be focused on possible differences between the confusion of certain faces with other faces rather than an overall difference in performance. No statistical significance was found in the confusion matrices in the scotoma-absent versus the scotoma-present conditions for either presentation time in all but one of the participants (only for the 0.35 s presentation time).

We also calculated the efficiency (see Materials and Methods) of the average of human observers in the 350 ms scotoma-absent and 350 ms scotoma-present free-viewing conditions compared with three different model observers; an ideal observer without foveation or a scotoma, an FIO without a scotoma, and an S-FIO. For the FIO and S-FIO, the fixation position used to calculate performance was taken from an average of the human preferred fixation points. In addition, the internal noise parameter in the FIO and S-FIO was set to zero to have a more accurate comparison with the ideal observer. For each model observer (ideal observer, FIO, and S-FIO), the efficiency was higher in the scotoma-absent condition (0.0018, 0.0036, and 0.0109, respectively) versus the scotoma-present condition (0.0014, 0.0026, and 0.0078, respectively), suggesting some degradation in the human ability to integrate information across facial features with the scotoma. Within each condition, the efficiency values increased from the ideal observer to the FIO and then to the S-FIO. This occurred because human performance became better relative to models with additional components that limit performance. The ideal observer was only limited by external noise. However, the foveation in the FIO added a source of suboptimality with the filtering operations across the visual field. The S-FIO had an additional source of suboptimality due to the presence of the scotoma, which decreased performance even further relative to the FIO.

Effects on first fixation. The results show that observers' initial point of fixation to faces with the scotoma differed significantly from that predicted by the S-FIO (Fig. 6*a*). Furthermore, we did not observe a significant difference in the initial point of fixation to faces with the inclusion of a scotoma (scotoma-absent, free-viewing vs scotoma-present, free-viewing before forced-fixation ($t_{(15)} = 1.54$, $p = 0.266$, two-tailed for the vertical spatial position; Fig. 6*c*; and $t_{(15)} = 1.285$, $p = 0.218$, two-tailed for the horizontal position) even after 3000 trials. Figure 6*d* also shows this in the form of a strong correlation ($r^2 = 0.86$) in the vertical position. The mean of the scotoma-present initial fixations did not reflect a bimodal distribution around the forehead and nose, but instead the individual fixations were clustered around the mean initial point of fixation (Fig. 6*e*, fixation distributions for each observer). There was also no observed significant difference between the mean fixation position across the scotoma-present blocks and the theoretical optimal point of fixation for the FIO ($t_{(15)} = 1.65$, $p = 0.12$, two-tailed, with an average magnitude difference of 0.48°; Fig. 6*f*).

Scotoma-present, forced-fixation

The free-viewing study showed no effect of the scotoma on the initial fixation position and also showed a discrepancy between the location where observers initially fixated (just below the eyes) and the theoretical prediction of the S-FIO (forehead or nose). This discrepancy with the S-FIO model might suggest an inability of humans to learn the new optimal points of fixation. However, it may also reflect that the S-FIO model prediction is simply wrong. In this view, the optimal strategy with the scotoma still remains to fixate just below the eyes, but the S-FIO simply incorrectly predicts a new optimal point of fixation. To test this possibility, in a third condition we assessed whether there was a difference in identification accuracy with a simulated scotoma at four different points of fixation (approximately corresponding to the forehead, eyes, tip of the nose, and mouth) by forcing observers to fixate each position during the duration of the trial using only the 350 ms presentation time (with feedback). We found that there was a general agreement with the S-FIO and a significant increase in performance when observers were forced to fixate the nose versus the eyes ($t_{(7)} = 4.309$, $p = 0.004$, one-tailed; 68% vs 52%) as well as the forehead versus the eyes ($t_{(7)} = 3.023$, $p = 0.02$, one-tailed; 65% vs 52%).

Comparison of Forced-fixation human data to S-FIO model

To quantitatively compare how human and S-FIO performance varied with fixation position, we varied the internal noise in the S-FIO model to degrade its performance to fit the human data (minimize the AIC; Akaike, 1974). All parameters related to the contrast sensitivity function remained the same as in Figure 4 based on fitting an independent set of observers participating in a forced-fixation study with no scotoma (see Materials and Methods). The continuous line in Figure 7*a* shows the predicted model performance of S-FIO, which successfully predicts the effect of the scotoma on the accuracy of the various points of fixation. Note that the level of internal noise shifted the curve downward or upward but did not significantly alter the shape of the curve and the relative rank order of accuracies across fixation points. These results suggest that the lack of change in human initial points of fixation with a scotoma reflects suboptimal oculomotor learning.

Scotoma-present free viewing after forced fixation

We repeated the scotoma-present, free-viewing condition to determine whether subjects would adapt their fixation strategy after

Figure 7. *a*, Plot of human performance while being forced to fixate at different points down the vertical mid-line of the face from the top of the forehead to the chin. Chance performance is at 0.1. The S-FIO model is fit using parameters from the FIO model, with only the level of noise being adjusted. The black bar shows 1 SD surrounding the average of observers' preferred point of fixation during the free-viewing scotoma-present post-forced-fixation condition. *b*, The progression of average vertical fixation locations for the face identification task, including data collected after the forced-fixation condition is shown. The scotoma is present in the shaded regions of the plot. *c*, The high correlation in vertical distance of degrees (deg) below the eyes between the scotoma-present conditions before forced fixation and the scotoma-present condition after forced fixation is shown.

being exposed to a significant increase in performance at another location (with trial feedback about correct decision). We found that subjects retained their initial preferred fixation positions (scotoma-absent, free-viewing condition vs current condition;

$t_{(7)} = 1.74$, $p = 0.126$, two tailed for the vertical position; and $t_{(7)} = 0.403$, $p = 0.699$, two tailed for the horizontal position) as seen in Figure 7*b*. Figure 7*c* also shows this in the form of a strong correlation ($r^2 = 0.92$) in the vertical position and horizontal

position ($r^2 = 0.62$) of fixation points in the scotoma-absent, free-viewing condition versus the post-forced fixation, scotoma-present, free-viewing condition.

Experiment 2: comparison of influence of simulated scotoma in face versus object-following/search tasks

In Experiment 1, we found that, even though the optimal point of fixation of humans (and the S-FIO) was altered by the presence of a simulated scotoma, observers did not compensate in their initial eye movements and kept directing their fovea toward their initial preferred point of fixation. To assess whether this inflexibility in eye movements is unique to the face identification task, a different set of observers in Experiment 2 alternated between the face identification task and two other tasks (the object-following and search tasks; Kwon et al., 2013) that have previously been shown to modify eye movement strategies in the presence of a simulated scotoma. See Figure 3 for a task time line. We compared changes in preferred fixation in the two tasks simultaneously by running them in alternating blocks (22 blocks each; 125 trials/block for the face task; 30 trials/block for the object-following and search tasks) and observing changes in eye movements. In the object-following task, participants followed a small random object around the screen as it moved to several different locations. In the search task, participants searched for the presence of a specific small object in an array of other small objects with a cluttered background and responded with a yes/no response.

Figure 8, *a* and *b*, shows retinal field plots that represent the locations on the retina that are used for fixating the preferred initial fixation location on the stimuli. This preferred stimulus location for each individual observer is used as the center reference point for the retinal field plots and was found from the control (no-scotoma) blocks. For the object-following subtask (Fig. 8*a*), the preferred stimuli location was an average of the fixation locations to objects (averaged across blocks and objects) in the object-following no-scotoma blocks (e.g., some observers may consistently look a little to the right of the object center, while others look a little below the object center). For the face task (Fig. 8*b*), the preferred location was an average of the fixation locations to faces in the no-scotoma blocks. Figure 8*a* shows a higher use of peripheral parts of the retina after the introduction of a simulated scotoma compared with control blocks. A measure of the bivariate contour ellipse area (Crossland et al., 2004) was used to quantify the use of peripheral locations to process task-relevant visual information. We found an increase in the contour ellipse area for the object-following task ($t_{(6)} = 8.94$, $p < 0.001$). For the face task, Figure 8*b* shows no change in the contour ellipse area ($t_{(6)} = 0.98$, $p = 0.36$), in agreement with Experiment 1, suggesting no change in eye movements for the face task when the simulated scotoma is introduced.

Figure 8*c* further shows that observers' fixation positions remain consistent (close to the identity line on the plot) between the no-scotoma and scotoma blocks in the face task ($t_{(6)} = 1.22$; $p = 0.27$), while this consistency is not seen in the object-following subtask ($t_{(6)} = 4.32$; $p < 0.01$). In the face task, the fixed reference point to measure the fixation locations was the position of the eyes, while in the object-following subtask the chosen reference point was the center of the object.

We also measured the increase in overall distance away from the initial preferred fixation point (Fig. 8*d*), which was the same reference point as used in the retinal field plots (Fig. 8*a,b*) and was taken from an average of the fixations in the control (no-scotoma) blocks, as explained above. As shown in the first exper-

iment, observers retained their preferred fixation positions to faces in the presence of a simulated scotoma when compared with control blocks as measured in fixation distance to the preferred point ($t_{(6)} = 1.4$, $p = 0.21$; Fig. 8*d*). However, in the object-following subtask the fixation distance to the preferred point of fixation to an object increased ($t_{(6)} = 20.98$, $p = 7.6 \times 10^{-7}$), suggesting that observers developed a peripheral viewing strategy after the simulated scotoma was introduced.

We also found evidence of a change of eye movement strategy for the search task. Both the number of eye movements (Fig. 8*e*) and search times (Fig. 8*f*) to find the target first increased with the introduction of the scotoma but then decreased with practice similarly as found by Kwon et al. (2013). Analyzing the influence of the scotoma on the actual retinal locations used to scrutinize the multiobject search task is problematic because observers often fixate in between objects increasing average distances of the fovea to the center of objects including the target (Fig. 8*e*). However, we still calculated a few metrics that suggested a change in eye movement strategy (not reported in Kwon et al., 2013). The average distance to the preferred point of fixation (the center of the target object in target-present trials) in the search subtask increased with the presence of the scotoma ($t_{(6)} = 7.76$, $p = 2.4 \times 10^{-4}$), suggesting compensatory eye movements. In addition, the distance of the closest saccade in a trial to the target object also increased ($t_{(7)} = 5.16$, $p = 2.1 \times 10^{-3}$) as well as the distance of every object center in both target-present and target-absent trials to its closest saccade ($t_{(7)} = 4.39$, $p = 4.58 \times 10^{-3}$). Together, the results suggest a dissociation in the observers' compensatory eye movements for faces versus other tasks (object-following and search tasks) in the presence of a simulated scotoma.

Experiment 3: human perceptual performance and eye movements with simulated central scotoma and small faces

One interpretation of the results in Experiment 2 is that there is a dissociation in adapting eye movements for face tasks versus other tasks in the presence of a simulated scotoma. However, an alternative explanation is that the dissociation may be related to the relative size of the scotoma compared with the stimuli rather than the type of task itself. For the object-following and search tasks, the scotoma completely covered the objects in the scene. However, for the faces task, the scotoma only covered part of the face. As a result, an alternative explanation may be that the learning of compensatory eye movements might be faster and easier when the scotoma covers the entire stimulus (object following and search) than when it covers a smaller part of the stimulus (faces). We assessed whether the size of the face stimulus played a role in the inability of observers to change their initial eye movements and keep directing their fovea toward their initial preferred point of fixation.

In Experiment 3, we repeated a short version of the free eye-movement condition from Experiment 1 with faces that are half the size (scotoma present vs scotoma absent) and six new observers. In the absence of the scotoma, initial fixations into the face were directed to a similar relative location (just below the eyes) within the face as with the larger faces (see also, Peterson and Eckstein, 2013a). We did not observe a significant difference in the initial point of fixation to faces with the inclusion of a scotoma (scotoma-absent vs scotoma-present, $t_{(5)} = 2.09$, $p = 0.09$, two-tailed for the vertical spatial position; Fig. 9*a*; and $t_{(5)} = 1.027$, $p = 0.352$, two-tailed for the horizontal position) after 750 trials. Figure 9*b* also shows this in the form of a strong correlation ($r^2 = 0.82$) in the vertical fixation position between the scotoma-absent and scotoma-present conditions. The correlation for the

Figure 8. *a, b*, Visual field plots for the object-following and face identification tasks are shown. The plots represent the density of retinal locations used to view an object relative to a preferred viewing point that was found from a control scotoma-absent condition. The gray spot in the center of each retinal field represents the size and location of the scotoma on the retina. *c*, A correlation plot of observers' preferred vertical fixation positions relative to a fixed reference point for the object-following and face tasks in the scotoma-absent vs scotoma-present conditions is shown. The fixed reference point for the face task is the position of the eyes, while for the object-following task the reference is the center of the object. *d*, The progression of the distance to the preferred point of fixation across blocks in all tasks is shown. The preferred point is taken from an average of the scotoma-absent blocks (not shaded in gray) for the object-following and face tasks. For the search subtask, the preferred point is taken as the center of the target object in target-present trials. The distance is found by taking an average of the distances of all fixation points in a trial to the center of the target object. Fixations that are outliers (defined as those in the top 10% of the distance distribution) are taken out. *e, f*, Changes in the number of saccades and reaction time used for the search task across blocks for correctly answered trials (both target present and target absent) is shown. Both metrics rise with the introduction of a scotoma and then fall back to being statistically indistinguishable from nonscotoma blocks.

horizontal position ($r^2 = 0.35$) was less strong. The initial fixation position still remained clustered around the eye region of the face, which is seen in example blocks in Figure 9c for each observer. In addition, the average number of saccades that landed inside the face region (the area not covered by the mask) was not significantly different between the scotoma-absent and scotoma-present conditions ($t_{(5)} = 0.879$, $p = 0.42$). The initial fixation position remained unchanged despite a large change in performance between the scotoma-absent and scotoma-present conditions ($t_{(5)} = 11.82$, $p = 7.68 \times 10^{-5}$), which is seen in Figure 9d. These

results suggest that the size of the scotoma relative to the stimulus cannot entirely account for the dissociation in compensatory eye movements measured in Experiments 1 and 2 for faces versus other tasks.

Discussion

There has been recent interest in how humans change their oculomotor plans in response to changes in front-end sensory processing due to retinal diseases such as macular degeneration (Crossland et al., 2005; Kwon et al., 2013; Walsh and Liu, 2014;

Figure 9. *a*, The progression of average vertical fixation locations for the face identification task with small faces is shown. The scotoma is present in the shaded regions of the plot. *b*, The high correlation in vertical distance of degrees below the eyes between the scotoma-absent condition and the scotoma-present condition is shown. *c*, Individual initial fixations for each observer are shown along with a block average (red) for the last block of the scotoma-absent condition and last block of the scotoma-present condition. The fixations for each observer represent a unimodal distribution rather than a bimodal one in both the scotoma-absent and scotoma-present conditions. Here, the pixel average of the 10 faces that were used in the experiment is used as the face in the background. *d*, A comparison of average performance during scotoma-absent and scotoma-present conditions is shown. Performance in the scotoma-present condition is significantly worse. Obs, Observer.

Janssen and Verghese, 2015). For simple tasks, there is ample evidence that patients experiencing central vision loss as well as observers with a simulated scotoma compensate their eye movements and develop a PRL to sample visual information. Here, we asked whether the learned oculomotor compensations might generalize to more complex perceptual tasks where the central scotoma does not occupy the entire visual stimulus. We chose face identification as one such task because of its ecological importance and focused on the initial fixation, which has been shown to be critical and sufficient to achieve close to 90% of asymptotic identification performance (Hsiao and Cottrell, 2008; Or et al., 2015).

Specific areas of the face differ in importance based on the information that needs to be gained for a specific task (i.e., gender discrimination vs identity; Schyns et al., 2002; Smith et al., 2005). For identification, the eyes are the most informative, followed by the nose and mouth (Peterson and Eckstein, 2012). However, the first saccade is driven to a position slightly below the eyes that optimizes perceptual performance by gathering information from several informative regions at once (Peterson and Eckstein, 2012). Although the information obtained from each region is somewhat degraded when fixating just below the eyes relative to foveating the regions individually, it allows an observer to maximize the joint acquisition of information across multiple features

of a face in a single fixation and is predicted by an optimal Bayesian observer constrained by a model of a foveated visual system. However, when the fovea was unavailable due to a visual impairment, the optimal strategy was not known.

We proposed a new theoretical model that takes into account the foveated nature of the visual system, the scotoma, and optimal decisions (S-FIO) to make novel predictions about the theoretical optimal point of fixation to faces for a human with central visual loss. In the presence of a simulated central scotoma, the new theoretical optimal point shifted from just below the eyes to one that is at the tip of the nose and to another that is at the top of the forehead. The fixation at the tip of the nose might seem surprising given that a lower density of retinal ganglion cells in the human upper visual field (Curcio et al., 1990) is often associated with a larger frequency of upward-directed saccades for visual search tasks (Najemnik and Geisler, 2009). The FIO incorporates an asymmetry in sensitivity with a steeper roll-off, with eccentricity at the upper visual field. Yet, the FIO optimal points of fixation are driven by the distribution of information in the faces, and fixating the lower tip of the nose places the scotoma lower in the face, uncovering the highly informative eyes.

Our experimental results confirmed that the model correctly predicts the human fixation points that maximized identification accuracy in the presence of a simulated scotoma. Face identification was improved by fixating, covering, and sacrificing the information at the nose region with the scotoma, while uncovering the highly informative eye region as well as the mouth. However, we found that even when the optimal point of fixation shifted, initial human saccades to faces did not vary significantly after 5000 trials with the simulated scotoma as well as experience in fixating at different points on the faces.

These results are in stark contrast with previous studies, which showed compensatory eye movements with simulated scotomas for tasks such as object following (Kwon et al., 2013), visual search (Walsh and Liu, 2014), and object comparison (Janssen and Verghese, 2015). However, the details of the implementation of the simulated scotoma differed between the current study and the previous study by Kwon et al. (2013). This might account for the dissociation of results between the face identification task and the previously studied object-following task. As a result, we conducted a second experiment in which we further evaluated whether we would observe a generalization of oculomotor learning across different perceptual tasks (face identification, object following, and search) using the same invisible simulated scotoma. Our results showed a dissociation between the eye movements for the face task and the object-following and search tasks. The development of the PRL was less strong and less focused than in the (Kwon et al., 2013) study. There are various possible reasons for this difference. It may be due to the nature of the simulated scotoma. Learning in the study by Kwon et al. (2013) initially consisted of a simulated scotoma with a sharp visible boundary, while in our study it was blurred into the surroundings (a gradient rather than clear borders). The initial training with the visible scotoma might have allowed for a clearer generalization of a PRL for the invisible scotoma conditions. Another possible reason for the weaker PRL development in our study could also be the fact that we interleaved the object-following task with the face identification task, which may have diminished the oculomotor compensation. Aside from these differences across studies, our results showed a strong dissociation in oculomotor compensation in initial eye movements across perceptual tasks.

One interpretation of the lack of oculomotor learning with the faces (Experiments 1 and 2) relates to their larger size, which,

unlike the stimuli in the object-following and search tasks, are not entirely covered by the simulated scotoma. In this view, the lack of learning with faces is not related to some domain-specific effect but to the brain having more difficulty in learning to compensate eye movements when the scotoma does not entirely cover the visual stimuli. Experiment 3 showed some evidence against this interpretation. Even when reducing the size of the faces so that the scotoma entirely covered the face, most observers still programmed their first eye movement toward the inner features of the face. Even though Experiment 3 consisted of fewer blocks of trials than Experiments 1 and 2, the number of trials used were sufficient for observers to change their eye movement strategies for the visual search and object-following tasks (Experiment 2).

Another interpretation of the dissociation in our results may be that compensatory oculomotor commands are not coded in a central brain area (e.g., within the brainstem), where they would be applied to different perceptual tasks. For example, there is evidence that the superior colliculus, a brainstem structure known to be critical for saccade programming, is involved in motor learning and sends signals that determine the direction of adaptive changes in saccade endpoints (Kaku et al., 2009). Instead, our results could suggest task and/or stimuli specificity in the learning of eye movements to compensate for central visual loss.

Face processing has its own dedicated system in the brain (Kanwisher et al., 1997). Face recognition is a highly practiced task for which humans develop fixation strategies that remain very consistent across time (Peterson and Eckstein, 2013b; Mehoudar et al., 2014) and show a consistent optimality and rigidity (Peterson and Eckstein, 2013b, 2014) that is rarely present in simpler artificial laboratory tasks (Verghese, 2012; Ackermann and Landy, 2013; Eckstein et al., 2015). The frequently used fixation strategy with faces might give rise to special fixation-specific representations in face-related brain areas. Recent studies using fMRI physiology (de Haas et al., 2016) and single-cell recordings in monkeys (Issa and DiCarlo, 2012) have shown higher neuronal activation in face-related areas when the features of the face are presented at their most frequent locations relative to the fixation point. A possible interpretation is that information about where to fixate faces may be represented in domain-specific areas. Such information might then be relayed to the superior colliculus, which has some selectivity for faces (Davies-Thompson and Andrews, 2012). In such a neural architecture, learning new eye movements to faces might require changes in the properties of fixation-specific neurons in face-selective areas, while a sole change in a more domain-general visuo-motor area such as the superior colliculus would not suffice.

Finally, some caution should be taken in the implications of our results for patients with age-related macular degeneration (AMD). The findings should not be interpreted as necessarily suggesting that patients will not adapt their initial eye movement strategy to faces. Although our observers completed 5000 trials that spanned 3 weeks, they experienced two daily hours of the simulated scotoma. AMD patients' constant experience with the scotoma might result in faster learning. Our findings do suggest that more complex and highly practiced tasks such as face recognition might have a different time course of oculomotor learning with respect to simpler tasks. For face identification, special training protocols may be necessary to train patients to overcome some of the limitations of their impairment by learning new eye movement strategies that improve the efficiency of sampling visual information from their environment.

References

- Ackermann JF, Landy MS (2013) Choice of saccade endpoint under risk. *J Vis* 13(3):27 1–27. [CrossRef](#)
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723. [CrossRef](#)
- Barlow HB (1980) The absolute efficiency of perceptual decisions. *Philos Trans R Soc Lond B Biol Sci* 290:71–82. [CrossRef Medline](#)
- Barrett HH, Yao J, Rolland JP, Myers KJ (1993) Model observers for assessment of image quality. *Proc Natl Acad Sci U S A* 90:9758–9765. [CrossRef Medline](#)
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300.
- Burgess AE (1994) Statistically defined backgrounds: performance of a modified nonprewhitening observer model. *J Opt Soc Am A Opt Image Sci Vis* 11:1237–1242. [CrossRef Medline](#)
- Burgess AE, Wagner RF, Jennings RJ, Barlow HB (1981) Efficiency of human visual signal discrimination. *Science* 214:93–94. [CrossRef Medline](#)
- Crossland MD, Sims M, Galbraith RF, Rubin GS (2004) Evaluation of a new quantitative technique to assess the number and extent of preferred retinal loci in macular disease. *Vision Res* 44:1537–1546. [CrossRef Medline](#)
- Crossland MD, Culham LE, Kabanarou SA, Rubin GS (2005) Preferred retinal locus development in patients with macular disease. *Ophthalmology* 112:1579–1585. [CrossRef Medline](#)
- Curcio CA, Sloan KR, Kalina RE, Hendrickson AE (1990) Human photoreceptor topography. *J Comp Neurol* 292:497–523. [CrossRef Medline](#)
- Davies-Thompson J, Andrews TJ (2012) Intra- and interhemispheric connectivity between face-selective regions in the human brain. *J Neurophysiol* 108:3087–3095. [CrossRef Medline](#)
- de Haas B, Schwarzkopf DS, Alvarez I, Lawson RP, Henriksson L, Kriegeskorte N, Rees G (2016) Perception and processing of faces in the human brain is tuned to typical feature locations. *J Neurosci* 36:9289–9302. [CrossRef Medline](#)
- Eckstein MP, Abbey CK, Bochud FO, others (2000) A practical guide to model observers for visual detection in synthetic and natural noisy images. In: *Handbook of medical imaging: physics and psychophysics*, Vol 1 (Beutell J, Kundel RL, Van Matter RL, eds), pp 593–628. Bellingham, WA: SPIE.
- Eckstein MP, Schoonveld W, Zhang S, Mack SC, Akbas E (2015) Optimal and human eye movements to clustered low value cues to increase decision rewards during search. *Vision Res* 113:137–154. [CrossRef Medline](#)
- Hsiao JH, Cottrell G (2008) Two fixations suffice in face recognition. *Psychol Sci* 19:998–1006. [CrossRef Medline](#)
- Issa EB, DiCarlo JJ (2012) Precedence of the eye region in neural processing of faces. *J Neurosci* 32:16666–16682. [CrossRef Medline](#)
- Janssen CP, Verghese P (2015) Stop before you saccade: looking into an artificial peripheral scotoma. *J Vis* 15(5):7, 1–19. [CrossRef Medline](#)
- Kaku Y, Yoshida K, Iwamoto Y (2009) Learning signals from the superior colliculus for adaptation of saccadic eye movements in the monkey. *J Neurosci* 29:5266–5275. [CrossRef Medline](#)
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311. [Medline](#)
- Kwon M, Nandy AS, Tjan BS (2013) Rapid and persistent adaptability of human oculomotor control in response to simulated central vision loss. *Curr Biol* 23:1663–1669. [CrossRef Medline](#)
- Legge GE, Klitz TS, Tjan BS (1997) Mr. Chips: an ideal-observer model of reading. *Psychol Rev* 104:524–553. [CrossRef Medline](#)
- Legge GE, Hoooven TA, Klitz TS, Stephen Mansfield JS, Tjan BS (2002) Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Res* 42:2219–2234. [CrossRef Medline](#)
- Mehoudar E, Arizpe J, Baker CI, Yovel G (2014) Faces in the eye of the beholder: unique and stable eye scanning patterns of individual observers. *J Vis* 14(7):6, 1–11. [CrossRef Medline](#)
- Najemnik J, Geisler WS (2005) Optimal eye movement strategies in visual search. *Nature* 434:387–391. [CrossRef Medline](#)
- Najemnik J, Geisler WS (2009) Simple summation rule for optimal fixation selection in visual search. *Vision Res* 49:1286–1294. [CrossRef Medline](#)
- Or CC-F, Peterson MF, Eckstein MP (2015) Initial eye movements during face identification are optimal and similar across cultures. *J Vis* 15(13):12, 1–25. [CrossRef](#)
- Paulun VC, Schütz AC, Michel MM, Geisler WS, Gegenfurtner KR (2015) Visual search under scotopic lighting conditions. *Vision Res* 113:155–168. [CrossRef Medline](#)
- Pelli DG, Burns CW, Farell B, Moore-Page DC (2006) Feature detection and letter identification. *Vision Res* 46:4646–4674. [CrossRef Medline](#)
- Peterson MF, Eckstein MP (2012) Looking just below the eyes is optimal across face recognition tasks. *Proc Natl Acad Sci U S A* 109:E3314–E3323. [CrossRef Medline](#)
- Peterson MF, Eckstein MP (2013a) First fixations during face identification are invariant to rotation and scale. Poster presented at 13th Annual Meeting of the Vision Sciences Society, Naples, FL, May.
- Peterson MF, Eckstein MP (2013b) Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychol Sci* 24:1216–1225. [CrossRef Medline](#)
- Peterson MF, Eckstein MP (2014) Learning optimal eye movements to unusual faces. *Vision Res* 99:57–68. [CrossRef Medline](#)
- Schyns PG, Bonnar L, Gosselin F (2002) Show me the features! Understanding recognition from the use of visual information. *Psychol Sci* 13:402–409. [CrossRef Medline](#)
- Seiple W, Rosen RB, Garcia PM (2013) Abnormal fixation in individuals with age-related macular degeneration when viewing an image of a face. *Optom Vis Sci* 90:45–56. [CrossRef Medline](#)
- Smith ML, Cottrell GW, Gosselin F, Schyns PG (2005) Transmitting and decoding facial expressions. *Psychol Sci* 16:184–189. [CrossRef Medline](#)
- Sunness JS, Schuchard RA, Shen N, Rubin GS, Dagnelie G, Haselwood DM (1995) Landmark-driven fundus perimetry using the scanning laser ophthalmoscope. *Invest Ophthalmol Vis Sci* 36:1863–1874. [Medline](#)
- Tjan BS, Braje WL, Legge GE, Kersten D (1995) Human efficiency for recognizing 3-D objects in luminance noise. *Vision Res* 35:3053–3069. [CrossRef Medline](#)
- Verghese P (2012) Active search for multiple targets is inefficient. *Vision Res* 74:61–71. [CrossRef Medline](#)
- Walsh DV, Liu L (2014) Adaptation to a simulated central scotoma during visual search training. *Vision Res* 96:75–86. [CrossRef Medline](#)
- Williams DA (1976) Improved likelihood ratio tests for complete contingency tables. *Biometrika* 63:33–37. [CrossRef](#)
- Yang N, Shafai F, Oruc I (2014) Size determines whether specialized expert processes are engaged for recognition of faces. *J Vis* 14(8):17, 1–12. [CrossRef Medline](#)
- Zhang Y, Pham BT, Eckstein MP (2004) Automated optimization of JPEG 2000 encoder options based on model observer performance for detecting variable signals in X-ray coronary angiograms. *IEEE Trans Med Imaging* 23:459–474. [CrossRef Medline](#)