# Supplementary Information: Saccade-confounded image statistics explain visual crowding

Anirvan S. Nandy[1,*] & Bosco S. Tjan[1,2]

[1]Department of Psychology and [2]Neuroscience Graduate Program

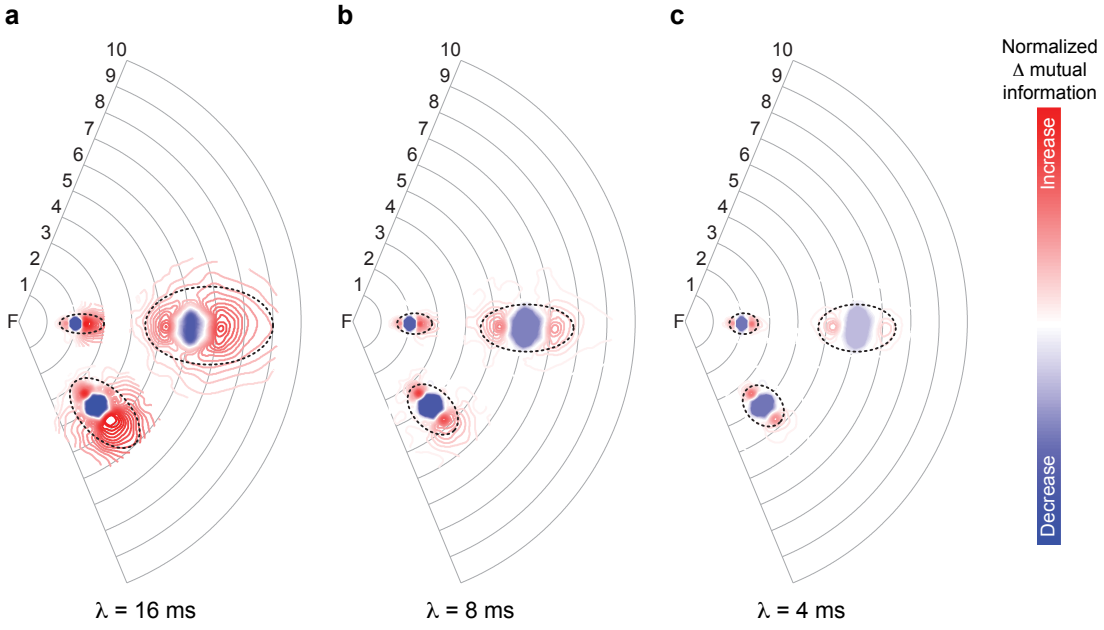
University of Southern California, Los Angeles, CA

[*]Current address: Systems & Computational Neurobiology Laboratories

The Salk Institute for Biological Studies, La Jolla, CA

## S.1 Exploration of the effect of $\lambda$ in determining the shape of the crowding zone

In our simulations, the temporal modulation of spatial attention (**Fig**. **2b**) was modeled as an exponential decay function with a time constant $\lambda$ (Equation 9, Online Methods). Here we explore the effect of the parameter $\lambda$ on the shape of the crowding zone (**Supp**. **Fig**. **1**). We see that the qualitative shape of the zones (radial elongation, proximal under-integration, distal over-integration) is preserved across values of $\lambda$.

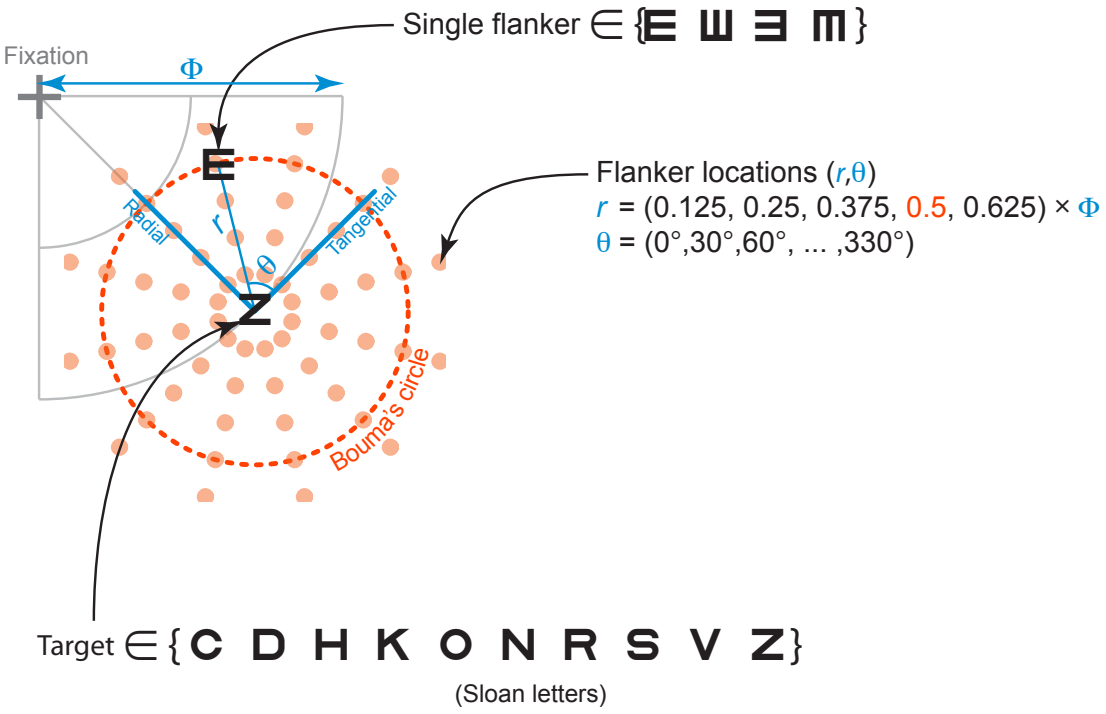



**Supp**. **Fig**. **1**. **Zones of inappropriate integration: effect of** $\lambda$. (a) The normalized difference between saccade-confounded (**Fig**. **4d**–**e**) and veridical (**Fig**. **4b**–**c**) pair-wise mutual information, pooled over all orientations (Equation 13, Online Methods), between a reference hypercolumn and neighboring hypercolumns is shown in visual space for three reference hypercolumns at 2°, 4° and 6°. The color-bar shows the magnitude and sign of the deviation from the veridical statistics, indicative of inappropriate integration: shades of red indicate that the mutual information between a reference hypercolumn and an adjacent hypercolumn is higher in the saccade-confounded statistics than expected; shades of blue indicate lower mutual information than expected. Elliptical fits (dotted lines at 40% of peak normalized difference) illustrate the elongated shape of the spatial extent of inappropriate integration. The time constant of the decay of spatial attention ($\lambda$) is 16ms. (b–c). The corresponding zones are shown for $\lambda$=8ms and $\lambda$=4ms respectively.

**S.2 Detailed Measurements of the crowding zones in the upper and lower visual fields**

Many of the current studies that map the shape of the crowding zone (**Fig**. **1c**) have two limitations: (a) flankers were displayed symmetrically on both sides of the target, but due to the inward-outward asymmetry, the measured zone was biased toward the fovea; and (b) performance was measured at only a few cardinal directions (typically 4 or 8) around the target. We wanted to obtain a dense and bias-free estimate of the crowding zone by (a) using a single flanker and (b) measuring performance at multiple locations around the target using the method of constant stimuli.

Human experiments were conducted in which observers had to perform a letter identification task in the presence of a single letter flanker. The target was one of 10 letters ('C', 'D', 'H', 'K', 'N', 'O', 'R', 'S', 'V', 'Z') rendered in Sloan font, while the flanker was the letter 'E' rotated at one of 4 orientations (0°, 90°, 180°, 270°). The flanker was presented at one of 12 azimuth angles around the target and at one of 5 separations from the target, thus yielding a total of 60 flanker positions (**Supp**. **Fig**. **2**). The center-to-center separation between the target and the flanker was at 0.125, 0.25, 0.375, 0.5 or 0.625 times the target eccentricity.




**Supp**. **Fig**. **2**. **Experimental design for measuring the fine structure of the crowding zone**. A single flanker ('E' at any one of 4 orientations) was placed at 60 different locations (orange dots) around a target (one of the 10 Sloan letters) at an eccentric location Φ. Performance (proportion correct target identification) was measured for each flanker location.

The stimuli were presented on a 19" CRT monitor (Sony Trinitron CPD-G400) at a viewing distance of 80 cm. In each trial, the observers had to fixate at a cross while the stimulus was presented according to the following temporal design: (1) a fixation beep immediately followed by a fixation screen for 250 ms, (2) stimulus presentation for 250 ms, (3)

observer response period (variable) with positive feedback beep for correct trials, and (4) 250 ms delay before onset of the next trial. Depending on the experimental condition, the target was presented at either 10° in the upper or 10° in the lower visual field. The target and the orientation of the flanker were randomly chosen in each trial. Both the target and the flanker were presented at 100% contrast. The response of the observer was stored in each trial as either correct or incorrect.

Each experiment consisted of 48 blocks, with 125 trials per block. In each block, the flanker was presented at a fixed azimuth angle with respect to the target but at any of the 5 separations. Within a block, there were 25 trials at each separation and the order of the presentations were randomly shuffled. There were 4 such blocks for each of the 12 azimuth angles. The blocks were randomly ordered with the constraint that the $n^{th}$ repeat of a particular azimuth occurred only after all azimuths had been tested for at least $n$–1 blocks.

Prior to the main experiment, letter acuity was measured at the desired target location. The observers were asked to identify any of the 10 target letters, while the size of the presented letter was varied using QUEST[1] to achieve an identification accuracy of 79%.

*Analysis Procedure*

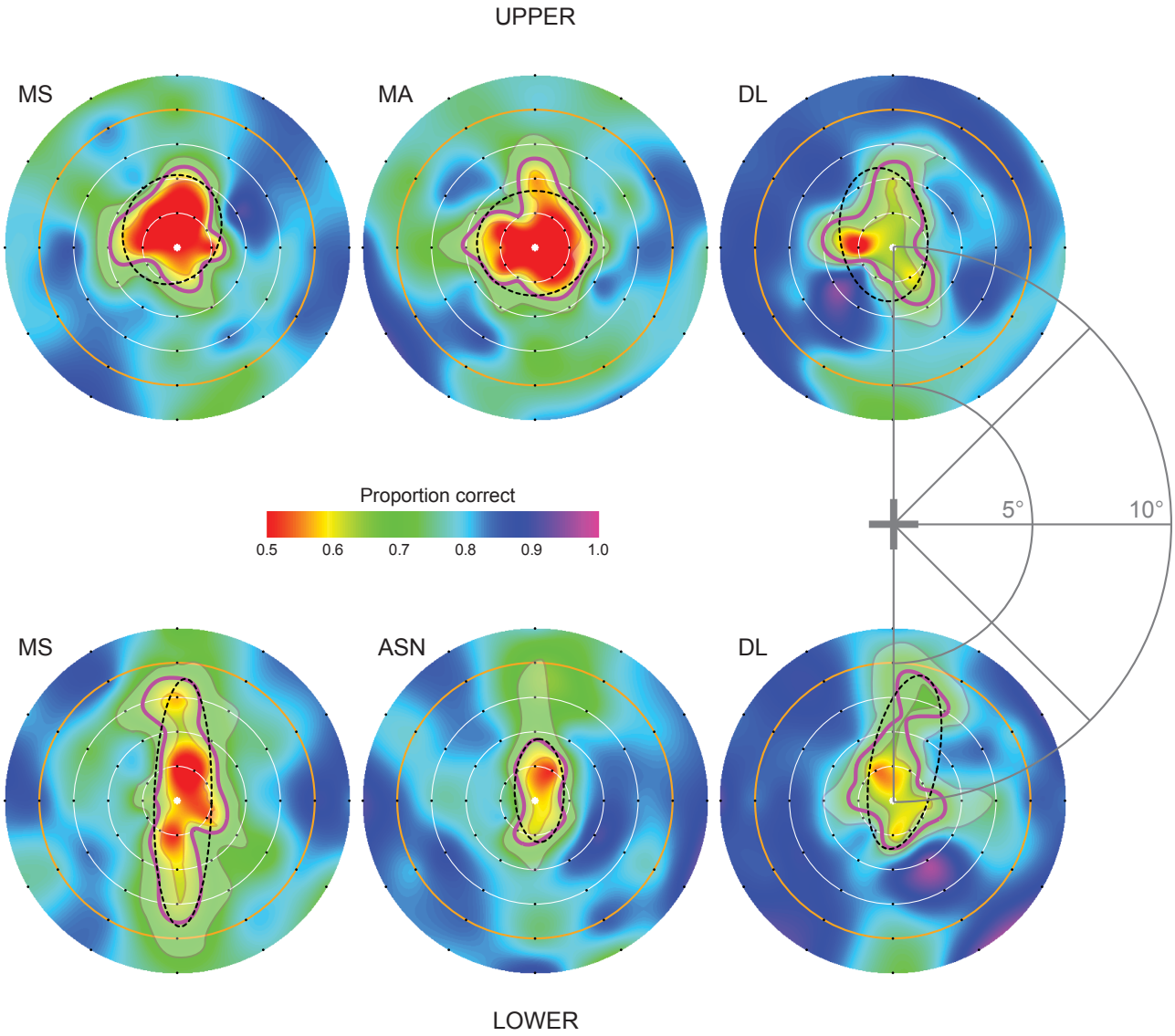
For each of the 60 flanker locations, we first obtained a percentage correct measure (100 trials per location) from the experimental data. We next performed a 2-D surface interpolation of the data points. This 2-D interpolated map gives a visual map of the crowding zone. To further quantify the shape characteristics, we obtained a bootstrapped estimate[2] of the 95% confidence interval for a particular map contour at 65% correct (chance performance was 10%). This was done by sampling with replacement the observed trials 100 times. For each such sample, we obtained an interpolated map as above and a contour at the same criterion. From the ensemble of such contours we calculated a confidence interval of the contour obtained from the raw data. Lastly, a least-squares elliptical fit to the contour measured from the raw data provides a summary of the shape of the zone at the particular criterion.

*Results*

A comparison of the crowding zone between the upper and lower visual fields is shown in **Supp**. **Fig**. **3**. The zones in the upper field are less elongated than those in the lower field. The aspect ratios of the fitted ellipses are shown in **Supp**. **Table 1**. The relationship between this empirical finding and our proposed model of crowding is discussed in the main text.

| Upper | Lower |
|---|---|
| 1.15 (MS) | 4.32 (MS) |
| 1.10 (MA) | 2.09 (ASN) |
| 1.53 (DL) | 2.65 (DL) |

**Supp. Table 1. Aspect ratios of elliptical fits to crowding zones.** Aspect ratios of the least squares elliptical fits to the contours at 65% correct (**Supp**. **Fig**. **3**) at 10° in the upper visual field (*left column*) and at 10° in the lower visual field (*right column*). Observer initials are given in parenthesis.




**Supp**. **Fig**. **3**. **Crowding zones in the upper and lower visual fields**. Interpolated maps from proportion correct data measured by placing a single letter target at 10° in the upper and lower visual fields (target location indicated by white dots) and a single letter flanker at one of sixty positions (black dots) around the target. A contour at 65% correct (magenta) and its 95% confidence interval (translucent bounds) are superimposed on the maps. The confidence intervals were computed using a bootstrap procedure. A least squares elliptical fit (dotted black ellipse) to the contour is also superimposed. The orange circle represents half the target eccentricity (Bouma's Law).

## S.3 Crowding zones with oriented flankers

Human experiments were conducted in which observers had to perform an orientation discrimination task in the presence of oriented flankers (**Supp**. **Fig**. **4a**). Gabor targets were presented either at 10° of visual angle to the right of fixation or at 10° below fixation. The targets were flanked by two Gabor distractors in either a radial or a tangential arrangement. The target Gabors were tilted at –5° or +5° from the base orientation that was either horizontal or vertical. The flanker Gabors shared the same base orientation as the target but their tilts were chosen independently and randomly in each trial from a uniform distribution between –

10° and +10°. Both target and flanker Gabors had a carrier spatial frequency of 2 cycles per degree. The phase of the carrier was randomized from trial to trial. The full width at 1/*e* height of the Gaussian envelope was 1.5°.

There were a total of 8 experimental conditions: 2 locations x 2 flanker arrangements x 2 target (and flanker) orientations. Observers viewed the stimuli with their left eye in a 2-alternative forced choice task to determine whether the target was tilted clockwise or counter-clockwise from the base orientation. The center-to-center separation between target and flankers was varied using QUEST[1], with $\beta$ set to 1.2 based on a pilot experiment, to achieve a performance accuracy of 84% correct.

The stimuli were presented on a 19" CRT monitor (Sony Trinitron CPD-G400) at a viewing distance of 80 cm. In each trial, the observers had to fixate at a cross while the stimulus was presented according to the following temporal sequence: (1) a fixation beep immediately followed by a fixation screen for 250 ms, (2) stimulus presentation for 250 ms, (3) observer response period (variable) with positive feedback beep for correct trials, and (4) 250 ms delay before onset of the next trial. Both the target and the flanker were presented at 50% contrast. The response of the observer was stored in each trial as either correct or incorrect.

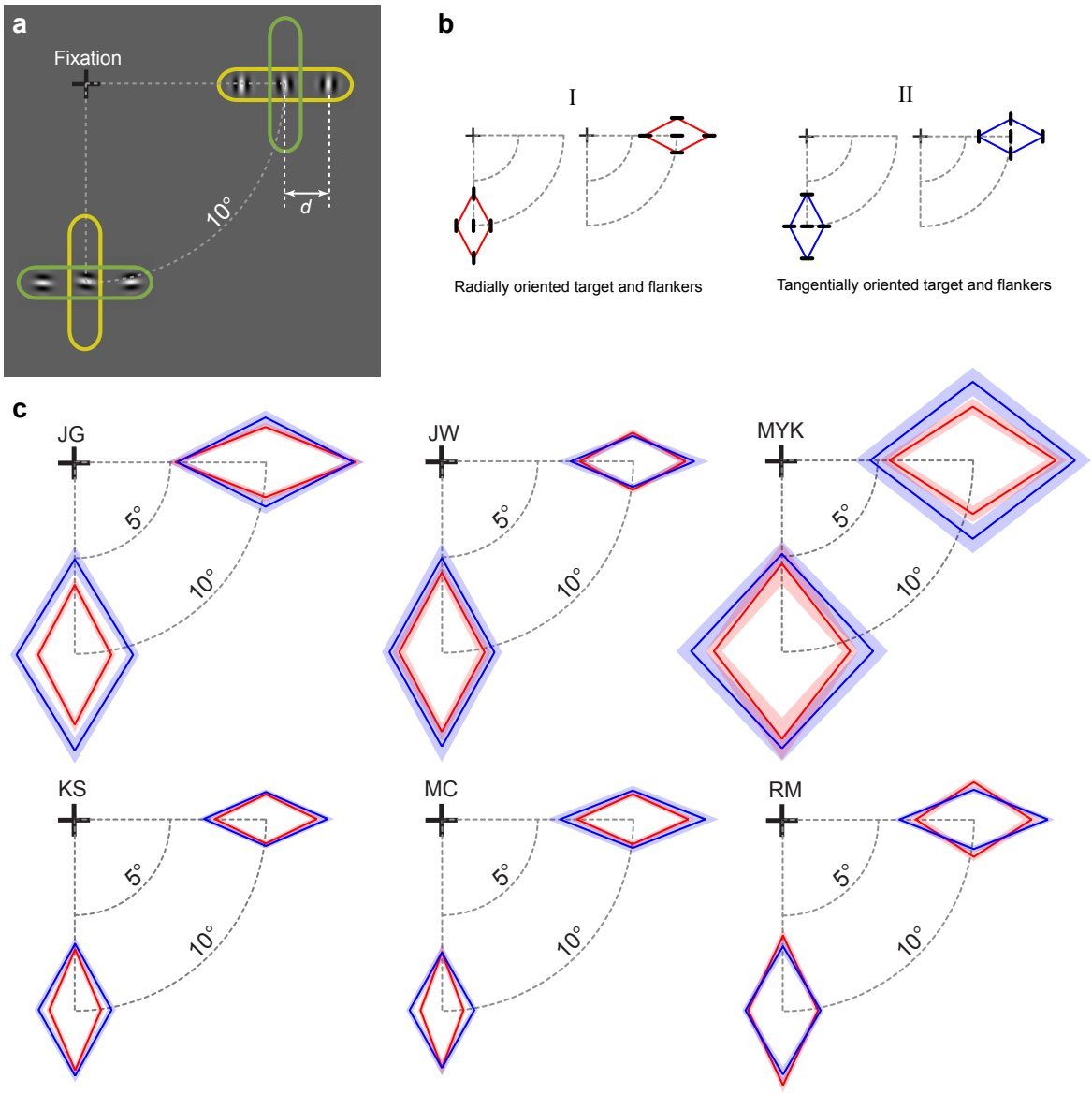
Each experimental condition was repeated 6 times for a total of 48 blocks, with 60 trials per block. Blocks were randomly ordered with the constraint that the $n^{th}$ repeat of a particular experimental condition occurred only after all conditions had been tested for $n$–1 blocks.

*Analysis Procedure*

For each observer and each of the 8 experimental conditions, we combined the data from all 6 blocks and estimated the mean and standard error of the target-flanker separation necessary to achieve a performance threshold of 84% correct with a bootstrap procedure[2]. This was done by sampling with replacement the observed trials 200 times. For each such sample, we first obtained an estimate of the threshold separation by fitting the Weibull psychometric function used in QUEST. The mean and standard error were then calculated from the ensemble of samples.

*Results*

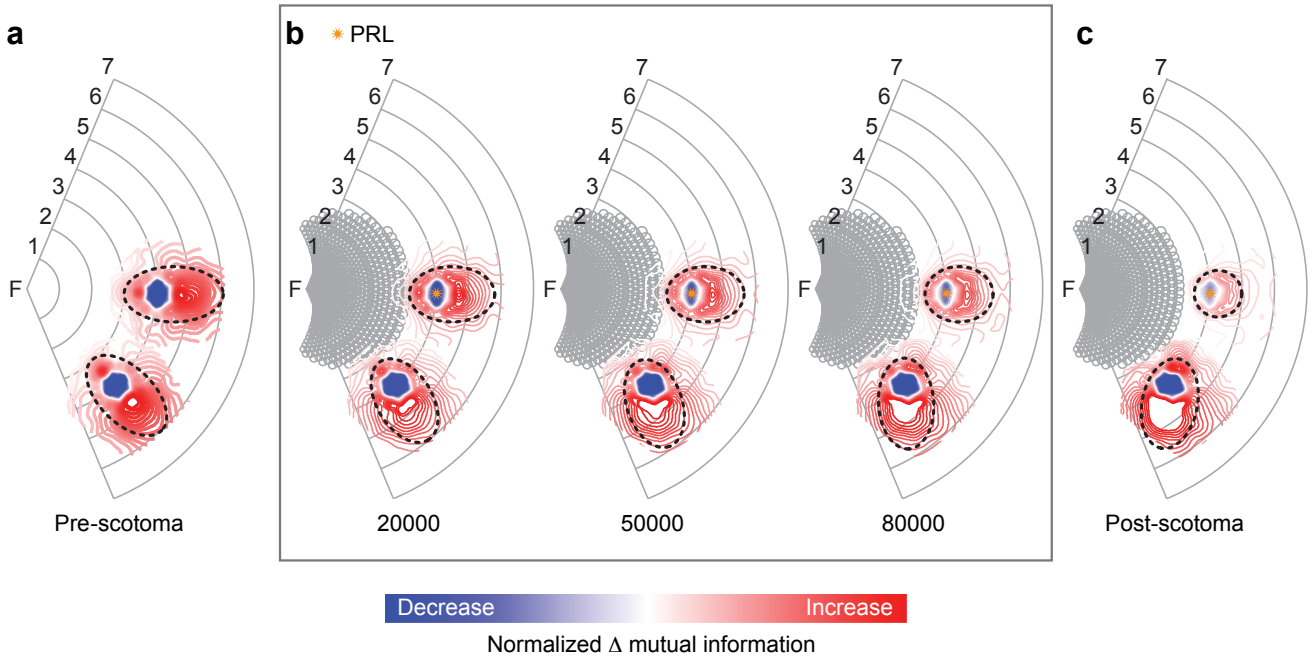
The 8 experimental conditions were grouped into two categories according to whether the flankers were oriented parallel to or orthogonal to the radial axis (**Supp**. **Fig**. **4b**). Data from the 6 individual observers are shown in **Supp**. **Fig**. **4c**. The red polygons show the crowding zones for flankers oriented parallel to the radial axis. Blue polygons show the corresponding zones for orthogonally oriented flankers. The results show that, in general, orthogonal flankers lead to a larger crowding zone irrespective of their location in the visual field and their spatial arrangement. The relationship between this empirical finding and our proposed model of crowding is discussed in the main text.




**Supp. Fig. 4. Crowding with oriented flankers**. (a) Gabor targets were presented at 10° eccentricity either to the right of or below the fixation mark. A target was flanked by two Gabor flankers in either a radial arrangement (demarcated by yellow rounded rectangles for illustration only) or a tangential arrangement (green rounded rectangles). The target Gabor was tilted either 5° or –5° from a base orientation that was either horizontal or vertical. The flanker Gabors shared the same base orientation as the target but their tilts were chosen independently and randomly from a uniform distribution between –10° and +10°. There were a total of 8 experimental conditions: 2 locations x 2 flanker arrangements x 2 orientations. Stimuli were viewed monocularly with the left eye. Observers' task was to determine whether the target was tilted clockwise or counter-clockwise from the base orientation. The center-to-center separation between target and flankers (d) was varied using an adaptive procedure (QUEST) to achieve a performance accuracy of 84% correct. (b) The 8 experimental conditions were grouped into 2 categories. I: target and flankers oriented parallel to the radial axis (line connecting fixation to the target location). II: target and flankers oriented orthogonal to the radial axis. (c) Crowding zones for 6 observers (color code as in b). Blue polygons show the crowding zones for target and flankers oriented perpendicular to the radial axis . Red polygons are the zones for target and flankers oriented parallel to the radial axis. Solid lines are bootstrapped estimates of the center-to-center separation of target and flanker (distance d in a) for a performance accuracy of 84%; translucent bounds are standard errors. The crowding zones with elements oriented perpendicular to the radial axis are significantly larger than those oriented parallel to the radial axis (19 out of 24, $p < 0.005$, sign test).

## S.4 Predictions of the shape of the crowding zone at and around the PRL

Many patients with central vision loss due to age-related macular degeneration (AMD) develop the use of a stable peripheral location in the retina for fixations during form-vision tasks. This is known as the preferred retinal locus (PRL) and is typically located just outside the central scotoma. Since the stable PRL is used for fixations, saccadic eye movements for such patients, if fully re-referenced, are now radial with respect to the PRL. We performed simulations of eye-movements (Online Methods) to predict the shape of the crowding zones at and around the PRL of a model with a simulated scotoma (**Supp**. **Fig**. **5**). Our simulations predict that with increasing exposure to PRL-centric saccades: (a) the crowding zone measured at the PRL should progressively become anisotropic since the PRL no longer experiences any radial bias in eye movements and (b) the elongated axes of the crowding zones at other peripheral locations should reorient and eventually point toward the PRL. Preliminary results from AMD patients measured with a scanning laser ophthalmoscope suggest that the zone of crowding measured at the PRL is indeed rounded (Chung, 2009, personal communication; Chung & Lin, *The Two-Dimensional Shape of Spatial Interaction Zones at the PRLs of Observers with Central Vision Loss*, 2008 ARVO abstract).




**Supp**. **Fig**. **5**. **Predicted zones of inappropriate integration at and around the PRL**. (a) Zones of inappropriate integration (normalized difference between pooled saccade-confounded and veridical image statistics) are shown at two locations in a normal visual field: 4° to the right and 4° to the lower right. (b) A central scotoma is introduced (zone demarcated by gray circles, indicating the receptive fields affected by the scotoma) and the retinal location at 4° to the right (starred location) is assumed to be the preferred retinal locus (PRL), serving as the center of gaze for all fixations subsequent to the development of the central scotoma. The number at the bottom of each panel indicates the number of simulated saccades. (c) The predicted post-scotoma zones of inappropriate integration at the two retinal locations. The elliptical fits (dotted lines) are at 40% of peak normalized difference. At the PRL, the zone progresses from being anisotropic to isotropic. At the other location, the zone progresses from fovea-centric anisotropy (elongation with the long axis pointing towards the fovea) to PRL-centric anisotropy.

## REFERENCES


1. Watson AB, Pelli DG. QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys.* **33**, 113–20 (1983).
2. Efron B, Tibshirani R. *An introduction to the bootstrap.* (New York: Chapman & Hall, Vol. 57. 1993).