

Similar Neural Representations of the Target for Saccades and Perception during Search

Miguel P. Eckstein,¹ Brent R. Beutter,² Binh T. Pham,¹ Steven S. Shimozaki,¹ and Leland S. Stone²

¹Vision and Image Understanding Laboratory, Department of Psychology, University of California, Santa Barbara, California 93106, and ²Human System-Integration Division, NASA Ames Research Center, Moffett Field, California 94035

Are the body's actions and the mind's perceptions the result of shared neural processing, or are they performed largely independently? The brain has two major processing streams, and some have proposed that this division segregates visual processing for action and perception. The ventral pathway is claimed to support conscious experience (perception), whereas the dorsal pathway is claimed to support the control of movement (action). Others have argued that perception and action share much of their visual processing within the primate cortex. During visual search, the brain performs a sophisticated deployment of eye movements (saccadic actions) to gather information to subserve perceptual judgments. The relationship between the neural mechanisms mediating perception and action in visual search remains unexplored. Here, we investigate the visual representation of target information in the human brain, both for perceptual decisions and for saccadic actions during visual search. We use classification image analysis, a form of reverse correlation, to estimate the behavioral receptive fields of the visual mechanisms responsible for saccadic and perceptual responses during the same visual search task. Results show that the behavioral receptive fields mediating the perceptual decisions are indistinguishable from those driving the oculomotor decisions, suggesting that similar neural mechanisms are responsible for both perception and oculomotor action during search. Diverging target representations would result in an inefficient coupling between eye movement planning and perceptual judgments. Thus, a common target representation would be more optimal and might be expected to have evolved through natural selection in the neural systems responsible for visual search.

Key words: visual search; eye movements; behavioral receptive fields; perception and action; decision; eye position

Introduction

Studies of the neurophysiology of the macaque monkey (Ungerleider and Mishkin, 1982; Sakata et al., 1997; Snyder et al., 1997) provide evidence for the existence of two functionally distinct neural streams in the brain that mediate the processing of visual information. Behavioral studies of patients with brain damage led to the proposal that perception is mediated by the ventral stream projecting from the primary visual cortex to the inferior temporal cortex, and that action is mediated by the dorsal stream projecting from the primary visual cortex to the posterior parietal cortex (Goodale et al., 1991; Goodale and Milner, 1992; Milner and Goodale, 1997). For motor control of eye movements, others have proposed that pathways for perception and oculomotor control largely overlap, leading to significant sharing of visual information (Krauzlis and Stone, 1999; Gegenfurtner et al., 2003; Stone and Krauzlis, 2003; Dassonville and Bala, 2004). Furthermore, in the case of saccadic eye movements, visual cortical pathways through the frontal eye field (Schall, 2002) and the lateral intraparietal cortex (Goldberg et al., 2002) play critical roles, as

do brainstem and cortical pathways through the superior colliculus (McPeck and Keller, 2004). In addition, recent studies have related areas in the ventral stream to target selection of saccades (Moore, 1999; Mazer and Gallant, 2003). Thus, considerable controversy exists concerning the extent to which the visual information available for perception can be used for saccadic targeting. If separate dorsal and ventral visual pathways control eye movements and perception, then different neural areas will control saccadic and perceptual decisions. These different neural areas will likely perform different computations and use neurons with different receptive fields. If so, the representations of the target for perception and visual saccades will differ. On the other hand, if information is shared across the dorsal and ventral streams and saccadic and perceptual search decisions are determined by a shared pool of neurons, then the processing and receptive fields used to support saccadic and perceptual decisions will be the same.

We used classification image analysis (Ahumada and Lovell, 1971; Gold et al., 2000; Eckstein and Ahumada, 2002; Neri and Heeger, 2002; Levi and Klein, 2003; Rajashekar et al., 2006), which allowed us to measure and compare the behaviorally defined, spatial receptive fields of the visual mechanisms responsible for saccadic and perceptual decisions. With this method, the entire brain is viewed as a single system in which input is the stimulus and output is either the saccadic or perceptual decision. Classification image analysis reveals the overall receptive field of

Received Sept. 12, 2006; revised Dec. 7, 2006; accepted Dec. 8, 2006.

This research was supported by National Science Foundation Grant NSF-0135118 and National Institutes of Health Grant 5R01EY15925-2.

Correspondence should be addressed to Miguel P. Eckstein, Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106. E-mail: eckstein@psych.ucsb.edu.

DOI:10.1523/JNEUROSCI.3975-06.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/271266-05\$15.00/0

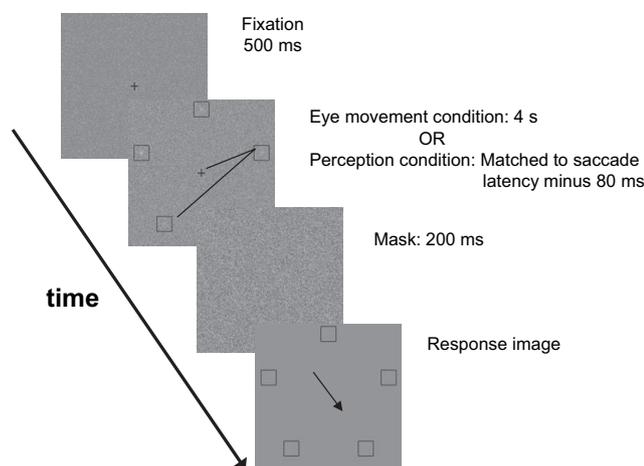


Figure 1. Visual task. The observers' task was to find the box containing the target. The timeline for an individual trial of the eye movement condition was as follows. A fixation image was presented. The trial was initiated by the observer. After 500 ms, the test image containing target, distractors, and noise was presented for 4 s. The observer searched freely for the target. After 4 s, an image containing the outline of the boxes was presented, and the observers used an arrow to choose a target location. The timeline for the perceptual decision condition was similar, except that the test image was presented for a brief time matching the latency of the corresponding saccade in the eye movement condition -80 ms.

the system based on the organism's choices in the visual task. Here, we applied the technique to find the behaviorally defined receptive fields underlying perceptual and saccadic decisions during search. We measured perceptual and saccadic decisions of six observers in separate visual search experiments that used the same set of stochastic stimuli.

Materials and Methods

Task. Six observers searched for a bright Gaussian-shaped target among four dimmer Gaussian-shaped distractors in a five-alternative forced-choice (5-AFC) task (Fig. 1). Spatially uncorrelated Gaussian noise was added to the stimulus. The locations containing the target and distractors were indicated with five black boxes, equidistant on an imaginary circle of radius of 7.9° . There were two experimental conditions. In the eye-movement condition, observers were given 4 s to freely search for the target. To allow a fair comparison between saccadic and perceptual performance, we matched the visual processing times for saccadic and perceptual decisions by measuring perception in a separate, short-duration condition. In the perception condition, we used the same visual stimuli, but each trial had a shorter duration that was chosen to match the visual processing time for the associated saccadic decision, i.e., the latency of the first saccade in the corresponding trial in the eye-movement condition -80 ms [the approximate time needed for motor programming of the saccade after the location decision has been made and visual processing is completed (Hooge, 1996; Caspi et al., 2004)]. Median latency across observers ranged from 178 ms (J.L.) to 326 ms (B.G.). The perception condition was always run after the eye-movement condition so that the perception-condition stimulus durations could be calculated for each observer and each trial. In the perception condition, the observers were instructed to fixate a central cross during the presentation of the test stimuli. Observers pressed a button to initiate the trial. Eye position was recorded to detect violations of central fixation. In both conditions, the stimulus was immediately followed by a high-contrast mask to minimize further visual processing, and observers used an arrow to indicate their 5-AFC perceptual decision about target location. In the eye-movement condition, we saved the oculomotor (first saccade) 5-AFC decision, whereas in the perception condition we saved the perceptual (button press) 5-AFC decision. Observers participated in 40 sessions of 100 trials for each of two conditions over a period of 20 business days (one session

for the eye movement condition and one session for the perceptual decision condition per day).

Display. Stimuli were viewed binocularly on an Image Systems ML21 monitor (Image Systems, Plymouth, MN) with linearized luminance. The mean luminance of the stimulus was 25 cd/m^2 . The five search-element locations were equally spaced at an eccentricity of 7.9° from the fixation point and were indicated by boxes ($2.07^\circ \times 2.07^\circ$). The target (and distractors) was a spatial Gaussian with an SD of 0.188° (4.0 pixels). The contrasts of the target and distractors (the pedestal) were chosen so that the contrast difference resulted in perceptual search accuracy of 65% (task signal-to-noise ratio, 2.835). The distractor contrast relative to the noise resulted in a signal-to-noise ratio of 6.24 (peak contrast, 0.156). For each trial, a different sample of zero-mean Gaussian spatial white noise (root mean square contrast of 0.1953) was added to the stimulus. Each stimulus image was stored and presented in both a perception and an eye-movement trial, albeit with a different duration.

Eye-position recording. An infrared video-based eye tracker that sampled at 250 Hz (SMI Eyelink; SensoMotoric Instruments, Teltow, Germany) and synchronized with the 60 Hz monitor was used to measure the position of the left eye. At the beginning of each session, calibration was performed with nine crosses that were arranged in a $12^\circ \times 12^\circ$ grid. To minimize head movements, observers were positioned on a chin rest and instructed to hold their head still. To control for residual head movements during the session, the SMI system measures head position (with a scene camera) and corrects the calibration for any head movements.

Eye-position analysis. Only eye movements that moved the eyes outside of a 1.7° window surrounding the fixation point were treated as saccadic decisions. Trials with anticipatory saccades (<90 ms latency) were discarded from the analysis, although they accounted for only $\sim 1\%$ of the trials. Saccades were detected when both eye velocity and acceleration exceeded a threshold (velocity $>35^\circ/\text{s}$; acceleration $>9500^\circ/\text{s}^2$).

Performance analysis. We converted the first saccade into a 5-AFC oculometric decision about target location by associating its endpoint with the closest display element. Thus, the oculometric decision was considered correct if the endpoint of the saccade was closest to the target location. The latencies of the first saccades were recorded for each trial.

Classification images. Classification image analysis relies on the use of stimuli containing external noise and is based on the logic that when a noisy distractor element is mistaken for the target, its noise must have randomly contained some luminance variation that fooled the decision mechanism into choosing it. The technique is closely related to the reverse correlation technique used with success to estimate single-cell receptive fields (Sutter, 1975; Marmarelis and Marmarelis, 1978). By averaging the noise fields for all the incorrect decisions, one can estimate the spatial information that is used to drive the decisions (Ahumada and Lovell, 1971; Abbey and Eckstein, 2002; Murray et al., 2002; Solomon, 2002). With the additional assumption that the observer is performing the task by using mechanisms that sum information linearly across space, the method generates an unbiased estimate of the linear receptive field of the visual mechanism responsible for performance (Ahumada and Lovell, 1971; Abbey and Eckstein, 2002) (for methods in the presence of nonlinearities, see Victor, 2005) [previous studies have shown that the linearity assumption (Victor, 2005) is a good approximation for the case of perceptual and saccadic decisions in contrast discrimination tasks of aperiodic, spatially compact targets such as Gaussians embedded in spatially uncorrelated Gaussian noise (Beutter et al., 2003)]. For each observer and condition (perception vs saccades), classification images were obtained by computing the mean of the 32×32 pixel noise fields ($1.5^\circ \times 1.5^\circ$) corresponding to locations that were incorrectly chosen (perception and first saccade, correspondingly, for each condition). Radial averages of classification images were obtained by averaging across all angles those pixels falling within the same distance interval from the center pixel (e.g., 1.5–2.5, 2.5–3.5 pixels). Statistical comparisons of the radial averages of classification images were performed with the two-sample Hotelling T^2 statistic given by (Harris, 1985) as follows:

$$T_2 = \frac{N_1 N_2}{(N_1 + N_2)} [\mathbf{x}_2 - \mathbf{x}_1]^T \mathbf{K}^{-1} [\mathbf{x}_2 - \mathbf{x}_1],$$

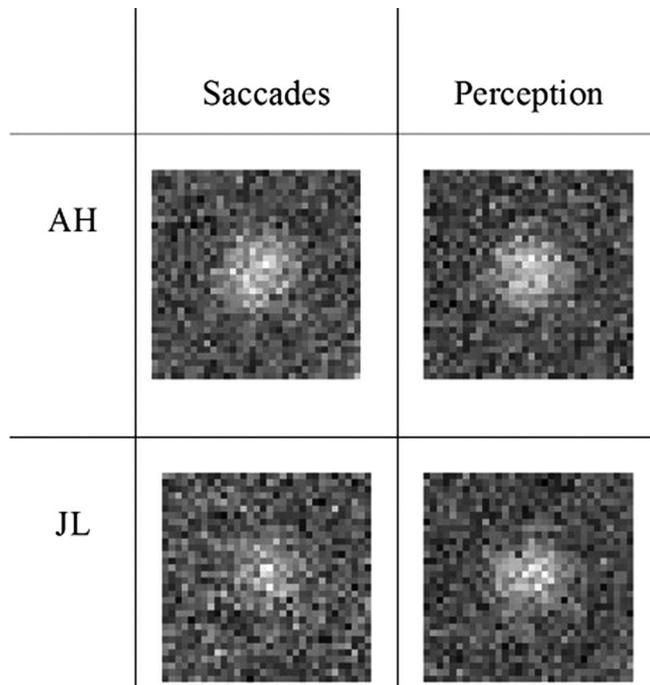


Figure 2. Classification images. Classification images were presented for two observers for saccades and perception. To calculate the classification images, we averaged noise fields from trials with an incorrect (1) perceptual decision (perception condition) or (2) first saccade decision (eye-movement condition).

where \mathbf{x}_1 and \mathbf{x}_2 are vectors containing the observed radial averages of the two classification images, and N_1 and N_2 refer to the number of observations or images used to compute the two classification images. For the two-sample test, a pooled covariance \mathbf{K} is computed combining the sum of square deviations and the sum of squared products from both samples. To test for significance, the two-sample T^2 statistic can be transformed to an F statistic by using the following relationship:

$$F = \frac{N_1 + N_2 - p - 1}{p(N_1 + N_2 - 2)} T^2,$$

where p is the number of dependent variables (number of vector elements in the radial average of the classification images); N_1 and N_2 are defined as before. The obtained F statistic can be compared with an F_{critical} with p degrees of freedom for the numerator and $N_1 + N_2 - p - 1$ degrees of freedom for the denominator.

Results

The average saccadic and perceptual decision accuracies (\pm SE, across observer) were 55 ± 6 and $66 \pm 3\%$, respectively; therefore, $\sim 45\%$ of the eye-movement condition trials and $\sim 34\%$ of the perception-condition trials, i.e., the erroneous ones, were included in our analysis. For correct saccadic decision trials, the average distance (\pm SE across observer) from the endpoint of the first saccade to the target center was $1.23^\circ \pm 0.07^\circ$. In $23 \pm 2.9\%$ of the trials, the second saccade was directed toward the target, whereas in the remaining 22% of the trials, the target was either fixated after a subsequent saccade or not fixated at all.

Figure 2 shows pairs of classification images for two different observers generated from erroneous trials. Although the images are noisy, each appears to contain a radially symmetric blob-like feature similar to the stimulus. For observer A.H., the amplitudes (represented by image intensity) were similar for the saccadic and

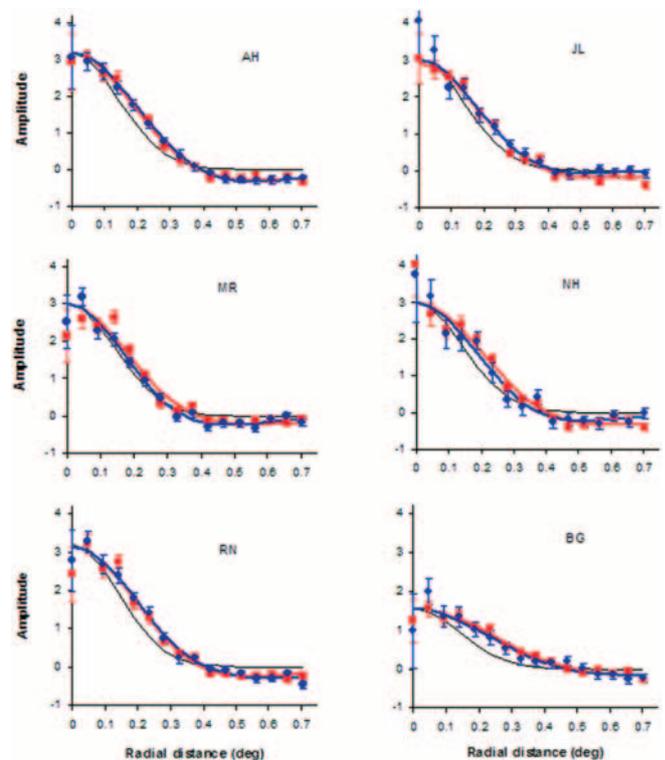


Figure 3. Radial profiles of estimated visual mechanisms: saccades (blue) and perceptual decisions (red). Data are shown for six observers. Symbols correspond to radial average values from classification images. Solid lines (red and blue) correspond to the best-fit DOG functions. Error bars indicate SEM. The black solid line shows the radial profile of the Gaussian target.

perceptual images, but for observer J.L., the amplitude of the saccadic classification image appears slightly lower.

For further analysis, the classification images were averaged across angles to generate a one-dimensional radial profile. To compare the spatial receptive-field shapes, we factored out amplitude differences and normalized the radial averages. Figure 3 plots these normalized radial profiles for the first saccade (blue circles) and for the perceptual decision (red squares) for all six observers. Note first that the saccadic and perceptual receptive-field profiles are indistinguishable for all six observers (Hotelling T^2 test; $p > 0.05$). In addition, comparison of the profile of the target (black line) reveals that the receptive fields for both saccades and perception appear to be slightly wider than the target and also show clear lateral inhibition typical of neural receptive fields. To quantify these differences, we fit the receptive-field profiles to difference of Gaussian (DOG) functions (Fig. 3, solid red and blue lines, and Table 1). The data were well fit by DOG functions (Table 1, χ^2 analysis). The mean (\pm SE across observer) SDs of the excitatory Gaussians ($0.201 \pm 0.012^\circ$ for perception; $0.199 \pm 0.009^\circ$ for saccades) were larger than the SD of the target (0.188°). This analysis shows that not only are the receptive-field profiles of the visual mechanisms for saccades and perception similar, they are more similar to each other ($t_{(5)} = 0.13$) than they are to the target profile itself ($t_{(5)} = 1.08$).

Discussion

Our results show that for a visual search task with matched processing times, there is little or no difference between the representations of the target used by the visual mechanisms responsible for target selection for saccadic and perceptual responses. This

Table 1. Fits to classification images: schematic of a difference of Gaussian function obtained by subtracting two Gaussian functions (G_e and G_i) with different amplitudes, C_e and C_i , and different SDs, σ_e and σ_i

Observer	First saccade					
	Difference of Gaussian					
	c_e	c_i	σ_e	σ_i	χ^2	χ^2_r
AH	4.2	1.15	4.4	8.4	10.29	0.858
JL	1.8	0.25	3.8	6.8	5.49	0.457
MR	4.1	1.2	3.8	6.4	16.03	1.336
NH	1.9	0.35	4.0	8.0	15.53	1.293
RN	3.0	0.35	4.0	15.0	12.32	1.026
BG	1.0	0.2	5.2	15.0	13.95	1.162

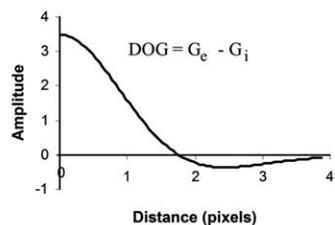
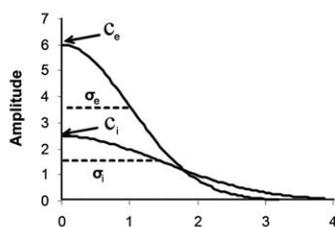
Observer	Perception					
	Difference of Gaussian					
	c_e	c_i	σ_e	σ_i	χ^2	χ^2_r
AH	3.7	0.4	3.8	15.0	11.16	0.929
JL	3.3	0.4	4.0	15.0	26.48	2.207
MR	4.2	1.15	4.0	6.4	28.99	2.415*
NH	4.1	1.2	4.6	8.2	16.49	1.375
RN	3.6	0.45	4.0	13.0	15.12	1.259
BG	1.9	0.35	5.2	13.6	13.95	1.162

The subscript "e" stands for excitatory and refers to the Gaussian with the higher amplitude and smaller SD, whereas the subscript "i" stands for inhibitory and refers to the Gaussian with the larger SD. Columns two through five of the table indicate the best-fit parameters of two-dimensional DOG functions for the saccade and perception classification images expressed in units of pixels (1 pixel = 0.047°). Columns six and seven indicate goodness of fit ($\chi^2_r = \chi^2/df$, where df is degrees of freedom = number of data points – fitting parameters). * $P(\chi^2_r \geq \chi^2_{obs}) < 0.05$, the probability of obtaining a χ^2_r that is larger than or equal to the observed χ^2_r given that the DOG model follows the data is less than 5%. The model can be rejected.

result is consistent with previous work showing that overall target localization performance for blurred discs and Gaussian blobs is similar over a wide range of signal strengths for saccades and perception in a visual search task (Beutter et al., 2003). The slightly lower overall accuracy of the saccadic decisions for some of our observers can be attributed to a difference in internal noise (Beutter et al., 2003; Gegenfurtner et al., 2003) in perceptual and saccadic decisions but not to a difference in the representation of the target.^a A possible objection to the present study is that the target we used is simple and thus it should be no surprise that both the saccade and perceptual systems have similar behavioral receptive fields that match the target; however, our results show that the behavioral receptive fields for the saccade and perception systems differed systematically from the target and were more similar to each other than to the target. The areas of the behavioral receptive fields (excitatory area ≈ 0.5 – 0.8 deg²; excitatory and inhibitory areas ≈ 0.9 – 1.2 deg²) are within the range of receptive field sizes at 8° eccentricity measured in area V1 of monkeys (areas ≈ 0.1 – 1.0 deg²) (Van Essen et al., 1984).

Our findings suggest that a shared neural locus encoding the spatial luminance profile of the target is responsible for both perceptual and oculomotor decisions. Furthermore, in the context of visual search, it is difficult to imagine the benefit of evolving multiple visual systems with differing target representations given that the purpose of eye movements is to reorient the high resolution fovea to maximize the acquisition of information for conscious perceptual judgment. A recent study (Najemnik and Geisler, 2005) has suggested that observers' eye movements during search are optimally planned to maximize the gathering of information to be used for a final perceptual decision. Indeed, in this ideal search model, the information used to plan saccades and the final perceptual decision arise from a common mecha-

^a It is also possible that lower saccadic accuracy is also related to the lack of explicit penalties for inaccurate saccades or to the time pressure given by the long duration (4 s) available for the search. A shorter presentation time or explicit penalties on saccade inaccuracy might result in higher observer saccade performance.



nism. Saccadic planning and perceptual decisions based on different target representations would lead to a disjointed strategy to plan saccadic eye movements to gather information about the location of the target and thus would lead to suboptimal performance. Given the importance of visual search in the foraging behavior of ancient humans, one might expect that the common visual processing for perception and saccades would be favored by evolution through the process of natural selection.

Finally, the classification images provide a receptive-field fingerprint for this shared perceptual–oculomotor mechanism, which will be useful for guiding future research examining the actual neural locus and mechanism responsible for this process. Our findings will also aid the development of models of the human saccadic and visual perceptual systems and their interrelationship. Indeed, any model saccadic mechanism constructed with signals from the frontal eye field, the lateral intraparietal cortex, the superior colliculus, and the ventral stream must not only simulate human saccadic performance, but must account for the observed classification images as well.

References

- Abbey CK, Eckstein MP (2002) Classification image analysis: estimation and statistical inference for two-alternative forced-choice experiments. *J Vis* 2:66–78.
- Ahumada Jr AJ, Lovell J (1971) Stimulus features in signal detection. *J Acoust Soc Am* 49:1751–1756.
- Beutter BR, Eckstein MP, Stone LS (2003) Saccadic and perceptual performance in visual search tasks. I. Contrast detection and discrimination. *J Opt Soc Am A* 20:1341–1355.
- Caspi A, Beutter BR, Eckstein MP (2004) The time course of visual information accrual guiding eye movement decisions. *Proc Natl Acad Sci USA* 101:13086–13090.
- Dassonville P, Bala JK (2004) Perception, action, and Roelofs effect: a mere illusion of dissociation. *PLoS Biol* 2:e364.
- Eckstein MP, Ahumada Jr AJ (2002) Classification images: a tool to analyze visual strategies. *J Vis* 2:1.
- Gegenfurtner KR, Xing D, Scott BH, Hawken MJ (2003) A comparison of pursuit eye movement and perceptual performance in speed discrimination. *J Vis* 3:865–876.
- Gold JM, Murray RF, Bennett PJ, Sekuler AB (2000) Behavioral receptive fields for visually completed contours. *Curr Biol* 10:663–666.
- Goldberg ME, Bisley J, Powell KD, Gottlieb J, Kusunoki M (2002) The role of lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. *Ann NY Acad Sci* 956:205–215.
- Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends Neurosci* 15:97–112.
- Goodale MA, Milner AD, Jakobson LS, Carey DP (1991) A neurological dissociation between perceiving objects and grasping them. *Nature* 349:154–156.
- Harris RJ (1985) A primer of multivariate statistics, pp 99–118. Orlando, FL: Academic.
- Hooge I (1996) Control of eye movement in visual search. PhD thesis, Helmholtz Instituut, School for Autonomous Systems Research, Universiteit Utrecht.
- Krauzlis RJ, Stone LS (1999) Tracking with the mind's eye. *Trends Neurosci* 22:544–550.
- Levi DM, Klein SA (2003) Noise provides some new signals about the spatial vision of amblyopes. *J Neurosci* 23:2522–2526.

- Marmarelis PZ, Marmarelis VZ (1978) Analysis of physiological systems: the white-noise approach. New York: Plenum.
- Mazer JA, Gallant JL (2003) Goal related activity in V4 during viewing visual search: evidence for a ventral stream visual salience map. *Neuron* 40:1241–1250.
- McPeck RM, Keller EL (2004) Deficits in saccade target selection after inactivation of superior colliculus. *Nat Neurosci* 7:757–763.
- Milner DA, Goodale MA (1997) The visual brain in action. Oxford: Oxford UP.
- Moore T (1999) Shape representations and visual guidance of saccadic eye movements. *Science* 285:1914–1917.
- Murray RF, Bennett PJ, Sekuler AB (2002) Optimal methods for calculating classification images: weighted sums. *J Vis* 2:79–104.
- Najemnik J, Geisler WS (2005) Optimal eye movement strategies in visual search. *Nature* 434:387–391.
- Neri P, Heeger DJ (2002) Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nat Neurosci* 5:812–816.
- Rajashekar U, Bovik AC, Cormack LK (2006) Visual search in noise: revealing the influence of structural cues by gaze-contingent classification image analysis. *J Vis* 6:379–386.
- Sakata H, Taira M, Murata A, Gallese V, Tanaka Y (1997) Parietal visual neurons coding 3-D characteristics of objects and their relation to hand action. In: Parietal lobe contributions to orientation in 3D space. (Thier P, Karnath HO, eds), pp 237–254. Heidelberg: Springer.
- Schall JD (2002) The neural selection and control of saccades by the frontal eye field. *Philos Trans R Soc Lond B Biol Sci* 57:1073–1082.
- Snyder LH, Batista AP, Andersen RA (1997) Coding of intention in the posterior parietal cortex. *Nature* 386:167–170.
- Solomon JA (2002) Noise reveals visual mechanisms of detection and discrimination. *J Vis* 2:105–120.
- Stone LS, Krauzlis RJ (2003) Shared motion signals for human perceptual decisions and oculomotor actions. *J Vis* 3:725–736.
- Sutter EE (1975) A revised conception of visual receptive fields based on pseudorandom spatiotemporal pattern stimuli. Paper presented at the First Symposium on Testing and Identification of Nonlinear Systems, Pasadena, CA, March.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: Analysis of visual behavior (Ingle DJ, Goodale MA, Mansfield RJW, eds), pp 549–586. Cambridge, MA: MIT.
- Van Essen DC, Newsome WT, Maunsell JHR (1984) The representation of the visual field in striate cortex of the macaque: anisotropies and individual variability. *Vis Res* 24:429–448.
- Victor JD (2005) Analyzing receptive fields, classification images and functional images: challenges with opportunities for synergy. *Nat Neurosci* 8:1651–1656.