

The footprints of visual attention in the Posner cueing paradigm revealed by classification images

Miguel P. Eckstein

Department of Psychology, University of California,
Santa Barbara, CA, USA



Steven S. Shimozaki

Department of Psychology, University of California,
Santa Barbara, CA, USA



Craig K. Abbey

Dept. of Biomedical Engineering, University of California,
Davis, CA, USA



In the Posner cueing paradigm, observers' performance in detecting a target is typically better in trials in which the target is present at the cued location than in trials in which the target appears at the uncued location. This effect can be explained in terms of a Bayesian observer where visual attention simply weights the information differently at the cued (attended) and uncued (unattended) locations without a change in the quality of processing at each location. Alternatively, it could also be explained in terms of visual attention changing the shape of the perceptual filter at the cued location. In this study, we use the classification image technique to compare the human perceptual filters at the cued and uncued locations in a contrast discrimination task. We did not find statistically significant differences between the shapes of the inferred perceptual filters across the two locations, nor did the observed differences account for the measured cueing effects in human observers. Instead, we found a difference in the magnitude of the classification images, supporting the idea that visual attention changes the weighting of information at the cued and uncued location, but does not change the quality of processing at each individual location.

Keywords: attention, computational modeling, ideal observer, noise, cueing paradigm

Introduction

An important paradigm for studying visual attention in the last two decades has been the Posner cueing paradigm (Posner, 1980). In this paradigm, a target can appear in one of two locations, and the observer reports whether the target is present (yes/no). Prior to the presentation of the stimulus, a cue (precue) indicates the probable location of the target (given that the target is present) with some validity (e.g., 80% of the trials). Those trials in which the cue correctly indicates the location of the target are known as the valid cue trials, whereas the trials in which the cue incorrectly indicates the location of the target are called the invalid cue trials. A classical result is that performance (measured with response times or target detection accuracy) is better in the valid cue trials versus the invalid cue trials. This result led Posner (Posner, 1980; Posner & Peterson, 1990) and many researchers in subsequent studies to conclude that the cue orients visual attention, which enhances processing at that cued (attended) location. An analogous interpretation of the result is that visual attention has limited resources that can be allocated at one of the locations. When the resources are allocated at the cued location, a performance benefit at the attended location arises.

The Bayesian Observer: Cueing Effects Without Capacity Limitations/Attentional Enhancement

Recently, an alternative approach has been proposed for the cueing paradigm in terms of a Bayesian observer. This model predicts a cueing effect without a change in the quality of processing at the attended and unattended locations (i.e., changes in the perceptual filters, internal noise, etc.). In this model, the observer monitors the responses of two equivalent perceptual filters¹ at the cued and uncued locations. Each of the perceptual filters linearly weights the luminance at the cued and uncued locations resulting in one scalar response for each of the locations. The scalar responses to the two locations are stochastic variables that vary from trial to trial due to internal noise in the observer (e.g., neural firing) and/or to luminance variability in the image (external noise).

The Bayesian observer calculates a likelihood of the scalar filter responses given target presence for each location. The model then optimally combines the two likelihoods across the cued and uncued locations. The likelihood from the cued location is weighted (w_c) by the prior probability of the target being present in that location (precue validity). The likelihood from the uncued location is weighted (w_u) by its corresponding

prior probability of target presence (1 minus precue validity). The result is an overall likelihood of the filter responses given target presence across the two locations. The Bayesian observer then calculates an overall likelihood of the data given target absence. Finally, the model computes a ratio of the likelihoods and makes a decision by comparing the likelihood ratio to a decision criterion or threshold. Figure 1 shows a schematic of the Bayesian observer for a task in which the signal is a Gaussian “contrast increment” embedded in white Gaussian noise. Appendix A summarizes the mathematical expressions describing the Bayesian observer for the Posner paradigm.

The optimal weighting of the likelihoods from the cued and uncued locations maximizes the overall hit rate given a false alarm rate across both types of signal trials: valid and invalid cue trials. The concept is easiest to understand for the extreme case of a cue that is 100% valid. In this case, the observer knows a priori that high evidence (likelihood) of target presence arising from the uncued location is due only to noise, and not to target presence (given that the uncued location never contains the target). Therefore, evidence of target presence arising from the uncued locations only contributes to generate errors (false alarm trials). As a result, for the particular case of a 100% valid cue, the optimal strategy is to completely ignore the information from the uncued location ($w_u = 0$ in Figure 1). For the more general case where the cue is valid a certain percent of the time (cue validity = 80%), the Bayesian observer simply gives more weight to evidence (or information) arising from the cued location.

A consequence of the higher weighting of information at the cued location is that the Bayesian observer will produce better performance (hit rate given a constant false alarm rate) for valid cue trials versus invalid cue without any difference in the quality of processing (e.g., difference in perceptual filters, internal noise, etc.) at the cued and uncued locations. Recently, Shimozaki, Eckstein, and Abbey (2001) have shown how a Bayesian observer can predict cue validity effects of the same or larger magnitude than human observers for a Gaussian blob detection task in one of two locations.

In this study, the Bayesian model can also quantitatively predict the cueing effect in a task where the target is a contrast increment in one of two Gaussian blobs. The probability of target presence is 50% and the cue validity is 80%. Figure 2 shows hit rate in this task for a Bayesian observer degraded with Gaussian internal noise in order to match approximately the false alarm rates of the human observers. The difference between the hit rate for the valid and invalid cue for the Bayesian observer is close to that of four human observers.

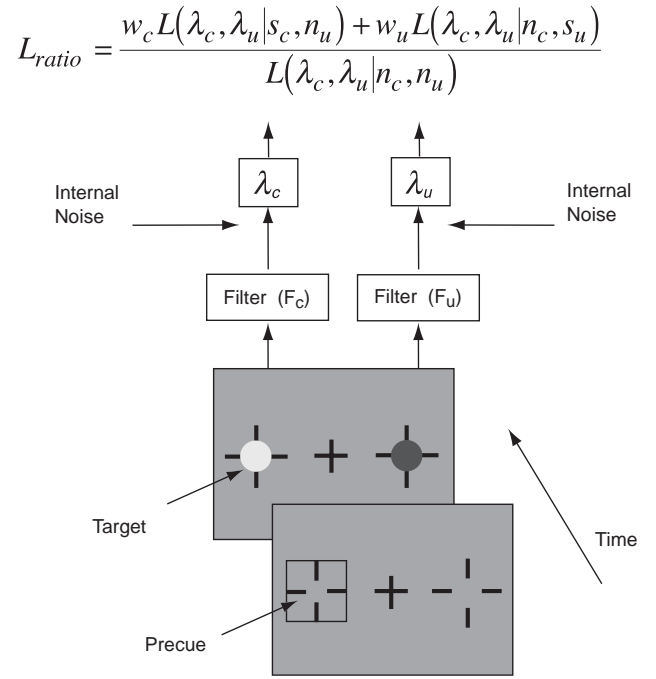


Figure 1. Schematic of a Bayesian observer in the Posner cueing paradigm. Stimuli are a simple schematic (actual experimental images contained added visual noise). The task of the observer is to determine whether a contrast increment is present at one of the two locations (yes/no task). In this study, the precue is valid 80% of the time.

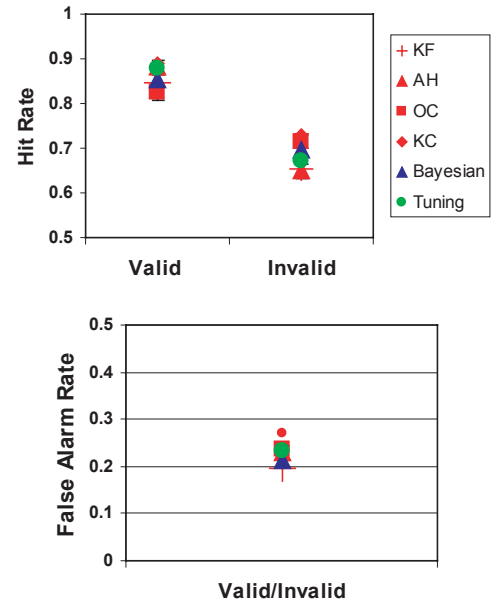


Figure 2. Upper graph: Hit rate for valid cue and invalid cue trials for four human observers (K.F., A.H., O.C., K.C.). Also plotted are a Bayesian observer (triangles) that simply optimally weights the likelihood from the cued and uncued locations and a Tuning model (circles) in which visual attention changes the tuning of the perceptual filter. Lower graph: False alarm rate for valid and invalid trials for the same four human observers and the two models.

Cueing Effects With Attentional Tuning of Perceptual Filters

Although the Bayesian observer can successfully predict human observers' cueing effects, there are other possible models that include attentional changes in the perceptual quality of the information at each location that could also predict the human cueing effects. For example, some previous studies have suggested that visual attention changes the tuning or shape of the perceptual filters. In another example, physiological studies have suggested that attention narrows the orientation tuning and color tuning of cells in V4 (Haenny and Schiller, 1988; Spitzer, Desimone, & Moran, 1988). Also, psychophysical studies using texture segmentation (Yeshurun and Carrasco, 1998, 1999) suggest that attention changes the spatial resolution of processing, which might translate to a change in the spatial frequency tuning of the perceptual filters. Lu and Doshier have used an extension of the linear amplifier model (the perceptual template model) and external noise with a cueing paradigm to show that in a number of tasks, attention increases the optimality of the perceptual filter (Lu & Doshier, 2000; Doshier & Lu, 2000)².

Figure 2 shows one example of a hypothetical model where visual attention at the cued location improves the tuning of the perceptual filter producing a cueing effect of the same size as observed in humans. The particular shapes of the filters used in this model are shown in Figure 6 (left column). The perceptual filter at the uncued location is a Difference of Gaussians (DOG) filter, and the perceptual filter at the cued (attended) location is a Gaussian that matches the signal. The likelihoods are equally weighted from each location to reach a decision. Independent Gaussian internal noise following the perceptual filters was used to degrade the model to match human performance levels. For this model, the cueing effect arises solely because visual attention changes the perceptual filter at the cued location to make it optimal. The lower performance in the invalid trials is due to the suboptimal nature of the perceptual filter at the uncued location. This example illustrates that a model with an attentional change of perceptual filters at the attended and unattended locations also can exhibit cueing effects similar to those measured in humans.

Tuning versus task performance-based tuning of perceptual filters

Although vision scientists commonly refer to the concept of perceptual tuning, the term is interpreted in different ways. Many investigators use the term to refer to the narrowing of the sensitivity (in orientation, space, color, etc.) of an inferred filter, a measured cell, or a population of cells. Another common use is to define the perceptual tuning in terms of how well the filter matches the signal to be detected. Our view is that changes in the

perceptual filters should also be judged in terms of the impact they have on performance in the task being studied. For example, there are tasks in which attention might narrow the tuning characteristics of the perceptual filter but might not enhance or might even degrade performance in the cued attended location. If so, those changes in the perceptual filter would not be able to account for a standard cueing effect in human performance. In this context, one can define the tuning of the perceptual filter in terms of performance in the relevant task.

For simple tasks in external noise, the optimal filters are known or computable. In these cases, one can define the perceptual tuning of a filter in terms of the ratio of signal energy (to achieve a given performance level, e.g., 80%) for the optimal filter and that of the human perceptual filter ($E_{\text{ideal filter}}/E_{\text{human filter}}$). This measure is known as the efficiency of the perceptual filter. For simple linear tasks in white Gaussian noise, the efficiency can be directly calculated by computing the squared correlation (match) between the perceptual filter and the optimal filter (which is the signal). However, when the external noise does not have equal power in all the frequencies (nonwhite noise), then the degree of match between the perceptual filter and the signal is not the sole factor determining the performance of the filter. In these cases, the optimal filter does not match the signal. For tasks such as the Posner paradigm where the decision is a nonlinear function of the data, no simple calculations are available and Monte Carlo simulations and/or numerical approximations are required to compute the task performance associated to a perceptual filter (Nolte & Jaarsma, 1967).

Classification Images as a Tool to Estimate Perceptual Filters

Given that two different models of visual attention (weighting of information with identical perceptual filters vs. change in perceptual filters) can predict cueing effects of the size observed in humans, there is a rationale to use more elaborate psychophysical techniques (beyond comparing model and human performance) to be able to distinguish different possible attentional modulations that mediate human visual performance in the Posner cueing paradigm. In this study, we use the technique known as classification images to distinguish the two different models of visual attention.

What is a classification image?

The classification image technique allows the investigator to directly estimate how the observer weights the information in the image to reach a decision. A related technique based on multiple linear regression was first applied by Ahumada and Lovell (1971) to audition. Ahumada (1996) and Beard and Ahumada (1998) used the classification image technique to study how observers

used visual information in a vernier acuity task. Ringach, Hawken, and Shapley (1997) used a related method to study the orientation tuning in the monkey primary visual cortex. Others have used the technique to look at illusory contours (Gold, Murray, Bennett, & Sekuler, 2000), stereo (Neri, Parker, & Blakemore, 1999), and off-frequency looking in nonwhite noise (Abbey & Eckstein, 2000)

For signals varying only in luminance, the main methodological requirement is to add random spatially uncorrelated luminance noise to the image. The investigator then keeps track of the noisy stimuli presented in the trials corresponding to the different human observer decisions: signal present trials in which the observer correctly responded “signal present” (hit trials), signal present trials in which the observer incorrectly responded “signal absent” (incorrect rejection or miss trials), signal absent trials in which the observer correctly responded “signal absent” (correct rejection trials), and signal absent trials in which the observer incorrectly responded “signal present” (false alarm trials).

The intuition behind classification images is best illustrated with the false alarm trials. In these trials, the investigator collects noise samples that did not contain the signal yet resulted in the observer responding that the signal was present. It follows that the random luminance perturbations in that trial must have contained some luminance pattern that corresponded to what the observer took as evidence of signal presence. Thus, the sample mean of all the noise images from the false alarm trials will reveal deviations in luminance that led the observer to respond that the target was present when it was not. For simple tasks, one can derive closed form expressions to show that the sample mean of the noisy images will accurately estimate a linear filter or template used to weight the information in the image to reach the decision. In statistics, one would refer to the classification image obtained by computing the sample mean of the noise images from false alarm trials as an unbiased estimator of the linear template or perceptual filter. Of course, for a yes/no task, there are four groups of noise samples that arise, one for each of the four types of decisions (correct detection or hit, correct rejection, incorrect detection or false alarm, and incorrect rejection or miss). For simple tasks, one can derive optimal methods to combine the noise samples arising from these four types of trials to optimally estimate classification images (Beard & Ahumada, 1998; Abbey & Eckstein, 2002, in this special issue). Two alternative forced choice tasks require taking the difference between the two noise images presented in each trial to compute the classification image (see Abbey & Eckstein in this issue for details on 2AFC classification image technique). If the added noise does not have a uniform power spectrum, then a more involved intermediate step is required to obtain an unbiased estimation of the linear filter (Abbey, Eckstein, & Bochud, 1999). For more

complex tasks in which decision rules are a nonlinear function of the image pixels, a derivation that shows that the classification image is an unbiased estimator of the perceptual linear filter does not yet exist (A. Ahumada, personal communication, 1999). However, Monte Carlo simulations can be used to determine whether the classification image arising from the signal absent trials is an unbiased estimator.

Assumptions of the classification image technique

An underlying assumption in the classification image technique is that the observer is monitoring a single perceptual filter to reach a decision. It is under these circumstances that the obtained classification image can be interpreted in terms of a single perceptual filter. When the observer is monitoring a number of perceptual filters and uses a nonlinear combination to reach a decision, caution is needed in the interpretation. A. Ahumada (personal communication, 1999) first noted that the classification images arising from the target present trials in tasks in which the observer is monitoring a number of filters per location (e.g., positional intrinsic uncertainty) may not accurately represent the linear perceptual filter or filters in the task. For these tasks, classification images from signal present trials can be misleading. In addition, the classification image obtained from signal absent trials cannot be interpreted in terms of single perceptual filter but a composite of many perceptual filters influencing the decision in some nonlinear fashion. One instance in which human observers monitor more than one perceptual filter per location is when they are uncertain about some parameter about the signal, such as position, spatial frequency, phase, etc. (Pelli, 1985). The presence of effects of intrinsic uncertainty can be diagnosed by measuring psychometric functions (accuracy vs. signal contrast) and/or by comparing classification images arising from the signal present and signal absent trials (A. Ahumada, personal communication, 1999). A difference in the classification images from signal present and signal absent trials points to a diagnosis of nonlinearity in some cases (Abbey and Eckstein, 2002, in this special issue). A good approach is to choose tasks that are known to show small effects of intrinsic uncertainty, such as contrast and size discrimination tasks where a linear observer is a good approximation to human performance (Burgess & Ghandeharian, 1984; Ahumada, 1987). It is under conditions in which intrinsic uncertainty has no effect that the classification image technique is most powerful in terms of information content (expressed as the signal to noise ratio of the classification image) and interpretation. On the other hand, tasks such as the detection of spatial and temporal periodic signals in noise typically show nonlinear psychometric function reflecting intrinsic uncertainty about phase and will not approximate the assumptions of the classification image technique.

Classification images for the Posner paradigm

For the Posner paradigm, the Bayesian observer nonlinearly combines the response of two perceptual filters to reach a decision. For this reason, we used only the false alarm trials arising from signal absent trials to derive classification images. To verify that the obtained classification images are unbiased estimators of perceptual filters, we implemented extensive Monte Carlo simulations with different versions of optimal and suboptimal Bayesian observers. The following section shows the results for these simulations and verifies the validity of the use of classification images to estimate perceptual filters for the cued and uncued locations of the Posner paradigm. The simulations also allow us to establish how the different models of visual attention in the Posner paradigm give rise to distinct classification image signatures. These signatures will be used to infer properties about visual attention from the human classification images described later.

Classification Image Signatures

Each of the models shown was generated by using implementation of the general model framework shown in Figure 1. The simulations were based on 13,000 trials, which is approximately the same number of trials performed by the human observers. The task used for the model simulations was identical to that one used for the

psychophysical experiments including the noise level, Gaussian pedestals, contrast increment of the Gaussian signal, and the cue validity. More details about the task and simulations are discussed in “Methods” and Appendix A.

1. Attention Changes the Weighting at Cued and Uncued Locations

These models assume that attention does not change the shape of the perceptual filter, but simply changes the weighting of information at the cued and uncued locations. We investigated three types of weighting of information at the cued and uncued locations corresponding to different attentional signatures: the optimal attentional weighting; attend both locations equally (equivalent weighting of each location); and, attend cued location only. These models are obtained by changing the weights of the likelihoods (w_c and w_u) in our Bayesian model (see Figure 1 and Appendix A).

Figure 3 shows the perceptual filters used in the simulations (optimal Gaussian filters for all conditions), the weights for the likelihoods for the cued and uncued locations, and the corresponding classification images obtained. The simulations show that the shape of the classification images match the shape of the model input perceptual linear filter scaled by a constant.

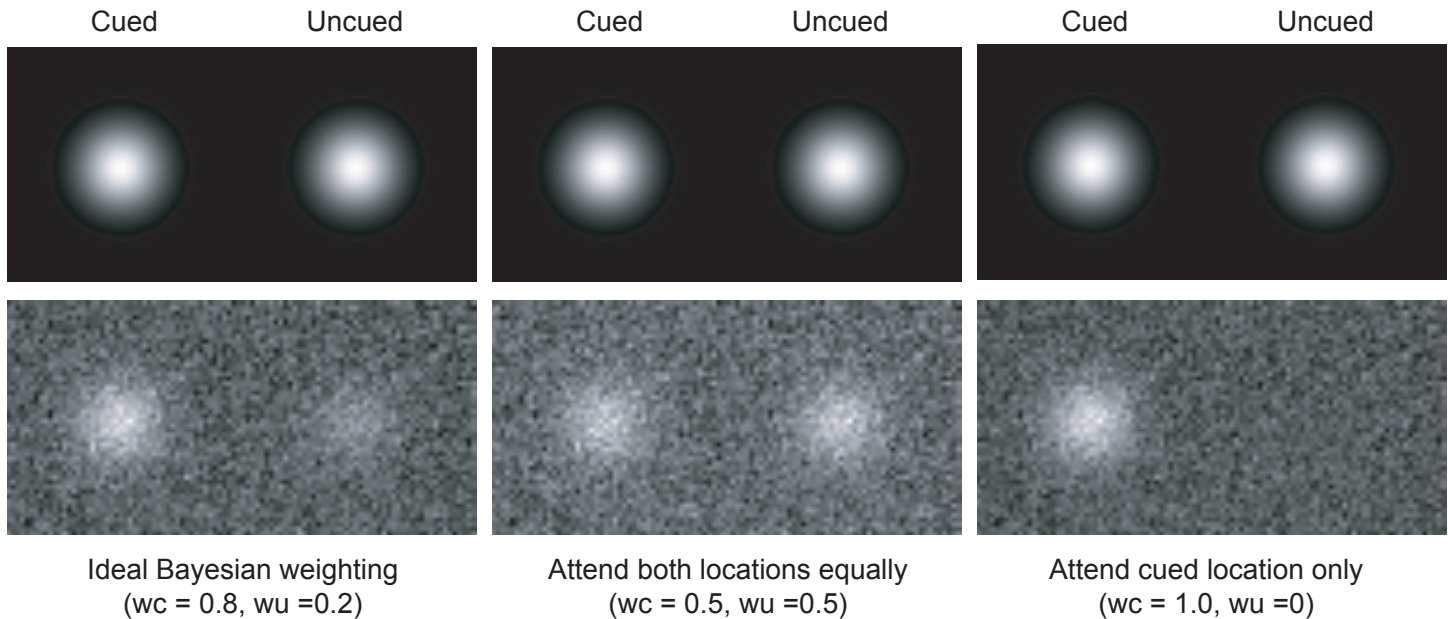


Figure 3. (a) Top row: Two equivalent perceptual filters (Gaussian filters that match the signal) at the cued and uncued locations for all three simulations. (b) Bottom row: The classification images from simulations associated to different weightings of the likelihood at the cued and uncued location. In all of these models, visual attention changes the weightings of the likelihood from the cued and uncued locations. The images shown here have been reduced (by a factor of 2 using bilinear interpolation) from the actual images.

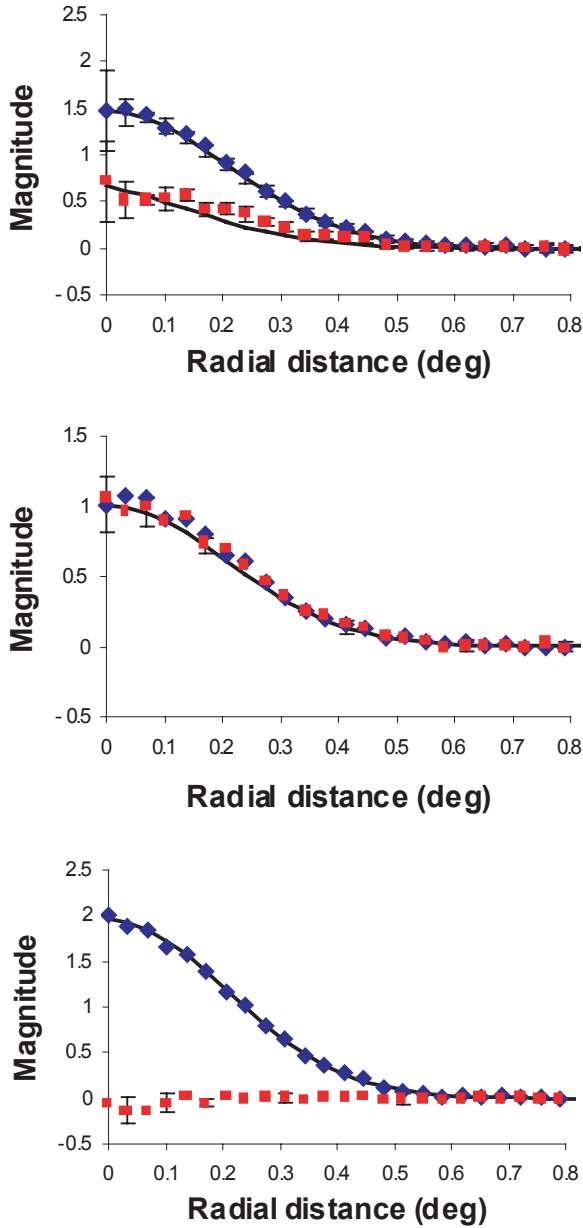


Figure 4. Radial averages of classification images from simulations for three different attentional weightings of the likelihood from the cued (blue) and uncued (red) locations. Solid curves are scaled versions of the perceptual filter used in the simulations. Top: Optimal weighting. Middle: Attend both locations equally. Bottom: Attend only cued location. Error bars are omitted when they are smaller than the symbol.

This point becomes more apparent if radial averages across angles are plotted for each classification image (Figure 4). The radial averages of the classification images are scaled versions of the perceptual filter used in the model simulation (a Gaussian). In addition, the input weighting of the cued and uncued location used in the model is reflected by the magnitude (or amplitude) of the classification images. For example, when the weightings of the two locations in the model are equal (attend both

locations equally), then the magnitudes of the classification images are the same (middle column in Figure 3; middle graph in Figure 4). When the model weights the cued location more heavily (and optimally) than the uncued location, the magnitude of the associated classification image for the cued location is also larger than that of the uncued location (Figure 3 left column; top graph in Figure 4). Finally, when the model solely weights the information from the cued location and ignores that from the uncued location, then no classification image is obtained at the uncued location (right column in Figure 3; bottom graph in Figure 4). In this way, if we obtain human observer classification images, we can potentially infer the observers' attentional weighting strategy.

Inferring the weighting across cued and uncued locations from the ratio of classification images

Although the previous section shows that the relationship between the magnitude of the classification images for the cued and uncued locations reflects the attentional weightings assigned to each of the two locations, it would be desirable to be able to directly relate the ratio of the magnitude of the classification images to the input model weights (w_c and w_u in Figure 1). Because of the nonlinear stage in the Bayesian observer in the Posner paradigm, the mathematical relationship between the weights used in the model for the likelihood for each location and the ratio of magnitudes of the obtained classification images is not easily derived analytically. We therefore performed extensive Monte Carlo simulations with the Bayesian observer with two Gaussian perceptual filters to empirically measure the relationship between these two. Figure 5 shows the ratio of magnitudes of the classification images and input weights used in the model (see "Methods" for technical details about fitting routine used)³. This relationship can potentially be used to infer the underlying weights used by the human observers for the cued and uncued locations from the obtained human classification images.

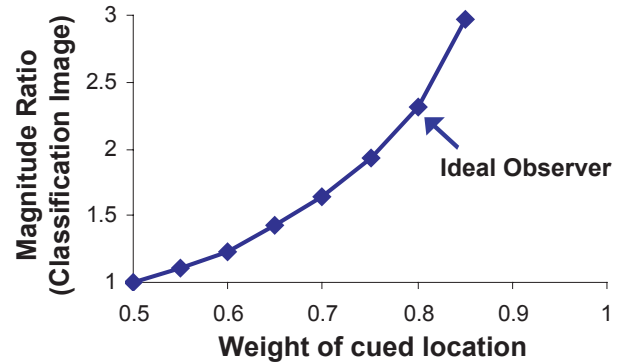


Figure 5. Relationship between the ratio of magnitudes of classification images and the input weight of the model for the cued location.

2. Attention Changes the Shape of the Perceptual Filter

The second type of attentional signatures we consider are those in which visual attention changes the tuning of the perceptual filter. Within the framework of the Bayesian model, one can hypothesize that attending to the cued location changes the tuning of the perceptual filter. For example, [Figure 6](#) shows a suboptimal DOG filter used at the uncued location and an optimal perceptual filter used at the attended location. Included in [Figure 6](#) are the perceptual filters for another scenario where the suboptimal perceptual filter at the uncued (unattended) location is wider than the optimal perceptual filter. The correspondence between the

original filters and their resulting classification images can be seen in [Figure 6](#). [Figure 7](#) shows the Fourier transform of the perceptual filters and the classification images. The results show that the obtained classification images for the cued and uncued locations match the shape of the underlying perceptual filters used in the model simulations. For example, the DOG filter gives rise to a noisy DOG filter in its associated classification image. The correspondence between the model's perceptual filter and the obtained classification image can be seen more easily in the plots of the radial averages ([Figure 8](#)). The simulations demonstrate that one can potentially infer the shape of the observers' perceptual filters from their classification images.

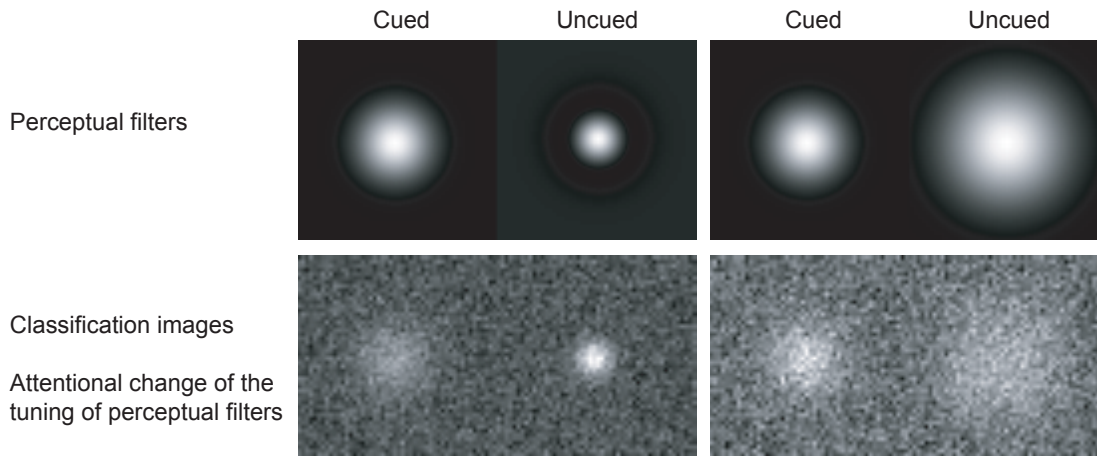


Figure 6. Top row: Perceptual filters for the cued and uncued locations. Bottom row: Classification images obtained through simulations. Left: Perceptual filter at the uncued location is a suboptimal Difference of Gaussians, whereas that for the cued location is an optimal Gaussian. Right: Perceptual filter at the uncued location is a suboptimal wide Gaussian, whereas that for the cued location is an optimal Gaussian.

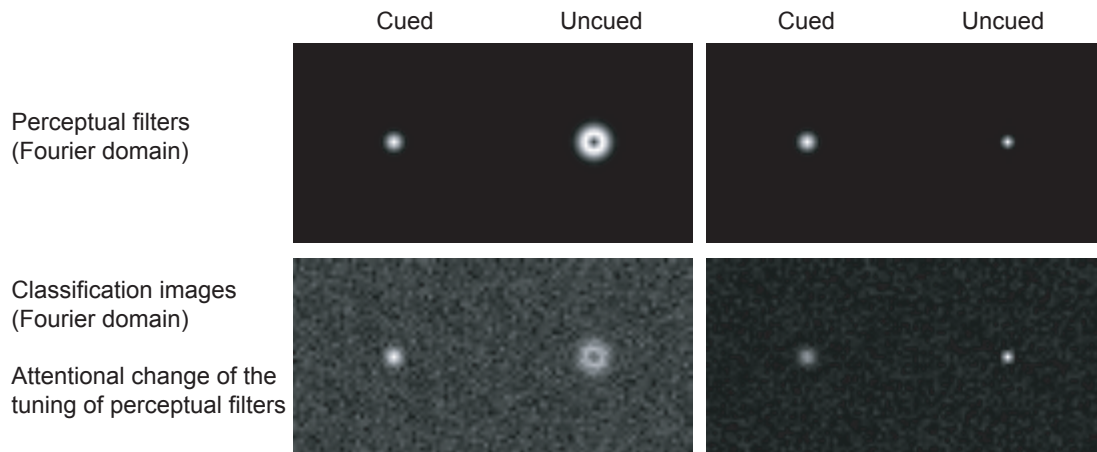


Figure 7. Top row: Fourier transform of the perceptual filters for the cued and uncued locations. Bottom row: Classification images obtained through Monte Carlo simulations. Radial distance from the center represents spatial frequency with the zero frequency at the center. Left: Perceptual filter at the uncued location is the Fourier transform of a suboptimal Difference of Gaussians, whereas that for the cued location is an optimal Gaussian. Right: Perceptual filter at the uncued location is the Fourier transform of suboptimal spatially wide Gaussian (and therefore narrower than the optimal filter in the Fourier domain), whereas that for the cued location is an optimal Gaussian.

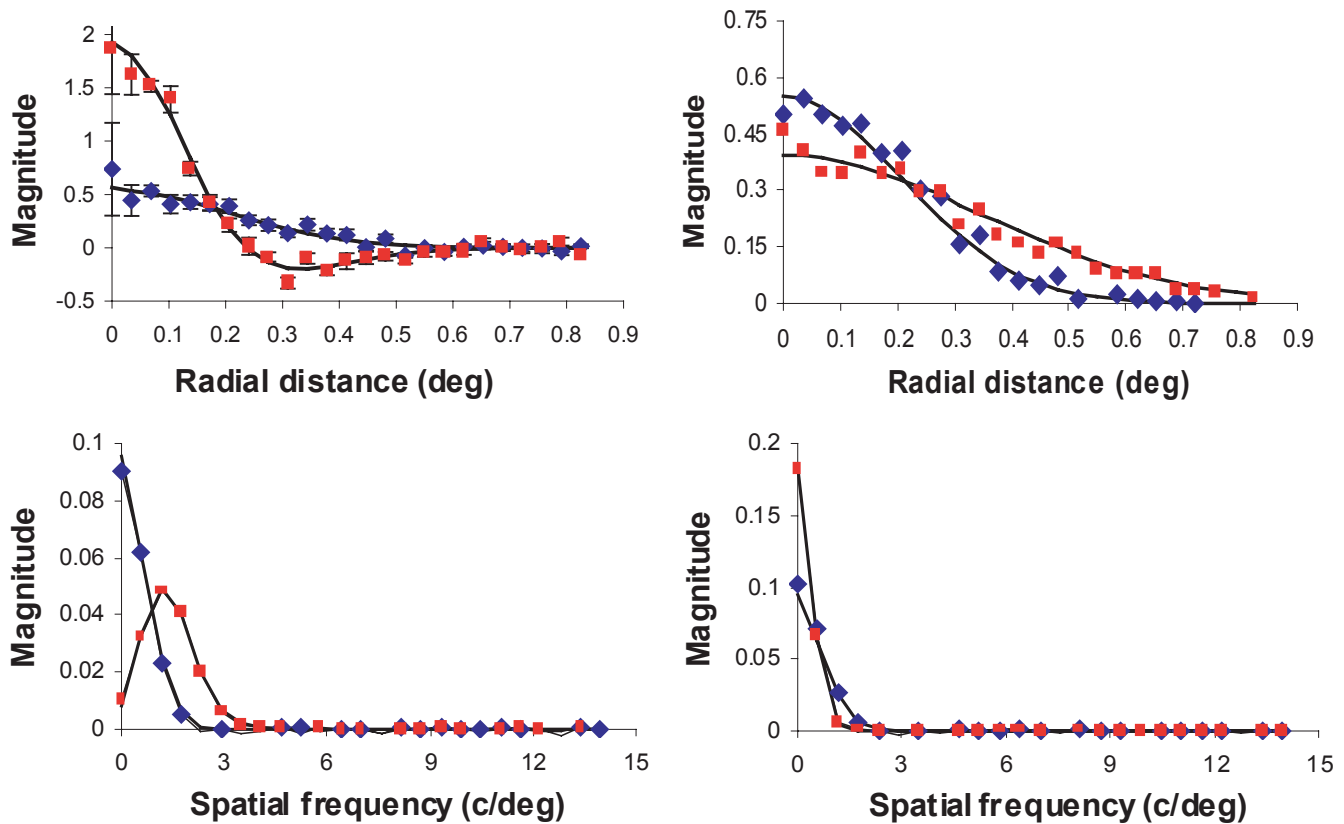


Figure 8. Radial averages of classification images (Figures 6 and 7) for simulations for two different examples of models where visual attention changes the shape of the perceptual filter at the cued locations. Blue symbols correspond to radial averages of classification images at the cued location, whereas the red symbols correspond to those from the uncued location. Solid lines correspond to the scaled radial averages of the perceptual filters used in the model simulations. Left column: Optimal Gaussian filter for the cued location and a Difference of Gaussians filter for the uncued location. Right column: Optimal Gaussian filter for the cued location and a spatially wider suboptimal Gaussian for the uncued location. Top row: Spatial domain. Bottom row: Fourier domain.

Methods

Psychophysical Task

The observers' task was to decide whether a contrast increment (4.69%) was present (yes/no) in one of two Gaussian pedestals (percentage root mean square [RMS] contrast = 6.25%). The two Gaussian pedestals were located to the right and left of a fixation point at an eccentricity of 2.5 degrees. White Gaussian luminance noise with a contrast of 0.117 was added to each image. Every image in every trial of the study had independent samples of noise. The viewing distance was 50 cm. The signal was present on 50% of the trials. The validity of the precue was 80% (i.e., the target was present in the precue location in 80% of the target present trials). Four naïve yet trained observers participated in the study. The observers participated in 50 sessions of 250 trials resulting in 12,500 trials. Stimuli were presented on an Image Systems monochrome monitor (Image Systems Corp.,

Minnetonka, MN). Each pixel subtended a visual angle of 0.03 degrees. The relationship between digital gray level and luminance was linearized using a Dome Md2 board (Imaging Systems, Waltham, MA) and a luminance calibration system.

Procedure

Observers started the trial with a key press. A fixation image was presented for 1 s. Observers were instructed to fixate a central cross at all times. Following a square precue (length of side = 2.5 degrees) appeared for 150 ms around one of the two possible target locations. The stimulus was then displayed for 50 ms. The short presentation of precue plus stimulus (200 ms) was chosen to preclude observers from executing a saccadic eye movement to fixate the cued location. A white noise mask with higher RMS contrast was then presented for 100 ms (same mean background luminance 24.8 cd/m²). The observers then pressed one of two keys on a computer keyboard to select their decision (signal present or signal absent). Feedback about the correct decision

was provided, but no feedback about the signal location was given.

Human and Model Performance

Performance for human observers was measured in terms of the proportion of signal present trials in which the observer correctly responded (hit rate). Hit rate was measured separately for the valid cue trials and the invalid cue trials. In addition, we determined the proportion of signal absent trials in which the observer incorrectly responded “signal present” (false alarm rate). Performance for the models was quantified using the same measures.

Classification Images

Classification images were obtained by computing the sample mean of the noise images presented in the signal absent trials in which the human and/or model observer incorrectly responded “signal present” (false alarm trials). The number of images used to compute the classification images was given by the number of signal absent trials \times false alarm rate. The actual number of images depended on the false alarm rate of each individual observer but was approximately 1,625 ($6,250 \times 0.26$). Radial averages across all angles were computed for each of the noise images. A sample mean, a standard deviation for each element of the radial averages, was computed, as well as the sample covariance between each element.

Statistical Inference for Classification Images

Although classification images can show the shape of the underlying perceptual filter, the images and radial averages contain a large amount of noise (statistical uncertainty). To make meaningful interpretations, statistical techniques are needed to test the different hypotheses. The Hotelling T^2 statistic is a generalization of the univariate t statistic to multivariate vectors, and can be used to test for differences between a sample multivariate vector and a population vector or between two-sample multivariate vectors. We used one-sample and two-sample Hotelling T^2 statistics (Harris, 1985) to do hypothesis testing of the radial averages of the classification images. The Hotelling T^2 statistic is

$$T^2 = N[\mathbf{x} - \mathbf{x}_0]' \mathbf{K}^{-1} [\mathbf{x} - \mathbf{x}_0]$$

where \mathbf{x} is a vector containing the observed radial average of the classification image, and \mathbf{x}_0 is either a population or a hypothesized radial average classification image. \mathbf{K}^{-1} is the inverse of the covariance matrix that contains the sample variance of each of the elements of the radial average classification images, and the sample covariance between them. To test for significance, the T^2 statistic

can be transformed to an F statistic using the following relationship:

$$F = \frac{N - p}{p(N - 1)} T^2$$

where p is the number of dependent variables (number of vector elements in the radial average of the classification images), and N is the number of observations (number of false alarm trials for our case). The obtained F statistic can be compared to an F_{critical} with p degrees of freedom for the numerator and $N - p$ degrees of freedom for the denominator.

To compare two-sample classification images, one can use the independent two-sample T^2 , which is given by the following expression:

$$T^2 = \frac{N_1 N_2}{(N_1 + N_2)} [\mathbf{x}_2 - \mathbf{x}_1]' \mathbf{K}^{-1} [\mathbf{x}_2 - \mathbf{x}_1]$$

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

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

where p , N_1 , and N_2 are defined before. The obtained F statistic can be compared to 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

Human Performance for Valid Cue and Invalid Cue Conditions

Table 1 shows the hit rates for valid cue and invalid cue trials, as well as the false alarm rate for the four human observers. The last column shows the size of the cueing effect computed as the difference of hit rates for the two types of cue trials. For all observers, we found a statistically significant cueing effect ($p < .001$), although the magnitude of the cueing effect varied from 0.108 to 0.23. Figure 2 presented in the “Introduction” plots the obtained performance results for the four observers.

Human Classification Images

Table 1. Hit rate for valid and invalid cue trials and false alarm rates for human observers in the contrast discrimination Posner task.

Observer	Hit rate (valid trials)	Hit rate (invalid trials)	False alarm rate (all trials)	Cueing effect ($HR_v - HR_{iv}$)
O.C.	0.824	0.716	0.235	0.108
K.F.	0.845	0.655	0.194	0.190
K.C.	0.890	0.729	0.270	0.160
A.H.	0.880	0.649	0.227	0.231

Figure 9 shows the classification images for four human observers obtained from the false alarm trials for the cued and uncued locations. Figure 10 shows the Fourier transform of the classification images. Overall, the classification images show a general similarity across observers, with a higher magnitude classification image for the cued locations with respect to the uncued location. Most classification images also show an inhibitory surround (Figure 9) that can be seen as a black hole at the center of the Fourier transform (Figure 10)

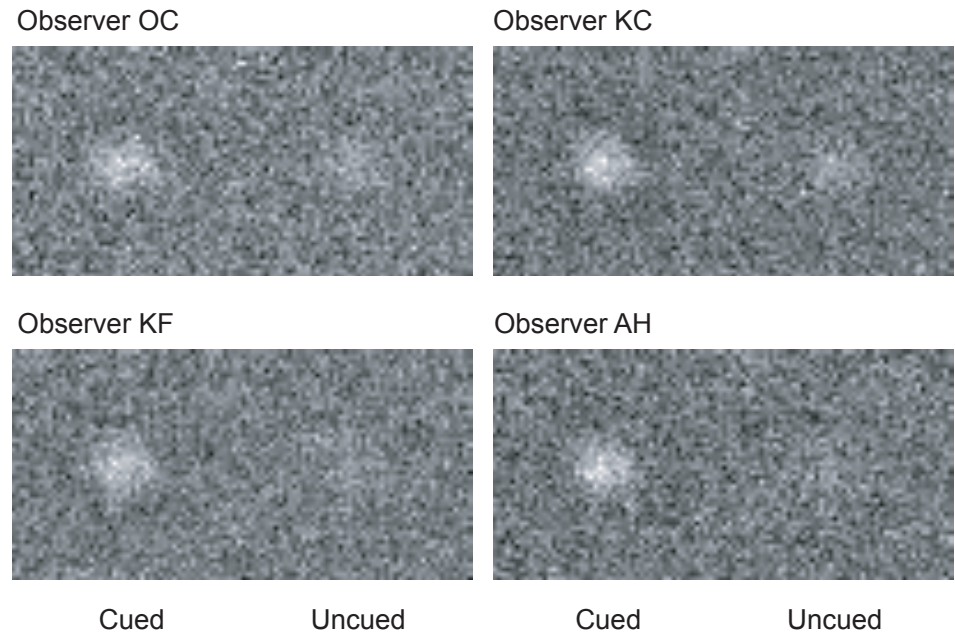


Figure 9. Human observer classification images for the cued and uncued locations.

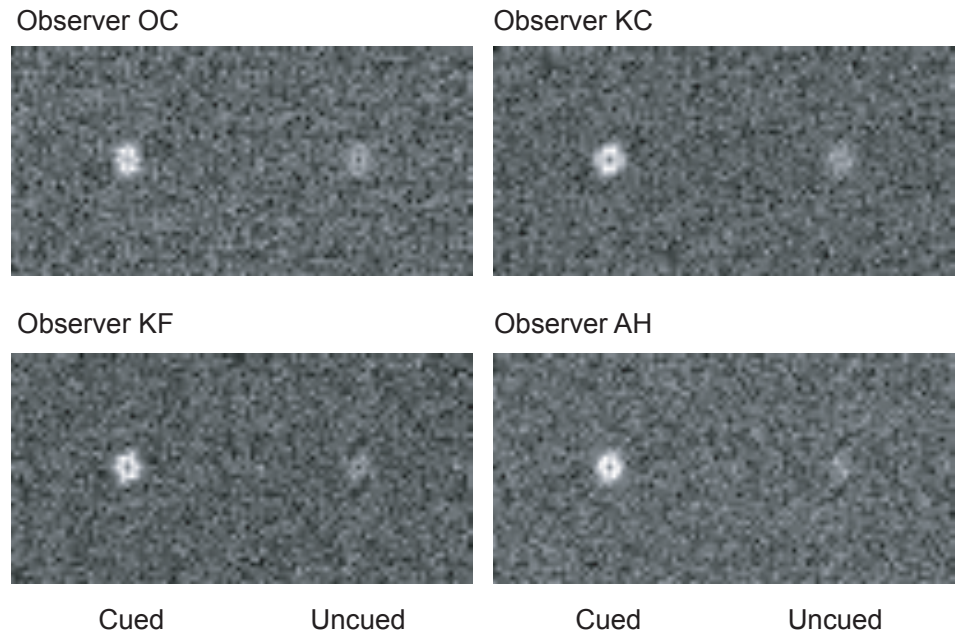


Figure 10. Human observer classification images in the Fourier domain (imaginary part discarded) computed separately for the cued and uncued locations. The Fourier origin is placed at the center of each image.

Radial Averages

Radial averages across all angles for each noise image from each false alarm trial were computed. Sample mean radial averages were then calculated for cued and uncued locations, as well as sample variances and covariances among the elements of the radial average vectors. Figure 11 shows the radial averages for the four observers. Figure 12 shows the radial averages in the Fourier domain. For reference each graph in Figure 12 shows the Fourier transform of the optimal filter (dotted lines). The one-sample Hotelling T^2 statistic was used to test whether the radial averages of the classification images were significantly different from a hypothesized null classification image (vector of zeros). All radial averages of the classification images were significantly different ($p < .01$) from the null classification image. The two sample

Hotelling T^2 statistic showed that the differences between the classification images at the cued and uncued locations were statistically significant for all four observers ($p < .001$).

Radial averages of classification images were fit with DOG functions with four fitting parameters: one amplitude for each of the two Gaussians (K_1 and K_2) and one standard deviation for each of the two Gaussians (σ_1 and σ_2). DOG is given by

$$DOG(x, y) = K_1 \exp(-x^2 / 2\sigma_1^2) - K_2 \exp(-x^2 / 2\sigma_2^2)$$

Table 2 shows the χ^2 best-fit parameters for the radial averages for the cued and uncued locations for all four human observers. The table also includes a χ^2 goodness of fit for each of the fits.

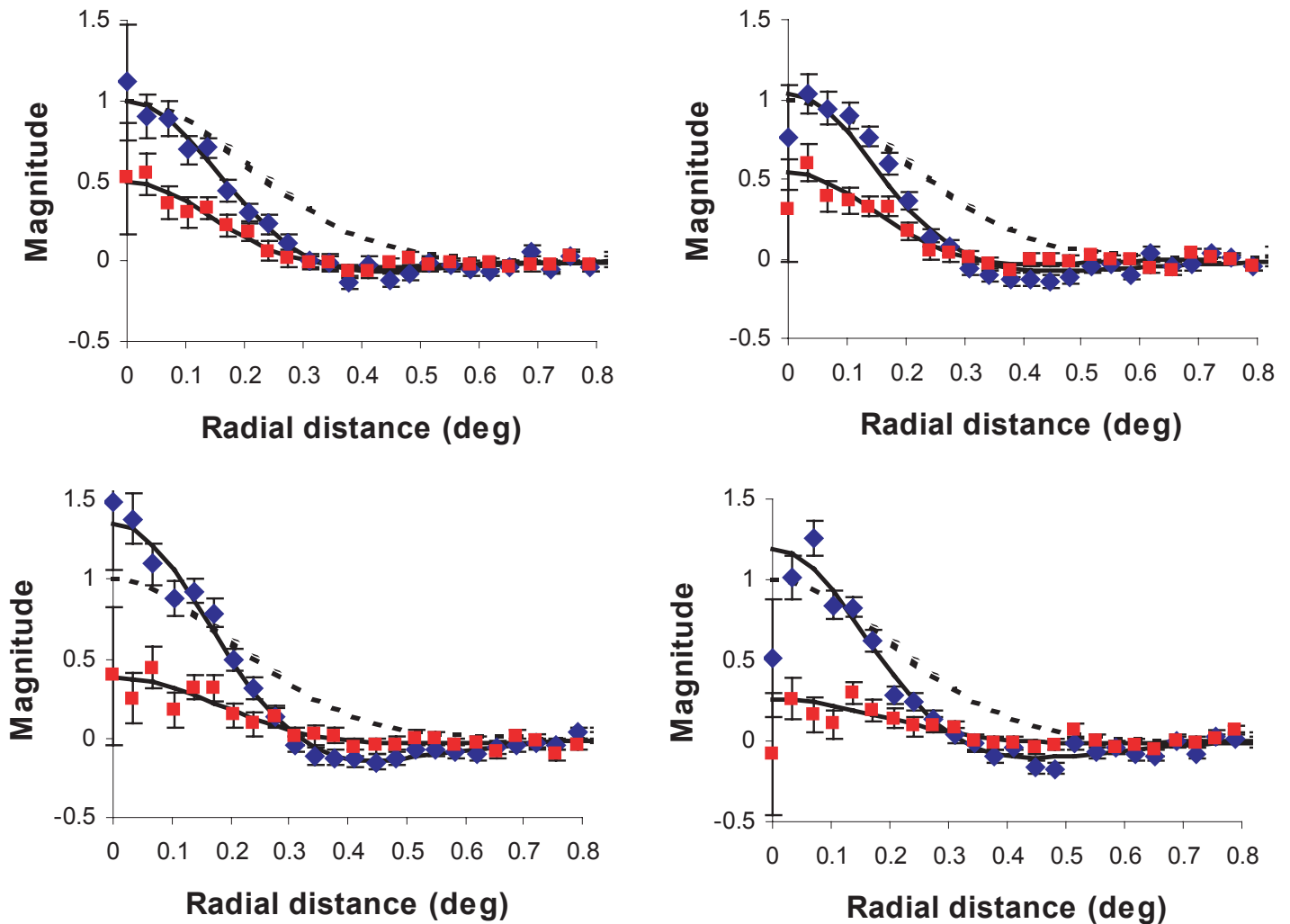


Figure 11. Radial averages (spatial domain) of the classification images for the four human observers. Top left: O.C. Bottom left: K.F. Top right: K.C. Bottom right: A.H. Blue symbols correspond to the cued locations and red symbols correspond to the uncued locations. Black solid lines are the best-fit Difference of Gaussians to the data. The dotted line corresponds to the radial profile of the optimal filter.

Table 2. Best-fit parameters for Difference of Gaussians to radial averages of human classification images with four fitting parameters. Goodness of fit and estimated weights are also given.

Observer	Perceptual filter at the cued location						Perceptual filter at the uncued location					
	K_1	K_2	σ_1	σ_2	χ^2	w_1	K_1	K_2	σ_1	σ_2	χ^2	w_2
O.C.	1.56	0.57	4.9	7.3	21.09	0.76	0.60	0.11	4.3	9.3	7.598	0.24
K.F.	2.14	0.79	5.1	8.2	19.07	0.84	0.46	0.08	5.3	14.0	25.07	0.16
K.C.	1.2	0.16	4.3	11.8	39.25	0.80	0.65	0.11	4.3	8.9	18.8	0.20
A.H.	1.7	0.51	4.9	8.4	33.83	0.88	0.76	0.5	7.3	6.4	20.71	0.12

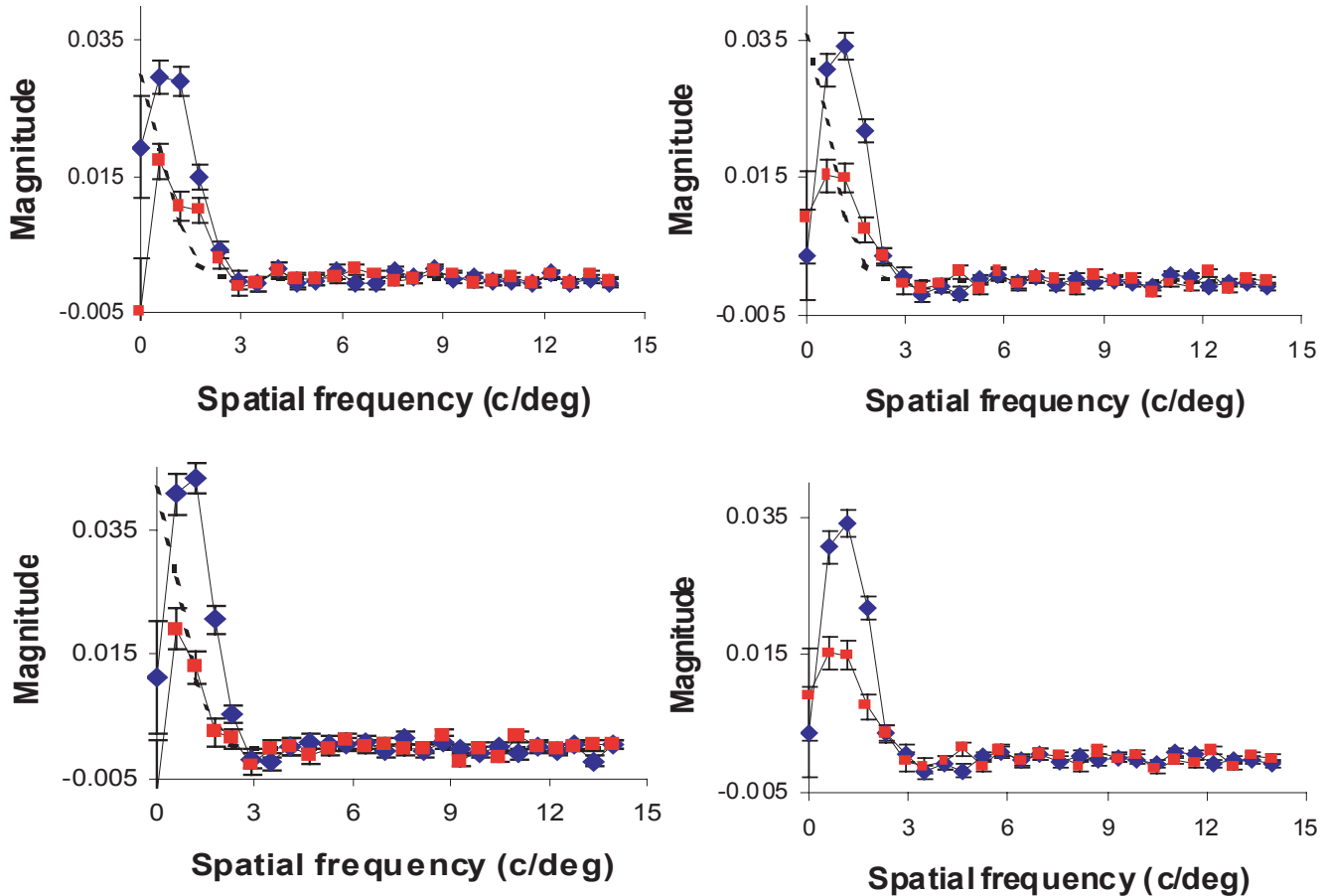


Figure 12. Radial averages (Fourier domain) of the classification images for the four human observers. Top left: O.C. Bottom left: K.F. Top right: K.C. Bottom right: A.H. Blue symbols correspond to the cued locations and red symbols correspond to the uncued locations. The dotted line corresponds to the Fourier transform of the optimal profile.

Scaled Perceptual Filters to Compare the Shape of the Filters

To compare the shape of the perceptual filters (in isolation from magnitude differences), we scaled the classification image for the uncued location to give the best fit to the classification image for the cued location. The fit for the scaling was performed by minimizing the error weighted inversely by the pooled sample variance across both classification images. In addition, the sample

covariance of the uncued classification image was scaled with the classification image.

Figure 13 shows the scaled perceptual filter at the uncued location to give the best fit to the unscaled perceptual filter at the cued location for all four observers. Error bars for observers A.H. and K.F. are larger due to the fact that the magnitudes of the classification images from the uncued location were lower, and had to be scaled by a larger constant. The larger scaling constant results in increased sample variance for A.H. and K.F. Results from two-sample

Hotelling T^2 were calculated (see “[Methods](#)”) to statistically compare the shape of the radial average of the perceptual filters for the cued and uncued locations. We

found no statistically significant difference between radial averages for the cued and the scaled uncued locations ($p > .01$).⁴

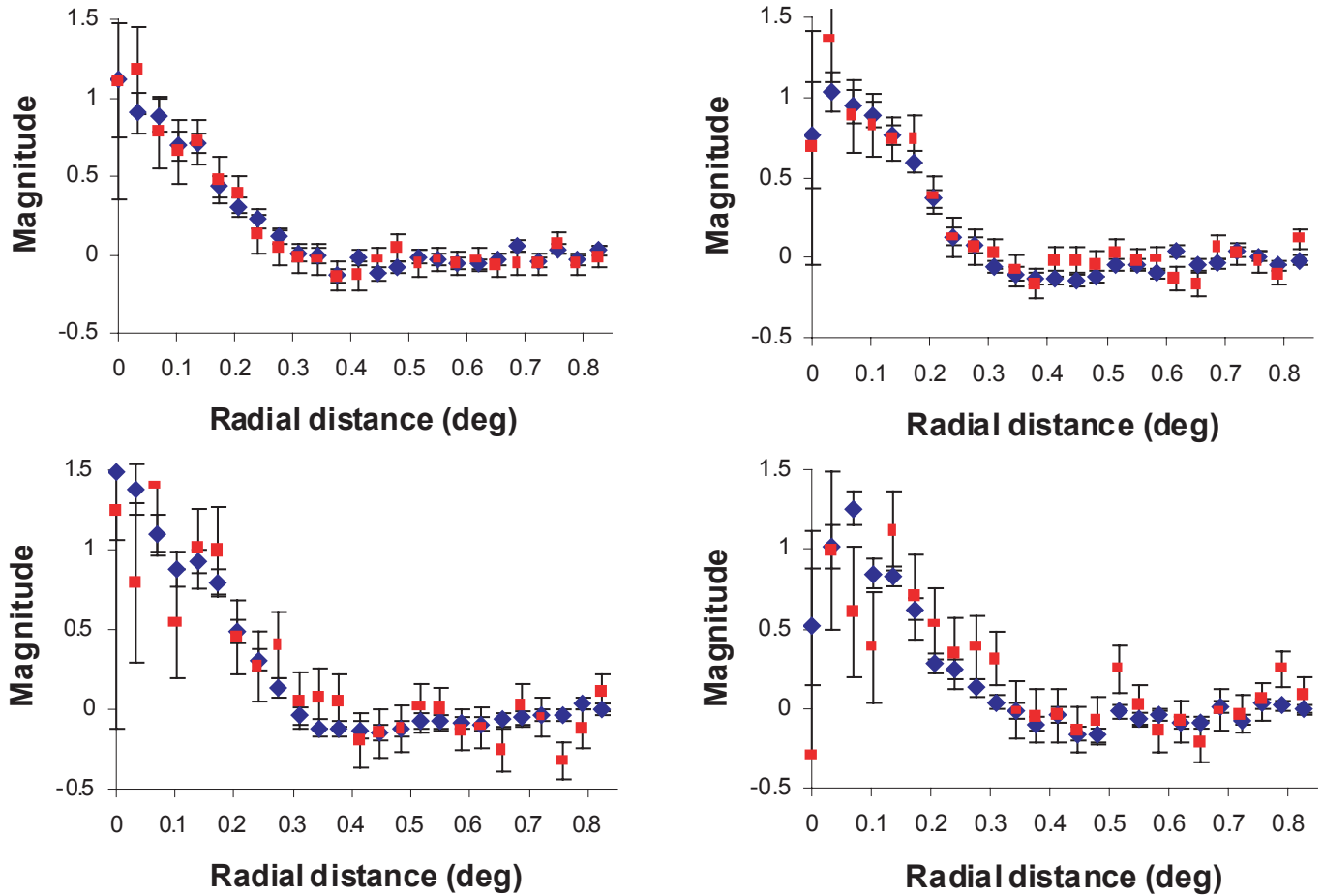


Figure 13. Radial averages (spatial domain) of the uncued location scaled (minimizing the weighted error) to match the radial average of the classification image for the cued location. Top left: O.C. Bottom left: K.F. Top right: K.C. Bottom right: A.H. Blue symbols correspond to the cued locations and red symbols correspond to the uncued locations.

Performance of the Human Classification Images

Although we did not find statistically significant differences across the shapes of the inferred perceptual filters at the cued and uncued locations, we evaluated the cueing effect that would arise from the observed differences in shape of the human perceptual filters. To do so, we used the best-fit DOG for each observer for the cued and uncued locations and performed computer simulations in the framework of our Bayesian model framework ([Figure 1](#)). To isolate cueing effects arising from the difference in shape of the filters from differential weighting of the cued and uncued locations, the simulations included equal weighting of the likelihood at each location. [Table 3](#) shows the obtained hit rates and false alarm rates for the best-fit DOG for

each observer. [Table 4](#) shows simulation results for the best-fit DOG for each observer for the case where internal noise (independent additive Gaussian noise added to the output of each filter) was added to match performance of the human observers. Both simulation results show that the cueing effects that arise from the observed differences in the shape of the perceptual filters are either too small (< 0.02 for K.C. and O.C.) or in the wrong direction (A.H. and K.F.) to explain the observed cueing effects in human observers ([Table 1](#)). [Figure 14](#) plots the cueing effects (hit rate for valid cue trials minus hit rate for invalid cue trials) measured for the human observers and those predicted from the differences between the inferred perceptual filters for the cued and uncued locations.

Table 3. Performance of the best-fit Difference of Gaussians with equal weighting of the likelihood of the cued and uncued locations.

Observer	Equal weighting			
	Hit rate (valid trials)	Hit rate (invalid trials)	False alarm rate (all trials)	Cueing effect ($HR_v - HR_{iv}$)
O.C.	0.939	0.919	0.053	0.02
K.F.	0.923	0.943	0.059	-0.02
K.C.	0.930	0.929	0.071	0.001
A.H.	0.930	0.969	0.050	-0.039

Table 4. Performance of the best-fit Difference of Gaussians with equal weighting of the cued and uncued locations. For these results, internal noise was injected to match the human performance levels.

Observer	Equal weighting			
	Hit Rate (valid trials)	Hit Rate (invalid trials)	False alarm rate (all trials)	Cueing effect ($HR_v - HR_{iv}$)
O.C.	0.819	0.799	0.231	0.02
K.F.	0.821	0.824	0.233	-0.03
K.C.	0.892	0.883	0.264	0.009
A.H.	0.880	0.924	0.229	-0.043

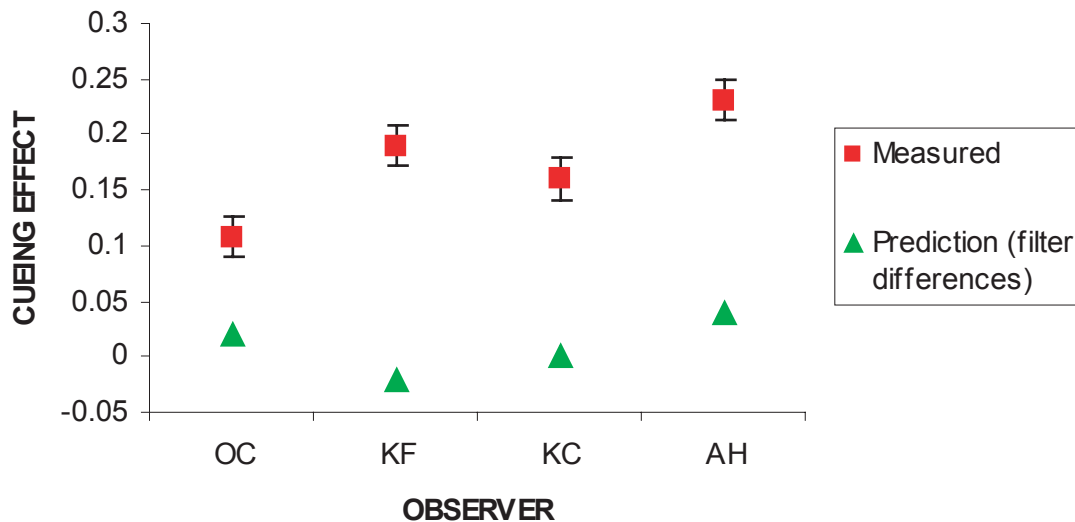


Figure 14. Cueing effect (hit rate for valid trials minus hit rate for invalid trials) measured in human observers (red symbols). Green symbols correspond to the cueing effect predicted by the differences in the inferred perceptual filters from the human classification images.

Inferring the Underlying Weights Used by the Observers From the Ratio of Magnitudes of the Classification Images

The scalar used to best fit the uncued human classification image to the cued human classification image was taken as the ratio of the magnitudes of the classification images for the cued and uncued locations. We then used computer simulations with the Bayesian model varying the input weights of the model to generate a lookup table between weights (w_c and w_u in Figure 1) and the ratio of the magnitude of the classification images obtained for the model (e.g., Figure 5). From this lookup

table we could then infer the weights used by the observers from the ratio of the magnitude of the human classification images. The simulations for the Bayesian model were performed by injecting internal noise and adjusting the criterion in order to match the false alarm rates observed in humans. The procedure was done separately for each human observer. The weights inferred for the cued location were: 0.76 (O.C.), 0.84 (K.F.), 0.8 (K.C.), and 0.88 (A.H.).

Discussion

Human Versus Optimal Perceptual Filters

For the special case in which the external noise is spatially uncorrelated (white) Gaussian noise, the perceptual filters of the ideal Bayesian observer match the signal. Comparison of the human classification images to the optimal perceptual filter (Figure 3 vs. Figure 9) shows that for all observers the human perceptual filters tend to be narrower in the spatial domain than the optimal Gaussian filter, and also have an inhibitory surround. The surround can be seen more clearly in the radial averages in Figure 11 and corresponds to a low spatial frequency suppression. The lower sensitivity to low spatial frequencies can be seen as a dark “hole” in the Fourier transformations of the classification images (Figure 10). The low-frequency suppression can also be seen as the decreased magnitude of the radial average of the Fourier transformations of the classification images (Figure 12). The Fourier transformation of the ideal perceptual filter corresponds to a Gaussian that is more compact than the human perceptual filter in the frequency domain (and more extensive in the spatial domain; see Figures 11 and 12). The inability of human observers to match the optimal profile when the signal is a Gaussian has been observed before by Abbey et al. (1999) for the detection of a Gaussian signal. The low frequency suppression might be explained in part by the decreased contrast sensitivity of the human visual system to low frequencies (i.e., the contrast sensitivity function).

Shape of Human Perceptual Filters at the Attended and Unattended Locations

A common explanation for the cueing effect is that visual attention enhances the quality of processing at the attended location. One possible mechanism suggested by previous studies is that attention changes the tuning of the perceptual filter at the attended location (e.g., Yeshurun & Carrasco, 1999; Doshier & Lu, 2000a, 2000b) so that it matches the signal more optimally. If so, the classification image technique should reveal a difference in the shape of the perceptual filters at the cued and uncued locations (see Figures 6, 7, and 8 for examples of possible classification image signatures for this scenario). Our results did not find statistical significance between the shape of the perceptual filters at the cued (attended) and uncued (unattended) locations for all four observers. Yet statistical significance should not be the only criterion to judge the differences across perceptual filters. It is plausible that if the number of trials were increased by a factor of 10, the differences in

shapes across perceptual filters would become statistically significant. Another important criterion is to determine how much of a cueing effect would be produced by the observed differences in the inferred shape of the perceptual filters. Monte Carlo simulations using the best-fit DOG (Table 2) to the observers' perceptual filters and equal weighting of information of both locations resulted in cueing effects ranging from -4% to +2%. The perceptual filters for observer A.H. resulted in a higher performance at the uncued location (-4% negative cueing effect). This result is consistent with her classification images (see Figures 9, 10, and 11) where the perceptual filter at the unattended location did not have the low-frequency suppression, and, therefore, better matched the optimal filter than the perceptual filter at the attended location. Overall, these findings suggest that even if the differences in shapes across the perceptual filters were assumed to be statistically significant, these differences by themselves would not be able to account for the large cueing effects measured on human observers, which are in the order of 10% to 23%. We therefore conclude that for the present task, visual attention does not change the tuning of the perceptual filter at the cued location sufficiently to account for the human observer cueing effects.

Visual Attention Changes the Weighting of Information at the Cued and Uncued Locations

Another explanation of the cueing effect is in terms of a differential weighting of information at the attended and unattended locations without resorting to a different quality of processing at each location. Kinchla, Chen, and Evert (1995) used a model that linearly weights information across both locations to fit to human data. Shimozaki et al. (2001) and this study used an optimal Bayesian observer with identical perceptual filters at both locations to predict the human cueing effect. This model predicts that the classification images for the cued and uncued location should differ in magnitude but not shape (Figures 3 and 4). We found that the human classification followed this pattern (Figures 10 and 12). These results support the idea that visual attention does change the weighting of information at the cued and uncued location.

In addition, we used simulations to infer the underlying weighting of information at each location (cued and uncued) used by the human observers from the ratio between the magnitudes of the human classification images. We obtained a range of weights (0.88, 0.85, 0.8, and 0.76) that were scattered around the optimal weighting (0.8). Note that the rank order of the weights for the observers is in agreement with the size of their observed cueing effect, as we would expect from the model described in Figure 1. The higher the weight

assigned to the cued location, the larger the cueing effect. In summary, the classification images support the idea that visual attention acts to more heavily weight the information at the cued location.

Attentional Weighting Versus Attentional Switching

An alternative model that is consistent with a difference in magnitudes for the classification images is one in which the observer monitors (attends) one location per trial and switches across trials by attending either the cued location or the uncued location with some probability. We refer to this model as the attentional switching model. A common assumption is that the attentional switching is determined by the prior probabilities of signal presence. Therefore, for our task, the model attends the cued location on 80% of the trials and the uncued location on 20% of the trials. This model will also yield classification images with a higher magnitude at the cued location than the uncued location. However, the model predicts (see [Appendix B](#)) cueing effects (of the order of 0.445), which are significantly larger than those measured for human observers and the attentional weighting model (Table 2). Therefore, the attentional switching model (as many other limited capacity attentional models) can be rejected because it predicts larger cueing effects than those present in human observers. Nevertheless, the fact that the attentional switching model generates classification image signatures that are similar to those of the attentional weighting model emphasizes the importance of considering both—classification images and task performance—when evaluating models.

Visual Attention: Selection and Combination of Information

Overall, our results support the idea that for the simple task studied, the cueing effect is due to the differential weighting of information at the cued and uncued location, and not due to a change in the shape of the perceptual filters at the attended and unattended locations. The concept that visual attention allows the observer to select and/or differentially weight information from different sources has been proposed before for the cueing paradigm ([Kinchla et al., 1995](#)). [Shaw \(1982\)](#), [Palmer \(1995\)](#), and others ([Sperling & Doshier, 1986](#); [Palmer, Verghese, & Pavel, 2000](#); [Verghese & Stone, 1995](#); [Eckstein, 1998](#); [Eckstein, Thomas, Palmer, & Shimozaki, 2000](#); [Verghese, 2001](#)) have also shown that human performance in simple visual search tasks can be accounted for in terms of visual attention as a selection mechanism and without resorting to a change in the quality of processing. These models have been successful in predicting many effects in visual search including set-size effects, distractor variability, search

asymmetries, and the feature/conjunction search dichotomy (see [Palmer et al., 2000](#), for a review)

However, more complex tasks ([Poder, 1999](#)) or those involving memory studies have shown that attending to a location will not simply allow the observer to select relevant information and ignore irrelevant information, but instead will improve the quality of processing at the attended location due to capacity limitations. In addition, the present results cannot explain cueing effects obtained in paradigms in which a 100% valid postcue (which can be localized by the observer) was presented in addition to the pre- or simultaneous cue ([Luck, Hillyard, Mouloua, & Hawkins, 1996](#); [Doshier and Lu, 2000a, 2000b](#); [Lu and Doshier, 2000](#)) and in tasks in which a noninformative precue was presented ([Henderson, 1991](#)).

A Hypothetical Experiment Where Attention Would Change the Shape of the Perceptual Filters Without Reflecting Limited Resources

It should be noted that, in theory, experiments could be designed so that attention has an effect on the shape of the perceptual filter used by the human observer. For example, one such task might be a detection task where the signal is a high-frequency windowed sine wave that might appear at one of two locations. Let us suppose that the precue is a high-contrast copy of the signal, appears directly below the probable signal location, and is in phase with the signal (when the signal is present). It is widely known that human observers have intrinsic uncertainty ([Pelli, 1985](#)) about the spatial phase of periodic signals ([Burgess & Ghandeharian, 1984](#)). In this case, the precue would provide not only information about the probable signal location (right vs. left location), but also information about the exact phase and/or position of the signal. Therefore, for this example, one might obtain a classification image for the uncued location that is not phase-coherent because the observer has intrinsic uncertainty about the phase of the signal, and therefore monitors many locations. On the other hand, for the attended location, the high contrast cue would provide the observer with information about the exact phase or position of the signal. In this case, the observer would monitor a single perceptual filter with the phase or position matching that of the reference. As a result, one would obtain a phase-coherent classification image for the cued/attended location. In fact, one could build a Bayesian model with intrinsic phase uncertainty that would predict the change in perceptual filters.

This example simply illustrates that one might find tasks in which the attended location changes the shape of the perceptual filter at the attended location. However, it should be clear that in this example the cue not only gives information about which of the two locations (right image vs. left image) has a higher probability of

containing the target but also provides information about the specific phase or position of the target within the cued location. Therefore, the cue also allows the observer to select one of many filters differing slightly in locations he/she is uncertain about within the right or left image. Therefore, the observed change in the perceptual filter would not be associated with a capacity limitation in visual attention, but instead the cue provides more/further information for the observer to select what is relevant and ignore what is irrelevant.

Classification Images Versus Other Methods to Infer Properties About Perceptual Filters

Variation of energy thresholds with external noise

A commonly used method to infer the ability of a perceptual filter to match the optimal filter is to vary the external noise and measure the signal energy required by a human observer to detect the signal at a given performance level. From the slope of the variation of energy with external noise (i.e., noise spectral density), one can infer what is known as the sampling efficiency of the perceptual filter (Burgess et al., 1981; Pelli, 1985). The sampling efficiency is a quantitative measure (squared correlation) of the match between the human perceptual filter and the optimal filter. As with the classification image technique, typically there is an underlying assumption that the observer is effectively monitoring a single filter to reach the decision. If the observer is monitoring more than one filter (e.g., the same filter but at different positions; i.e., spatial uncertainty) and combining the responses of the filter nonlinearly or when the filter response goes through a transducer nonlinearity, then a more complex analysis is required to obtain the sampling efficiency (Eckstein, Ahumada, & Watson, 1997; Lu & Doshier, 1999). Although the sampling efficiency is a very useful measure, it has the limitation that it does not provide information about the shape of the perceptual filter. In fact, perceptual filters with a variety of different shapes can have identical sampling efficiencies. In this respect, the sampling efficiency estimation technique could be combined with the classification image technique to provide the investigator with information about the shape of the perceptual filter.

Bandpass noise-masking experiments

Another method that has been used to infer the underlying tuning of the spatial frequency or orientation of the perceptual filters has been the bandpass noise-masking paradigm. In this paradigm, the frequency content of the noise is systematically varied so that the noise contains power in different frequency bands in each particular condition. The investigator then measures the

energy threshold to detect the signal as a function for the different noise frequency bands. From the effect of the different noise frequency bands on human observers' threshold elevation, the investigator infers the sensitivity to a given spatial frequency of the perceptual filter used to perform the task. The basic idea is that noise frequencies that do not affect performance correspond to spatial frequencies to which the human perceptual filter is not sensitive. On the other hand, noise frequency bands that drastically elevate the threshold energy for detection correspond to spatial frequencies to which the human perceptual filter is highly sensitive. Thus, one can derive mathematical methods to derive the frequency tuning of the perceptual filters (e.g., Solomon & Pelli, 1994). The main limitation of the bandpass noise-masking technique is that it assumes that the observer always monitors the same perceptual filter in the different bandpass noise conditions. An optimal Bayesian observer would change the perceptual filter to avoid regions of high noise to optimize performance. The ability of a model or human observer to modify the perceptual filter as a function of the frequency content of the noise is referred to as prewhitening and/or off-frequency looking. It has been shown that in many instances human observers are able to do off-frequency looking and/or prewhitening (Burgess, Li, & Abbey, 1997; Burgess, 1999; Abbey and Eckstein, 2000; Solomon, 2000). In these cases, use of the bandpass noise-masking technique to derive an underlying single fixed perceptual filter can result in misleading results. Because the classification image technique does not change the frequency content of the noise, it does not present the problem of off-frequency looking.

Conclusions

We have applied the classification image technique to determine how attention affects the processing of information at the attended and unattended locations in the Posner cueing paradigm. Our results show that, for the contrast discrimination task studied, changes in the shape of the perceptual filters were neither statistically significant nor were the small changes in the shapes of the perceptual filters able to account for the size of the cueing effect measured for human observers. On the other hand, the human classification image signatures corresponded to the concept that visual attention weights the information at the attended location more heavily. The Bayesian model explored here is analogous to the Bayesian or quasi-Bayesian (i.e., approximations to Bayesian models) models used previously to explain various results in visual search, such as set-size effects and the dichotomy between feature and conjunction searches. Thus in the greater context, our findings suggest that for simple tasks, the Posner cueing paradigm now joins another influential attentional paradigm, visual search, that can be explained in terms of a Bayesian observer. In

this framework, visual attention allows the observer to select or differentially weight information at different locations but does not change the perceptual quality of the processed information at each of the possible locations.

Appendix A

Ideal and Suboptimal Bayesian Observer

The perceptual filters at the cued and uncued locations are given by $F_c(x,y)$ and $F_u(x,y)$ and are normalized to have unit length. The image at the cued and uncued locations is given by $g_{c,i}(x,y)$ and $g_{u,i}(x,y)$. The first subscript refers to the locations ("c" for cued and "u" for uncued), whereas the second subscript refers to the i^{th} trial.

The images for signal-present valid cue trials are given by

$$\begin{aligned} g_{c,i}(x,y) &= s(x,y) + p(x,y) + n_{c,i}(x,y) \\ \text{and} \\ g_{u,i}(x,y) &= n_{u,i}(x,y) + p(x,y) \end{aligned} \quad (\text{A.1})$$

where $s(x,y)$ is the signal luminance profile, $p(x,y)$ is the pedestal that has the same spatial profile as the signal, and $n_{c,i}(x,y)$ and $n_{u,i}(x,y)$ are the external image noise samples at the cued and uncued locations, which are independently sampled.

For signal-present invalid cue trials the images are given by

$$\begin{aligned} g_{c,i}(x,y) &= n_{c,i}(x,y) + p(x,y) \\ \text{and} \end{aligned} \quad (\text{A.2})$$

$$g_{u,i}(x,y) = s(x,y) + n_{u,i}(x,y) + p(x,y).$$

Finally for signal absent trials the images are given by

$$\begin{aligned} g_{c,i}(x,y) &= n_{c,i}(x,y) + p(x,y) \\ \text{and} \\ g_{u,i}(x,y) &= n_{u,i}(x,y) + p(x,y). \end{aligned} \quad (\text{A.3})$$

The response of each of the perceptual filters ($\lambda_{c,i}$ and $\lambda_{u,i}$) to the stimuli in the i^{th} trial is given by

$$\lambda_{c,i} = \iint F_c(x,y)g_{c,i}(x,y)dxdy + \varepsilon_{c,i} \quad (\text{A.4})$$

$$\lambda_{u,i} = \iint F_u(x,y)g_{u,i}(x,y)dxdy + \varepsilon_{u,i} \quad (\text{A.5})$$

where $\varepsilon_{c,i}$ and $\varepsilon_{u,i}$ is a random scalar corresponding to internal noise, which is independently sampled for each trial and location (cued and uncued) from a Gaussian distribution with standard deviation σ_{int} .

The Bayesian model calculates the likelihood of the responses ($\lambda_{c,i}$ and $\lambda_{u,i}$) given that the signal is present at the cued location, $L(\lambda_c, \lambda_u | s_c, n_u)$, and a likelihood of the responses given that the signal is present at the uncued location $L(\lambda_c, \lambda_u | n_c, s_u)$. The model then computes an overall likelihood of the responses given that the signal is present by weighting the individual likelihood from each location by a weight (w_c and w_u):

$$L(\lambda_c, \lambda_u | s) = w_c L(\lambda_c, \lambda_u | s_c, n_u) + w_u L(\lambda_c, \lambda_u | s_u, n_c) \quad (\text{A.6})$$

The optimal weights are those that match the prior probability of the signal appearing at the locations given by the precue validity. Next the model computes a likelihood of the responses given signal absence, $L(\lambda_c, \lambda_u | n_c, n_u)$. Finally, the Bayesian model computes the ratio of the likelihood for signal presence and signal absence:

$$L_{\text{ratio}} = \frac{w_c L(\lambda_c, \lambda_u | s_c, n_u) + w_u L(\lambda_c, \lambda_u | n_c, s_u)}{L(\lambda_c, \lambda_u | n_c, n_u)} \quad (\text{A.7})$$

The model makes a decision by comparing the likelihood ratio (L_{ratio}) to a decision threshold or criterion:

If $L_{\text{ratio}} > \text{threshold}$, then respond "signal present,"; otherwise respond "signal absent."

For the specific case where the filter responses at each location are Gaussian distributed, the individual likelihood of the filter responses given the signal presence and absence is given by

$$L(\lambda_c, \lambda_u | n_c, n_u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda_c^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda_u^2}{2}} \quad (\text{A.8})$$

and,

$$\begin{aligned} L(\lambda_c, \lambda_u | s) &= w_c \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\lambda_c - d'_c)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda_u^2}{2}} \\ &+ w_u \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\lambda_u - d'_u)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{\lambda_c^2}{2}} \end{aligned} \quad (\text{A.9})$$

where d'_u and d'_c are defined as the mean response of the perceptual filter to the signal present location minus the response to the signal absent location divided by the standard deviation of the response (including the effects of external and internal noise):

$$d'_c = \frac{\langle \lambda_{c,s} \rangle - \langle \lambda_{c,n} \rangle}{\sqrt{\sigma_{\lambda_c}^2 + \sigma_{\text{int}}^2}} \quad (\text{A.10})$$

where, $\langle \lambda_{c,s} \rangle$ is the expected value of these responses of the perceptual filter at the cued location when the signal is present; $\langle \lambda_{c,n} \rangle$ is the expected value of the response of the perceptual filter at the cued location when the signal is absent; σ_{λ_c} is the standard deviation of the response due to external noise; and, σ_{int} is the standard deviation of the additive internal noise. Similarly, d'_u is given by

$$d'_u = \frac{\langle \lambda_{u,s} \rangle - \langle \lambda_{u,n} \rangle}{\sqrt{\sigma_{\lambda_u}^2 + \sigma_{\text{int}}^2}} \quad (\text{A.11})$$

When the noise is white, one can calculate d'_c and d'_u directly from the perceptual filter, $F(x,y)$, the signal, $s(x,y)$, and external image noise (pixel standard deviation given by σ_e):

$$d'_c = \frac{\iint F_c(x,y)s(x,y)dxdy}{\left[\sigma_e^2 \iint F_c^2(x,y)dxdy + \sigma_{\text{int}}^2 \right]^{1/2}} \quad (\text{A.12})$$

$$d'_u = \frac{\iint F_u(x,y)s(x,y)dxdy}{\left[\sigma_e^2 \iint F_u^2(x,y)dxdy + \sigma_{\text{int}}^2 \right]^{1/2}} \quad (\text{A.13})$$

This general framework of the Bayesian observer becomes the ideal observer for the case of white noise when the filters at the locations match the optimal filter (the signal for the case of white noise), the weighting of the cued and uncued likelihoods are determined by the precue validity (0.8 for the cued location and 0.2 for the uncued location in the present study), and there is no internal noise.

Monte Carlo Simulations of Models

The model outlined was implemented in Interactive Data Language (IDL). In the computer implementation, continuous integrals in the above equations were replaced by summations. The different models of attentional weightings were implemented by changing the weights in Equation A.7. The different models that assumed that attention changes the shape of the perceptual filters at the cued location were implemented by changing the filters in Equations A.4 and A.5. The decision threshold of the model was also adjusted to match the false alarm rate of the human observers. The internal noise was adjusted to match human performance.

Appendix B

Single Perceptual Filter Model With Attentional Switching Determined by Prior Probabilities

Here we derive the performance predictions for a model that monitors a single perceptual filter that is switched from the cued location to the uncued location from trial to trial (attentional switching). The frequency with which the model monitors the perceptual filter at the cued location is matched to the prior probability of the signal being present (0.8 for the cued location and 0.2

for the uncued location). In this treatment, the attentional switching model is developed in the context of signal detection theory where the responses to each location are stochastic (due to the external and internal noise).

The first stages of the model remain the same as those described for the Bayesian model. The observer is assumed to have two perceptual filters (Equations A.4 and A.5), and their responses are perturbed by internal noise. The difference between the Bayesian model and the attentional switching model is that the latter model monitors only one perceptual filter on each trial to reach a decision. The model computes the likelihood of the response of a single perceptual filter given that the signal is present, the likelihood given that the signal is absent, and computes a likelihood ratio. This decision rule results in identical performance to comparing the response of the single perceptual filter to a decision criterion (the likelihood is a monotonic function of the filter response).

The hit rate for the model in the valid cue trials is calculated by considering the 0.8 proportion of the valid cue trials in which the observer will correctly monitor the cued location and the 0.2 proportion of the valid cue trials in which the observer incorrectly monitors the uncued location (i.e., the signal is at the cued location but he/she is monitoring the response arising from the uncued location).

The hit rate for the valid cue trials is therefore given by the probability that the filter response exceeds the decision criteria (th) in these two circumstances:

$$H_{vc} = 0.8 \cdot G(th - d') + 0.2 \cdot G(th) \quad (\text{B.1})$$

where G is the cumulative Gaussian, d' is the index of detectability, which is given by Equations A.11 and A.12 and th is the decision criteria.

Similarly, the hit rate for the invalid cue trials is given by

$$H_{ic} = 0.2 \cdot G(th - d') + 0.8 \cdot G(th) \quad (\text{B.2})$$

The false alarm rate for both types of trials is simply given by the probability of the response exceeding the decision criteria in signal absent trials:

$$F = 0.2 \cdot G(th) + 0.8 \cdot G(th) = G(th) \quad (\text{B.3})$$

To obtain the predictions of the attentional switching model comparable to the levels of human performance obtained in this study, the decision criteria and the internal noise were adjusted to match the false alarm rate and hit rate in the valid cue condition of the human observers. Performance was calculated from Equations B.1, B.2, and B.3 using the cumulative normal functions of IDL.

Acknowledgments

This work was supported by a National Aeronautics and Space Administration grant (NASA NAG-1157) and a National Institutes of Health grant (NIH-HL 53455). The authors would like to thank Albert Ahumada Jr. for insight in the topic of classification images and Charlie Chubb for a careful review and insightful comments. The authors also thank Kristine Fazio, Audrey Hill-Lindsey, Oriana Chavez, and Kathy Chong for participating as observers in the study. Some of the results in this paper were previously presented at the Annual meeting of the Vision Science Society, Sarasota, FL, 2001. Commercial Relationships: None.

Footnotes

¹In this paper we use the term filter to refer to a template that is applied to individual locations of the image and not to a kernel that is convolved with the image.

²In the perceptual template model (PTM) model, attention changes what is referred to as the external noise exclusion, which is identical to what traditionally is known as the sampling efficiency in the linear template model (Burgess, Wagner, Jennings, & Barlow, 1981).

³Our simulations show that the relationship depends on the decision threshold (or criterion) used by the model. It is therefore important when inferring the human weights to adjust the model threshold to match the measured false alarm rates in the individual human observers.

⁴A potential problem is that the two sample Hotelling T^2 assumes equal covariance. This is clearly not true, at least for observers A.H. and K.F., where the covariance for the uncued location was scaled by a constant, resulting in higher variance than for the cued location. Ito and Schull (1964) have shown that the Hotelling T^2 statistic is robust to violations of the equal covariance when N is large. We believe our case, N of approximately 1,625, to be sufficiently large.

References

- Abbey, C. K., Eckstein, M. P., & Bochud, F. O. (1999). Estimation of human-observer templates for 2 alternative forced choice tasks. *Proceedings of SPIE*, 3663, 284-295.
- Abbey, C. K., & Eckstein, M. P. (2000). Estimates of human-observer templates for simple detection tasks in correlated noise. *Proceedings of SPIE* 3981, 70-77.
- Abbey, C. K., & Eckstein, M. P. (2002). Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments. *Journal of Vision*, 2(1), in press.
- Ahumada, A., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 49, 1751-1756.
- Ahumada, A., Jr. (1987). Putting the visual system noise back in the picture. *Journal of the Optical Society of America A*, 4, 2372-2378. [PubMed]
- Ahumada, A., Jr. (1996). Perceptual classification images from Vernier acuity masked by noise [Abstract]. *Perception*, 26, European Conference on Visual Perception, Supplement 18.
- Beard, B. L., & Ahumada, A., Jr. (1998). Technique to extract relevant image features for visual tasks. *Proceedings of SPIE, Human Vision & Electronic Imaging*, 3299, 79-85.
- Burgess, A. E., Wagner, R.F., Jennings, J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science*, 214, 93-94. [PubMed]
- Burgess, A. E., & Ghandeharian, H. (1984). Visual signal detection. I. Ability to use phase information. *Journal of the Optical Society of America A*, 1, 900-905. [PubMed]
- Burgess, A. E. (1999). Visual signal detection with two-component noise: Low-pass spectrum effect. *Journal of the Optical Society of America A*, 16, 694-704. [PubMed]
- Burgess, A. E., Li, X., & Abbey, C. K. (1997). Visual signal detectability with two noise components: Anomalous masking effects. *Journal of the Optical Society of America A*, 14, 2420-2442. [PubMed]
- Doshier, B. A., & Lu, Z. L. (2000a). Noise exclusion in spatial attention. *Psychological Science*, 11, 139-146. [PubMed]
- Doshier, B. A., & Lu, Z. L. (2000b). Mechanisms of perceptual attention in precueing of location. *Vision Research*, 40, 1269-1292. [PubMed]
- Eckstein, M. P., Ahumada, A. J., & Watson, A. B. (1997). Visual signal detection in structured backgrounds. II. Effect of contrast gain control, background variations and white noise. *Journal of the Optical Society of America A*, 14, 2406-2419. [PubMed]
- Eckstein, M. P. (1998). The lower efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science*, 9, 111-118.

- Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, 62, 425-451. [PubMed]
- Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, 10, 663-666. [PubMed]
- Harris, R. J. (1985). A primer of multivariate statistics (pp. 99-118). Orlando, Florida: Academic Press.
- Ito, K., & Schull, W. J. (1964). On the robustness of the T^2 test in multivariate analysis of variance when variance-covariance matrices are not equal. *Biometrika*, 51, 71-82.
- Haenny, P. E., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research*, 69, 225-244. [PubMed]
- Henderson, J. M. (1991). Stimulus discrimination following covert attentional orienting to an exogenous cue. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 91-106. [PubMed]
- Kinchla, R. A., Chen, Z., & Evert, D. (1995). Precue effects in visual search: Data or resource limited? *Perception and Psychophysics*, 57, 441-450. [PubMed]
- Lu, Z. L., & Doshier, B. A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1534-1548. [PubMed]
- Lu, Z. L., & Doshier, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A*, 16, 764-778. [PubMed]
- Luck, S. J., Hillyard, S. A., Mouloua, M., & Hawkins, H. L. (1996). Mechanisms of visual-spatial attention: Resource allocation or uncertainty reduction? *Journal of Experimental Psychology: Human Perception and Performance*, 22, 725-737. [PubMed]
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401, 695-698. [PubMed]
- Nolte, L. W., & Jaarsma, D. (1967). More on the detection of one of M orthogonal signals. *Journal of the Acoustical Society of America*, 41, 497-505.
- Palmer, J. (1995). Attention in visual search: Distinguishing four causes of set-size effects. *Current Directions in Psychological Science*, 4, 118-123.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40, 1227-1268. [PubMed]
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, 2, 1508-1530. [PubMed]
- Poder, E. (1999). Search for feature and for relative position: Measurement of capacity limitations. *Vision Research*, 39, 1321-1327. [PubMed]
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3-25. [PubMed]
- Posner, M., & Peterson, S. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13, 25-42. [PubMed]
- Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, 387, 281-284. [PubMed]
- Shaw, M. L. (1982). Attending to multiple sources of information. I. The integration of information in decision-making. *Cognitive Psychology*, 14, 353-409.
- Shimozaki, S. S., Eckstein, M. P., & Abbey, C. K., (2001). Cueing effects without capacity limitations [Abstract]. *Investigative Ophthalmology and Visual Science*, 42, 5867.
- Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, 369, 395-397. [PubMed]
- Solomon, J. A. (2000). Channel selection with non-white-noise masks. *Journal of the Optical Society of America A*, 17, 986-993. [PubMed]
- Sperling, G., & Doshier, B. A. (1986). Strategy optimization in human information processing. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1. Sensory processes and perception* (pp. 2-1 – 2-65). New York: John Wiley and Sons.
- Spitzer, H., Desimone, R., & Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, 240, 338-340. [PubMed]
- Verghese, P., & Stone, L. S. (1995). Combining speed information across space. *Vision Research*, 35, 2811-2823. [PubMed]
- Verghese P. (2001). Visual search and attention: A signal detection theory approach. *Neuron*, 31, 523-535. [PubMed]
- Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. *Vision Research*, 39, 293-306. [PubMed]
- Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, 396, 72-76. [PubMed]