If I saw it, it probably wasn't far from where I was looking

Eli Brenner
Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Pascal Mamassian
CNRS, LPP Université Paris Descartes, Paris, France

Jeroen B. J. Smeets
Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

People are most likely to see something if their gaze is directed at it. Thus if they saw something they may be biased towards believing that they had been looking at it. In order to examine whether this is so we asked participants where a target that jumped to a new position every 250 ms had been at a moment indicated by a flash or a tone. The jumping introduced uncertainty about where the target was at the indicated moment, giving room for biases to be expressed. Participants showed a clear preference to select positions that were nearer to where they were looking.

Keywords: flash lag, temporal resolution, spatial vision, localisation

Citation: Brenner, E., Mamassian, P., & Smeets, J. B. J. (2008). If I saw it, it probably wasn't far from where I was looking. *Journal of Vision, 8*(2):7, 1–10, http://journalofvision.org/8/2/7/, doi:10.1167/8.2.7.

Introduction

How people interpret the light reaching their eyes is influenced by the assumptions that they make about the environment. For instance, when the motion parallax that people create by moving around is used to judge objects' shapes, it is assumed that the objects themselves are static and rigid. Similarly, when people rely on shading to judge an object's shape they are making assumptions about the reflectance of the object's surface and about the illumination. Without such assumptions the possibilities to interpret visual stimulation would be very limited, so it makes sense to accept assumptions that are seldom violated in daily life. And indeed, people readily assume that surfaces' textures are homogeneous or isotropic in order to judge their slants (Knill, 2003) and that the illumination is from above in order to distinguish bumps on the surface from dents in the surface (Mamassian & Goutcher, 2001), in accordance with the texture of many natural surfaces being isotropic and the illumination usually being from above.

When an assumption is less unlikely to be violated it is less certain that the assumption should be accepted, so the likelihood that it is correct to do so must somehow be considered. This likelihood is expressed in the extent to which people rely on information based on an assumption to make a certain judgement, in comparison with alternative sources of information that do not rely on the same assumption, and in comparison with prior knowledge of the likelihood of certain judgements being true. These comparisons can readily be modelled within a Bayesian framework (for a review see Kersten, Mamassian, & Yuille, 2004). Doing so allows one to attribute biases in people's judgments to priors that are directly related to the assumptions that people make. This method can reveal less self-evident assumptions, such as that motion is likely to be slow (Weiss, Simoncelli, & Adelson, 2002).

The extent to which one relies on various assumptions presumably arises through interactions with the environment, during development and through evolution, just as various other aspects of the visual system develop or evolve to suit the environment (e.g., for colour see Brenner & Cornelissen, 2005; Párraga, Troscianko, & Tolhurst, 2002; Purves, Lotto, Williams, Nundy, & Yang, 2001). Thus one may expect to find priors that are related to one's interactions with the environment as well as ones related to the statistics of the environment itself. An obvious example is the viewpoint from which an object is seen. A circular outline in the distance is more likely to be judged to originate from a sphere than from a rod oriented exactly along one's line of sight, despite the fact that both would give this outline. The reason is clear: if it is a rod it is quite unlikely that it should be seen from exactly this angle. Thus even if it is just as likely to find a rod in that place as it is to find a sphere, the likelihood that the image is caused by a rod is smaller. The assumption is that one is not looking at the object from a special position.

In a recent study, moving targets were judged to be too near to where the observer was looking (Brenner, van Beers, Rotman, & Smeets, 2006). The bias was only along the target's path. This raised the suggestion that people are

doi: 10.1167/8.2.7 Received July 13, 2007; published February 21, 2008 ISSN 1534-7362 © ARVO

biased towards believing that they are looking at what they see. The reasoning that was presented was similar to that given above: people are more likely to see something if their eyes are directed towards it, so if they saw something it is likely to have been close to where they were looking. A bias towards localising targets near where one is looking is equivalent to a bias towards small eccentricities. In this case the origin of the bias is not to be found in the statistics of the scenes that we encounter in daily life. Its origin lies in the way in which visual information is processed within the eye and brain, with most neuronal resources being devoted to a small area on the retina, and eye movements directing this part of the retina (the fovea) towards selected parts of the scene. In the present study we demonstrate that people have such a bias.

Since we do not expect a strong bias towards small retinal eccentricities, we can only expect such a bias to become apparent when there is considerable uncertainty about targets' positions. We did not want to achieve such uncertainty by using targets that are difficult to detect, because it is obvious that detection thresholds increase with eccentricity. We could have tried to correct for differences in detection across the retina, but that would require knowledge of the particular aspect that is limiting for our task (Raninen, Franssila, & Rovamo, 1991), and we could never be certain that the bias that we are looking for did not influence the tasks on which the correction is based. We therefore used clearly visible targets, and introduced temporal rather than spatial uncertainty about the target's position. Our experiment is somewhat similar to Murakami's (2001) experiment in which he asked subjects to indicate whether a flash was to the left or to the right of a jumping target. He found that people systematically related the position of the flash to a slightly later position of the target. We asked subjects to indicate the position at which a jumping target had been at the moment of a flash or tone, so we expect to find a similar systematic temporal error. Our main interest, however, was whether we would also find a spatial bias.

Methods

The experiments were conducted in a dimly illuminated room. Stimuli were presented on a 47.3 × 30.0 cm CRT screen (resolution: 1280 × 800 pixels; 120 Hz) that was 150 cm from the subject. The background on the screen was grey (about 15 cd/m^2). Subjects fixated a 0.4 cm (about 0.15 deg) diameter black disk at the centre of the screen. The target was a 1 cm diameter (0.4 deg; about 30 cd/m2) green disk that jumped to a new position every 250 ms (see Figure 1). The positions at which the target could appear formed five concentric rings, 2, 4, 6, 8 and 10 cm from the fixation point; i.e., the target was always at an eccentricity of about 0.8, 1.5, 2.3, 3.1 or 3.8 deg. Consecutive positions could be at any eccentricity other than the same one, and the angle between two consecutive points in relation to the fixation point was never within 60 degrees. The moment of interest was

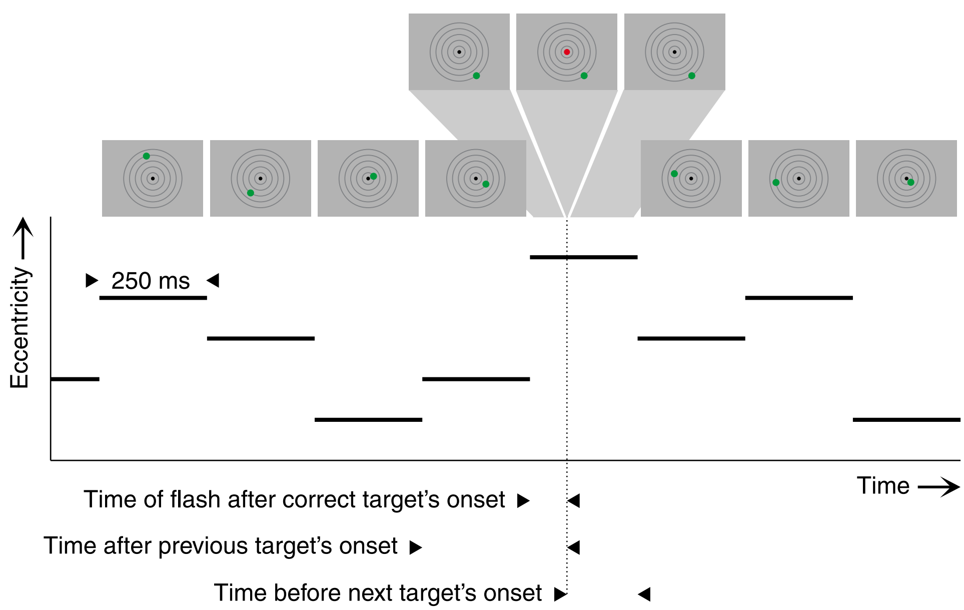

Figure 1. Experimental design. A green target jumped to a new position on one of five (invisible) concentric circles around the fixation point every 250 ms. At some moment a red disk flashed over the black fixation point for one frame. The subject had to indicate where the target was at the moment of the flash. We examined how the timing of the flash influenced the subject's response. Time was defined relative to the onset of the *chosen* target. In the second experiment the flash was replaced by a 20 ms, 8000 Hz tone (time was measured relative to tone onset).

either indicated by a flash at the fixation point or by a tone. The flash that indicated the moment of interest was a 1 cm diameter (0.4 deg; about 20 cd/m^2) red disk that covered the fixation point during one frame. This could only be seen if one was looking at the fixation point as instructed. The tone that indicated the moment of interest was a computer-generated 8000 Hz tone of which the amplitude decreased linearly to zero within 20 ms.

Subjects

Eight subjects each took part in two sessions. Two of the subjects were authors. The other six were unaware of the hypotheses under study. They were all colleagues with experience in psychophysics. The only difference between the two sessions was that in one session the target was a dim flash, whereas in the other it was a tone. The research was carried out at the ErasmusMC in Rotterdam in accordance with the local guidelines.

Procedure

The flash or tone occurred between 1250 and 2500 ms after the target started jumping around. Within this interval the exact frame on which the flash occurred or the tone presentation started was determined at random, so it could occur at any time relative to the moment that the target position changed. As soon as subjects detected a flash or tone they were to move the computer mouse. Moving the mouse more than 1 cm made the target disappear and a cursor appear from beneath the fixation point in accordance with how the mouse was moved. The cursor was identical to the jumping target (i.e., it was a 1 cm diameter green disk). The subjects' task was to place this cursor at the position at which the jumping target had been at the moment of the flash or tone. As soon as they pressed the mouse button to indicate that the cursor was at the appropriate position the cursor disappeared and the jumping target appeared again. If subjects missed the flash (or tone) they did not move the mouse so the target kept jumping. In that case a new flash or tone appeared 2500 ms after the previous one. Each session continued until subjects had made 250 responses.

Data analysis

For each response we stored information about the exact moment of the flash or tone with respect to target onset, the target's position, and the target's previous and next positions (the target was always at least one position further by the time the subject responded). We consider two kinds of errors: selecting the wrong position (a temporal error) and misjudging that position (a spatial error). In order to decide which position had been selected we first determined the position that was closest to the indicated position. This could be the actual position at that moment or the previous or next one. We removed all three positions associated with a response if there was an angle of more than 60 degrees between the indicated and the nearest position (in relation to the fixation point), because in such cases we doubt whether we have correctly identified the selected target.

There are two possible manifestations of a bias towards small eccentricities, associated with the two kinds of errors mentioned above. The first is a tendency to select the closer of two targets when in doubt about which was present at the moment of the flash or tone. The second is a tendency to underestimate the eccentricity of the selected target when in doubt about its position. Since the uncertainty in this study was primarily temporal, we mainly expected to find the former kind of errors. However we also examined whether the latter were evident by plotting histograms of the eccentricity of the indicated positions for each selected target eccentricity.

A first indication of whether there is a bias towards selecting targets at small eccentricities was obtained by simply counting the number of times that positions at each eccentricity were chosen. We express this as a percentage of the number of times that positions at the eccentricity in question were present at the time of the flash or tone.

If there is a bias towards small eccentricities, we expect its influence to depend on the timing of the flash or tone: its influence will be strongest when people are least certain about whether the target had been at either of two positions. We therefore determined both the number of times that the target was presented at each eccentricity for every moment at which the flash or tone occurred, and the number of times that the target at that eccentricity was selected for each moment of the flash or tone. Beside times that correspond with selecting the correct position, we also considered ones that correspond with selecting the previous and following positions, in which case 250 ms was added or subtracted from the presentation time. Since we defined time from the appearance of the correct target this gave a range of times from −250 to 500 ms, with the interval between 0 and 250 ms representing correct responses. We determined the percentage of times that a target position was chosen as a function of the time at which the flash or tone was presented. This was done separately for each of the five eccentricities (for both the flash and the tone). We did so for each subject separately, as well as for the pooled data of all eight subjects. These sets of percentages were used to model the data.

Model

We propose a simple spatio-temporal model with which to quantify the potential bias towards small eccentricities. This spatial bias, the uncertainty about which target was present at the time of the flash or tone, and the expected systematic timing error are all modelled within a Bayesian framework. Let us assume that the probability of selecting the target that is present at a certain moment relative to the moment of the flash or tone onset can be described by a Gaussian distribution $G_{[\delta,\ \sigma]}(t)$ with mean δ and standard deviation σ. A positive value for δ indicates that subjects systematically select a target position at a later time than the actual moment of the flash or tone (Murakami, 2001). The standard deviation σ represents the participants' temporal precision. The target stays at each location for T = 250 ms. The likelihood of selecting the *correct* target $\lambda_c(t)$ when the flash or tone was presented a time t after the beginning of the interval is simply the area under the Gaussian distribution over the duration of the interval containing the target:

$$\lambda_c(t) = \int_0^T G_{[\delta+t,\sigma]}(s)\mathrm{d}s \quad (1)$$

Similarly, we can determine the likelihoods $\lambda_a(t)$ and $\lambda_b(t)$ of selecting the targets that belong to the intervals that occur respectively just *after* and *before* the correct interval:

$$\begin{cases} \lambda_a(t) = \int_T^{2T} G_{[\delta+t,\sigma]}(s)\mathrm{d}s \\ \lambda_b(t) = \int_{-T}^{0} G_{[\delta+t,\sigma]}(s)\mathrm{d}s \end{cases} \quad (2)$$

In the spatial domain, all five eccentricities are equally likely to be presented, so the spatial likelihood is uniform and equal to one fifth. Given that the time and space dimensions are independent in our experiment, the overall spatio-temporal likelihood is simply the product of the temporal and spatial likelihoods.

The flash or tone was equally likely to be presented while the target was at each of the five positions between 1250 and 2500 ms after the target appeared. We will therefore assume that the temporal prior probability is uniform[1]. In contrast, we expect that small eccentricities will be favoured over large ones, so we expect the spatial prior probability $\pi_s(e)$ of selecting eccentricity e to decrease with eccentricity. We do not propose a specific function to describe $\pi_s(e)$. Only four degrees of freedom are necessary to characterize the five measured values of this function, because the prior probabilities must sum to one. Again, given that the time and space dimensions are independent in our experiment, the overall spatio-temporal prior is simply the product of the temporal and spatial priors.

We can now compute the posterior probability of selecting the eccentricity of the correct interval given the eccentricity e of the target and the time t since the onset of the interval containing the target $p(c|e,\ t)$. Using Bayes' rule, this posterior probability is proportional to the product of the likelihood and prior. Omitting the terms that are constant:

$$p(c|e,t) \propto \lambda_c(t)\pi_s(e) \quad (3)$$

The normalizing term that is necessary to make the posterior a probability distribution consists of all possible target choices weighted by their respective priors. Since the experiment was designed to never have two consecutive targets at the same eccentricity, the *average* probability when the participant does not choose the correct interval depends on the eccentricity of the correct target, because the other two relevant eccentricities can each be any of the other four eccentricities. The posterior can therefore be written as:

$$p(c|e,t) = \frac{\lambda_c(t)\pi_s(e)}{\lambda_c(t)\pi_s(e) + (1-\lambda_c(t))(1-\pi_s(e))/4} \quad (4)$$

Similarly, we can write for the posteriors $p(a|e,\ t)$ and $p(b|e,\ t)$ for selecting the target that belongs to the interval that occurs respectively just *after* or *before* the correct interval[2]:

$$\begin{cases} p(a|e,t) = \dfrac{\lambda_a(t)\pi_s(e)}{\lambda_a(t)\pi_s(e) + (1-\lambda_a(t))(1-\pi_s(e))/4} \\ p(b|e,t) = \dfrac{\lambda_b(t)\pi_s(e)}{\lambda_b(t)\pi_s(e) + (1-\lambda_b(t))(1-\pi_s(e))/4} \end{cases} \quad (5)$$

The model therefore has eight parameters: separate parameters for the systematic temporal error δ and precision σ in the flash and tone conditions, and four parameters to specify the five values of the spatial prior $\pi_s(e)$. We expect δ and σ to differ for the flash and tone conditions, but the spatial bias to be the same in both conditions. We fit this model to the 880 data points (88 moments of the flash and 88 moments of the tone for targets at each of the five eccentricities). We used the corrected Akaike information criterion (AICc) to further evaluate whether our choice of this particular model is justified. To do so we compared the fit with that to several simple modifications of the model.

Control experiment

More eccentric targets could be chosen less frequently because they are more difficult to detect. We used clearly visible targets to prevent them from going by unnoticed, but to make completely sure that this was not the issue we also examined whether varying the visibility (well above

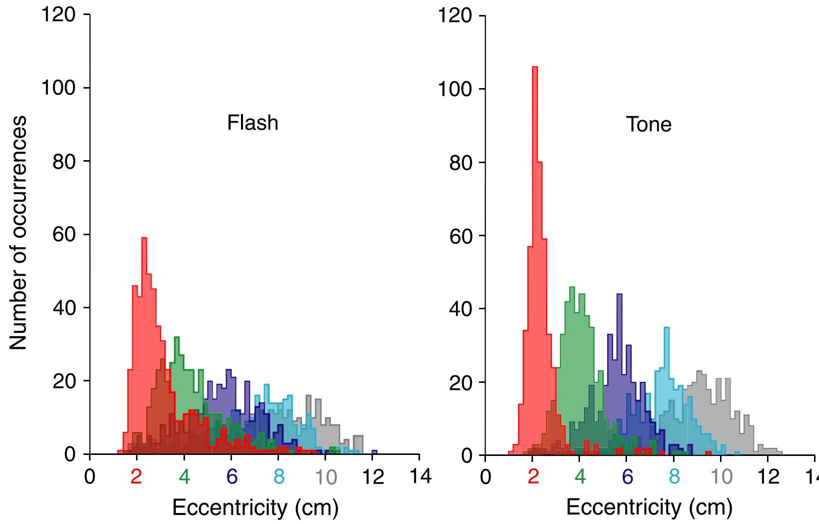

Figure 2. Histograms of indicated eccentricity (2 mm bins). Pooled data of eight subjects. Different colours represent the five actual eccentricities of the target that was considered to have been chosen.

detection threshold) makes a difference. Eight subjects performed the same task, but the target was now always at an eccentricity of 6 cm (2.3 deg) and it was either twice as large (2 cm, 0.8 deg diameter) or half the size (0.5 cm, 0.2 deg diameter) of the original target. Its size was chosen at random for every position. It was equally likely to be large or small. Otherwise the stimuli (including the cursor size), procedure and task were identical to those for the main session in which a flash was used to signal the moment of interest. The analysis was also the same except that the model was modified by replacing the prior for selecting targets at a certain eccentricity $\pi_s(e)$ by a prior for selecting targets of a certain size π_{large} and π_{small}. We now fit 3 parameters (σ, δ, π_{large}; $\pi_{\text{small}} = 1 - \pi_{\text{large}}$) to twice 88 data points.

Results

We first used the positions that the subjects indicated to determine which target the subject judged to have been present at the moment of the flash or tone. Since the targets were quite conspicuous we did not expect subjects to indicate positions that were very far from one of the true positions. Figure 2 shows histograms of the eccentricities at which subjects reported having perceived the targets. The colours indicate the true eccentricity of the target that we considered the subject to be indicating. Subjects had a slight tendency to overestimate the eccentricity of the nearest targets (2 cm) and to underestimate the eccentricity of the farthest ones (10 cm). They probably quickly became aware of the range of possible eccentricities, and were biased towards responding within this range (Körding & Wolpert, 2004). If there is any tendency to underestimate the eccentricity it is clearly negligible in comparison with this bias. Subjects were slightly more accurate spatially when the moment of interest was indicated by a tone.

Obviously, due to the way we designed the experiment, the subjects' uncertainty about where to indicate was mainly in choosing between the possible target positions, rather than in knowing where the target had been. A first indication of a preference for selecting targets at certain eccentricities when determining which target had been present at the moment of the flash or tone is obtained by dividing the number of times that targets at each eccentricity were chosen by the number of times that doing so was the correct choice (Table 1). These ratios suggest that people are biased towards small eccentricities: people selected targets at small eccentricities more frequently than they should have (values above 100%).

Figure 3 shows four examples of the way in which the probability of choosing a target position depends on the moment of the flash or tone and on the eccentricity of the target. Note that the timing was more precise (sharper transitions) for the tone than for the flash, and that a position had the highest chance of being chosen if the flash or tone was presented soon after the target

Eccentricity		2	4	6	8	10
Preference (%)	Flash	140	112	93	85	74
	Tone	117	108	90	96	89

Table 1. Preference for each eccentricity. Trials were categorised by the eccentricity of the target at the moment of the flash or tone onset. The preference is obtained by dividing the number of times that a target at a given eccentricity was chosen by the number of trials in which the correct target was at that eccentricity.

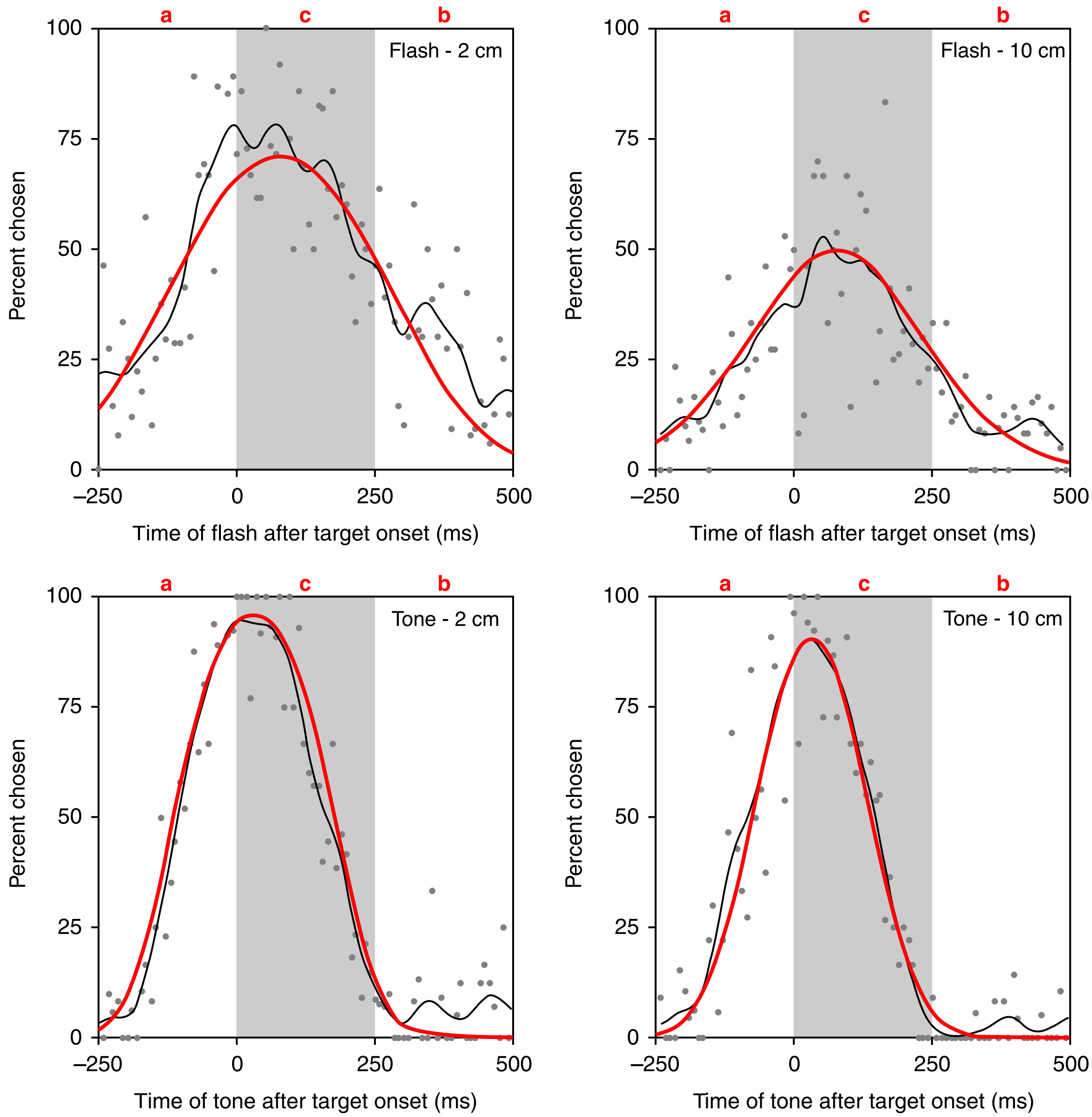

Figure 3. The eight subjects' pooled data for the positions nearest to (2 cm) and furthest from (10 cm) fixation. Dots show the proportion of presentations on which that target position was chosen. Those within the shaded area are correct responses. The thin black curve is a smoothed version of the data (convolution with a normal distribution with a standard deviation of 20 ms). The thick red curves show the best fit of our 8-parameter model to the data for all five eccentricities, considering instances on which the correct target was chosen (c), as well as ones in which subjects picked the target after (a) or before (b) the correct one.

appeared at that position (peaks are at the left side of the grey areas) and if the position itself was near fixation (peaks are higher for the smaller eccentricity of 2 cm). The thick red curves show the best fit of our model that combines temporal uncertainty with a spatial bias.

Figure 4 shows the model parameters of the best fit (the parameters that give the smallest sum of squares of the residuals). Based on this fit, the likelihoods of selecting the five positions is 0.298, 0.218, 0.174, 0.164 and 0.146 for eccentricities of 2, 4, 6, 8 and 10 cm, respectively. In other words, in the absence of any evidence subjects are twice as likely to choose a target at an eccentricity of 2 cm than one at 10 cm. There was also a systematic temporal error of 48 ms for the flash and 94 ms for the tone. The temporal uncertainty (standard deviation) is 151 ms for the flash and 67 ms for the tone (note that this is the combined uncertainty of the moment that the target position changes, the moment that the flash or tone occurs, and comparing these moments). In addition to fitting the model to data pooled across all subjects we also fit the model to each subject's data separately. The average values of the fits for individual subjects are very similar to the values of the fits to the pooled data, and are shown by the symbols in Figure 4.

Trials were removed from further analysis if we were uncertain about which position had been chosen: if the

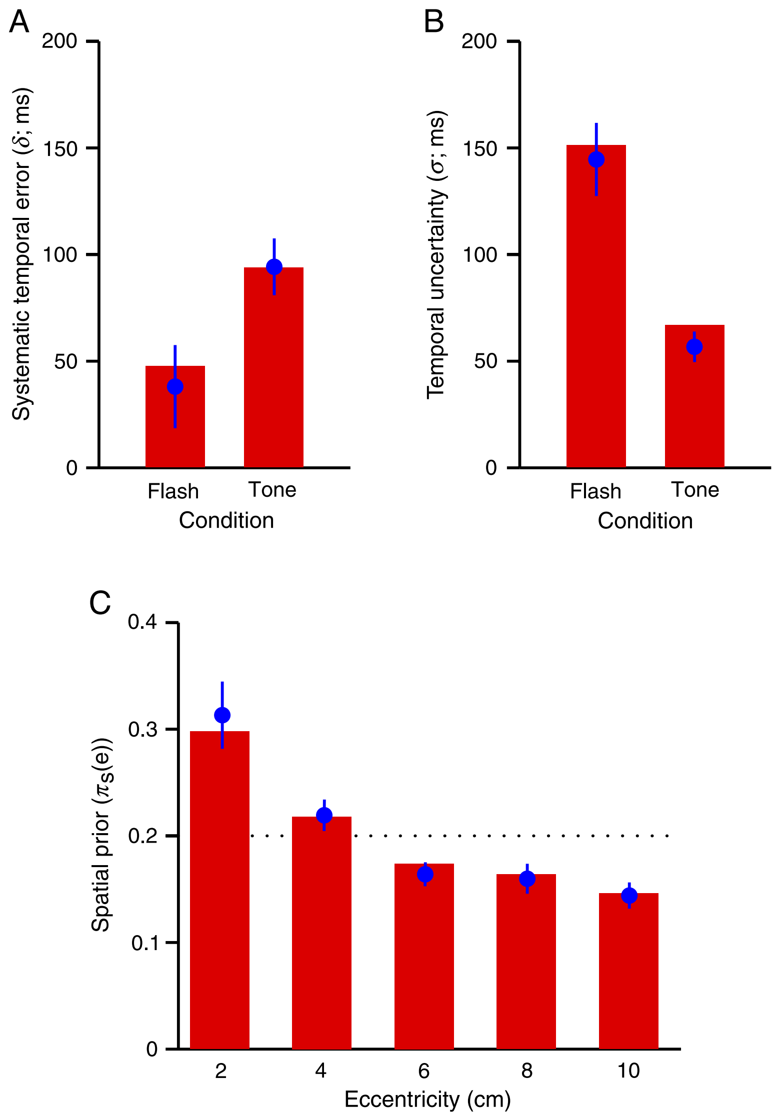

Figure 4. Values of the model's eight parameters that give the best fit. The red bars are values for the fit to the pooled data as shown in Figure 3. The blue symbols show average values for similar fits to individual subjects' data (with standard errors across subjects). A. The systematic temporal error: how much before target onset that the flash or tone has to appear for it to be judged to appear at the same time as the target. B. The temporal uncertainty: the standard deviation in judgments of the timing of target onset relative to the flash or tone. C. The spatial bias: the prior probability of selecting each eccentricity. These five values are defined by four parameters because the sum of the values is one. The dotted line indicates what the values of $s(e)$ would be if there were no bias. There is a clear bias to select targets that are nearer to fixation.

angle between the nearest target position to the indicated position, the fixation point, and the indicated position itself was more than 60 degrees (Table 2). One reason for this happening is subjects choosing a target position that was presented earlier or later than the temporal range that we considered: more than 250 ms earlier or later than the correct one. Figure 3 suggests that this could be an issue, especially for the flash, so we estimated how often it is expected to have happened from our model: this corresponds with the part of the surface under the thick red curve that falls outside the Figure. We did so separately for each subject on the basis of the parameters of the best

Eccentricity		2	4	6	8	10
Removed (%)	Flash	6.9	6.7	6.5	8.4	6.5
	Tone	2.0	2.0	2.0	2.4	3.3

Table 2. Removed trials. Trials were categorised by the eccentricity of the target at the moment of the flash or tone onset. They were removed if neither the correct target nor the preceding or next target was within 60 deg of the indicated position (in relation to the fixation point).

Eccentricity		2	4	6	8	10
Expected to be outside temporal range (%)	Flash	4.4	4.1	3.8	3.7	3.7
	Tone	0.3	0.3	0.3	0.2	0.3

Table 3. The percentage of trials for which our model predicts that a target will be chosen that is beyond the considered temporal range. The percentages were calculated on the basis of individual subjects' fits. Values are averages across subjects.

fit for that subject, and then averaged these percentages across subjects (Table 3). This is an overestimate of the expected number of trials removed due to temporal errors of more than 250 ms, because on some of these trials one of the three targets within the considered time interval will fulfil the criteria for being considered to have been chosen. These latter trials may be responsible for some of the very incorrect indicated eccentricities in Figure 2. Comparing Tables 2 and 3 we see that large temporal errors cannot account for all the trials that were removed from further analysis.

The model fits in Figure 3, and the consistency between individual subjects' parameters and the parameters for the pooled data in Figure 4, suggest that our 8-parameter model does quite a good job of fitting the data.

The sum of squares of the residuals is only 4.5% larger for the eight-parameter model than if we fit three parameters (δ, σ and πs(e)) to each condition separately (30 parameters in total). On the basis of AICc this means that our 8-parameter model is 42 times more likely to be true than such a 30-parameter model. Reducing the number of parameters from eight to seven by using the same temporal uncertainty (standard deviation of the cumulative Gaussian) for the flash and the tone increased the residual variability by 53%. Using the same temporal bias for the flash and the tone increased the residual variability by 12%. Reducing the number of parameters from eight to four by removing the spatial biases increased the residual variability by 20%. In all three cases AICc indicates that our 8-parameter model is more than a million times more likely to be true than the model with fewer parameters. Thus we consider our 8-parameter model to give a good description of the data.

We find a systematic tendency to select the nearer target position (higher values for smaller eccentricities in Table 1; higher peaks for the smaller target eccentricity in Figure 3; larger values for the prior at smaller eccentricities in Figure 4C) rather than to underestimate the selected target's eccentricity (Figure 2). This was what we intended to achieve by presenting a conspicuous target that was changing its position at a high rate (4 Hz), so that the uncertainty would primarily be temporal. Nevertheless, to make completely sure that our interpretation is correct, rather than the bias arising from targets nearer to fixation being easier to detect, we conducted a control experiment in which targets of two sizes were presented at a single eccentricity. Figure 5 shows that a factor 4 in target diameter (16 in surface) makes no difference at all to subjects' likelihoods of selecting the target. Subjects also set the same average eccentricity for both target sizes (5.53 cm).

Discussion

Our most important result is that we found a clear bias towards selecting target positions near fixation. This was already evident from comparing how often subjects chose targets at each eccentricity with how often they should have done so (Table 1) and is confirmed by our more elaborate analysis. From Table 1 one may get the impression that the spatial bias is stronger when a flash is used to indicate the moment of interest, which would run counter to our conclusion that the spatial and temporal aspects of the task are independent. However, it is in fact consistent with our model because the spatial bias should be more pronounced if there is more temporal uncertainty, which is the case for the flashes (see Figure 4B).

We interpret the tendency to select positions near fixation as a bias in people's choices when in doubt.

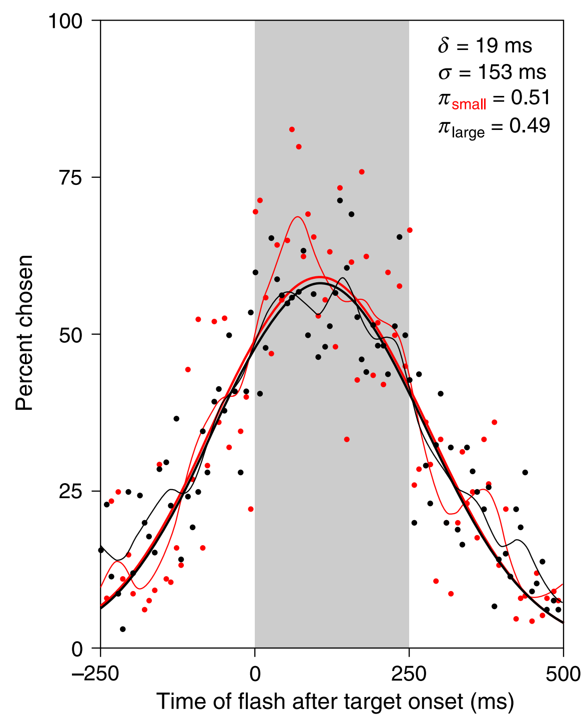

Figure 5. Eight subjects' pooled data for a control experiment with large (2 cm diameter; black symbols and curves) and small (0.5 cm diameter; red symbols and curves) targets, all at an eccentricity of 6 cm. Dots show the proportion of presentations on which that target was chosen. The thin curves are smoothed versions of the data. The two thick curves show the best fit of a 3-parameter model (values given in the top right corner). Clearly a 16-fold change in target surface makes very little difference.

However subjects were not extremely precise in indicating the selected position, especially at large eccentricities (Figure 2). What if subjects underestimated large eccentricities to a greater extent than would appear from Figure 2, and on some trials in which the eccentricity was clearly underestimated a different target position was nearer the indicated position, so such trials were removed because the nearest target was not within 60 deg of the indicated position (remember that consecutive positions were never presented within 60 deg of each other). Part of the reason for having to remove more trials than we would expect on the basis of the temporal properties alone (compare Tables 2 and 3), especially at large eccentricities, could be that subjects sometimes made large errors in indicating the selected position. However the difference in the number of removed positions between the eccentricities is much too small to interfere with our conclusions.

We found a larger systematic temporal error and less temporal uncertainty when the moment of interest was indicated by a tone than when it was indicated by a flash (Figures 4A and 4B). The lower uncertainty could just be because the temporal resolution of auditory stimuli is higher. However, the modelled uncertainty is a combination of uncertainty about the moment that the target jumps, uncertainty about the moment of the flash or tone, and uncertainty about synchronising the two. It is therefore also possible that the lower uncertainty with the tone is partly due to it being easier to synchronise the change in target position with a tone, than with another visual stimulus (the flash). It may also have to do with the fact that the flash was quite faint whereas the tone was clearly audible. Thus we cannot really conclude anything definite from these values.

The systematic temporal error may also be influenced by how conspicuous the flash was relative to the tone. However it is surprising to see that although the tone was more conspicuous and was associated with less temporal uncertainty, the systematic error was larger for the tone. We have no explanation for this, but it is reassuring to find that the systematic temporal errors that we found (about 50 ms for the flash and slightly more for the tone) are similar to those that have been found in other studies in which targets had to be localised at the moment of a flash (Brenner & Smeets, 2000; Murakami, 2001; Nijhawan, 1994) or tone (Alais & Burr, 2003).

In the introduction we proposed that it would be logical to be biased towards localising objects where we are looking, because we are most likely to see objects if we are looking at them. The fact that the bias towards the fixation point is as strong for the tone as for the flash makes it unlikely that it has anything to do with the relevance of the fixation point for the task at hand, because the region of the fixation point is much more relevant to the task when a flash indicates the moment of interest (the flash is at the fixation point) than when a tone does so. Therefore we feel that we can conclude that people are indeed biased towards believing that things that they saw were where they were looking.

Acknowledgments

Commercial relationships: none.
Corresponding author: Eli Brenner.
Email: e.brenner@fbw.vu.nl.
Address: Faculteit der bewegingswetenschappen, Vrije Universiteit, Van der Boechorstraat 9, NL-1081 BT Amsterdam, The Netherlands.

Footnotes

[1]This is probably not completely true, because for instance the fifth target position (1000–1250 ms) is sometimes chosen although it is never the correct choice.

[2]For simplicity we ignore the fact that subjects may—and probably do—occasionally consider a target that is further removed from the correct one than the one immediately before or after it, and the fact that the target before and after the correct one could be at the same eccentricity, because neither can be expected to make much difference.

References

Alais, D., & Burr, D. (2003). The "Flash-Lag" effect occurs in audition and cross-modally. *Current Biology, 13,* 59–63. [PubMed] [Article]

Brenner, E., & Cornelissen, F. W. (2005). A way of selectively degrading colour constancy demonstrates the experience dependence of colour vision. *Current Biology, 15,* R864–R866. [PubMed] [Article]

Brenner, E., & Smeets, J. B. (2000). Motion extrapolation is not responsible for the flash-lag effect. *Vision Research, 40,* 1645–1648. [PubMed]

Brenner, E., van Beers, R. J., Rotman, G., & Smeets, J. B. (2006). The role of uncertainty in the systematic spatial mislocalisation of moving objects. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 811–825. [PubMed]

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55,* 271–304. [PubMed]

Knill, D. C. (2003). Mixture models and the probabilistic structure of depth cues. *Vision Research, 43,* 831–854. [PubMed]

Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning. *Nature, 427,* 244–247. [PubMed]

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition, 81,* B1–B9. [PubMed]

Murakami, I. (2001). A flash-lag effect in random motion. *Vision Research, 41,* 3101–3119. [PubMed]

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370,* 256–257. [PubMed]

Párraga, C. A., Troscianko, T., & Tolhurst, D. J. (2002). Spatiochromatic properties of natural images and human vision. *Current Biology, 12,* 483–487. [PubMed] [Article]

Purves, D., Lotto, R. B., Williams, S. M., Nundy, S., & Yang, Z. (2001). Why we see things the way we do: Evidence for a wholly empirical strategy of vision. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 356,* 285–97. [PubMed] [Article]

Raninen, A., Franssila, R., & Rovamo, J. (1991). Critical flicker frequency to red targets as a function of luminance and flux across the human visual field. *Vision Research, 31,* 1875–1881. [PubMed]

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604. [PubMed]