The Role of Uncertainty in the Systematic Spatial Mislocalization of Moving Objects

Eli Brenner, Robert J. van Beers, Gerben Rotman, and Jeroen B. J. Smeets
Erasmus MC

It only makes sense to talk about the position of a moving object if one specifies the time at which its position is of interest. The authors here show that when a flash or tone specifies the moment of interest, subjects estimate the moving object to be closer to where it passes the fixation point and further in its direction of motion than it really is. The authors propose that these biases arise from a combination of a large temporal uncertainty, a temporal asymmetry related to sampling the moving object's position, and a bias toward believing that one is looking at what one sees.

Keywords: timing, compression, flash-lag, neuronal delays, spatial localization

Studies of human perception usually either deal with perceived positions of static targets or with perceived velocities of moving targets. Yet in daily life, if we want to move our hand to intercept a moving object, knowing precisely how fast it is moving and in which direction is not enough. We need to know where it will be at a certain time, so we must also estimate its position at some moment. Intercepting moving targets is not the only or even the most frequently occurring situation in which we need to judge the position of an object that is moving relative to ourselves. When we reach out to pick up a cup, to press a button, or to manipulate any object in any way, we normally move more than just our arm. In most cases in which our body moves, our head and therefore our eyes will also move with respect to the surrounding. Thus in terms of relative motion between the eyes and the object of interest in the surrounding, and therefore in terms of visual processing, very similar issues arise when interacting with a static object while moving one's body, as when interacting with moving objects. Judging the position of moving targets or of targets that change position relative to us as we move introduces a fundamental problem to studies of human perception because the precise timing of the percept becomes a critical issue. This problem is the basis of the present study.

Although we have the impression that we know where moving objects are, it only makes sense to talk about the position of a moving object if one specifies when one wants to know this position. When interacting with objects, the moment of interest is usually specified implicitly by the purpose of the interaction. However, for perceptual judgments the time has to be indicated explicitly. Many studies have used flashes or tones to indicate the moment of interest. In such cases people generally locate the moving target too far along its path (e.g., Alais & Burr, 2003; Baldo, Kihara, Namba, and Klein, 2002; Brenner & Smeets, 2000; Chappell, Hine & Hardwick, 2002; Ishii, Seekkuarachchi, Tamura, and Tang, 2004; Murakami, 2001b; Nijhawan, 1994; Ogmen, Patel, Bedell, and Camuz, 2004; Watanabe, Nijhawan, Khurana, and Shimojo, 2001; Whitney, Murakami, and Cavanagh, 2000).

One explanation for moving targets appearing to be further than they really are is that the target's position must be "sampled" in response to the flash (Brenner & Smeets, 2000). This sampling process can be thought of as finding the neurons that are most strongly activated by the moving target at the moment of interest. Since the peak in activation is constantly shifting and the moment of interest is not defined until the flash occurs, we find a position that corresponds to where the moving target is some time after the flash. This *sampling hypothesis* is different from the differential latency account of such mislocalization (Murakami, 2001a,b; Ogmen et al., 2004; Purushothaman, Patel, Bedell, and Ogmen,1998; Whitney and Murakami, 1998; Whitney, Murakami, and Cavanagh, 2000) in that the error results from the time taken to sample the position after having detected the flash rather than from differences between the delays of the signals involved. The critical assumptions are that the amount of time that the sampling process (determining the position of the peak in activation) itself takes is not negligible and that the human brain does not correct for this delay (because it has no way to look back at the activity at a certain moment; see Pöppel, 1997).

That any neuronal process must take time is obvious, but whether sampling a moving object's position would really take the tens of milliseconds that it would have to take for this to account for the observed errors in localizing moving objects is pure speculation. The idea that the brain does not bother to correct for mismatches in timing is slightly less speculative in that it receives some support from studies in which the mismatch has a very

Eli Brenner, Robert J. van Beers, Gerben Rotman, and Jeroen B. J. Smeets, Department of Neuroscience, Erasmus MC, Rotterdam, The Netherlands.

Eli Brenner is now at the Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands; Robert J. van Beers is now at the Department Physics of Man, Utrecht University, Utrecht, The Netherlands; Gerben Rotman is now at the Department of Psychology, Queen's University, Kingston, Ontario, Canada; and Jeroen B. J. Smeets is now at the Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands.

Correspondence concerning this article should be addressed to Eli Brenner, Faculty of Human Movement Sciences, Vrije Universiteit, Van der Boechorststraat 9, 1081 BT Amsterdam, The Netherlands. E-mail: e.brenner@fbw.vu.nl

different origin. Auditory information takes longer to reach us than visual information because sound travels less fast than light. The time difference depends on the distance, so all that one would need to compensate for the mismatch is an estimate of the distance. Lewald and Guski (2004) found that people do not compensate for the distance when judging simultaneity between sound and vision.

Perhaps the reason that no effort is made to compensate for mismatches in timing is that the delay between a stimulus and its neuronal response depends on so many factors that it is impossible to consider them all, making it rather pointless to consider any. Delays differ for different properties and brain regions, they are influenced by factors such as stimulus intensity and contrast (Oram, Xiao, Dritschel, and Payne, 2002; Schmolesky et al., 1998), and they can even change with age (Wang, Zhou, Ma, and Leventhal, 2004). Moreover neurons code information in firing rates, so neuronal responses are necessarily spread over an extended time period. Even responses to a very short flash consist of a series of action potentials rather than a single action potential at one instant. For moving objects, the responses are also extended in time as the moving object's image moves across the receptive field. Thus localizing moving objects must rely on some form of spatiotemporal averaging (Lappe & Krekelberg, 1998; Krekelberg & Lappe, 1999), whereby the precise timing is lost anyway.

Another consideration for not compensating for mismatches in timing is that the benefit of doing so is not really clear. Usually, when synchronicity is an issue, simultaneity can be assumed on the basis of causal relationships. For instance, when we watch a tennis game on television, the distance between the players' rackets and ourselves is no longer defined in any simple manner. Nevertheless we have the impression that we hear the racket hitting the ball at the same moment that we see the racket do so (also see Bertelson & Aschersleben, 2003). Naturally, events must occur at about the same time for them to be perceived as being simultaneous, but the resolution for detecting errors in synchronicity can probably be quite poor because we seldom need to rely on an ability to detect such errors in daily life.

The core assumption of our sampling hypothesis is that the relevant position of the moving target is determined in response to the flash or the tone, rather than synchronicity between the response to the flash or the tone and the response to the target at a certain position being detected. We therefore started our study by comparing the errors that subjects made when localizing a moving target at the moment of a flash with the errors that the subjects made when localizing a moving target at the moment of a tone, and with the errors that they made when synchronizing a flash with a tone. In the first part of this paper we report a systematic difference between the perceived timing of the flash and the tone in the synchronization task, which was not evident in the localization tasks (contrary to findings by Hine, White & Chappell, 2003). There was also more variability in the synchronization task than in the localization tasks. In the localization tasks, there was a bias toward perceiving the moving target to be close to the position at which it came closest to the fixation point. In the second and third parts of the study we examine some of the properties of this bias. In the fourth part we develop a model that can account for our findings as well as for several previous findings. This model also reconciles the different amounts of variability in the three tasks of the first part of the study. Finally we discuss our model's implications for localization under various other conditions.

Part 1: Comparing Errors in Synchronization and Localization

The main purpose of the first part of this study was to compare the errors that people make when synchronizing two signals with the errors that they make when localizing a moving object at the time of one of those signals. If the errors in all three tasks arise from a failure to compensate for differences in the delays of the underlying signals, we should be able to directly compare the three tasks in terms of timing errors. If the sampling hypothesis is correct, then the localization errors may be incompatible with the synchronisation errors.

Materials and Methods

Three of the authors and three of our colleagues (all men; 23 to 45 years of age) each performed three tasks in a fixed order. Only the authors were aware of the purpose of the study. Each subject first synchronized a flash with a tone (25 trials). Next the subject indicated the position of a moving target at the time of a flash (100 trials). Finally the subject indicated the moving target's position at the time of a tone (100 trials). This study is part of an ongoing research program that has been approved by the local ethics committee.

Stimulus Presentation

The stimuli were presented on a SONY G200 Monitor that was viewed from a distance of 75 cm in a normally illuminated room. The screen resolution was 1024 × 768 pixels (32 × 24 cm; 24 × 18 deg) and the refresh rate was 120 Hz. The tone was generated by the built-in loudspeaker of the Apple G4 computer that was standing near the monitor.

In all three tasks subjects fixated a 0.2° diameter red dot at the center of a white computer screen. The flashes were 1.2° diameter black disks superimposed onto the fixation point at the center of the screen for one frame. The amplitude of the 4500 Hz tones declined linearly to about half the initial value during the 230-ms duration of their presentation. Subjects were explicitly instructed to always consider the onset of the flash or tone if the presentations appeared to last for some time.

The moving target was a 0.2° diameter red dot that moved horizontally at 4 pixels per frame (about 11.2 deg/s) from left to right across the screen center. It could move along one of two paths, passing either 0.35° above or 0.35° below the fixation point. It was always at one of four points on the path (0.35° or 1.05° to the left or right of the fixation point) when the flash or tone occurred. The positions on the path at which the moving target appeared and disappeared were selected to ensure that the target would be visible for between 500 and 1,000 ms before the flash or tone and for between 250 and 500 ms after the flash or tone (see Figure 1). The durations were chosen at random for each trial from within these limits. Only one of the subjects (an author) was aware that there were only eight possible target locations at the moment of the flash or tone (the crosses in Figure 1).

Experiment 1. The subjects' first task was to synchronize the occurrences of flashes and tones. Both the flashes and the tones were repeated at a rate of 1 Hz. The subjects could manipulate the timing of the tones relative to that of the flashes by moving the computer mouse laterally. Their task was to align the relative timing so that the flashes and tones appeared to occur at the same time. The presentations continued until the subjects indicated that they were content with the set relative timing by pressing the button of the computer mouse. The initial relative timing was randomized on each trial.

Experiment 2. The subjects' second task was to indicate the moving target's position at the moment of a flash. After the presentation of the motion, a new 0.2° diameter red dot appeared at a random position on the
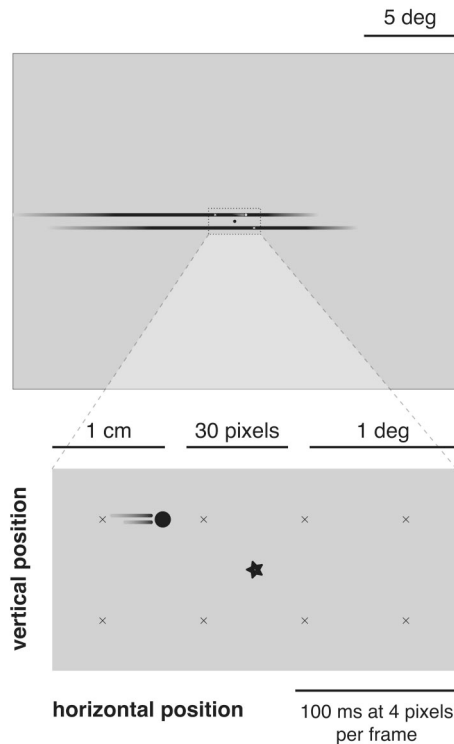
Figure 1. Schematic representation of the stimulus on the computer screen in Experiments 2 and 3. Top: Horizontal lines indicate the paths' positions on the screen. The gradients at the ends represent the variability in the starting points and endpoints of the motion for the leftmost (upper path) and rightmost (lower path) position at the moment of the flash or tone. Bottom: Enlarged view of the central area indicated by the dotted lines. This part of the screen is portrayed in Figures 2–6. The black circle represents the target moving to the right along the upper path. The crosses show the four positions on each of the two horizontal paths at which the target could be at the moment of the flash or tone. The star shows the fixation point, which is also the position of the flash in Experiment 2.

screen. It could be repositioned as a mouse cursor. The subjects' task was to place this "cursor" at the position at which the moving target had been at the moment of the flash. Subjects had to fixate the central fixation point while the target was moving but were free to move their eyes as they liked when setting the cursor. Again subjects indicated that they were content with their settings by pressing the button of the computer mouse.

Experiment 3. The subjects' third task was identical to the second, but the moment of interest was indicated by a tone instead of a flash. Subjects had to place the cursor at the position at which the moving target had been at the moment of the tone.

Analysis

We were interested in both the variability and any biases in our subjects' performance. For the synchronization task (Experiment 1), we analyzed the time that the flash was set to appear relative to the tone. We averaged the data in two steps. First, we calculated the means and the standard deviations of the set time intervals for each subject. We then calculated the average of these means as our value for the bias and the average of the standard deviations as our value for the variability within individual subjects' settings. We determined the standard deviation of the distribution of the means and of the standard deviations to get an estimate for the variability between subjects.

For the two localization tasks (Experiments 2 and 3), we determined the average set horizontal and vertical positions and the standard deviations of the set horizontal and vertical positions for each of the real target positions. We did so for each subject separately. The difference between the average set position and the real position is our measure of the bias. The standard deviations across replications are our measures for the variability. To get a single value for each subject, we averaged the biases and standard deviations across the eight real positions. We then calculated the average and the standard deviation of both the bias and the variability across the six subjects to get an estimate for the variability between subjects.

Results and Discussion

Figure 2 and Table 1 summarize the results of all three tasks (Experiments 1 through 3). The thick bar in the upper panel of Figure 2 shows that subjects consistently set the flash to appear before the tone in the synchronization task (Experiment 1; $t_5 = 4.2$; $p < .01$). The tone had to be presented about 56 ms later than the flash for it to appear to be at the same moment as the flash. The scale of the figure is such that if there had been a similar bias in timing in the localization tasks, the black symbols (flash; Experiment 2) in the lower panel would have been shifted to the right by the length of the thick bar in the upper panel relative to the corresponding white symbols (tone; Experiment 3). This is clearly not the case, so we cannot interpret our results in terms of differences between the neural latencies for flashes, tones, and moving stimuli. A possible interpretation that suits our sampling hypothesis is that localization only depends on the moment that subjects
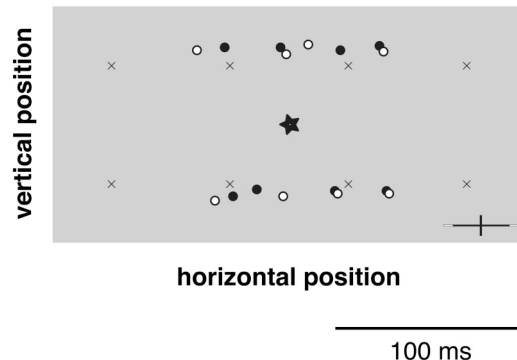
Figure 2. *Upper panel.* Average set timing difference when trying to synchronize the flash with the tone (Experiment 1). *Lower panel.* Average set position when the moving dot was at each of the eight positions on the screen (crosses) at the moment that the flash (black circles; Experiment 2) or tone (white circles; Experiment 3) occurred. The horizontal scales in the two panels are equivalent in terms of misjudgments of time. The bars in each panel's lower right corner show the average standard deviation of replications by the same subject (both in the direction of motion and in the orthogonal direction). In the lower panel black bars are for the flash and white ones are for the tone.

Table 1
Average Bias and Variability in the First Three Experiments

Experiment	Task	Bias (ms)	Variability (ms)
1	Synchronization	56 ± 32	86 ± 23
2	Localization at flash	9 ± 17	31 ± 17[a]
3	Localization at tone	6 ± 18	40 ± 12[a]

Note. For the localization tasks, each subject's bias and variability was first determined for each of the eight positions, and then averaged across positions. The presented values are the means ± standard deviations across subjects. We considered setting a later position of the moving target and presenting the tone later than the flash as positive biases. The variability is the standard deviation between settings on identical trials.

[a] Only the horizontal variability is considered. The spatial variability was converted into milliseconds on the basis of the target's velocity.

first detect the flash or tone, whereas synchronization also depends on the perceived duration (despite our instructions to synchronize the onsets; see Jaskowski, 1991). Perhaps brief targets appear to occur simultaneously when the peaks in the neuronal responses that they generate coincide in time, rather than when the responses reach the detection threshold at the same moment. An analogous difference could explain why simple reaction times did not reproduce the systematic differences between stimuli that were found in a synchronization task or in the flash-lag effect (Nijhawan, Watanabe, Khurana, and Shimojo, 2004; Stone et al., 2001).

The variability between the settings was much larger for the synchronization task than for the two localization tasks (Table 1 and horizontal bars in the lower right corners of the two panels of Figure 2). The variability between replications was more than twice as large when synchronizing two signals (Experiment 1) than when localizing a moving object at the time of these signals (Experiments 2 and 3), despite the fact that subjects could see and hear the stimuli as often as they wanted in the synchronization task, whereas the stimuli were only presented once in the localization tasks. The variability between subjects (both in the bias and in the variability) was also larger in the synchronization task than in the localization tasks. The larger variability is consistent with synchronization being determined by the moment of the peak in the neuronal response, while localization is determined by the moment that a neuronal response is reliably detected, because for responses that last for some time the onset will be more sharply defined than the peak. The larger variability in our synchronization task is therefore consistent with the proposal that in Experiments 2 and 3 localization was based on sampling the continuously changing position signal in response to detecting the flash or tone, rather than being based on synchronizing the peaks of the neural responses to the two signals (the flash or tone and the moving target).

The settings in the localization tasks also show a strong compression in the horizontal direction (see Figure 2). There was no such compression in the vertical direction, where the settings may even be slightly further apart than the target positions. We had expected to see a shift in the direction of motion (i.e., to the right), but what we saw was mainly a horizontal shift toward the fixation point. This compression of the range of perceived positions was clearly present in all six subjects' data. To estimate the extent of the compression for each subject, we fit a straight line to the relationship between their average set horizontal position and the

actual horizontal position at the moment of the flash or tone. The compression is the extent to which the slope of this line is lower than 1. The compression was 58 ± 11% for Experiment 2 and 54 ± 21% for Experiment 3 (average ± standard deviation across subjects). Without knowing the origin of this compression, it is impossible to interpret the results. In particular, we need to know whether the compression could have decreased the variability and the shift in the direction of motion. We therefore decided to look for the origin of this compression.

Part 2: Why the Compression?

We only found compression in the horizontal direction. The variability in the subjects' settings was also much larger in the horizontal than in the vertical direction, as can be seen from the error bars in the lower right corner of the lower panel of Figure 2. The difference in variability is understandable in the light of temporal uncertainty. Uncertainty about the precise moment of interest is not expected to affect judgments of the vertical position, because the vertical position does not change when the target moves along either of the clearly identifiable horizontal paths. Uncertainty about the moment of interest obviously does affect judgments of the horizontal position, because the horizontal position is constantly changing. To determine whether the compression could also be related to temporal uncertainty, perhaps through its influence on spatial uncertainty, we asked the same six subjects to perform a number of additional experiments.

We first confirmed that the compression was related to the movement of the target, rather than to remembering the target's position (Sheth & Shimojo, 2001) or to moving the cursor to indicate its position. Next we examined whether the compression was toward the direction of gaze or the center of the screen, and whether it was influenced by clear visual landmarks (it could not be toward the position of the flash, because we found the same compression when a tone was used as when a flash was used). Finally we examined how the compression depends on the target's velocity.

Materials and Methods

Part 2 of the study consisted of five experiments that were based on Experiments 2 and 3 of Part 1 of the study. The experiments and analyses were identical to those of Experiments 2 and 3 except for the details mentioned in the next paragraphs, and that the cursor always appeared at the same place in Part 2, rather than at a random position on the screen. It emerged from "behind" the fixation point as soon as the subject moved the mouse (but only if the target had reached the end of its path; mouse movements before that time were ignored). The same six subjects who had participated in Part 1 also participated in Part 2 of the study, except that one subject was replaced in Experiment 4 (because the original subject was not available at the time). Experiments 4, 5 and 6 were conducted on separate days. Experiments 7 and 8 were conducted consecutively within one session on another day.

Experiment 4. Experiment 4 was designed to confirm that the compression is related to the target's motion and to obtain a measure of the spatial accuracy in the absence of such motion. In Experiment 4 the target jumped at random between positions on the central 3° of the two "paths." The target jumped to a new position every 500 ms for between 1.5 and 3 s (3 to 6 positions). All targets except for the last could be anywhere within the above-mentioned range. The last target was always at one of the eight positions at which the moving targets had been at the moment of the flash

or tone in the previous localization experiments (Experiments 2 and 3; crosses in Figure 3).

After spending 500 ms at the last position the target disappeared. As in all the localization experiments (Experiments 2 through 8), the subjects had to keep fixating the fixation point until the target disappeared, but in Experiment 4 they had to indicate its final position rather than its position at the time of a flash or a tone. We presented the target for quite a long time at each position to be sure to eliminate any temporal uncertainty. Each subject indicated the positions of 100 targets.

Experiments 5 and 6. Experiment 5 was designed to determine whether the compression that was found in Experiments 2 and 3 was toward the center of the screen or toward the direction of gaze. It consisted of two blocks of 100 trials. Each block was an exact replication of Experiment 2 except for the position of the fixation point (and therefore also the position of the flash), which was to the left of the center of the screen in the first block and to the right of the center of the screen in the second block. (For the precise positions see Figure 4.)

To determine whether a conspicuous visual reference on the target's path would influence the localization, we repeated Experiment 5 with a high-contrast vertical black line on the screen (Experiment 6). The line was slightly to the right of the center of the screen if the fixation point was on the left and slightly to the left of the center of the screen if the fixation point was on the right (see Figure 5).

Experiments 7 and 8. Experiments 7 and 8 were exact replications of Experiments 2 and 3, respectively, except for the fact that the target could now move at one of three different velocities: the same velocity as in Experiments 2 and 3 or at half or one quarter of that velocity (4, 2, and 1 pixel per frame, or 11.2, 5.6 and 2.8 deg/s). We repeated the conditions of Experiments 2 and 3 in case randomly varying the velocity influences the magnitude of compression. There were 100 trials at each velocity, and they were interleaved in a random order.

Results and Discussion

Figure 3 shows the settings for the jumping target (Experiment 4). The settings were quite accurate. Most importantly, there was no systematic compression. Thus the compression in Experiments 2 and 3 must indeed be related to the fact that the target was moving. Figure 4 shows how a slightly eccentric fixation affects the subjects' settings (Experiment 5). The compression is clearly primarily horizontally toward fixation rather than toward the center of the screen or toward the average real position. The settings are further to the right when the fixation point is on the right than they are to the left when the fixation point is on the left, suggesting that there is also a modest shift to the right (as predicted by our
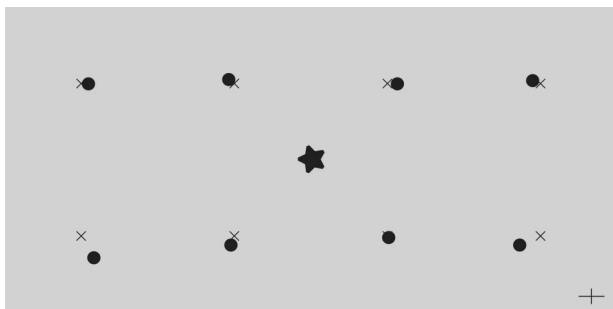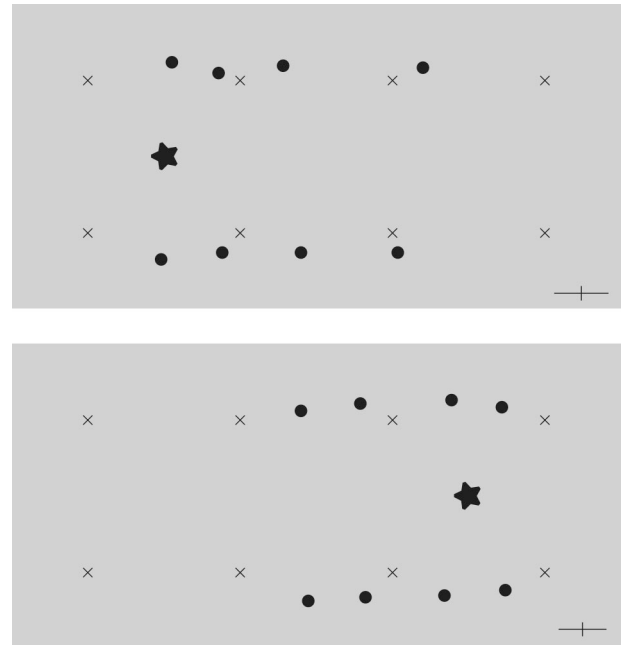

Figure 4. Experiment 5: Eccentric fixation. The average set positions (circles) when the moving dot was at each of the eight positions indicated by the crosses at the time that the flash occurred. The stars indicate the position of the fixation point and of the flash. The bars in each panel's lower right corner show the average standard deviation between replications by the same subject.

sampling hypothesis). These findings are confirmed in Figure 5 (Experiment 6), which shows that adding a conspicuous visual reference has very little effect on the compression and the rightward shift. Our subjects could not even reliably tell whether the target had or had not crossed this clearly visible landmark at the time of the flash: they usually indicated that it was at the same side of the line as the fixation point, even if it was not. Note that only the position of the fixation point was different in the two conditions of Experiment 5, and only the positions of the fixation point and the line were different in the two conditions of Experiment 6, so that in terms of positions on the screen or along the moving target's path, the correct answers in the two conditions of each experiment were identical. We conclude from these experiments that the compression must be related to *retinal* eccentricity.

Figure 6 shows the results for the three different target velocities (Experiments 7 and 8). The lowest panel is for the same velocity that was used in Experiments 2 and 3, so we expected the same results. The results are indeed very similar, but this time the compression was slightly stronger when the flash indicated the moment of interest than when a tone indicated the moment of interest. Note that although the compression was different for the tone and the flash, again the settings for the tone were not shifted relative to those for the flash in the manner that one would expect from the synchronization task in Experiment 1. The most important finding was that the compression was smaller when the target moved more slowly.

The spatial variability does increase with target velocity, but the increase is modest (see Figure 6). When converted into variability

Figure 3. Experiment 4: Jumping dot. The average set positions (circles) for the eight final target positions (crosses). The star indicates the fixation point. The bars in the lower right corner show the average standard deviation between replications by the same subject.
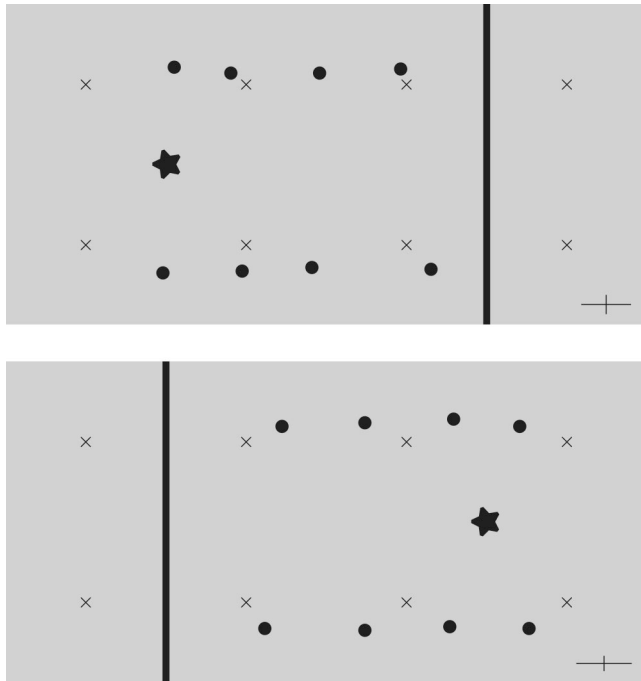
Figure 5. Experiment 6: Visual reference. Same format as in Figure 5, but the moving targets' paths crossed a conspicuous vertical line.

in timing (see Table 2), the standard deviation clearly decreases with increasing target speed, indicating that the errors cannot be purely temporal. Again we estimated the extent of the compression (for each experiment and velocity) by fitting a straight line to the relationship between the average set horizontal position and the actual horizontal position at the moment of the flash or tone. The compression is the extent to which the slope of this line is lower than 1. Table 2 shows that the decrease in temporal variability for high target velocities is accompanied by an increase in compression and a decrease in bias (in ms).

Part 3: Eccentricity and Direction of Motion

The results of Part 2 strongly suggest that the compression is toward the fovea (fixation). However, since we always used horizontal motion, we cannot exclude the possibility that the compression that we found is toward the intersection between the left and right visual fields rather than toward the fovea. It could, for instance, somehow be related to the fact that the moving target stimulates neurons in different hemispheres of the brain when the target passes from the left side to the right side of the fixation point (as has even been shown to occur in the parietal cortex; Medendorp, Goltz, Vilis, and Crawford, 2003). Similarly, we attributed the lack of vertical compression to the fact that temporal uncertainty does not influence the vertical position. But it could also be due to the fact that the vertical eccentricity was always the same. To better describe the spatial characteristics of the compression, we conducted two final experiments. In these experiments we varied the direction of motion and the (minimal) distance of the moving target from the fixation point.

Materials and Methods

The new experiments were based on those of the second part of the study. The methods and analyses were identical to those used for the previous experiments in which the moment of interest was signaled by a flash, except for the details mentioned in the next paragraphs. Three of the six subjects who had taken part in the first two parts also took part in Experiment 9 (two authors; one naïve; all male; 32–48 years old), as did three additional subjects who had not taken part in any of the previous experiments (one author; two naïve; all male; 25–37 years old). Six other naïve subjects who had not taken part in any of the previous experiments took part in Experiment 10 (3 men, 3 women; 22–39 years old).

Experiment 9. In this experiment the stimuli were presented on a larger monitor than in Experiments 1 through 8 (SONY GDM-FW900; 1280 × 800 pixels; 48 × 30 cm; 35° × 23° at the 75 cm viewing distance; 120 Hz). The fixation point was now a 0.2° diameter greenish disk at the center of a gray screen. This disk turned black for one frame to produce a faint "flash." Our reason for making the flashes much less conspicuous was that this ensured that the subjects were fixating correctly. If they were not, they were sure to miss the flash. When this happened, the subject pressed the space bar and the trial was repeated. Subjects were instructed to only repeat the trial if they completely missed the flash.

The moving target was a 0.2° diameter reddish dot that moved at either 2 or 8 pixels per frame (about 6.4 or 26 deg/s). The distances between the
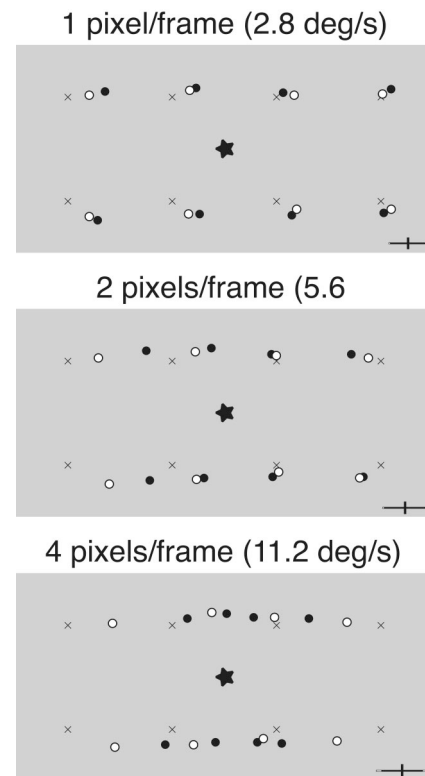

Figure 6. Experiments 7 and 8: Target velocity. The average set positions when the moving dot was at each of the eight positions indicated by the crosses at the time that the flash (black circles; Experiment 7) or tone (white circles; Experiment 8) occurred. The bars in each panel's lower right corner show the average standard deviation between replications by the same subject. Black bars are for the flash and white ones are for the tone. The stars at the centers indicate the fixation position, which was also the position of the flash. Each panel represents one target velocity.

Table 2
Average Horizontal Bias, Variability, and Compression in Experiments 5 through 8

Experiment	Condition	Velocity (pixels/frame)	Bias (ms)	Variability (ms)	Compression (%)
5	Flash / range asymmetric	4	9 ± 7	21 ± 7	51 ± 12
6	Flash / asymmetric with line	4	7 ± 5	21 ± 7	49 ± 15
7	Flash	1	40 ± 28	87 ± 30	11 ± 17
		2	22 ± 17	51 ± 20	35 ± 23
		4	7 ± 11	25 ± 11	62 ± 13
8	Tone	1	42 ± 43	93 ± 21	4 ± 26
		2	15 ± 25	54 ± 11	28 ± 30
		4	9 ± 13	31 ± 10	51 ± 30

Note. Each subject's bias and variability was first determined for each of the eight positions (and two fixation conditions in Experiments 5 and 6) and was then averaged across positions (and fixation conditions). These values were converted into milliseconds on the basis of the target's velocity. We considered setting a later position of the moving target as a positive bias. The variability is the standard deviation between settings on identical trials. Each subject's compression was estimated from the slope of their average set positions (for each real horizontal position) as a function of the real position. All the presented values are means ± standard deviations across subjects.

positions at which the moving target could be at the moment of the flash were twice as large as in Parts 1 and 2: they were 1.4° apart instead of 0.7°. The only other difference between the trials in this ninth experiment and those in Experiment 2 was where the target moved. Due to the large number of trials, the experiment was conducted in two sessions on different days.

In the first session the target passed the fixation point at three different distances: 0, 0.7, and 1.4 deg. It always moved from left to right. There were 10 trials for each horizontal position of the target at the moment of the flash (for each subject, target velocity and distance). Unless the target moved through the fixation point (distance = 0), the target moved above the fixation point on five of these trials and below the fixation point on the other five (see Figure 7).

In the second session the target always passed 2.1 deg from the fixation point, but it moved in four different directions: rightward above fixation, downward to the left of fixation, or diagonally (rightward and downward), either passing above and to the right or below and to the left of the fixation point. In each case, the four possible target positions at the moment of the flash were arranged symmetrically with respect to the position at which the target passed closest to the fixation point. There were 10 trials for each position on each path. The vertical path keeps the target in the left visual field (right hemisphere of the brain). The diagonal paths separate crossing the horizontal midline of the screen or visual field, from crossing the midpoint of the set of possible positions or the position closest to the fixation point.

Within each session the different kinds of trials were presented in a random order. The procedure and analysis were identical to those of the previous parts of the study. We still distinguish between variability along the moving target's trajectory and variability orthogonal to its trajectory, but these obviously do not correspond with the horizontal and vertical variability for the diagonal paths.

Experiment 10. To examine whether our findings also hold for larger eccentricities, we back-projected stimuli onto a large screen (115 × 85 cm; 75° × 59° at the 75-cm viewing distance; 800 × 600 pixels; 96 Hz) using a CRT projector (SONY VPH 1271QM) and moved the fixation point away from the center of the display (15° down and 4° to the right). Otherwise the trials were like those with horizontal paths in Experiment 9, except for the following details: The target moved to the right at 5 pixels per frame (about 43 deg/s). The distances between the positions at which the moving target could be at the moment of the flash were 2.8° apart (they were centered on the screen so that the rightmost position was aligned with the fixation point). The target's path was either 30° above the fixation point

or 0.8° below the fixation point (see Figure 8). The diameter of the greenish fixation point was increased to 0.8°. The diameter of the reddish moving dot was also 0.8° when its path was just below fixation, but it was 3.2° when its path was far above fixation (because otherwise the dot was hard to see). For each subject there were 10 trials for each position on each path (presented in random order).

Results and Discussion

Figure 7 shows the average set positions and the variability for the seven kinds of paths and two target speeds of Experiment 9. Irrespective of both the eccentricity and the direction of the path, there was a modest compression (about 10%) for the slower velocity (6.4 deg/s) and a considerable compression (about 50%) for the faster velocity (26 deg/s). Estimates of the extent of the compression for each path and velocity are given in Table 3.

The compression was slightly weaker than in Experiment 7: the same magnitude of compression was found for slightly faster targets. This was not a consequence of having three different subjects because it was just as evident for the three subjects who had also taken part in the previous experiments. Systematically varying the attributes that we had changed in pilots for Experiment 10 suggested that the larger distance between the positions at which the target could be at the moment of the flash was responsible for the compression being weaker (and not the fact that the target was 50% brighter rather than 20% darker than the background in Part 3, or that the flash was much less conspicuous). This was confirmed in Experiment 10, where the distances were twice as large as in Experiment 9 (four times as large as in the other experiments), and we found a similar compression at about the same eccentricity for an even higher target velocity. The compression in Experiment 10 was smaller when the target moved farther from fixation (at 30° eccentricity; compare the points along the two paths in Figure 8 or the values in Table 3). This implies that we are not really dealing with a uniform compression. (We will return to this issue below and when discussing the model in the next part of the study.) However, the compression is clearly not restricted to certain parts of visual space or to certain directions of target motion.
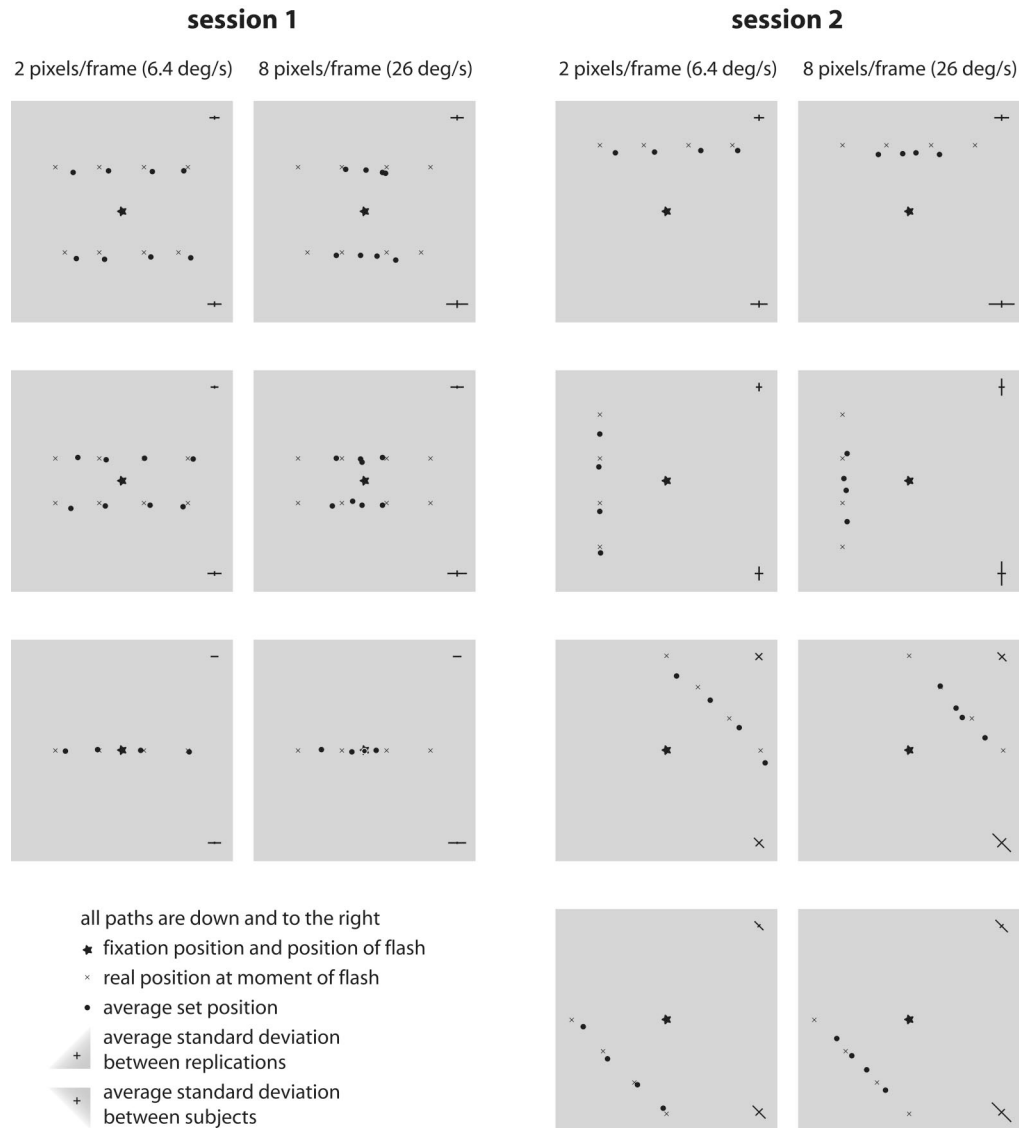
Figure 7. Localization errors for targets moving at various distances from fixation and in various directions (Experiment 9). The black circles show the average set positions when the moving dot was at each of the positions indicated by the crosses at the time that the flash occurred.

As we saw before, the variability is mainly in the direction of target motion. In Experiment 9 the spatial variability was not four times as large when the target moved four times as fast (see error bars in lower right corners of Figure 7). Thus, when converted into temporal variability (on the basis of the target's speed), the variability between replications by the same subject was smaller for the higher target velocity (see Table 3). The variability in the bias between subjects was also smaller for the higher target velocity, as was the bias itself (to the extent that it was not even consistently in the direction of target motion). At the same time, the compression was larger for the higher target velocity. Altogether the results of Experiment 9 are similar to those (for a slightly lower velocity) in Experiment 7, irrespective of the target's path.

In Experiment 10 we found less compression at a larger eccentricity. As was to be expected, the variability in the settings was larger when the target's path was never closer than 30° from fixation than when it passed 0.8° from fixation. It was larger both in the direction of target motion and in the orthogonal direction and both within replications by the same subject and between the subjects' average values (see Figure 8). This combination is inconsistent with any simple interpretation whereby a larger spatial uncertainty is directly responsible for the compression. However, we are not confident that the stronger compression when the target passed at 0.8° eccentricity is related to the reduced variability at the lower eccentricity. Perhaps the target partly passing behind the fixation point when at 0.8° eccentricity somehow helped the subjects to determine that the flash did not occur after the target passed, and thereby increased the compression. This possibility is supported by the fact that the distribution of settings was clearly skewed for the targets that passed at 0.8° eccentricity (as shown for
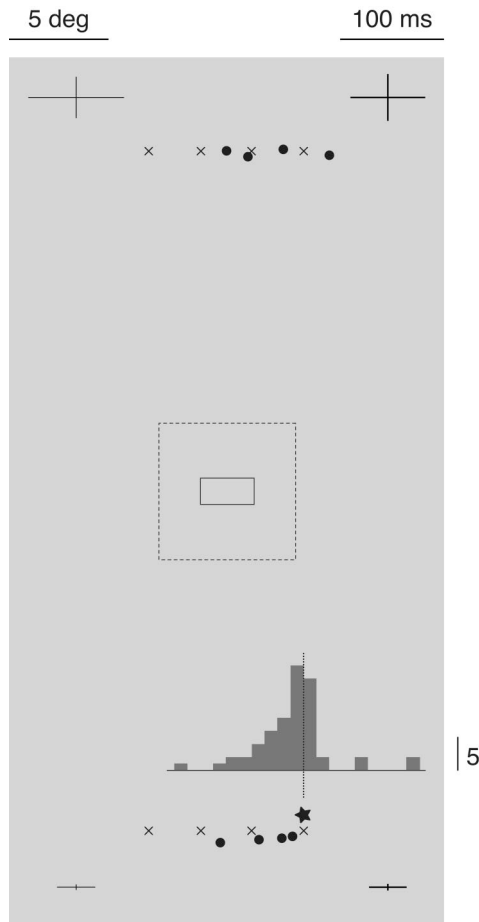
Figure 8. Localization errors for targets moving at larger distances from fixation (Experiment 10). The black circles show the average set positions when the moving dot was at each of the positions indicated by the crosses at the time that the flash occurred. The star indicates the position of the fixation point. The rectangle and dotted square in the center show the dimensions of the parts of the screen portrayed in Figures 2–6 and Figure 7 (gray areas), respectively. The bars in the corners show the standard deviations between replications (thick bars; right corners) and between subjects (thin bars; left corners). The histogram shows the distribution of settings for the rightmost position on the lower path (same horizontal scale; dotted line indicates the correct position).

the rightmost target in the histogram in Figure 8) but not for those that passed at 30° eccentricity (not shown).

A clear conclusion that we can draw from this part of the study is that the compression has nothing to do with the connections between the two hemispheres of the brain. From the diagonal paths we can also conclude that the compression is toward the position at which the path comes closest to fixation rather than toward the vertical or horizontal meridian. Despite the different paths (directions and eccentricities) on different trials, the compression is always along the path, indicating that we were correct in assuming that subjects know where the path is but are uncertain about where on the path the target was at the moment indicated by the flash. Of course, there may also be a bias toward the center of the screen (or of the path), but such a bias must be small in comparison with that

toward the fixation point. The shift in the direction of motion (which we attribute to a sampling delay) is evident from the settings for the rightmost position on the more eccentric path being to the right of both the correct position and the fixation point.

Part 4: The Model

Now that we have characterized the compression, it is time to look at the relationship between the compression, the bias in the direction of target motion, and the variability between replications of the same setting. We saw that as the target velocity increases, the compression increases to the extent that it completely overshadows the bias. The spatial variability also increases slightly with target velocity, but the increase is much smaller than one would expect if the errors were purely temporal. To try to reconcile these findings, we decided to model them within the context of the sampling theory that was described in the Introduction section.

Let us start with the velocity-dependent compression toward the direction of gaze. Why should there be such a bias toward the fovea? We propose that when faced with a high degree of uncertainty about a target's position, people assume that they are looking at it. This is a reasonable assumption because there is certainly a higher probability of seeing things that are where one is looking than of seeing things that are where one is not looking. Thus there is an increased probability that objects that one sees are where one is looking. If we accept this proposal we can model the compression as the result of a Bayesian prior. (For a more thorough explanation of the underlying reasoning as applied to object and motion perception, see Kersten, Mamassian & Yuille, 2004 and Weiss, Simoncelli & Adelson, 2002, respectively). According to our prior, positions closer to the fovea are more likely than positions that are further away. This means that the less reliable the retinal information is about the target's position, the more the target's position will be misjudged toward the fovea. This may explain why *flashed* targets are also sometimes perceived too close to the fovea (Musseler et al., 1999).

Modeling the compression in this way can explain why the compression is only evident in the direction of motion: in the orthogonal direction the prior receives little weight because there is very little uncertainty. Assuming that temporal uncertainty is the limiting factor (for localization in the direction of motion), it is evident that the spatial uncertainty will increase with the target's speed, so the relative weight assigned to the prior will increase with target speed. The standard deviations in the settings for the different velocities do not appear to support the assumption that the uncertainty is primarily temporal in origin, but we will show that this could be because the expected increase in spatial variability is partly compensated for by the spatial compression.

Figure 9 explains schematically how our model works. The horizontal axis is the position along the target's path, and the vertical axis is time. The slanted line represents the moving target, and the horizontal line represents the moment of the flash or tone. The horizontal position at which the two lines intersect is the actual target position at the time of the flash or tone (x). The curve on the vertical (time) axis represents the uncertainty in the timing (see Murakami, 2001a). Both the delay (D) and the uncertainty (σ_t) are shown. The delay is visible as an upward shift of the peak relative to the time of the flash or

Table 3
Average Bias, Variability, and Compression in Experiments 9 and 10

Experiment	Velocity (deg/s)	Eccentricity (deg)	Direction	Bias (ms)	Variability (ms)	Compression (%)
		0	Horizontal	5 ± 19	62 ± 13	7 ± 9
		0.7	Horizontal	31 ± 20	67 ± 25	14 ± 5
		1.4	Horizontal	36 ± 26	68 ± 25	14 ± 15
	6.4		Horizontal	49 ± 37	82 ± 30	7 ± 11
			Vertical	47 ± 24	67 ± 10	9 ± 11
		2.1	Diagonal above fixation	73 ± 31	67 ± 15	7 ± 13
			Diagonal below fixation	20 ± 51	86 ± 30	14 ± 12
9		0	Horizontal	-14 ± 5	21 ± 6	60 ± 17
		0.7	Horizontal	-7 ± 11	23 ± 8	66 ± 8
		1.4	Horizontal	4 ± 7	26 ± 9	62 ± 7
	26		Horizontal	0 ± 15	30 ± 10	55 ± 13
			Vertical	6 ± 15	29 ± 6	51 ± 20
		2.1	Diagonal above fixation	11 ± 10	30 ± 10	51 ± 16
			Diagonal below fixation	-5 ± 16	32 ± 8	47 ± 19
10	43	0.8	Horizontal	*	36 ± 17	54 ± 27
		30	Horizontal	*	72 ± 15	34 ± 30

Note. A positive bias means that a later position of the moving target was indicated. The variability is the standard deviation between settings on identical trials. Both values were converted into milliseconds on the bias of the target's velocity. The compression is estimated from the slope of the average set position along the path as a function of the real position. The given eccentricity is at the moment that the target passes the fixation point. Each subject's bias and variability was determined for each position of the target at the moment of the flash and was then averaged across positions. The presented values are the means ± standard deviations of the six subjects' (average) values. For the precise paths of the moving targets, see *Methods* and *Figures 7 and 8*.
* Because of the asymmetry in the display we cannot provide an assumption-free measure of the bias for Experiment 10.

tone. The uncertainty is represented by the width of the (normal) distribution. The prior is represented by a normal distribution (standard deviation σ_p) centered on the fovea (dotted, leftmost curve on horizontal axis). We propose that the localization is based on a Bayesian (optimal) combination of the prior and the target's position at the estimated moment of interest (in Bayesian terms the evidence or the spatial likelihood distribution). For more about such optimal combination, see
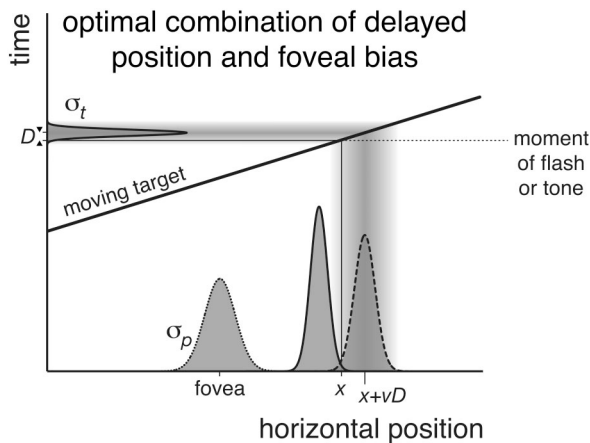

Figure 9. Schematic representation of the sampling model with delay *D*. The temporal uncertainty (σ_t) leads to a velocity-dependent spatial uncertainty (dashed curve), which is combined with the foveal prior (dotted curve; σ_p) to obtain a distribution of perceived positions (central solid curve).

Ernst and Banks (2002). Since the relationship between uncertainty about the moment of interest (curve on vertical axis) and the spatial likelihood distribution (dashed, rightmost curve on horizontal axis) depends on the target's speed, the optimal combination of the spatial likelihood distribution and the prior will also depend on the target's speed. The optimal combination is represented by the central solid curve on the horizontal axis in Figure 9.

So, can this model account for all the biases in our data? To find out we have to fit the model to the data. We will focus on Experiments 7 and 8 because a critical test of the model requires variability in target speed, and we used three speeds in Experiments 7 and 8. First we must convert the hypothetical estimated moment of the flash or tone into a spatial likelihood distribution. The peak is at the position that the target reaches after a delay *D* relative to the moment of the flash or tone. Thus, if the real horizontal target position at the time of the tone or flash was *x*, the peak will be at $x + vD$, where *v* is the target's speed. The uncertainty is $v\sigma_t$. If we define the horizontal positions with respect to the fovea, then the position of the peak (*p*) of the posterior distribution is:

$$p = \frac{\sigma_p^2}{\sigma_p^2 + (v\sigma_t)^2}(x + vD) \qquad (1)$$

Where $\frac{\sigma_p^2}{\sigma_p^2 + (v\sigma_t)^2}$ is the weight given to the evidence. The remaining weight $\frac{(v\sigma_t)^2}{\sigma_p^2 + (v\sigma_t)^2}$ is given to the prior, but since the position is measured relative to the fovea, so that the position indicated by the prior is zero, this term can be omitted. Equation

1 has three unknown parameters: D, σ_p and σ_t. However, only D and the ratio between the last two parameters are relevant for finding the position of the peak p, so if we define a reliability ratio r as:

$$r = \frac{\sigma_t}{\sigma_p} \qquad (2)$$

Equation 1 can be rewritten as:

$$p = \frac{x + vD}{1 + (vr)^2} \qquad (3)$$

This equation was fit to the average set horizontal positions for the three target speeds (thus averaged across subjects and vertical positions). Figure 10 shows these average positions together with the fit lines. The fits are quite good considering that there were only two free parameters (D and r) for each set of three lines. The fit was conducted separately for the flash and the tone, but the results were obviously similar in both cases because the data were similar. The values for the delay (D; 30 ms for the flash and 22 ms for the tone) are quite modest. They do not differ by the 56 ms that we found for the direct comparison in the synchronization task. We also fit the model to the individual subjects' data. This gave an average value of D of 36 ms for the flash and 25 ms for the tone (with standard errors of 16 and 11 ms respectively). The average values for r were respectively 135 ($\pm$ 21) and 131 ($\pm$ 38) ms/deg. Fitting the model to the average data of individual subjects for all the conditions in Experiment 9 gave a value of D that was not significantly different from zero and a value of r of 42 ($\pm$ 2) ms/deg. The smaller value was to be expected because, as we already noted, the compression was weaker (for the same velocity).

Equation 3 gives us the most likely position at which subjects will see the target. However we can also try to predict the variability in the settings. To do so we have to separate the reliability ratio (r; see Equation 2) into its two components. This is only possible if we have some absolute measure of variability. Since our model is based on the assumption that most of the uncertainty is temporal in origin, we decided to use the variability that we had found in the synchronization task (Experiment 1) for this, because

it is the only purely temporal estimate that we have. This variability (σ_{synch}; see equation 4) can be combined with the two fit values of r (see equations 5 and 6) to estimate a value of σ_t for the flash ($\sigma_{t\text{-}flash}$) and one for the tone ($\sigma_{t\text{-}tone}$). There is no reason to expect the prior (i.e., the value of σ_p) to depend on whether a tone or a flash was used, so we can write three equations with three unknown parameters:

$$\sigma_{synch}^2 = \sigma_{t\text{-}flash}^2 + \sigma_{t\text{-}tone}^2 \qquad (4)$$

$$r_{flash} = \frac{\sigma_{t\text{-}flash}}{\sigma_p} \qquad (5)$$

$$r_{tone} = \frac{\sigma_{t\text{-}tone}}{\sigma_p} \qquad (6)$$

where r_{flash} and r_{tone} are the values that we obtained from fitting equation 3 to the average set positions and σ_{synch} is the average standard deviation of replications in the synchronisation task (85 ms; see Table 1). From equations 4 through 6 we can then determine the values of $\sigma_{t\text{-}flash}$, $\sigma_{t\text{-}tone}$ and σ_p. We find values of 67 ms, 53 ms and 0.53°, respectively (for a similar estimate of $\sigma_{t\text{-}flash}$ see Murakami 2001a). Combining the information optimally, in accordance with the reasoning underlying equation 3, gives us predictions for the combined spatial variability (see van Beers, Sittig & Denier van der Gon, 1996). Since there is also variability in people's responses that is unrelated to the target's motion (see Experiment 4), we also added the modest spatial variability in the responses for the jumping targets ($\sigma_{jump} = 0.12°$), to get:

$$\sigma_{flash}^2 = \frac{\sigma_p^2(v\sigma_{t\text{-}flash})^2}{\sigma_p^2 + (v\sigma_{t\text{-}flash})^2} + \sigma_{jump}^2 \qquad (7)$$

$$\sigma_{tone}^2 = \frac{\sigma_p^2(v\sigma_{t\text{-}tone})^2}{\sigma_p^2 + (v\sigma_{t\text{-}tone})^2} + \sigma_{jump}^2 \qquad (8)$$

Figure 11 shows the spatial variability in the settings for each target speed, together with the variability that is predicted by equations 7 and 8. It is not at all obvious that the variability predicted by the model should even be close to the values measured in Experiments 7 and 8, because the overall level of variability is derived from the synchronization task (Experiment 1), with a small additional contribution from the task with the jumping targets (Experiment 4). From the variability in Figure 2 we got the impression that the synchronization task was completely unrelated to the localization task. Similarly, from the modest increase in spatial variability with the speed of the moving targets (Figure 6; reproduced as measured values in Figure 11) we got the impression that the variability must mainly be spatial. Nevertheless, basing our model on the temporal variability in the synchronization task gives values that are quite close to the measured ones. The implications of this finding are discussed in the following paragraphs.

The spatial variability only increases modestly with target speed, despite its temporal origin, because the compression toward the fovea also compresses the variability, and this compression is larger for faster targets. Admittedly, the dependence on target speed should have been larger (see Figure 11), but the overall prediction is very good considering that the predictions are made without considering any measured variability with moving targets.
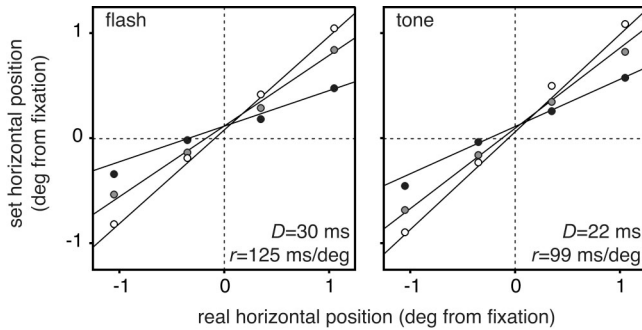

Figure 10. Fitting the model to the data. The symbols show the average set horizontal positions for each horizontal target position at the moment of the *flash* (Experiment 7) or *tone* (Experiment 8), for each of the three target speeds (darker symbols for faster targets). The set of lines (one for each target speed) is the best fit of equation 3.
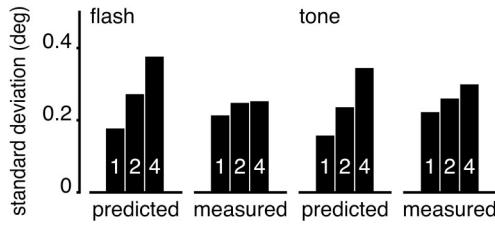
Figure 11. Predicted (equations 7 and 8) and measured (Experiments 7 and 8) spatial variability. The variability is shown separately for the flash and tone, and for the three target speeds (indicated within the bars in pixels per frame). Although the prediction is not perfect, the values are quite close.

Moreover, in our analysis we completely ignored several issues that should not simply be ignored but with which we do not yet know how to deal.

The most obvious example of such an issue is that our estimate of the variability in timing is presumably too low because subjects were allowed to see the stimuli as often as they liked in the synchronization task (Gengel & Hirsh, 1970). We do not know how our subjects combined information from repeated presentations or even how many presentations they considered when making each setting, so we cannot correct for this.

A second example of an issue that we ignored in our analysis of the variability is that we took the variability in timing from the synchronization task without accounting for the bias that we found in that task. If the bias were due to a difference between the neuronal processing time for detecting the flash and the tone, we should have found the same difference in our estimates of the temporal delay (*D*) in the localization tasks. A possible explanation for not having found this (as already mentioned) is that the synchronization task may be based on the relationship between the peaks of the neuronal responses, whereas the localization is based on the onsets of the responses. The relative timing of the peaks and onsets need not be the same. Perhaps the response to the flash starts at the same time but increases more gradually than that to the tone, so that it peaks about 50 ms later. If so, the variability in the timing of the peak may not provide a good estimate of the variability in the timing of the onset. Considering the larger rate of change at onset, an estimate of the accuracy that is based on the variability in the timing of the peak may be too high. However, this interpretation may be incorrect, and even if it is correct, we have no way to determine how the variability in onset time is related to the variability in the peak of the response.

Based on these issues, the comparison that is presented in Figure 11 cannot really be considered as a critical test of our model. However, it is encouraging to see that the values are at least of about the same magnitude as our model predicts. One more issue that may need some explanation is that we found a difference between the two delays. Since we attributed the delay (*D*) to the time needed to sample the moving target's position, one may expect this to be independent of the signal that indicates that it is time to start sampling. However, this obviously cannot be true, because the signal must be detected first, and differences in the time needed to detect the signal (relative to the time needed to register the moving target's position) will obviously also contribute to the delay. The values of *D* obtained by fitting the model

therefore imply that the tone that we used was detected 8 ms faster than the flash. The additional delay related to sampling is responsible for the general bias toward selecting a later position of the moving target.

A final issue is that the magnitude of the compression appears to depend on the position along the path. The compression is largest for positions near where the path passes closest to the fixation point (see separation between central points for fast targets in Figure 8) and is smaller when the positions are further apart along the path (as discussed in relation to the speeds used in Experiments 9 and 10). According to our model, the compression should be independent of the position along the path (see equation 3). Thus if our model is to accommodate all of our data, it has to be modified so that the influence of the prior is smaller if the evidence suggests that the target is farther from the point at which it passes the fixation point. There are several ways in which this could be achieved (e.g., by applying the model to a different space than the distance on the screen), but considering the limited data that we have concerning this issue, we will stick to our model as a reasonable approximation.

General Discussion

In the first part of this study we found an unexpectedly strong compression along the moving target's path. In the second and third parts of the study we established that this compression was primarily toward the position at which the target's path came closest to the fixation point. This led us to propose that people are biased toward believing that they are looking at what they see. This was a considerable sidestep from our original goal of testing the sampling hypothesis for the origin of flash-lag, but it turned out to be quite fruitful for developing a more detailed, quantitative model of the sampling hypothesis in the fourth part of the study. This obviously included additional factors to explain the bias toward the fovea. Since we think that uncertainty about the position plays an important role in both aspects (sampling and compression), we think it makes sense to combine them into one model in this way. The resulting model rests on three assumptions.

1. *Sampling.* You can only judge a moving target's position if the moment of interest is specified. We assume that people only start determining the position when an event indicates that it is time to do so. This sampling process takes time, so it introduces a localization bias in the direction of motion (see bias columns in Tables 1 through 3).

2. *Temporal resolution.* The limiting factor in determining the moving target's position is the temporal resolution of the underlying signals. Thus the spatial uncertainty increases with target speed (see Figures 6 and 11).

3. *Foveal prior.* People are biased toward localizing visual targets where they are looking. For static objects this bias is too weak to introduce noticeable errors, but for moving targets the temporal uncertainty can be large enough for the bias to no longer be negligible (see compression columns in Tables 2 and 3).

A model based on these assumptions gives a consistent description of most aspects of our data. In the next few paragraphs we will show that it is also consistent with many other findings.

Flash-lag

In our experiments, subjects had to indicate where the moving target was on the screen. In flash-lag experiments, subjects have to indicate where it is relative to the flash. The fact that we found a systematic bias in the direction of motion confirms that it is the moving target that is mislocalized. However, the value of the temporal delay (*D*) that we derived from our model is smaller than the value that is normally found in flash-lag experiments. According to our interpretation, this difference implies that the sampling process takes less time in our task. This could be because in our task subjects only needed to *detect* the flash or tone, whereas in flash-lag experiments they also have to *localize* the flash. Perhaps localizing the flash takes additional time.

On the other hand, the finding that there is no flash-lag if the subject's eyes are pursuing the moving target (Nijhawan, 2001) shows that the spatial uncertainty must be due to uncertainty in the retinal position rather than in the position in space, suggesting that the flash-lag effect results from a judgment of the retinal separation between the moving target and the flash (rather than from a judgment of the egocentric position of each; Brenner & Cornelissen, 2000). Perhaps localizing a single target in space, as subjects must have done in the present study, takes less time than determining the retinal separation between two targets.

Moreover, it is important to realize that the delay in our model is not a direct measure of the sampling time, because it also depends on the time that it takes for the different signals to reach the relevant parts of the brain. This is evident from the effects that the many factors that influence the magnitude of such delays have on the flash-lag effect: attention (Baldo, Kihara, Namba, and Klein, 2002), predictability (Brenner & Smeets, 2000), color and luminance (Chappell, Hine & Hardwick, 2002; Kerzel, 2003; Ogmen, Patel, Bedell, and Camuz, 2004), eccentricity (Kanai, Sheth & Shimojo, 2004), sound (Kerzel, 2003; Vroomen & de Gelder, 2004) and spatial configuration (Watanabe et al., 2001; Watanabe, 2004).

Our model readily explains why the flash-lag effect is larger for targets approaching the fovea than for ones moving away (van Beers, Wolpert & Haggard, 2001; Mateeff & Hohnsbein, 1988). For targets approaching the fovea, the bias toward the fovea and the delay in sampling the position of the moving target add up because they are in the same direction. For targets moving away from the fovea, the bias toward the fovea and the delay in sampling the position of the moving target cancel because the former is in the opposite direction than the motion.

Our model also explains why the most consistent evidence for a constant timing error is found when the target moves on a circular path around the fovea (as in Nijhawan, 1994). The circular path does not reduce the temporal uncertainty, but compression toward the fovea would not affect the perceived position on this path. Moreover, since the path itself is evident, we expect to find no compression toward the fovea, just as we found no compression orthogonal to the motion in the present study. Thus, for circular motion the spatial error that arises from a constant temporal error will be directly proportional to the target velocity.

It has previously been suggested that the mislocalization in conventional flash-lag experiments only depends on the target's motion after the flash (Brenner & Smeets, 2000; Whitney & Murakami, 1998; Whitney, Murakami & Cavanagh, 2000). If the speed, or direction of motion, changes near the time of the flash, the *average* localization error depends on the speed and direction of motion after the flash. However, we estimate the temporal uncertainty to be so large (also see Murakami 2001b) that even for the longer delays that are typically found in flash-lag experiments, the target's motion before the flash could influence the perceived position to some extent. One condition in which we could expect such an effect, and in which such an effect has indeed been found, is if the target is visible but static before the flash. In that case the tendency to see it at the starting position is larger than if it suddenly appears at that position at the time of motion onset (Chappell & Hine, 2004).

We have developed and discussed our model in relation to a moving target. Of course, the foveal bias is only relevant when determining a visually perceived position. However, the temporal uncertainty and the sampling delay can just as readily be used to account for errors in judging the values of other changing properties at the time of an unpredictable event (Bachmann & Poder, 2001; Murakami 2001a; Nijhawan & Kirschfeld, 2003; Sheth, Nijhawan & Shimojo, 2000).

Peri-saccadic Compression

Other experiments in which a compression of space has been reported are experiments in which subjects have to localize a target that is flashed near the time of a saccade. If the flash is presented near enough to the time of the saccade, people tend to see it at the position toward which the saccade is made, irrespective of where it really was (Ross, Morrone & Burr, 1997). The compression is predominantly in the direction of the saccade (i.e., horizontal for horizontal saccades; Morrone, Ross & Burr, 1997) and is absent if the experiment is conducted in complete darkness (Honda, 1993; Lappe, Awater & Krekelberg, 2000). Thus the compression may be a consequence of subjects localizing the flash with respect to the image that is shifting rapidly across the retina.

Extrapolating our model predictions to saccadic velocities would give rise to an extreme compression of space. However, compression is also reported during the periods before and after the saccade, during which there is no retinal motion. Presumably this is because the temporal uncertainty extends beyond the saccade, as is evident from the fact that targets flashed up to 100 ms before a saccade are mislocalized in the direction of the saccade (Mateeff, 1978; Matin, Matin & Pola, 1970; Schlag & Schlag-Rey, 1995, 2002). This is presumably inevitable if the uncertainty about the moment of the flash (in addition to uncertainty about the moment of the saccade) is as large as our study suggests.

Consequences for Our Daily Life

What implications does all this have for our everyday interactions with moving objects or with objects while we ourselves move? Normally, the kind of compression that we found is probably irrelevant because we keep our eyes on (i.e., pursue) an object with which we want to interact. The additional neuronal delay associated with sampling a position in response to a signal is probably also irrelevant because it is an artifact of the kind of task that we used: It results from having to explicitly specify the moment of interest. Normally, visual judgments of the target's position will be available continuously because—as mentioned in

the introduction—the moment is specified implicitly by the purpose of the interaction. Thus, although our results may help us to understand phenomena such as the flash-lag effect and perisaccadic compression, they have the rather disappointing implication that such phenomena are irrelevant for our everyday interactions with objects, rather than them having relevance for compensating for the neuromuscular delays associated with our actions. In terms of relevance for daily life, the most important aspect of our study is probably that we provide evidence that the temporal uncertainty about neuronal events is quite large (standard deviations in the order of 50 ms). This has strategic implications for the visual guidance of our actions because we can expect people to use strategies that minimize the effects of temporal uncertainty (see Brenner & Smeets, 2005).

Conclusions

We conclude that people's accuracy in localizing moving objects at an indicated moment is severely limited by their poor temporal resolution. When in doubt people are biased toward localizing objects where they are looking. This foveal bias is normally too weak to influence the positions at which things are seen, but its influence emerges if the task's temporal demands introduce a lot of spatial uncertainty. People try to circumvent their poor temporal resolution by reacting as quickly as possible to a relevant signal rather than trying to determine when it occurred. This usually leads them to localize a moving object at a position that it reaches slightly after the signal because they can only start estimating the position after the moment of interest is indicated.

References

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

Bachmann, T., Poder, E. (2001). Change in feature space is not necessary for the flash-lag effect. *Vision Research, 41,* 1103–1106.

Baldo, M. V. C., Kihara, A. H., Namba, J., Klein, S. A. (2002). Evidence for an attentional component of the perceptual misalignment between moving and flashed stimuli. *Perception, 31,* 17–30.

Bertelson, P., Aschersleben, G. (2003). Temporal ventriloquism: Cross-modal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. *International Journal of Psychophysiology, 50,* 147–155.

Brenner, E., Cornelissen, F. W. (2000). Separate simultaneous processing of egocentric and relative positions. *Vision Research, 40,* 2557–2563.

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

Brenner, E., Smeets JBJ (2005). Intercepting moving targets: Why the hand's path depends on the target's velocity. *Human Vision and Electronic Imaging X (Eds. BE Rogowitz, T. N. Pappas, SJ Daly) Proc. SPIE-IS&T Electronic Imaging, SPIE* 5666:374–384.

Chappell, M., Hine, T. J. (2004). Events before the flash do influence the flash-lag magnitude. *Vision Research, 44,* 235–239.

Chappell, M., Hine, T. J., Hardwick, D. (2002). The flash-lag effect and equiluminance. *Clin Experiment Ophthalmol, 30,* 213–216.

Ernst, M. O., Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433.

Gengel, R. W., Hirsh, I. J. (1970). Temporal order: The effect of single versus repeated presentations, practice, and verbal feedback. *Perception & Psychophysics, 7,* 209–211.

Hine, T. J., White, A. M., Chappell, M. (2003). Is there an auditory-visual flash-lag effect? *Clin Experiment Ophthalmol, 31,* 254–257.

Honda, H. (1993). Saccade-contingent displacement of the apparent position of visual stimuli flashed on a dimly illuminated structured background. *Vision Research, 33,* 709–716.

Ishii, M., Seekkuarachchi, H., Tamura, H., Tang, Z. (2004). 3D flash lag illusion. *Vision Research, 44,* 1981–1984.

Jaskowski, P. (1991). Perceived onset simultaneity of stimuli with unequal durations. *Perception, 20,* 715–726.

Kanai, R., Sheth, B. R., Shimojo, S. (2004). Stopping the motion and sleuthing the flash-lag effect: Spatial uncertainty is the key to perceptual mislocalization. *Vision Research, 44,* 2605–2619.

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

Kerzel, D. (2003). Asynchronous perception of motion and luminance change. *Psychol Research, 67,* 233–239.

Krekelberg, B., Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research, 39,* 2669–2679.

Lappe, M., Awater, H., Krekelberg, B. (2000). Postsaccadic visual references generate presaccadic compression of space. *Nature, 403,* 892–895.

Lappe, M., Krekelberg, B. (1998). The position of moving objects. *Perception, 27,* 1437–1443.

Lewald, J., Guski, R. (2004). Auditory-visual temporal integration as a function of distance: No compensation for sound-transmission time in human perception. *Neuroscience Letters, 357,* 119–122.

Mateeff, S. (1978). Saccadic eye movements and localization of visual stimuli. *Perception & Psychophysics, 24,* 215–224.

Mateeff, S., Hohnsbein, J. (1988). Perceptual latencies are shorter for motion towards the fovea than for motion away. *Vision Research, 28,* 711–719.

Matin, L., Matin, E., Pola, J. (1970). Visual perception of direction when voluntary saccades occur: II. Relation of visual direction of a fixation target extinguished before a saccade to a subsequent test flash presented before the saccade. *Perception & Psychophysics, 8,* 9–14.

Medendorp, W. P., Goltz, H. C., Vilis, T., Crawford, J. D. (2003). Gaze-centered updating of visual space in human parietal cortex. *Journal of Neuroscience, 23,* 6209–6214.

Morrone, M. C., Ross, J., Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *The Journal of Neuroscience, 17,* 7941–7953.

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

Murakami, I. (2001b). The flash-lag effect as a spatiotemporal correlation structure. *Journal of Vision, 1,* 126–136.

Musseler, J., van der Heijden, A. H., Mahmud, S. H., Deubel, H., Ertsey, S. (1999). Relative mislocalization of briefly presented stimuli in the retinal periphery. *Perception & Psychophysics, 61,* 1646–1661.

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

Nijhawan, R. (2001). The flash-lag phenomenon: Object motion and eye movements. *Perception, 30,* 263–282.

Nijhawan, R., Kirschfeld, K. (2003). Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: Evidence from motor flash-lag. *Current Biology, 13,* 749–753.

Nijhawan, R., Watanabe, K., Khurana, B., Shimojo, S. (2004). Compensation of neural delays in visual-motor behaviour: No evidence for shorter afferent delays for visual motion. *Visual Cognition, 11,* 275–298.

Ogmen, H., Patel, S. S., Bedell, H. E., Camuz, K. (2004). Differential latencies and the dynamics of the position computation process for moving targets, assessed with the flash-lag effect. *Vision Research, 44,* 2109–2128.

Oram, M. W., Xiao, D., Dritschel, B., Payne, K. R. (2002). The temporal resolution of neuralcodes: Does response latency have a unique role? *Phil Trans R. Soc Lond B, 357,* 987–1001.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences, 1,* 56–61.

Purushothaman, G., Patel, S. S., Bedell, H. E., Ogmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396,* 424.

Ross, J., Morrone, M. C., Burr, D. C. (1997). Compression of visual space before saccades. *Nature, 386,* 598–601.

Schlag, J., Schlag-Rey, M. (1995). Illusory localization of stimuli flashed in the dark before saccades. *Vision Research, 35,* 2347–2357.

Schlag, J., Schlag-Rey, M. (2002). Through the eye, slowly: Delays and localization errors in the visual system. *Nature reviews, Neuroscience, 3,* 191–200.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology, 79,* 3272–3278.

Sheth, B. R., Nijhawan, R., Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nature Neuroscience, 3,* 489–495.

Sheth, B. R., Shimojo, S. (2001). Compression of space in visual memory. *Vision Research, 41,* 329–341.

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 268,* 31–38.

van Beers, R. J., Sittig, A. C., Denier van der Gon, J. J. (1996). How humans combine simultaneous proprioceptive and visual position information. *Experimental Brain Research, 111,* 253–261.

van Beers, R. J., Wolpert, D. M., Haggard, P. (2001). Sensorimotor integration compensates for visual localization errors during smooth pursuit eye movements. *Journal of Neurophysiology, 85,* 1914–1922.

Vroomen, J., de Gelder, B. (2004). Temporal ventriloquism: Sound modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 513–518.

Wang, Y., Zhou, Y., Ma, Y., Leventhal AG (2004). Degradation of signal timing in cortical areas V1 and V2 of senescent monkeys. *Cerebral Cortex*

Watanabe, K. (2004). Visual grouping by motion precedes the relative localization between moving and flashed stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 504–512.

Watanabe, K., Nijhawan, R., Khurana, B., Shimojo, S. (2001). Perceptual organization of moving stimuli modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 879–894.

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

Whitney, D., Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1,* 656–657.

Whitney, D., Murakami, I., Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40,* 137–149.

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://watson.apa.org/notify/ and you will be notified by e-mail when issues of interest to you become available!