Vision Research 50 (2010) 1101–1108

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

How well can people judge when something happened?

Eli Brenner *, Jeroen B.J. Smeets

Faculty of Human Movement Sciences, VU University, Van der Boechorststraat 9, NL-1081 BT Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 3 November 2009
Received in revised form 4 March 2010

Keywords:
Temporal processing
Synchronization
Temporal order judgments
Colour change
Motion
Predictability

ABSTRACT

One way to estimate the temporal precision of vision is with judgments of synchrony or temporal order of visual events. We show that irrelevant motion disrupts the high temporal precision that can be found in such tasks when the two events occur close together, suggesting that the high precision is based on detecting illusory motion rather than on detecting time differences. We also show that temporal precision is not necessarily better when one can accurately anticipate the moments of the events. Finally, we illustrate that a limited resolution of determining the duration of an event imposes a fundamental problem in determining when the event happened. Our experimental estimates of how well people can explicitly judge when something happened are far too poor to account for human performance in various tasks that require temporal precision, such as interception, judging motion or aligning moving targets spatially.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The standard deviation in the timing of the moment of contact when intercepting moving virtual targets with a hand-held stylus is less than 20 ms (Brenner & Smeets, 2009). People can time the swing of a bat to hit a falling ball to within about 6 ms (standard deviation estimated from data provided in McLeod, McLaughlin, & Nimmo-Smith, 1985). In sports situations, top sportsmen may be even more precise (Regan, 1997). If people can be so precise in interception tasks, in which one must not only judge when the target will be where, but also arrive there at the right moment oneself, one may expect people to be even better at judging whether two visual events happened at the same time. However, standard deviations of between 40 and 50 ms have typically been reported for visual synchronicity (e.g. Virsu, Oksanen-Hennah, Vedenpää, Jaatinen, & Lahti-Nuuttila, 2008) and temporal order (e.g. Nava, Bottari, Zampini, & Pavani, 2008) judgments. Such poor temporal resolution appears to be incompatible with the above-mentioned performance in interception, even if one were to assume that the perceptual judgments are based on completely independent estimates of the moments of the two events (each with a standard deviation of about 30 ms), although obviously only the relative timing really matters for perceptual judgments, so common fluctuations are irrelevant and the appropriate standard deviation for a comparison with interception is likely to be larger. When judging a moving object's position at the time of a cue, the standard deviation in identifying the correct moment is also very poor (Brenner, van Beers, Rotman, & Smeets, 2006; Linares, Holcombe, & White, 2009).

Why are visual judgments so imprecise? Could something about the conditions used in the purely visual studies be responsible for the lack of precision? One issue that has been reported to be important for temporal order judgments is the separation between the targets. A standard deviation of about 5 ms was found when the two targets in question were very close together (estimated from the data for GW at a separation of 4 min of arc in Westheimer & McKee, 1977). This is about one tenth of the standard deviation for similar targets at larger separations. Similarly accurate temporal order judgments were found when a variety of line configurations were used as targets (Westheimer & McKee, 1977) and even for targets with different shapes presented at corresponding positions in the two eyes (Robinson, 1967). Westheimer (1983) showed that it is the separation that matters rather than the retinal eccentricity, although the optimal separation does depend on the eccentricity.

The resolution of (apparent) motion perception also depends on the separation between the signals, with the smallest time differences for small separations (Koenderink, van Doorn, & van de Grind, 1985). Allik and Kreegipuu (1998) and Victor and Conte (2002) have shown that judgments of temporal order are more accurate under conditions that give rise to apparent motion. The accuracy of temporal order judgments is poor (standard deviation of about 50 ms) for complex interleaved patterns (Eskes, Klein, Dove, Coolican, & Shore, 2007), despite the small spatial separations, possibly because the direction of apparent motion is ambiguous when surfaces are interleaved. If people achieve a high temporal resolution with small separations between two targets by responding to perceived motion rather than to when things

* Corresponding author.
E-mail address: e.brenner@fbw.vu.nl (E. Brenner).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.03.004

happened, then performance under such conditions should be disrupted by irrelevant motion. In our first experiment we examine whether this is the case.

Perhaps the difficulty in judging exactly when something happened is especially evident in typical synchronicity and temporal order tasks because one cannot accurately anticipate when the stimuli will occur. In temporal order and synchronicity judgment tasks the targets are usually flashes of light or some other abrupt changes. In interception tasks people can normally constantly see the moving object as it approaches the point of interception. If being able to continuously predict when the moments of interest will occur increases the temporal resolution, then making it easier to predict the moment of the change should improve judgments of synchrony in visual tasks as well. If people have a fundamental problem in judging the timing of events, then providing ample opportunity to anticipate the moments of interest should not help. In our second experiment we examine whether making it easier to anticipate when two events will take place allows people to synchronise them more precisely.

It is also possible that the human visual system simply does not have specialised pathways for evaluating when things happened (at least not to within a few milliseconds; see Battelli, Walsh, Pascual-Leone, & Cavanagh, 2008 for a review of evidence that specialised 'when' pathways do exist for much longer time scales). A possible reason for not having specialised pathways for determining (in retrospect) exactly when things happened, even when their occurrence is predictable, is that the information required to make such judgments is already lost within the retina. In analogy to metamerism in colour vision, two signals with different intensities and durations that give rise to the same pattern of retinal cone or ganglion cell responses will not be distinguishable. To what extent the patterns of responses have to be exactly identical for this to happen is not certain, but it is long known that it is difficult to distinguish short flashes of various durations that have the same total physical intensity (e.g. Kietzman & Sutton, 1968). This in itself does not necessarily limit the temporal resolution, but if the delay between the signal and the response is different for the two stimuli then the timing of the stimulus cannot be known without knowing what stimulus was presented. In our final experiment, we examine whether this could indeed be a fundamental problem for the visual system.

2. Methods

In the first two experiments the task was to synchronise changes in colour (Fig. 1). We used changes in colour because colour can be changed quite independently of other attributes. In the first experiment we examined how well people can synchronise changes in colour when the targets are close together, whether precision is reduced if one of the targets is moving, and if so how this depends on the moving target's speed. In the second experiment we examined whether precision is higher when the moment of the change in colour is predictable. In the third experiment we investigated whether there are fundamental difficulties in judging when something happened. To do so we asked subjects to *synchronize* and to *discriminate between* flashed targets of different durations and intensities.

To make completely sure to give our subjects the opportunity to perform as well as they possibly could, we allowed them to look wherever they liked and to see the stimuli as often as they liked before making a decision. Allowing subjects to base each of their judgments on several presentations undoubtedly improved the matches, so considering the set values as representing the resolution for single presentations clearly overestimates performance. Thus if subjects perform unexpectedly *well* in comparison with

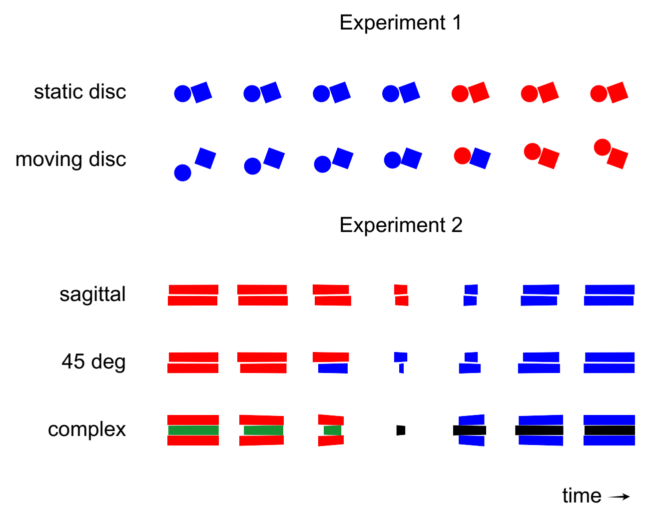

Fig. 1. Stimuli of the first two experiments. Schematic representations of seven frames (at 30-frame intervals) related to a single pair of changes in colour. In the upper example for Experiment 1, both targets are static and change colour at the same time. In the lower example the disc is moving upwards and slightly to the right and the square changes colour too late. In Experiment 2, rectangles rotate in depth in opposite directions. In the upper example the rectangles change colour simultaneously when they are both orthogonal to the screen. In the central example their colours change when another 45 deg of rotation will bring them into the sagittal plane. The upper rectangle is lagging behind the lower one and changes colour later. In the lower example there are three rectangles. The central rectangle changes colour slightly too early. The other two rectangles are not visible in the central frame because that is the moment of the change in colour (when the rectangle is orthogonal to the screen). For QuickTime movies of examples of such stimuli (with synchronous and asynchronous changes) see Supplementary material at doi:10.1016/j.visres.2010.03.004. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

previous studies, we will have to make sure that this is not just due to the repetition. The reason to nevertheless use this method is that it reduces the likelihood that subjects perform poorly for reasons that are unrelated to our question (for instance by not directing their gaze optimally on some trials or by missing occasional events as a result of blinks or lapses of attention).

2.1. Subjects

The two authors and four of our colleagues each took part in six sessions: two for Experiment 1 and four for Experiment 2. The two authors and seven colleagues (including three of the four who had taken part in Experiments 1 and 2) took part in the two sessions of Experiment 3. They all had ample experience with psychophysical experiments, but only the authors were aware of the hypotheses under study. The study is part of a programme that has been approved by the local ethical committee.

2.2. Equipment

All images were presented on a 47.8 by 30.4 cm CRT screen (1096 by 686 pixels; 160 Hz; 8 bits per gun) in a normally illuminated room. Anti-aliasing techniques were used to present targets at sub-pixel resolution. Colour and luminance were measured with a Minolta CS-100A Chroma Meter (Minolta Camera Co., Ltd., Japan). Subjects sat about 150 cm from the screen in Experiment 1, about 50 cm from the screen in Experiment 2, and about 70 cm from the screen in Experiment 3. They moved the computer mouse laterally to change the relative timing of the changes in colour (Experiments 1 and 2) or of the flashes (Experiment 3), and pressed the mouse button when they were content that the changes occurred simultaneously. In the detection task of Experiment 3 they pressed the '1' or '3' key of the keyboard to indicate which of three flashes looked different. Moving the mouse controlled the timing of the change in colour and of the flash onset in steps of 6 ms (1 frame), and controlled the orientation of the rotating object in Experiment 2 in steps of 0.05 deg (corresponding with 0.3 ms).

2.3. Stimuli

In *Experiment 1*, the stimuli were a 2 cm (0.8°) diameter disc and a square with sides of 2 cm (0.8°) on a white screen (CIE_{xyY} = [0.31, 0.34, 82 cd/m^2]). On alternating presentations the disc and square were either initially both red (CIE_{xyY} = [0.58, 0.36, 24 cd/m^2]) and turned blue (CIE_{xyY} = [0.22, 0.18, 24 cd/m^2]), or they were initially both blue and turned red. This procedure meant that the colour did not have to change between presentations (which could be confusing). The disc changed colour every second (1 Hz). Within each trial it did so at the same position on all presentations. The disc was either static at this position, or moved upwards parallel to the square's left edge so that it reached this position after between 500 and 800 ms (chosen at random for each presentation). The square never moved. Its colour changed from the same colour as the initial colour of the disc to the same colour as the final colour of the disc at some time between 200 ms before the disc changed colour and 200 ms after it did so (depending on the position of the computer mouse). After the change in colour the disc remained visible until a new presentation started, so if it had been moving it continued to do so at the same constant velocity for between 200 and 500 ms before jumping to its new starting position. A trial lasted until the subject pressed the mouse button to indicate that he or she was satisfied that the disc and square changed colour simultaneously.

The square had a fixed position and orientation for all presentations of each trial, but the position and orientation varied across trials. The square could be anywhere within the central 8.7 cm (10°) of the screen, and could be tilted (and the disc's path with it) by up to 30 deg in either direction (to avoid confusion we will use *deg* to indicate rotations of the targets and ° to indicate degrees of visual angle). The position at which the disc changed colour was determined at random for each trial from the 2 cm range for which the disc 'touches' the left edge of the square. Subjects manipulated the moment at which the square's colour changed by moving the computer mouse: the position of the mouse (relative to a random initial value for each trial) determined the temporal offset between the two changes in colour.

In the first session the subjects first synchronised 25 pairs of static targets, and then 25 pairs of targets when the disc was moving at three pixels per frame (about 8 °/s). In the second session the subjects again synchronised 25 pairs of targets when the disc was moving at three pixels per frame, and then another 25 pairs of targets (each) when the disc was moving at half, quarter, one eighth and one sixteenth of that speed.

In *Experiment 2* the stimuli were two or three simulated flat rectangular objects with different colours on their two surfaces. They rotated at a constant velocity of 160 deg/s around a vertical axis at their – and the screen's – centre. When there were two objects (36 by 7 cm; 36° by 7° when viewed frontally) they always rotated in opposite directions. When there were three objects, the top and bottom objects (36 by 7 cm) rotated together in a counter-clockwise direction (as viewed from above). The central object (36 by 9 cm; 36° by 10° when viewed frontally) rotated in the opposite direction. We used perspective projection and a small viewing distance so that the direction of motion was evident from the changing image shape despite the images being constructed for a single viewing point between the subjects' eyes (and being presented to both eyes without the appropriate binocular disparities).

When there were two objects we used the same colours as in Experiment 1, and the task was again to make the objects change between the same two colours at the same time. When there were three objects only the red colour was the same as in Experiment 1. In that case the top and bottom objects were either red (CIE_{xyY} = [0.58, 0.36, 24 cd/m^2]) or green (CIE_{xyY} = [0.29, 0.58, 57 cd/m^2]), and the central object was either blue (CIE_{xyY} = [0.17, 0.10, 10 cd/m^2]) or black (CIE_{xyY} = [0.40, 0.41, 4 cd/m^2]). The task was to synchronize the changes between blue and black with those between red and green.

The change in colour could take place when the objects' simulated surfaces were aligned with the screen (frontal), slanted by 45 deg with respect to the screen, or orthogonal to the screen (in the sagittal plane). When the change took place at the moment that the objects' surfaces were orthogonal to the screen the stimulus was simply a simulation of thin rotating rectangular objects with different colours on their two sides. The colour depended on which side was visible and the change in colour coincided with the moment that the object's size was minimal. In the other cases the surfaces appeared to change colour.

The position of the computer mouse determined the phase of the upper of the two objects, or of the central of the three objects, relative to that of the other object or two objects. This determined the relative timing of the changes in colour because the objects changed colour when they reached a given angle. Subjects could change the angular difference between the surfaces by 30 deg in either direction by moving the mouse, so the variable object could change colour at any moment between188 ms before and 188 ms after the other object(s) did so.

One way to determine whether the ability to predict when the object's colours will change affects the accuracy with which the changes can be synchronised is by comparing changes when the objects have different orientations. The extent to which motion signals arising from asynchronous changes in colour are masked is also likely to depend on the orientation. We expect it to be more difficult to predict precisely when the surfaces will reach 45 deg (in opposite directions so that the angle between them is 90 deg) than to predict when they will both be orthogonal to the screen. There is likely to be less masking when the speed of the motion on the screen is lowest, which is when the surface is in the plane of the screen (frontal).

We also varied the predictability of the changes in colour more directly by allowing the angle at which the colour changed to vary by up to 5 deg or up to 15 deg in either direction (a temporal range of 63 or 188 ms). For each change, a new random orientation was chosen from within this range for both or for all three rectangles, so the time of the next change could not be anticipated as accurately as without such variations, but the relative timing of the changes in colour did not change (unless the subject moved the mouse). Thus, if the mouse was not moved and the change on a given presentation occurred when the objects were oriented 4 deg further in the direction of motion than on the previous presentation, then both objects changed colour 25 ms after they reached the orientation at which the colour changed on the previous presentation, but their relative timing was unaffected.

Each subject took part in four sessions. In the first session there were three objects. The change always took place exactly when the objects were in the sagittal plane. We pointed out to the subjects that the colour changed when the simulated object was orthogonal to the screen, so that they could also synchronise the moments that the image sizes were minimal, rather than when the colours changed, because the two obviously happened at the same time. This is the condition in which the timing is most predictable. In the second and third sessions there were two objects. In the second session the change took place when the objects were either in the frontal plane, at an angle of 45 deg with the frontal plane, or at a random angle between 30 and 60 deg (45 ± 15 deg) with the frontal plane. In the third session the change took place when the objects were in the sagittal plane, at a random angle between −5 and 5 deg from the sagittal plane, or at a random angle between −15 and 15 deg from the sagittal plane. The fourth session was identical to the third, except that there were three objects (that changed between different colours). There were 25 trials for each condition within each session (except session 2 in which there were accidentally 33 trials for each condition), and the trials within each session were presented in random order.

Experiment 3 consisted of two parts. In each trial of the first part, subjects saw three grey 4.4 cm (3.6°) diameter discs presented sequentially (at 500 ms intervals) on an 82 cd/m^2 grey background. Two were either 58 cd/m^2 (*low contrast*) or 10 cd/m^2 (*high contrast*) discs that lasted for one frame. The third, that was always presented either first or last, had some duration between 1 and 12 frames, and a luminance that ensured that the average luminance across the duration of the flash would be the same as that of a one-frame flash and the background at the same location if measured during the same time interval. Thus, for instance, a high contrast three-frame disc's luminance would be 58 cd/m^2 on each of the three frames (the average of the 10 cd/m^2 of the 1-frame target and two frames of the 82 cd/m^2 background luminance). The subject's task was to identify whether the first or last flash was different from the other two (by pressing the corresponding key on the keyboard). There were 20 trials for each of the 12 flash durations. In the second part subjects had to make two discs appear to flash at the same time. The targets were identical to those that subjects had to distinguish between in the first part. Both were presented at 2 Hz. There was a 9.7 cm (7.1°) horizontal separation between them. Subjects manipulated the relative timing of the two flashes by moving the computer mouse. The flash on the left always lasted for one frame, while the duration of that on the right varied across trials (but was the same for all presentations within each trial). There were nine trials for each of the 12 flash durations.

2.4. Analysis

For Experiments 1 and 2, we calculated the standard deviation in the matches for each subject in each condition after removing any trial for which the set moment of synchrony was more than four standard deviations from the average (when the mean and standard deviation was determined without that trial). We report the averages of these standard deviations with the associated standard errors across subjects. For the first part of Experiment 3, we determined the percentage of trials in which each subject correctly identified whether the first or last flash was different (for each flash duration). For the second part, we determined the median matched onset asynchrony between the two flashes for each subject and flash duration. For these measures we also report averages with the associated standard errors across subjects.

3. Results

3.1. Experiments 1 and 2

On average a setting in the first experiment took 25 s (20 s for static targets and 25 s for moving targets in the first session; 21, 27, 29, 31 and 28 s for targets with increasing speeds in the second session). Those in the second experiment took 24 s (no effect of number of objects but about 23 s when there was little or no random variability in when the change occurred, and 27 s when the timing of the changes was most variable). Altogether five matches were removed in Experiment 1 and four in Experiment 2. In four of these cases the removed trials corresponded with ones for which subjects had spontaneously reported after the session having accidentally pressed the mouse button. In the first experiment subjects had a slight (but significant) tendency to set the adjustable change to occur too early. This tendency increased with the disc's velocity from zero to almost 11 ms. A tendency to set the adjustable change to occur too early was also found in the second experiment (about 6 ms irrespective of the number of objects, the added variability in timing and the angle at which the change took place).

Fig. 2A summarises the standard deviations of Experiment 1. Open symbols show the data of the first session and closed symbols those of the second. It is evident that subjects matched the timing of the changes in colour less precisely when the disc was moving, despite the changes always having occurred when the targets were adjacent to each other. For the fastest targets there was even a tendency (not shown) to synchronize the change in the colour of the square with the disc reaching a fixed position (the disc's position – rather than its colour – was given a weight of up to 40%, as judged from the slope of the relationship between the set asynchrony for the colour changes and the position along the 2 cm range at which the disc's colour changed).

Fig. 2B shows the standard deviations of Experiment 2. It is evident from the data that making the moment of the change in colour more predictable does not make the matches more precise: the standard deviation in the settings did not increase when the orientation at the time of the change was varied. The largest standard deviation was found when three objects changed between two colours and two different colours at the moments that they were in the sagittal plane, which is the condition in which the occurrence of the changes was most predictable. Subjects were slightly more precise if the change in colour occurred when more of the surface was visible and the motion on the screen was slower (i.e. orientations closer to frontal) and slightly more precise when there were only two objects and they had the same colours (rather than three with two different sets of colours), but these differences are quite modest. A comparison of the two identical conditions from the first and last session (red and black discs) shows that the subjects' performance improved with practice. However even after synchronizing hundreds of changes in colour the standard deviation in this condition was well above 30 ms.

3.2. Experiment 3

Fig. 3 shows the results of Experiment 3. A difference in duration of tens of ms was required for the longer flash to be distinguished reliably from the single-frame flash, especially at low contrast. The values in Fig. 3A probably overestimate people's ability to judge differences in flash duration, because a larger total intensity is needed for a longer flash to look equally bright (Graham & Kemp, 1938), so differences in perceived luminance may also have revealed which flash was different. The first frame of a longer, dimmer flash had to be presented earlier for the two disks to appear to have been presented at the same time (Fig. 3B). The standard deviations in the individual subjects' matches (for given contrast and flash durations) were about 19 ms (increasing slightly with the difference in flash duration).

4. Discussion

4.1. Experiment 1

It is clear from Experiment 1 that irrelevant motion makes synchronizing changes in colour less precise. Even very slow motion had quite a clear effect. For the lowest velocity, the disc only moved a bit more than half the length of the side of the square during each 1 s presentation. The second lowest velocity (1 °/s) is equivalent to someone walking by at a distance of about 80 m. Since we allowed our subjects to see the pairs of changes as often

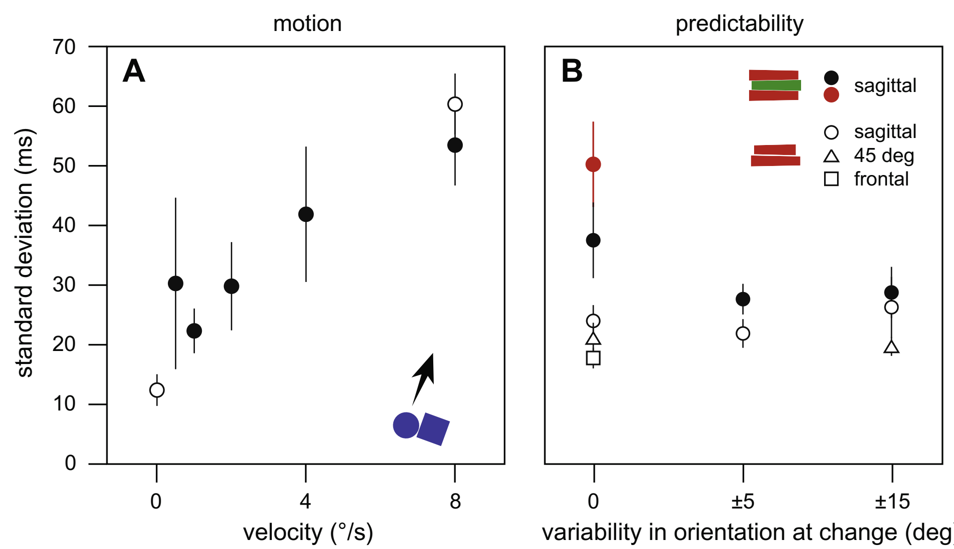

Fig. 2. Standard deviations in subjects' matches of the moments the colours changed. Each symbol shows the mean value for the six subjects (with the associated standard error). Open and solid symbols in (A) represent conditions of the first and second session of Experiment 1. In (B) the red disc shows the data of the first session, the triangles and square those of the second session, the circles those of the third session and the black discs those of the last session. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

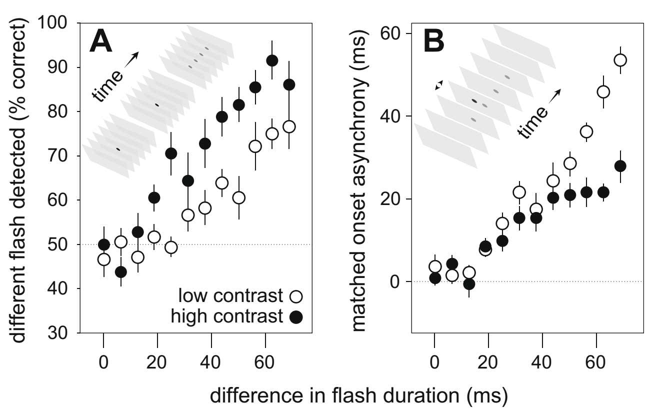

Fig. 3. Performance in Experiment 3. Percentage of trials in which subjects correctly identified whether the first or last flash was different (A) and how much after a flash of the indicated duration a one-frame flash had to be presented (median matched value) for the two to appear to be synchronous (B). Each symbol shows the mean value for the nine subjects (with the associated standard error). The insets in the panels' top left corners illustrate the sequence of frames with a 4-frame flash.

as they liked, rather than only once, a direct comparison with the values mentioned in the introduction is not completely appropriate. Nevertheless, the poorest value that we found in Experiment 1 is similar to the values that others found when judging synchrony with large separations (Nava et al., 2008; Virsu, Oksanen-Hennah, Vedenpää, Jaatinen, & Lahti-Nuuttila, 2008; Westheimer & McKee, 1977). Our average value for static targets is not quite as precise as that reported for optimal conditions by Westheimer and McKee (1977), but our best subject matched colour changes as well as their two subjects matched luminance changes (standard deviation of about 4 ms).

When the disc was moving fast, subjects partly relied on when it was aligned with the square for their synchronicity judgments, although the disc was at different positions relative to the square when it changed colour on different trials. This indicates that the ability to judge whether the changes happened at the same time is quite poor, because otherwise subjects would not revert to using alternative information. Using incorrect information obviously will not give rise to very good performance, but apparent motion as a result of the changes in colour provides reliable alternative information that is presumably detected very accurately by specialised mechanisms.

4.2. Experiment 2

In Experiment 2, the fact that the standard deviation was largest for the condition in which the moment of the change in colour was easiest to anticipate shows that predictability is not a major factor in determining the precision with which changes in colour can be synchronised. Testing this directly by manipulating the variability in the objects' orientations at the time of the change (across the individual changes within a trial) confirms that predictability makes very little difference (also see Experiment 3 in Linares, Holcombe, & White, 2009).

The differences in performance between the conditions in Experiment 2 are consistent with performance depending on how readily apparent motion (between the two changes in colour) can reveal which target appeared first. This is true both for the influence of the number of objects and for the influence of their orientations at the time of the change in colour. When the targets are not properly synchronised, subjects will see abrupt apparent motion in a single vertical direction (beside the continuous apparent motion of the rotating surfaces) if there are two objects, but apparent motion in opposite vertical directions for the upper and lower pair of three objects. It is reasonable to expect the latter motion to be more difficult to detect and interpret. Moreover, apparent motion may be more difficult to detect and interpret when different objects switch between different colours (and luminances) than when both objects change in an identical manner. Systematic differences between the latencies for detecting the four changes in colour (as demonstrated for targets with different luminances and durations in Experiment 3) could also contribute to the larger standard deviations in the conditions with three objects.

The standard deviations in Experiment 2 were smallest when the objects changed colour when they were oriented in the frontal plane. This is consistent with apparent motion determining performance because this is the orientation at which the objects' images were largest and their edges were moving slowest. That the standard deviations in the last session (black discs in Fig. 2B) were lower with added variability in the orientation at the moment of the change supports this explanation, because the added variability means that the change did not happen exactly when the object's size (and therefore the apparent motion signal) was minimal and the retinal image speed (and therefore the masking of apparent motion) maximal.

4.2.1. More than only colour changes and apparent motion

Changing the colour at the moment that the image size is minimal obviously has the disadvantage that the change in colour itself is less visible. A new frame was presented every 6 ms. (The resolution of the set angles was not limited to discrete frames, because moving the mouse changed the angle not the time of the change, but of course the presentations did occur in discrete frames.) Within one frame, the near edge of the rotating object moved about 3 mm, which corresponds with about half a degree of visual angle. Having to synchronise colour changes when the objects' images are so small may be responsible for the poor performance in the sagittal condition in which there was no additional variability. However, in this condition the subjects did not have to judge when the colour changed because they could also have synchronised the moments that the objects' sizes were minimal, so it is not evident that performance should be poor for this condition.

The reduction in the standard deviation with practice that can be observed in the data of the condition that was performed in both the first and the last session of Experiment 2 (the red and black discs at zero in Fig. 2B) could be due to learning to pick up suitable movement artefacts. However, several subjects reported having learnt not to just press the mouse button once the colours appeared to change simultaneously, but to look for the centre of the range of mouse positions for which this was so. Switching to such a strategy may have improved their performance. The fact that subjects used the time at which the disc crossed the square to estimate when its colour changed in Experiment 1 shows that subjects considered other information than only the changes in colour. That they only used it when the disc moved fast in Experiment 1 is logical if one considers the extent to which using such information could improve performance (rather than only adding variability). The error of up to 1 cm that one makes if one assumes that the disc changes colour when it is aligned with the centre of the square corresponds with a time interval of 50 ms at the highest speed. At lower velocities the same spatial error obviously corresponds with a longer time. In Experiment 2, also considering the orientation (and image size) would always improve performance, because the colour changed when the object reached a specific orientation so the two are perfectly correlated.

An alternative interpretation of the influence of motion on synchronicity judgments is that the transient retinal stimulation caused by the motion makes it more difficult to identify the transients that are caused by the changes in colour (in analogy to the argumentation in Terao, Watanabe, Yagi, & Nishida, 2008). This could account for the influence of velocity in Experiment 1 (Fig. 2A), and for the reasonably small standard deviations when the change in colour took place when the rectangles were oriented in the frontal plane so that the motion and associated transients were minimized and the transient caused by the change in colour maximized (square in Fig. 2B). It is not consistent with the almost as small standard deviations when two rectangles changed colour as they passed the sagittal plane, or near the time they did so, because at that time the retinal motion is even faster than the fastest velocity of Experiment 1 and the surface that changes colour is small and changing in size as well as colour (circles in Fig. 2B). In that case subjects may base their judgments on the surfaces' orientations rather than their colours, but that would not explain why performance is worse with three objects (discs in Fig. 2B). Moreover, this interpretation of the current data would not explain why temporal order judgments depend on the separation between the objects (Allik & Kreegipuu, 1998; Westheimer, 1983; Westheimer & McKee, 1977). If it were true, this interpretation would imply that we cannot know precisely when something happened if other things are happening nearby at about the same time.

4.2.2. Why such poor performance?

Considering all the above reasoning we are inclined to take the worst value that we found in Experiment 2 to be our best estimate of the temporal resolution of judging when something happened. We have described many reasons why people might have performed better in various tasks that involve timing than they can judge when changes happen. For instance, they probably used apparent motion to improve their performance in some conditions of our experiment. We can think of no reason for them to systematically perform worse than their temporal resolution of judging when something happened allows. Furthermore, even the many standard deviations between 20 and 30 ms (corresponding with individual judgments with standard deviations between 14 and 21 ms) cannot account for performance in interception, especially if one considers that this resolution is achieved after multiple presentations. So why are synchronisation judgments so imprecise in comparison with performance in interception?

One possibility is that like motion perception, the judgments that are required for interception have access to signals with a higher temporal resolution than explicit judgments of temporal order or synchrony. Judging motion is not the only task for which human performance is clearly based on a better temporal resolution than what we found for explicit temporal judgements. People can align two bars moving together at the same velocity very precisely. For instance, Chung, Levi, and Bedell (1996) determined alignment thresholds for bars moving laterally at various velocities. Our crude analysis of their data shows that it is consistent with a spatial uncertainty σ_s of 12 arc sec and an independent temporal uncertainty σ_t of 5 ms, that are combined into an overall spatial uncertainty $\sigma_{\text{overall}} = \sqrt{\sigma_s^2 + (v\sigma_t)^2}$, where v is the velocity at which the bars are moving. Similar data for alignment in depth reported by Ramamurthy, Bedell, and Patel (2005) (vertical separation of 5 min arc) is consistent with a spatial uncertainty σ_s of 20 arc sec and an independent temporal uncertainty σ_t of 3 ms. Thus the temporal resolution is also an order of magnitude higher when comparing time-varying spatial signals than when making temporal order or synchrony judgments.

The substantial differences in temporal resolution between judgments in diverse tasks suggest that a high temporal resolution can only be achieved with specialised mechanisms (see overview in Holcombe, 2009). Perhaps the human brain does not have specialised mechanisms for reliably detecting synchrony (other than specialised mechanisms for judging asynchrony in the form of motion) because processing times differ by tens of ms between attributes (e.g. Schmolesky et al., 1998; Veerman, Brenner, & Smeets, 2008), and even between different cues for the same attribute (van Mierlo, Louw, Smeets, & Brenner, 2009), and because factors such as contrast, luminance and eccentricity all influence the latencies (e.g. Ogmen, Patel, Bedell, & Camuz, 2004; Prestrude, 1971; Roufs, 1963). It may therefore often be more important to tolerate differences in latency than to detect them (van Mierlo, Brenner, & Smeets, 2007).

4.3. Experiment 3

The first part of Experiment 3 confirms that it is difficult to distinguish flashes of various durations that have the same total physical intensity (Kietzman & Sutton, 1968). This is undoubtedly a result of the low-pass temporal characteristics of early visual processing (see Burkhardt, Fahey, & Sikora, 2007; Dunn, Lankheet, & Rieke, 2007; Lankheet, Molenaar, & van de Grind, 1989; Scheich & Korn, 1971; van Hateren, 2007). The response to a shorter, brighter flash starts sooner after flash onset, and lasts longer after the flash has terminated. If the low-pass-filtered responses to flashes of different durations are indistinguishable, then there can be no question about how to align them. Aligning the overall responses (i.e. maximising the overlap between the responses) is probably responsible for the onset asynchronies in the second part of Experiment 3. Although the responses are indistinguishable, the response to the shorter, brighter flash starts sooner after the true flash onset, so the true onset of the brighter flash has to be later than that of the longer, dimmer flash for the two responses to occur at the same time, and therefore for the two flashes to appear to occur simultaneously (Fig. 3B). Similar asynchronies between flash onsets are found for two flashes with longer – but different – durations (Jaskowski, 1991) and for flashes with the same duration (5 ms) but different intensities (Allik & Kreegipuu, 1998). The perceived duration of the flash also depends on various stimulus parameters (Terao et al., 2008).

If an inability to account for how miscellaneous stimulus parameters modify the latency of responses is responsible for the brain not having specialised mechanisms for explicit temporal judgments, it is not unreasonable to find specialised mechanisms with much higher temporal resolution for motion judgments (and spatial alignment of moving stimuli), because for moving stimuli it is reasonable to assume that the parameters are identical at consecutive moments, so it makes sense to compare low-pass-filtered responses across short periods of time (and across short separations in space). Allik and Kreegipuu (1998) have shown that temporal order judgments are less sensitive to differences in luminance between the stimuli involved if the stimuli give rise to an impression of motion, so there may even be specific processing within specialised motion mechanisms to make motion detection less sensitive to variations in contrast.

4.3.1. A problem with using short flashes

Apart from contributing to a low temporal precision, uncertainty about flash durations may also give rise to systematic errors, because flashes may generally be considered to have lasted longer than their true, extremely short duration. In many studies extremely short flashes of light are used to precisely specify the moment of interest. Usually the obtained results are consistent with flashes appearing to have occurred later than that moment. A target that is flashed exactly as a moving object passes, appears to be flashed when the object has already passed (Nijhawan, 1994). The same happens if the object does not move smoothly but changes gradually in colour, luminance or entropy (Sheth, Nijhawan, & Shimojo, 2000) or jumps to new positions (Brenner, Mamassian, & Smeets, 2008; Murakami, 2001), and if it is not an object that is moving but the eyes (Brenner, Smeets, & van den Berg, 2001; Matin, Matin, & Pola, 1970; Rotman, Brenner, & Smeets, 2004; Schlag & Schlag-Rey, 2002). It being impossible for subjects to judge the duration of the flash, whereas the data are interpreted in terms of the true, extremely short duration, may be responsible for various reported flash-related errors (also see Krekelberg & Lappe, 2000; Pola, 2004). A longer response to visual stimulation by brighter stimuli (considering the adaptation of the eye; Bowen, Pola, & Matin, 1974; Nisly & Wasserman, 1989) may even explain why a short dim flash is sometimes judged to have occurred before a brighter one when they were actually presented at the same time (Bachmann, Põder, & Luiga, 2004).

4.3.2. Concluding remarks

The arguments in the preceding paragraphs imply that unless one knows a lot about the stimulus, it is impossible to be very precise about when it occurred. The results of Experiment 2 suggest that one cannot even be very precise about when things happened when one could theoretically know enough about the stimulus (including being able to predict when the change will occur). The human visual system probably does not have special sophisticated mechanisms for evaluating synchrony (correcting for the many delays), because such mechanisms would seldom be useful in daily life. For other tasks the visual system probably does have special mechanisms. When evaluating the position of a moving target one might rely on the initial response at each position, rather than on the peak in a low-pass-filtered response, which is likely to be less variable. When aligning two moving lines, a task for which the temporal resolution is relatively high (Chung et al., 1996), some delays are not compensated for: a line with a lower contrast and therefore a longer delay is perceived to lag behind one with a higher contrast (Hess, 1904; White, Linares, & Holcombe, 2008). Thus the mechanism involved has a high temporal precision but tolerates systematic errors.

In interception tasks with repeated trials one could compensate for systematic errors that arise from not taking the contrast into account by adapting ones actions on the basis of feedback on previous trials (see de Lussanet, Smeets, & Brenner, 2001). Since one anyway needs to compensate for neuromuscular delays by predicting where the target will be some time in the future, adjusting this time on the basis of recent experience with similar targets could help circumvent the need to know all the different delays. Alternatively, one may avoid many of the above-mentioned issues by pursuing the target with the eyes, so that the target's retinal position hardly changes, and eye orientation signals provide information about its position and motion. Indeed, subjects normally pursue targets with their eyes if they intend to intercept them (Brenner & Smeets, 2009).

Relying on initial responses or eye orientation signals to improve the temporal precision are just speculations. The present study cannot explain how the high temporal precision in interception is achieved; it only suggests that interception must be based on mechanisms that do not require explicit judgments of when things happened, because the latter is not known even nearly accurately enough.

Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2010.03.004.

References

Allik, J., & Kreegipuu, K. (1998). Multiple visual latency. *Psychological Science, 9*, 135–138.

Bachmann, T., Põder, E., & Luiga, I. (2004). Illusory reversal of temporal order: The bias to report a dimmer stimulus as the first. *Vision Research, 44*, 241–246.

Battelli, L., Walsh, V., Pascual-Leone, A., & Cavanagh, P. (2008). The 'when' parietal pathway explored by lesion studies. *Current Opinion Neurobiology, 18*, 120–126.

Bowen, R. W., Pola, J., & Matin, L. (1974). Visual persistence: Effects of flash luminance, duration and energy. *Vision Research, 14*, 295–303.

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), 1–10.

Brenner, E., & Smeets, J. B. J. (2009). Sources of variability in interceptive movements. *Experimental Brain Research, 195*, 117–133.

Brenner, E., Smeets, J. B. J., & van den Berg, A. V. (2001). Smooth eye movements and spatial localisation. *Vision Research, 41*, 2253–2259.

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

Burkhardt, D. A., Fahey, P. K., & Sikora, M. A. (2007). Retinal bipolar cells: Temporal filtering of signals from cone photoreceptors. *Visual Neuroscience, 24*, 765–774.

Chung, S. T., Levi, D. M., & Bedell, H. E. (1996). Vernier in motion: What accounts for the threshold elevation? *Vision Research, 36*, 2395–2410.
de Lussanet, M. H. E., Smeets, J. B. J., & Brenner, E. (2001). The effect of expectations on hitting moving targets: Influence of the preceding target's speed. *Experimental Brain Research, 137*, 246–248.
Dunn, F. A., Lankheet, M. J., & Rieke, F. (2007). Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature, 449*, 603–606.
Eskes, G. A., Klein, R. M., Dove, M. B., Coolican, J., & Shore, D. I. (2007). Comparing temporal order judgments and choice reaction time tasks as indices of exogenous spatial cuing. *Journal of Neuroscience Methods, 166*, 259–265.
Graham, C. H., & Kemp, E. H. (1938). Brightness discrimination as a function of the duration of the increment in intensity. *The Journal of General Physiology, 21*, 635–650.
Hess, C. (1904). Untersuchungen über den Erregungsvorgang im Sehorgan bei kurz- und bei längerdauernder Reizung. *Pflügers Archiv – European Journal of Physiology, 101*, 226–262.
Holcombe, A. O. (2009). Seeing slow and seeing fast: Two limits on perception. *Trends in Cognitive Sciences, 13*, 216–221.
Jaskowski, P. (1991). Perceived onset simultaneity of stimuli with unequal durations. *Perception, 20*, 715–726.
Kietzman, M. L., & Sutton, S. (1968). The interpretation of two-pulse measures of temporal resolution in vision. *Vision Research, 8*, 287–302.
Koenderink, J. J., van Doorn, A. J., & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America, A2*, 252–259.
Krekelberg, B., & Lappe, M. (2000). A model of the perceived relative positions of moving objects based upon a slow averaging process. *Vision Research, 40*, 201–215.
Lankheet, M. J. M., Molenaar, J., & van de Grind, W. A. (1989). The spike generating mechanism of cat retinal ganglion cells. *Vision Research, 29*, 505–517.
Linares, D., Holcombe, A. O., & White, A. L. (2009). Where is the moving object now? Judgments of instantaneous position show poor temporal precision (SD = 70 ms). *Journal of Vision, 9*(13), 1–14.
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 and Psychophysics, 8*, 9–14.
McLeod, P., McLaughlin, C., & Nimmo-Smith, I. (1985). Information encapsulation and automaticity: Evidence from the visual control of finely timed actions. In Marin Posner (Ed.), *Attention and performance XI* (pp. 391–406). Hillsdale: Lawrence Erlbaum Associates.
Murakami, I. (2001). A flash-lag effect in random motion. *Vision Research, 41*, 3101–3119.
Nava, E., Bottari, D., Zampini, M., & Pavani, F. (2008). Visual temporal order judgment in profoundly deaf individuals. *Experimental Brain Research, 190*, 179–188.
Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*, 256–257.
Nisly, S. J., & Wasserman, G. S. (1989). Intensity dependence of perceived duration: Data, theories, and neural integration. *Psychological Bulletin, 106*, 483–496.
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.
Pola, J. (2004). Models of the mechanism underlying perceived location of a perisaccadic flash. *Vision Research, 44*, 2799–2813.
Prestrude, A. M. (1971). Visual latencies at photopic levels of retinal illuminance. *Vision Research, 11*, 351–361.
Ramamurthy, M., Bedell, H. E., & Patel, S. S. (2005). Stereothresholds for moving line stimuli for a range of velocities. *Vision Research, 45*, 789–799.
Regan, D. (1997). Visual factors in hitting and catching. *Journal of Sports Sciences, 15*, 533–558.
Robinson, D. N. (1967). Visual discrimination of temporal order. *Science, 156*, 1263–1264.
Rotman, G., Brenner, E., & Smeets, J. B. J. (2004). Mislocalization of targets flashed during smooth pursuit depends on the change in gaze direction after the flash. *Journal of Vision, 4*, 564–574.
Roufs, J. A. J. (1963). Perception lag as a function of stimulus luminance. *Vision Research, 3*, 81–91.
Scheich, H., & Korn, A. (1971). Timing properties and temporal summation in the retina. *Pflugers Archiv, 327*, 16–36.
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., et al. (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.
Terao, M., Watanabe, J., Yagi, A., & Nishida, S. (2008). Reduction of stimulus visibility compresses apparent time intervals. *Nature Neuroscience, 11*, 541–542.
van Hateren, J. H. (2007). A model of spatiotemporal signal processing by primate cones and horizontal cells. *Journal of Vision, 7*(3), 1–19.
van Mierlo, C. M., Brenner, E., & Smeets, J. B. J. (2007). Temporal aspects of cue combination. *Journal of Vision, 7*(7), 1–11.
van Mierlo, C. M., Louw, S., Smeets, J. B. J., & Brenner, E. (2009). Slant cues are processed with different latencies for the online control of movement. *Journal of Vision, 9*(3), 1–8.
Veerman, M. M., Brenner, E., & Smeets, J. B. J. (2008). The latency for correcting a movement depends on the visual attribute that defines the target. *Experimental Brain Research, 187*, 219–228.
Victor, J. D., & Conte, M. M. (2002). Temporal phase discrimination depends critically on separation. *Vision Research, 42*, 2063–2071.
Virsu, V., Oksanen-Hennah, H., Vedenpää, A., Jaatinen, P., & Lahti-Nuuttila, P. (2008). Simultaneity learning in vision, audition, tactile sense and their cross–modal combinations. *Experimental Brain Research, 186*, 525–537.
Westheimer, G. (1983). Temporal order detection for foveal and peripheral visual stimuli. *Vision Research, 23*, 759–763.
Westheimer, G., & McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research, 17*, 887–892.
White, A. L., Linares, D., & Holcombe, A. O. (2008). Visuomotor timing compensates for changes in perceptual latency. *Current Biology, 18*, R951–953.