Predictive eye movements in natural vision

Mary M. Hayhoe,
Center for Perceptual Systems, University of Texas at Austin, 1 University Station, #A8000, Austin, TX 78712-0187, USA

Travis McKinney,
Center for Perceptual Systems, University of Texas at Austin, 1 University Station, #A8000, Austin, TX 78712-0187, USA

Kelly Chajka, and
School of Optometry, SUNY, New York, NY, USA

Jeff B. Pelz
Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, USA

Mary M. Hayhoe: mary@mail.cps.utexas.edu

Abstract

In the natural world, the brain must handle inherent delays in visual processing. This is a problem particularly during dynamic tasks. A possible solution to visuo-motor delays is prediction of a future state of the environment based on the current state and properties of the environment learned from experience. Prediction is well known to occur in both saccades and pursuit movements and is likely to depend on some kind of internal visual model as the basis for this prediction. However, most evidence comes from controlled laboratory studies using simple paradigms. In this study, we examine eye movements made in the context of demanding natural behavior, while playing squash. We show that prediction is a pervasive component of gaze behavior in this context. We show in addition that these predictive movements are extraordinarily precise and operate continuously in time across multiple trajectories and multiple movements. This suggests that prediction is based on complex dynamic visual models of the way that balls move, accumulated over extensive experience. Since eye, head, arm, and body movements all co-occur, it seems likely that a common internal model of predicted visual state is shared by different effectors to allow flexible coordination patterns. It is generally agreed that internal models are responsible for predicting future sensory state for control of body movements. The present work suggests that model-based prediction is likely to be a pervasive component in natural gaze control as well.

Keywords

Saccadic eye movements; Prediction; Internal models; Squash; Gaze pursuit

Correspondence to: Mary M. Hayhoe, mary@mail.cps.utexas.edu.

Introduction

Given the existence of significant sensory and motor processing delays, there are many instances where it is useful to predict sensory state ahead of time in order to allow actions to be planned in advance of an event. In the case of eye movements, it is well known that the pursuit system involves a predictive component (Kowler et al. 1984; Becker and Fuchs 1985). Similarly, prediction is well known to occur with saccadic eye movements (Findlay 1981; Shelhamer and Joiner 2003). Less is known, however, about the mechanisms underlying this prediction, although it appears to be based on memory for stimulus motion accumulated over previous experience (Barnes and Collins 2008; Tabata et al. 2008; Barborica and Ferrera 2003; Xaio et al. 2007), and it has been suggested that it is based on an internal visual model of stimulus motion (Ferrera and Barborica 2010). It is generally accepted that the proprioceptive consequences of a planned body movement are predicted ahead of time using stored internal models of the body's dynamics (e.g., Wolpert et al. 1998; Mulliken and Andersen 2009). Subjects also use internal models or representations of the physical properties of objects in order to plan and control grasping (e.g., Johansson 1996; Flanagan and Wing 1997; Kawato 1999). Internal models have also been proposed for manual interceptions (Zago et al. 2009). Thus, sensory information from past experience can be used to predict future sensory state so that actions can be initiated before those state changes occur.

What kind of representations might allow prediction of visual state? That is, what is the nature of the proposed internal visual models? A number of studies have demonstrated neural responses in parietal cortex and frontal eye fields representing trajectories of moving targets during occlusions (Assad and Maunsell 1995; Xaio et al. 2007). In a study by Ferrera and Barborica (2010), the moving target changed direction during the occluded portion, as might occur, for example, when an object reflects off a surface. In these cases, the predictions involve fairly simple interpolation across occluded portions of the 2D trajectory using memory from previous trials. However, it seems likely that humans need to represent a variety of properties of the visual environment in order to anticipate changes in the visual scene. For example, subjects take account of gravitational acceleration when intercepting falling objects (Zago et al. 2004, 2008). When gravitational acceleration is removed from a virtual falling object, subjects move to intercept the object too early. This suggests that subjects learn the statistics of natural environments reflecting the pervasive influence of gravitational acceleration. Work by Battaglia et al. (2005) and Lopez-Moliner et al. (2007) shows that subjects take into account familiar size of objects during interceptive behavior. The recent development of eye trackers mounted on the head has allowed the study of eye movements in natural settings, where a wide variety of natural coordinated behaviors are possible (Land 2004; Hayhoe and Ballard 2005). In the context of natural behaviors such as playing cricket and table tennis, predictive saccadic eye movements have also been observed. In cricket, batsmen made a saccade to the anticipated bounce point of the ball (Land and McLeod 2000). In table tennis, subjects also anticipate the bounce point (Land and Furneaux 1997). Predictive eye movements are not limited to skilled performance. In an easy ball catching task, untrained subjects anticipate the bounce point when catching a ball and appear to take into account the elastic properties of the ball (Hayhoe et al. 2005). The ability to predict where the ball will be in all these cases depends on previous experience with the way that balls with particular dynamic properties typically bounce on a given surface, following a range of different trajectories, speeds, etc. Thus, internal visual models may involve a variety of different kinds of information based on extensive experience with the statistics of the natural world. In the present paper, we provide further evidence for the complexity of internal visual models for predictive control of gaze.

There are advantages to investigating natural behavior when trying to understand the characteristics of visual prediction. In natural contexts, the subject can plan eye and body movements in advance, rather than waiting for a stimulus presentation determined by the experimenter (Aivar et al. 2005). Dynamic environments such as sports, where subjects intercept rapidly moving objects, are also useful because the temporal constraints place a premium on predictive behaviors. We therefore investigated eye movements while playing squash. In this game, ball speeds are often as great as 125 mph, and the interval between successive hits approximately 1–2 s. Unlike cricket and table tennis, the player makes large body movements during interception. Thus, it is a very active and challenging environment. Given the difficulty of tracking eye movements in such an active task, little is known about eye movements during racquet sports. (Eye movements of passive observers have been examined (Abernethy 1990) but not those of actual players.) The lightweight nature of the headgear we used (see below) allowed us to track the eye movements in the presence of the large head and body movements involved in the game. Not unexpectedly, prediction was a pervasive component of gaze behavior in this context. We show in addition that these predictive movements are highly precise and operate continuously in time across multiple trajectories and multiple movements. This suggests that prediction is based on complex dynamic visual models of the way that balls move, accumulated over extensive experience. It also suggests that prediction based on internal models is likely to be a pervasive component in natural gaze control. It is also consistent with recent applications of statistical decision theory in visually guided reaching, which reveal that current and prior information is combined in a Bayesian manner in the control of targeting movements (e.g., Koerding and Wolpert 2004, 2006; Tassinari et al. 2006; Schlicht and Schrater 2007).

Methods

The movements of one eye of each of four players were recorded using an RIT wearable tracker shown in Fig. 1. This is a light, fully portable, infrared, video-based tracker that measures pupil and corneal reflection (Babcock and Pelz 2004). The system uses frames from very lightweight safety glasses with an off-axis infrared illuminator and a miniature, IR-sensitive CMOS camera to capture a dark-pupil image of the eye. A second miniature camera is located just above the eye and is used to capture the scene from the subject's perspective. The field of view of the scene camera is 65° horizontally and 47° vertically. The glasses frame is stable with respect to the head during subject movement, so that reliable eye tracking can be performed even when the subject is in motion. Players reported that they felt comfortable wearing the tracker while playing and they were not aware of any significant effect on their performance. The tracker is designed for off-line analysis; raw video of the eye and scene are captured using a video splitter that combines eye and scene images into a single video image, while calibration and analysis are carried out after data collection. The accuracy of the tracker is about 1° of visual angle. Following calibration, which allows the generation of a video record of the scene camera with crosshairs indicating eye position, the data were analyzed frame-by-frame from the 30 Hz video records to identify location, initiation, and termination of fixations and pursuit movements, and the times when the ball contacted the wall or the racquet.

A stationary video camera was mounted with a view of the players and the court and was used to capture events that occurred off the screen of one of the players' scene camera. For example, if a player's racquet was not on the screen of the head mounted scene camera at the time of contact, the stationary scene camera was used to determine the precise time of contact so that we could determine the location of gaze at the time of contact. This results in three video records for each playing session, one for each of the two players with gaze location superimposed on the scene video record, and a third camera filming the action externally. The external camera was consulted only when ambiguity existed on the player's

camera. The cameras were synchronized in time by having each of them look at the experimenter at the beginning of the game, while she clapped her hands once. Tests with a common time reference verified that the mean error associated with this procedure was 16 ms, and the maximum error was 33 ms.

We analyzed video records from four skilled squash players (two sets of two) during play using the mobile eye tracker. Two of the subjects were male players on the University of Rochester squash team. The other two were University of Rochester staff members who were skilled squash players. Each of these pairs played a game lasting about 15 min. All subjects signed an informed consent form. The only instruction given to players was to play as they normally would during a game of squash. They followed the rules of squash, but formal scorekeeping was not kept. All subjects had a brief warm up period while wearing the tracker, before play was initiated. The subjects were instructed to look at each of nine points in the form of a calibration grid on the front wall, 5.5 m from the observers, at the start and end of each playing session. Calibration was subsequently performed off-line using ISCAN software. Since calibration data were recorded at the end of each session, it was possible to confirm the validity of the calibration throughout the session. Thus, mis-calibration resulting from any slippage relative to the head or other disturbance was easily detectable. Calibrations were found to be almost always quite stable. Since the scene camera is placed 2.4 cm above the eye being tracked, there is a small vertical parallax error. The parallax error was less than 0.5° at all distances greater than 2 m. When the ball is on the racquet, approximately 1 m from the eyes, the parallax error grows to a maximum of 1.1°.

Since there is no separate numeric eye movement record, and no separate measurement of head position, data analysis was performed by manual frame-by-frame coding of the video images. The scene image, with eye cursor superimposed directly provides the point of gaze in the world and allows straightforward interpretation of the data. When eye-in-head position is recorded as well as head-in-space, it is possible to compute eye-in-space, or gaze position. In a real scene, however, the position of the ball in space is typically not known, so the object that gaze is falling on is hard to calculate, especially when the object moves in the scene. This information is directly available in the video image of the scene plus eye cursor, however. The frame at which events of interest occurred (saccade start or end, gaze location, pursuit movements, distance between eye and ball, and ball bounces) was logged, and statistics were extracted from these raw data. When it was necessary to take account of head movements, a feature in the scene image (such as a mark on the wall, or a corner) was chosen to allow estimation of the image displacement across several frames. Trials in which the eye track was temporarily lost, as occasionally occurred during a very large movement, were not used in the analysis. Otherwise coding began at the beginning of play, until 10–15 instances of the event of interest had been recorded. Although the temporal sampling is 30 Hz, averaging over 10–15 instances allows an unbiased estimate of the mean value, although it adds a few msec to the variance. Apart from occasional track losses, there was no selection of trials. An image of the eye was superimposed on the scene camera record to help validate the analysis based on the gaze cursor.

Results

Even though the players were given no instructions on what to do while they were playing, the players all demonstrated similar, stable, eye movement patterns. For the most part, the sequence of movements was constant across subjects, and only small differences arose when comparing the specific timing and accuracy of the movements. The dominating feature of performance was that subjects kept gaze on the ball almost continuously. This was achieved by head and body movements as well as by eye movements. They rarely fixated the opponent directly except when the opponent was hitting the ball, and the opponent's body

occluded the ball. Thus, information about the opponent was obtained from peripheral vision most of the time. We will describe the primary features of the cycle illustrated in Fig. 2. This sequence takes about 2 s, between one player's hit to the next player's hit. We first describe the general features of the cycle and then describe the quantitative features of performance. An example video is shown in Video 1 of the Supplementary Materials. The events are as follows: *Phase A*. As the opponent is hitting the ball, the player makes one or two large saccades to the front wall, with a combined eye and head movement, arriving shortly before the ball hits the wall. *Phase B*. The player pursues the ball shortly after it comes off of the front wall. Pursuit often involves large head movements and lasts until the ball reaches a high angular velocity as it approaches the player. The player then makes a saccade ahead of the ball prior to the bounce on the floor or wall. *Phase C*. The ball bounces on the floor and the player hits the ball at this point, or else the ball continues and then bounces off the side or back wall. Prior to the bounce, players typically stop pursuit and make a saccade ahead of the ball to a location through which the ball will pass shortly after reflecting off the wall or floor. Shortly after the ball passes through this location, the player resumes pursuit of the ball. *Phase D*. Just before contact between the player's racquet and the ball, the player stops tracking and holds gaze steady at a location in space until the racquet contacts the ball. Almost simultaneous with contact with the ball, the player initiates a saccade to the front wall, and the sequence repeats, except that now it is the opponent's turn to hit the ball. We will examine features of the four stages of this sequence first, when it is the player's turn to hit the ball, and then, we point out some of the differences when it is the opponent's turn. In each case, we recorded accuracy and timing of the movements over 10–15 instances of the event.

Phase A

When the opponent hit the ball, players made a large gaze shift involving one saccade, or occasionally two saccades, to the front wall. This gaze shift typically involved a large head and/or body rotation. The ball typically took about 300–500 ms to reach the front wall following a hit. Gaze arrived at the front wall 100–200 ms in advance of the ball. The latency between arrival of gaze and impact of the ball on the front wall is shown in Fig. 3a for the four subjects. Over the four players, the average arrival time in advance of the ball was 153 ms (±28). The error between gaze and the ball ranged between 6° and 13°, averaging 9° (±2) for the four players. This is shown in Fig. 3b. To give a sense of the accuracy of this movement, the front wall subtends about 60° when viewed from midcourt. Occasionally the player made a small corrective movement between the time gaze arrived on the front wall and the time of impact. In these cases, error was calculated based on gaze location at the initial fixation location, not the corrected location, since the correction was presumably based on visual information about the ball's trajectory acquired during the fixation. Earlier arrival did not appear to translate into greater accuracy, nor did it appear to correlate with skill level. However, with only four subjects such a relationship would be suggestive at best. It is unclear exactly what the advantage of early arrival is, although presumably it facilitates subsequent pursuit of the ball in some way. The initial saccade to the front wall was initiated either around the time the opponent was hitting the ball, or shortly after the hit, and so may be based partly on the opponent's posture and swing, in addition to the early part of ball's trajectory following the hit.

Phase B

Pursuit Performance—Following fixation on the front wall, players invariably pursued the ball as it came off the front wall, prior to either a bounce or to hitting it. The latency of this pursuit movement was on average 184 ± 17 ms after the ball bounced off the wall. The pursuit invariably was achieved with a large head/body rotation plus translation, in addition to eye rotation in the head. This has been referred to as "gaze pursuit" (Fukushima et al.

2009). Horizontal head rotations were often as large as 60°. Vertical movements were typically smaller, between 10 and 20°. It was commonly the case that the head rotation was comparable in magnitude to the eye-in-space movement, sometimes being smaller, and sometimes larger than the total gaze displacement. When the head movement was larger than the gaze movement, the eye was counter-rotated in the orbit a small amount. On average, over the 4 subjects, the head accounted for 88% of the total gaze change (between subjects SEM was ±15%). Presumably during these large head movements, the vestibular-ocular reflex is at least partially suppressed, as suggested by Collins and Barnes (1999), and the gain of the pursuit movement relative to the head modulated in a flexible manner to make up the difference. The existence of some trials where the head movement exceeds the total gaze change indicates the VOR is not completely suppressed.

Given the nature of the data and the large head and body movements, it is difficult to compute a measure of pursuit gain. In a sample case, we calculated angular ball and gaze velocity during the pursuit movement. This is shown in Fig. 4. To calculate angular ball velocity relative to the observer from the video image, we measured the ball position with respect to a fixed feature in the video image of the scene, because the intervening head movement shifts the image on the camera frame. Gaze velocity was computed similarly. Note that we need to assume that the head does not translate but only rotates, for this calculation to be accurate. This is almost certainly wrong in general, but the plot gives an indication that the observer is doing a good job of keeping gaze close to the ball. In Fig. 4, zero is the time the ball leaves the wall. By viewing the frame-by-frame record and the spatial separation of the ball and gaze, it can be seen that in this case, the eye appears to move in space about the same time as the ball comes off the wall. In general, eye and ball velocity are closely matched for about 400 ms. There is a small dip in eye velocity around 200 ms, but the spatial distance between eye and ball remained less than 2.5°. Gaze stays close to the ball until $t = 430$ ms where the eye lags the ball, and then, a large saccade is made ahead of the ball. The ball bounced on the floor at $t = 730$ ms. In order to measure of how well subjects were able to pursue the ball, we measured the fraction of the time between the front wall and the floor bounce that the player was able to maintain accurate pursuit. Pursuit was easy to distinguish from saccades despite the limited temporal resolution, as the gaze cursor was close to the ball (within 2.5°) and moved in step with the ball in each frame. A video record of this sequence can be seen in Video 2 (supplementary material). Players maintain pursuit for about half of the trajectory. The time between bouncing on the front wall and bouncing on the floor typically ranged between 400 and 800 ms. On average, players maintained gaze on the ball 54 ± 5% of this time period. Pursuit was typically terminated by a saccade. We estimated the 2D velocity of the ball in space (projected onto the head camera plane) at the point where pursuit ended, estimated from the preceding two frames of the video. Players were able to maintain gaze within 2.5° of the ball, up to velocities between 125 and 150°/s for three players. The best player (who is nationally ranked) was able to maintain gaze on the ball up to the highest velocities (186°/s). Although these velocities are in the saccadic range, the continuity of the movement clearly identifies it as a pursuit movement. The ability to maintain combined eye and head pursuit up to such high velocities implicates a significant predictive component to the control of both head and eye components of the pursuit movement, given that the ball travels several meters within the 100–200 ms period that would be required for visually guided corrections. Typically, the pursuit movement was terminated by a saccade ahead of the ball prior to a bounce, and then, pursuit was resumed following the bounce. However, on some other occasions, players appeared to briefly maintain a pursuit movement with rapidly decreasing gain when the ball was approaching the floor. In these instances, gaze appeared to keep step with the horizontal position of the ball but did not follow the vertical movement of the ball toward the floor. These brief periods were then terminated by a saccade ahead of the ball, as described above. This particular strategy may be an efficient one, allowing the player to keep pace with the

horizontal movement without making a vertical movement down then up again after the ball bounces on the floor. This is somewhat speculative, however, and we did not examine these trials in detail.

Phase C

Predictive Saccades—If the ball is not hit on a volley, it bounces either on the floor or off the side or back wall (or both). We separated the instances when the ball bounced off the side or back wall from those when it bounced on the floor. We chose wall bounces because they were easier to see on the video than floor bounces, which often occurred below the bottom of the scene camera image. We observed that players typically made a saccade ahead of the ball to a location through which the ball passed shortly after bouncing off the wall. This saccade occurred prior to the bounce and appeared to be directed to a point in space where ball would pass shortly after the bounce, rather than the actual point at which the ball made contact with the surface. Given that only one eye is tracked, we do not know exactly what depth plane the observer is converged on. However, it appeared that the ball trajectory came very close to the stationary point of gaze shortly after the bounce. Immediately after the ball passed through the closest point, the pursuit movement was resumed (see below), so it seems reasonable to assume that players were indeed aiming for a point after the bounce. This interpretation is also consistent with data from cricket and ball catching (Land and McLeod 2000; Hayhoe et al. 2005). An example of such an anticipatory saccade is shown in Video 3, Supplementary Materials. We measured both the accuracy and timing of these anticipatory saccades. First, we found the point at which the ball came closest to the stationary gaze point. At this point, the eye was within $2.5 \pm 0.5°$ of the ball in the two-dimensional image, averaged over the four subjects. The time at which gaze arrived at this point, prior to the ball, averaged 186 ± 36 ms. Accuracy for the individual subjects is shown in Fig. 5. For each subject, the standard error over 10 trials ranged from 0.3 to 0.7°, so the behavior was quite reproducible both within and between subjects. This precision is quite extraordinary, given that the player's entire body is moving during the period when the saccade is planned and executed and that the estimate of the ball's future location will depend on a number of factors such as angle and speed of impact, and elasticity of the ball. (Note that players usually saccade ahead of the ball when it bounces on the floor as well, but these episodes were not analyzed in detail.)

Once the ball passed close to the center of gaze, players continued to pursue the ball for a short period, prior to hitting it. This pursuit movement occurred, on average, 46 ± 7 ms after the ball passed through this closest point. This short latency suggests that subjects were planning the pursuit movement during the period following the reflection off the wall. By making the anticipatory saccade, players were able to hold gaze steady while the ball trajectory went through a discontinuity, and this may have facilitated the resumption of pursuit after the bounce. Some evidence that the saccade ahead of the ball facilitates subsequent pursuit is that there is a negative correlation between the lead-time of the saccade and the latency of the subsequent pursuit ($R^2 = 0.86$). This is shown in Fig. 6 for the 4 players. With only 4 data points, this relationship is at best suggestive, but players who arrive earlier appear to begin pursuit earlier. It is possible that the timing of the anticipatory saccade reflects some kind of a trade-off between two factors: if the saccade is made too early, the player will have less immediate information about the ball's trajectory, leading to greater inaccuracy; if the saccade is made too late, there is less time to get information about trajectory after the bounce, and the subsequent pursuit movement will be delayed.

Phase D

As discussed above, following a bounce off the wall or floor, players pursued the ball for a short period before their racquet made contact with the ball. Players did not continue pursuit

of the ball onto the racquet, however. Instead, they stopped pursuit and held gaze (usually both eye and head) stationary for the last 160 ± 6 ms of the ball's flight (see Fig. 7). Because of visuo-motor delays, visual information acquired after this point is unlikely to influence the arm movement. Saunders and Knill (2003) found that the minimum time for a target step to result in a correction in the pointing movement was about 160 ms. Almost simultaneously with the hit, the player initiated a saccade to the front wall. The timing of the saccade relative to the hit averaged 18 ± 11 ms after contact for the four players. The individual players' latencies ranged from 45 ms after contact for Subject 1, to 6 ms before contact for Subject 4 (the best player). Standard errors for the individual subjects ranged from 6–24 ms. Presumably the timing of this movement is based on efferent information about the arm movement, together with a prediction about the time of impact. Again, the precision of this prediction is impressive. When this saccade arrives at the front wall, we are back to Phase A of the sequence in Fig. 2, but now it is the opponent's turn to hit the ball.

Dependence on task context: observing the opponent's play

When the player looks to the front wall at the time of his own hit, as discussed in Phase D above, the arrival time at the wall was a little later than for an opponent's hit. This is shown in Fig. 8. For all four subjects, gaze arrived at the front wall a little later when they had hit the ball themselves, as opposed to the opponent. The magnitude of this difference depended on the player. Even though the eye was only just ahead of the ball, prediction is still required. Since the saccade was initiated approximately simultaneously with the hit (see above), it must have been programmed prior to the time the ball was actually in flight. Presumably holding gaze steady just prior to impact is more important than early arrival at the front wall, when it is the opponent's turn to hit. There is some suggestion that accuracy is slightly better (7° on average) when the player was the one who hit the ball, as opposed to the opponent, although the difference is significant for only one of the subjects.

After the ball bounced off the front wall and continued along its trajectory toward the opponent, the player pursued the ball until it bounced on the floor or one of the other walls, even though it was the opponent's turn to hit the ball. The player began pursuit 210 ± 32 ms after it bounced off the front wall. Given that gaze arrived at the front wall simultaneously with the ball, the initiation of the pursuit movement was probably not predictive, as the latency falls well within normal pursuit reaction times. Once pursuit began, the player maintained the pursuit movement for $47\% \pm 6$ of the path between wall and floor. Pursuit began slightly later and occurred for a shorter percentage of the flight, when it was the opponent's turn, but neither of these differences was significant. Moreover, players were able to pursue the ball at the same velocities regardless of whether they were preparing to play or watching an opponent.

Discussion

Performance in squash reveals a pervasive influence of prediction. Predictive effects are maintained across sequences of different actions during a playing cycle (Fig. 2) and involve eye, head, arm, and body movements. Thus, movements following an event such as a bounce or hit reveal evidence of predictive control initiated prior to that event. Additionally, since each cycle of events, from one hit to the next, involves different ball trajectories, the basis for prediction must derive from a long-term history of experience with the statistics of the ball trajectories.

To review the observations, first, players anticipated where the ball would hit the front wall with moderate accuracy. We did not analyze in detail when this movement was initiated. In some cases, it was likely based on the early part of the trajectory of the ball as it came off

the opponent's racquet (less than 100 ms). In many instances, however, the saccade was initiated prior to, or around the time of the hit, and may therefore be based, at least partly, on such factors as the opponent's body posture and the direction of the swing. Given the time taken to program the saccade, only minimal trajectory information was involved in the prediction, even in those cases when the movement was initiated following the hit. This may account for the fact that the spatial accuracy is only moderate. Evidence for use of information prior to release of a ball in predicting its trajectory has also been demonstrated by Bahill et al. (2005). Another possible factor in the accuracy of this prediction is that the point of impact will depend on the distance of the ball from the wall. Since players typically hit the ball from mid- or back-court, the magnitude of the predictive gaze movement was very large, of the order of 90°, so would have greater absolute error than when predictions were made over a shorter distance.

Perhaps the most impressive instance of prediction occurred when players made a saccade ahead of the ball when it was about to bounce off a side or back wall. The accuracy of this movement is very high, especially given that it is executed while the entire body is in motion. Since a discontinuity in the path is involved, the prediction must be based on more complex factors than simple extrapolation from the current trajectory and is presumably based on extensive experience with the ball that takes into account its trajectory and speed and the elasticity of ball and the nature of the reflective surface (floor, side, or back wall). An interesting feature of this saccade is that it appears to be for the purpose of facilitating a pursuit movement after the ball has reflected off the surface. The initiation of the pursuit movement occurs less than 50 ms after the ball passes through the gaze point. Although there were only four subjects, it appeared that this pursuit latency was reduced when the saccade had been initiated earlier. Earlier arrival presumably allowed the subject to accrue more information about the ball's path and plan the subsequent movement more effectively.

Another unexpected result was, contrary to standard coaching, gaze did not follow the ball into the racquet, but paused about 160 ms before contact. This was followed by a saccade to the front wall that was initiated almost simultaneously with contact of the ball on the racquet. Presumably the saccade was programmed during this 160 ms interval and must be based on the ball's trajectory prior to contact in addition to efferent (or proprioceptive) information about the swing. The fact that this movement occurred almost simultaneously with the hit suggests that players also have a precise estimate of the time at which the contact event will occur. Accurate timing of predictive saccades to repeated externally generated events has been demonstrated previously in a number of studies (e.g., Findlay 1981). Manual tasks also reveal accurate timing estimates, and subjects modulate the timing of the movement to achieve interception depending on the required precision and the quality of the sensory information (Tresilian and Houseman, 2005; Battaglia and Schrater 2007). In the present case, knowledge of the time of impact appears to control the timing of the eye movement. A similar sensitivity to timing is indicated by the saccade prior to a wall bounce, discussed in the preceding paragraph. Since players resumed pursuit so soon after the ball passed through the center of gaze, it appears that both position and timing of the anticipatory saccade are tightly controlled. It seems likely that the primary objective of the subject is to keep gaze on the ball and that the role of saccades is to facilitate resumption of pursuit following a path discontinuity. The timing of the saccades appears to be tightly orchestrated to facilitate pursuit.

As mentioned above, it is well known that pursuit has a substantial predictive component, and the performance of these players in tracking the ball at such high, continuously increasing, velocities is consistent with such prediction. Pursuit at similarly high velocities, up to 150°/s for gaze pursuit, has previously been observed by Bahill and Laritz (1984) in baseball. In their case, however, the head movements were much smaller, and most of the

pursuit movement was accomplished by the eye, which moved at a peak velocity of 120°/s. In the present situation, most of the movement appears to be accomplished by the head (and body) rotation. Thus, the predictive component must be used to control both eye and head jointly. The predictive component of head movements during gaze pursuit has been investigated by Collins and Barnes (1999) and also by Fukushima et al. (2009). Collins and Barnes proposed that visual information from previous trials provided a common store that could be used separately by different effectors (eye and head in this case) either to coordinate those movements or to control them separately. They raise the possibility that an important role for a common visual representation of predicted target motion is the coordination and orchestration of the different effectors. The observations reported here support this hypothesis. Thus, in addition to mitigating the problem of visual delays, another value of experience-based internal visual models is that they allow flexible, task-specific coordination of eye, head, arm, and body movements. In the situation examined here, where body, arms, head, and eyes are all making movements within a very short time period, the value of a common visual model for this purpose is clear.

What is the nature and locus of the model on which prediction is based? In the case of pursuit movements, it seems likely that it represents a visual memory signal rather than memory for eye velocity (Barnes and Collins 2008; Madelain and Krauzlis 2003; Churchland et al. 2003). The overlap in circuitry between pursuit and saccades suggest that they share common target selection mechanisms (Krauzlis 2005; Joiner and Shelhamer 2006). There is also evidence for a common predictive mechanism for both movements, possibly located in the Supplementary Eye Fields (Nyffeler et al. 2008). In the present context, not only are pursuit and saccades coordinated in a predictive manner, but also the head, hand, and body. This suggests that the basis of prediction may be an abstract model of the moving object that can be used by multiple effectors. Such a model would need to represent knowledge of the elastic properties of the ball, contingent on speed and angle of incidence, in addition to gravitational effects and the properties of the reflective surfaces in the court. The information also appears to be represented in an allocentric, court-based reference frame, since the eye movements are executed on the moving platform of the body. Prediction also appears to be modulated by the behavioral context, since players modulate the timing of arrival at the front wall depending on whether or not it is their turn to hit. In a simple saccade-timing task, Isotalo et al. (2005) also showed that predictive behavior is strikingly modified by the specific interpretive and motivational state of the subject. Thus, the predictive performance we observe here is considerably more complex than many of the conditions under which predictive saccades have been examined, which typically involve simple timing anticipation (Shelhamer and Joiner 2003) or 2D trajectory interpolation (Ferrera and Barborica 2010). It is presumably based on extensive experience with the ball and involves more than simple extrapolation based on cues within a single trajectory segment (Zago et al. 2009).

A number of different brain areas are likely to be involved in the internal models postulated to underlie the complex predictive behavior we observe. Since much of the critical information for prediction is visual, it seems likely that at least some component of the internal model is stored in high-level visual cortex, and the existence of neural signals in parietal cortex corresponding to occluded target motion (Assad and Maunsell 1995) is consistent with a visual component in prediction. Other areas are likely to be involved, in addition to visual areas, however. For example, Shichinohe et al. (2009) identified a visual memory component for predictive pursuit in SEF, but a visual memory component was not observed in the discharge of caudal FEF pursuit neurons (Fukushima et al. 2008). In the case of saccades, the frontal eye fields appear to represent a visual memory component in addition to an internally generated error signal (Barborica and Ferrera 2003; Xaio et al. 2007; Ferrera and Barborica 2010). Evidence from fMRI in humans implicates dorso-lateral

prefrontal cortex (Pierrot-Deseilligny et al. 2005) and a distributed cortico-subcortical memory system including prefronto-striatal circuitry (Simo et al. 2005). Since the cerebellum has long been implicated in internal proprioceptive models, it is likely that visual signals in the cerebellum also play a role in internal models (Blakemore et al. 2001).

In summary, performance in squash reflects continuous prediction of visual state across extended time periods and multiple movements. As discussed previously, the need for internal models for control of body movements is well established (Shadmehr and Mussa-Ivaldi 1994; Wolpert et al. 1998; Wolpert 2007; Kawato 1999; Mulliken and Andersen 2009), although the precise nature of these models is unclear. Similarly, internal models of some kind appear to play an important role in the control of eye movements as well as in the coordination of eye and body movements (Ferrera and Barborica 2010; Collins and Barnes 1999; Chen-Harris et al. 2008; Nyffeler et al. 2008). In previous experimental work, prior visual experience is controlled by the experimental procedure, and the subjects accumulate visual evidence across trials. In the present context, however, subjects must have accumulated visual information about ball dynamics as a result of extensive experience with the statistical properties of the environment. Thus, internal models can reflect both long-term and short-term experience. The complexity of the information required to predict ball behavior in the current context seems to rule out simple models such as visual interpolation or extrapolation. Instead, elasticity, 3D velocity, angle of incidence, surface reflectance properties, and gravity all are likely to be taken into account. Predictive saccades have previously been observed in other natural behavior such as cricket and table tennis (Land and McLeod 2000; Land and Furneaux 1997). The present results extend these findings to continuous sequences of eye movements and wider ranges of body movements, as well as providing more extensive analysis of performance.

Another important consideration is the way that internal models influence the movement. In the present context, movements are based on some combination of the current visual information with the model-based prediction, since any given instance of prediction is based on the properties of the current trajectory that must constrain the set of stored distributions that are relevant for prediction in any given instance. In reaching, there is evidence for the optimal Bayesian integration of current visual information with stored visual models or visual priors, (Koerding and Wolpert 2004; Brouwer and Knill 2007; Tassinari et al. 2006). The present results raise the issue of whether a similar optimal weighting occurs with saccadic eye movements. Note that the proposition that the targeting mechanism computes a weighted combination of visual and memory signals differs from the idea that predictive saccades and visually guided (reactive) saccades derive from separate control strategies (Findlay 1981; Shelhamer and Joiner 2003). It should also be noted that the existence of pervasive prediction in eye movement targeting is in marked contrast to image-based models of target selection, such as those based on image salience (e.g., Itti and Koch 2001 and many others). Although the current visual image obviously has an important role to play, predictive eye movements reveal the observer's best guess at the future state of the environment based on learnt representations of the statistical properties of dynamic visual environments (Land and Furneaux 1997; Jovancevic-Misic and Hayhoe 2009). A more complete understanding of the selection of gaze targets requires more complicated models where target selection is based on internally driven behavioral goals, such as the Reinforcement Learning model of Sprague et al. (2007).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abernethy B. Expertise, visual search, and information pick-up in squash. Perception. 1990; 19:63–77. [PubMed: 2336337]

Aivar P, Hayhoe M, Mruczek R. Role of spatial memory in saccadic targeting in natural tasks. J Vis. 2005; 5:177–193. [PubMed: 15929644]

Assad JA, Maunsell JH. Neuronal correlates of inferred motion in primate posterior parietal cortex. Nature. 1995; 373(6514):518–521. [PubMed: 7845463]

Babcock, JS.; Pelz, JB. Building a lightweight eyetracking headgear; ACM SIGCHI eye tracking research & applications symposium; 2004. p. 109-114.

Bahill AT, LaRitz T. Why can't batters keep their eye on the ball? Am Sci. 1984; 72:249–253.

Bahill AT, Baldwin DG, Venkateswaran J. Predicting a baseball's path. Am Sci. 2005; 93:218–225.

Barborica A, Ferrera VP. Estimating invisible target speed from neuronal activity in monkey frontal eye Weld. Nat Neurosci. 2003; 6:66–74. [PubMed: 12483216]

Barnes G, Collins C. Evidence for a link between the extra-retinal component of random-onset pursuit and the anticipatory pursuit of predictable object motion. J Neurophysiol. 2008; 100:1135–1146. [PubMed: 18596183]

Battaglia PW, Schrater P. Humans trade off viewing time and movement duration to improve visuomotor accuracy in a fast reaching task. J Neurosci. 2007; 27(26):6984–6994. [PubMed: 17596447]

Battaglia, PW.; Schrater, P.; Kersten, D. Auxiliary object knowledge influences visually-guided interception behavior; Proceedings of 2nd symposium APGV; 2005. p. 145-152.

Becker W, Fuchs A. Prediction in the oculomotor system: smooth pursuit during the transient disappearance of a visual target. Exp Brain Res. 1985; 57:562–575. [PubMed: 3979498]

Blakemore SJ, Frith CD, Wolpert DW. The cerebellum is involved in predicting the sensory consequences of action. Neuroreport. 2001; 12(9):1879–1884. [PubMed: 11435916]

Brouwer A, Knill DC. The role of memory in visually guided reaching. J Vis. 2007; 7(5):6. (1–12). [PubMed: 18217846]

Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R. Adaptive control of saccades via internal feedback. J Neurosci. 2008; 28(11):2804–2813. [PubMed: 18337410]

Churchland M, Chou I, Lisberger S. Evidence for object permanence in the smooth pursuit movements of monkeys. J Neurophysiol. 2003; 90:2205–2218. [PubMed: 12815015]

Collins C, Barnes GR. Independent control of head and gaze movements during head-free pursuit in humans. J Physiol. 1999; 515.1:299–314. [PubMed: 9925900]

Ferrera V, Barborica A. Internally generated error signals in monkey frontal eye Weld during an inferred motion task. J Neurosci. 2010; 30(35):11612–11623. [PubMed: 20810882]

Findlay JM. Spatial and temporal factors in the predictive generation of saccadic eye movements. Vision Res. 1981; 21(3):347–354. [PubMed: 7269312]

Flanagan J, Wing A. The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. J Neurosci. 1997; 17:1519–1528. [PubMed: 9006993]

Fukushima K, Akao T, Shichinohe N, Nitta T, Kurkin S, Fukushima J. Predictive signals in the pursuit area of the monkey frontal eye Welds. Prog Brain Res. 2008; 171:433–440. [PubMed: 18718338]

Fukushima K, Kasahara S, Akao T, Kurkin S, Fukushima J, Peterson B. Eye-pursuit and reafferent head movement signals carried by pursuit neurons in the caudal part of the frontal eye Welds during head-free pursuit. Cereb Cortex. 2009; 19:263–275. [PubMed: 18483002]

Hayhoe, M. Visual memory in motor planning and action. In: Brockmole, J., editor. Memory for the visual world. UK: Psychology Press; 2009. p. 117-139.

Hayhoe M, Ballard D. Eye movements in natural behavior. Trends Cogn Sci. 2005; 9(4):188–193. [PubMed: 15808501]

Hayhoe, M.; Mennie, N.; Sullivan, B.; Gorgos, K. The role of internal models and prediction in catching balls; Proceedings of AAAI fall symposium series; 2005.

Isotalo E, Lasker AG, Zee DS. Cognitive influences on predictive saccadic tracking. Exp Brain Res. 2005; 165:461–469. [PubMed: 16025290]

Itti L, Koch C. Computational modeling of visual attention. Nat Rev Neurosci. 2001; 2:194–203. [PubMed: 11256080]

Sensory control of dextrous manipulation in humans. In: Johansson, RS., editor; Wing, A.; Haggard, P.; Flanagan, J., editors. Hand and brain: the neurophysiology and psychology of hand movements. San Diego: Academic Press; 1996. p. 381-414.

Joiner W, Shelhamer M. Pursuit and saccadic tracking exhibit a similar dependence on movement preparation time. Exp Brain Res. 2006; 173:572–586. [PubMed: 16550393]

Jovancevic-Misic J, Hayhoe M. Adaptive gaze control in natural environments. J Neurosci. 2009; 29(19):6234–6238. [PubMed: 19439601]

Kawato M. Internal models for motor control and trajectory planning. Curr Opin Neurobiol. 1999; 9(6):718–727. [PubMed: 10607637]

Koerding K, Wolpert D. Bayesian integration in sensorimotor learning. Nature. 2004; 427:244–247. [PubMed: 14724638]

Koerding K, Wolpert D. Bayesian decision theory in sensorimotor control. Trends Cogn Sci. 2006; 10:319–326. [PubMed: 16807063]

Kowler E, Martins A, Pavel M. The effect of expectations on slow oculomotor control- IV Anticipatory smooth eye movements depend on prior target motions. Vis Res. 1984; 24:197–210. [PubMed: 6719834]

Krauzlis R. The control of voluntary eye movements: new perspectives. Neuroscientist. 2005; 11(2): 124–137. [PubMed: 15746381]

Land, M. Eye movements in daily life. In: Chalupa, L.; Werner, J., editors. The visual neurosciences. Vol. vol 2. Cambridge: MIT Press; 2004. p. 1357-1368.

Land MF, Furneaux S. The knowledge base of the oculomotor system. Philos Trans R Soc Lond B Biol Sci. 1997; 352:1231–1239. [PubMed: 9304689]

Land MF, McLeod P. From eye movements to actions: how batsmen hit the ball. Nat Neurosci. 2000; 3:1340–1345. [PubMed: 11100157]

Lopez-Moliner J, Field DT, Wann JP. Interceptive timing: prior knowledge matters. J Vis. 2007; 7:1–8.

Madelain L, Krauzlis RJ. Effects of learning on smooth pursuit during transient disappearance of a visual target. J Neurophysiol. 2003; 90:972–982. [PubMed: 12904499]

Mulliken, G.; Andersen, R. Forward modles and state estimation in parietal cortex. In: Gazzaniga, MS., editor. The cognitive neurosciences IV. Cambridge: MIT Press; 2009. p. 599-611.

NyVeler T, Rivaud-Pechoux S, Wattiez N, Gaymard N. Involvement of the supplementary eye Weld in oculomotor predictive behavior. J Cog Neurosci. 2008; 20(9):1583–1594.

Pierrot-Deseilligny, Ch; Müri, A.; NyVeler, B.; Milea, D. The role of the human dorsolateral prefrontal cortex in ocular motor behavior. Ann NY Acad Sci. 2005; 1039:239–251. [PubMed: 15826978]

Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control fast reaching movements. Exp Brain Res. 2003; 152:341–352. [PubMed: 12904935]

Schlicht E, Schrater P. Reach-to-grasp trajectories adjust for uncertainty in the location of visual targets. Exp Brain Res. 2007; 182:47–57. [PubMed: 17503025]

Shadmehr R, Mussa-Ivaldi FA. Adaptive representation of dynamics during learning of a motor task. J Neurosci. 1994; 14:3208–3224. [PubMed: 8182467]

Shelhamer M, Joiner WM. Saccades exhibit abrupt transition between reactive and predictive, predictive saccade sequence have long-term correlations. J Neurophysiol. 2003; 90:2763–2769. [PubMed: 14534279]

Shichinohe N, Akao T, Kurkin S, Fukushima J, Kaneko C, Fukushima K. Memory and decision making in the frontal cortex during visual motion processing for smooth pursuit eye movements. Neuron. 2009; 62:717–732. [PubMed: 19524530]

Simo LS, Krisky CM, Sweeney JA. Functional neuroanatomy of anticipatory behavior: dissociation between sensory-driven and memory-driven systems. Cereb Cortex. 2005; 15:1982–1991. [PubMed: 15758195]

Sprague N, Ballard D, Robinson A. Modeling embodied visual behaviors. ACM Trans Appl Percept. 2007; 4(2):11.

Tabata H, Muira K, Kawano K. Trial-by-trial updating of the gain in preparation for smooth pursuit eye movement based on past experience in humans. J Neurophysiol. 2008; 99:747–758. [PubMed: 18077667]

Tassinari H, Hudson TE, Landy MS. Combining priors and noisy visual cues in a rapid pointing task. J Neurosci. 2006; 26:10154–10163. [PubMed: 17021171]

Tresilian JR, Houseman JH. Systematic variation in performance of an interceptive action with changes in the temporal constraints. Q J Exp Psychol. 2005; 58A:447–466.

Wolpert D. Probabilistic models in human sensorimotor control. Hum Mov Sci. 2007; 26:511–524. [PubMed: 17628731]

Wolpert D, Miall C, Kawato M. Internal models in the cerebellum. Trends Cogn Sci. 1998; 2:338–347. [PubMed: 21227230]

Xaio Q, Barborica A, Ferrera V. Modulation of visual responses in macaque frontal eye Weld during covert tracking of invisible targets. Cereb Cortex. 2007; 17(4):918–928. [PubMed: 16723405]

Zago M, Bosco G, MaVei V, Iosa M, Ivanenko Y, Lacquaniti F. Internal models of target motion: expected dynamics overrides measured kinematics in timing manual interceptions. J Neurophysiol. 2004; 91:1620–1634. [PubMed: 14627663]

Zago M, McIntyre J, Patrice Senot P, Lacquaniti F. Internal models and prediction of visual gravitational motion. Vis Res. 2008; 48:1532–1538. [PubMed: 18499213]

Zago M, McIntyre J, Patrice Senot P, Lacquaniti F. Visuo-motor coordination and internal models for object interception. Exp Brain Res. 2009; 192(4):571–604. [PubMed: 19139857]
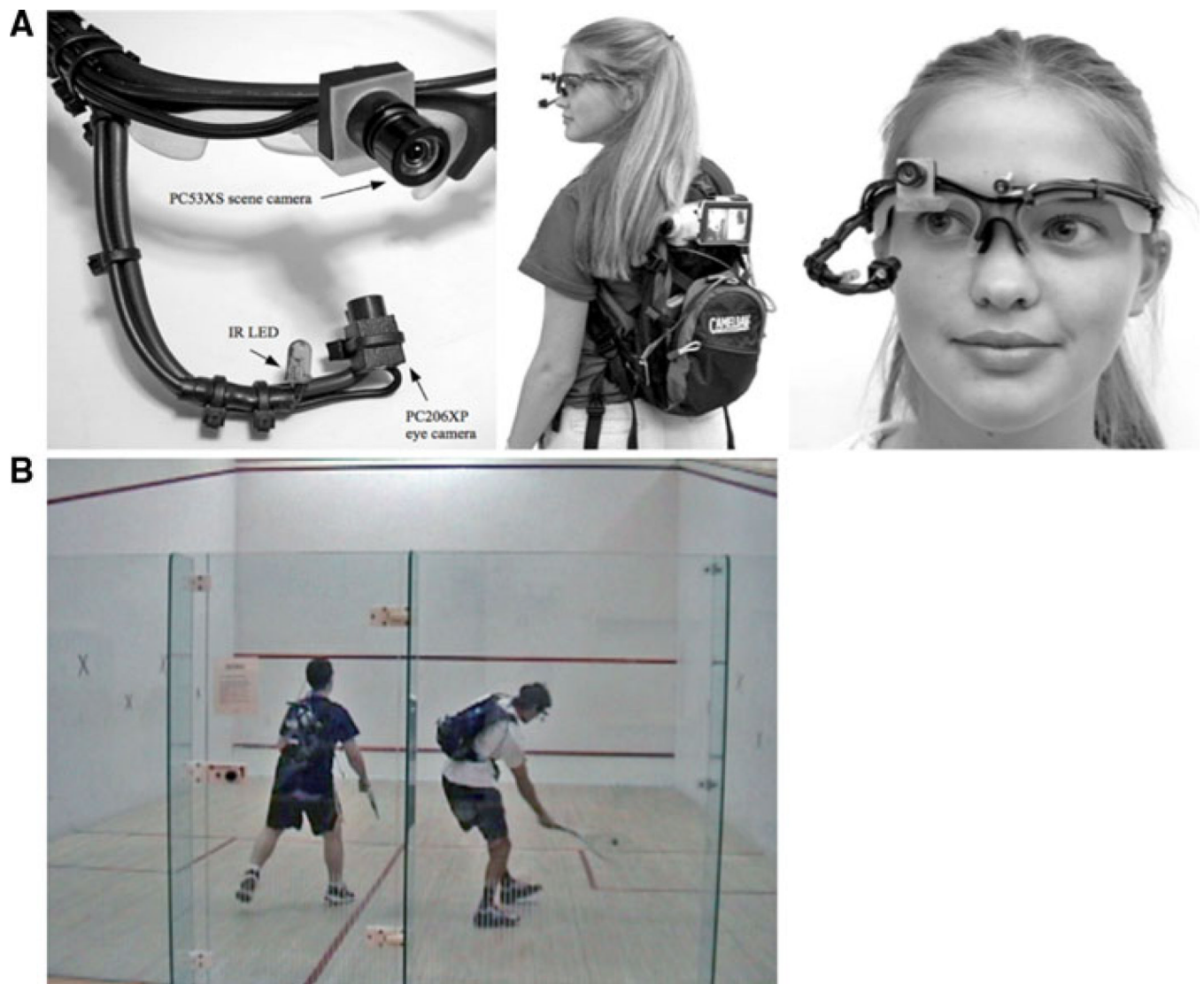
Fig. 1.
a The RIT wearable tracker, showing eye and scene cameras and IR illuminator, backpack with video camera and split-screen image, and lightweight head mount. **b** View of two players wearing the RIT trackers from the stationary scene camera
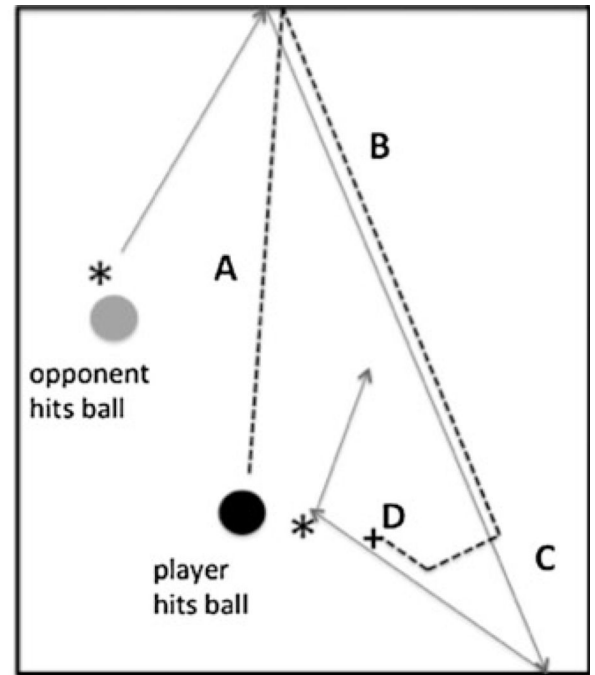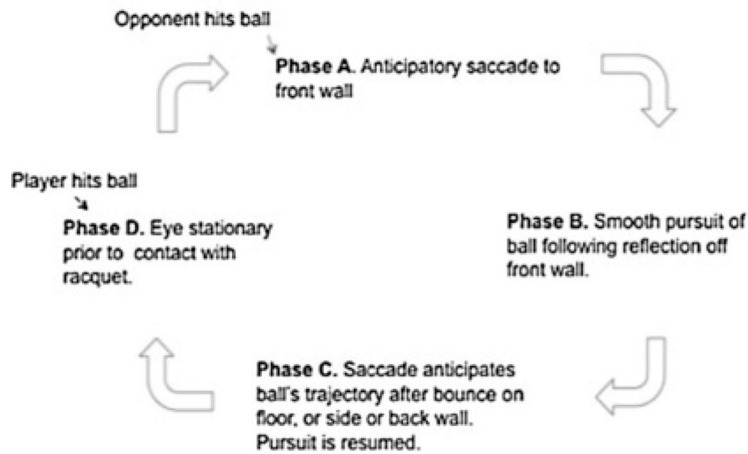
Fig. 2.
The sequence of eye movements during play, as described in the text. Beginning when the opponent hits the ball, the player saccades to the front wall (*Phase A*), then pursues the ball off the wall (*Phase B*), saccades ahead of the ball when it bounces on floor or wall (*Phase C*), then pursues briefly before holding gaze stationary shortly before hitting the ball (*Phase D*). On the *right*, the *solid line* indicates the path of the ball, and the *dashed line* indicates gaze direction. The *asterisks* indicate when the ball was hit, first by the opponent, then by the player whose eye movements are analyzed. The 'plus' sign indicates fixation prior to contact with the racquet
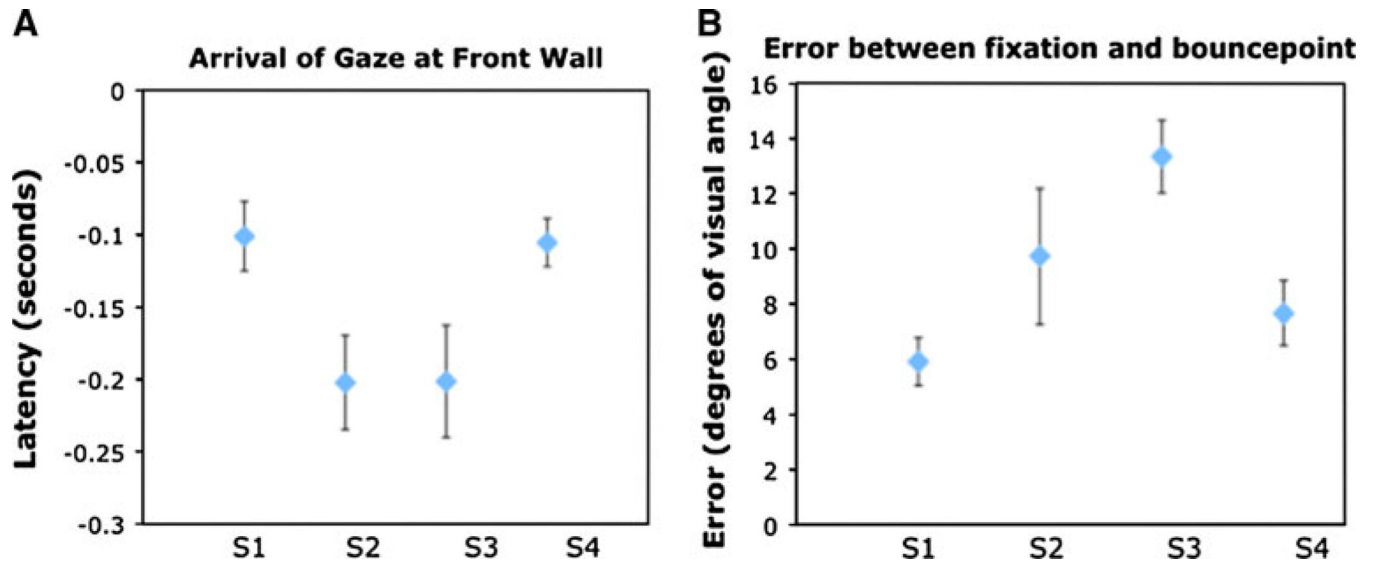
Fig. 3.
a Latency of saccade to the front wall relative to the impact of the ball. *Negative values* of latency indicate that the saccade arrived prior to the ball. **b** Distance between gaze location and point of impact. Data are for four subjects, S1–S4. *Errors bars* are ±1 SEM
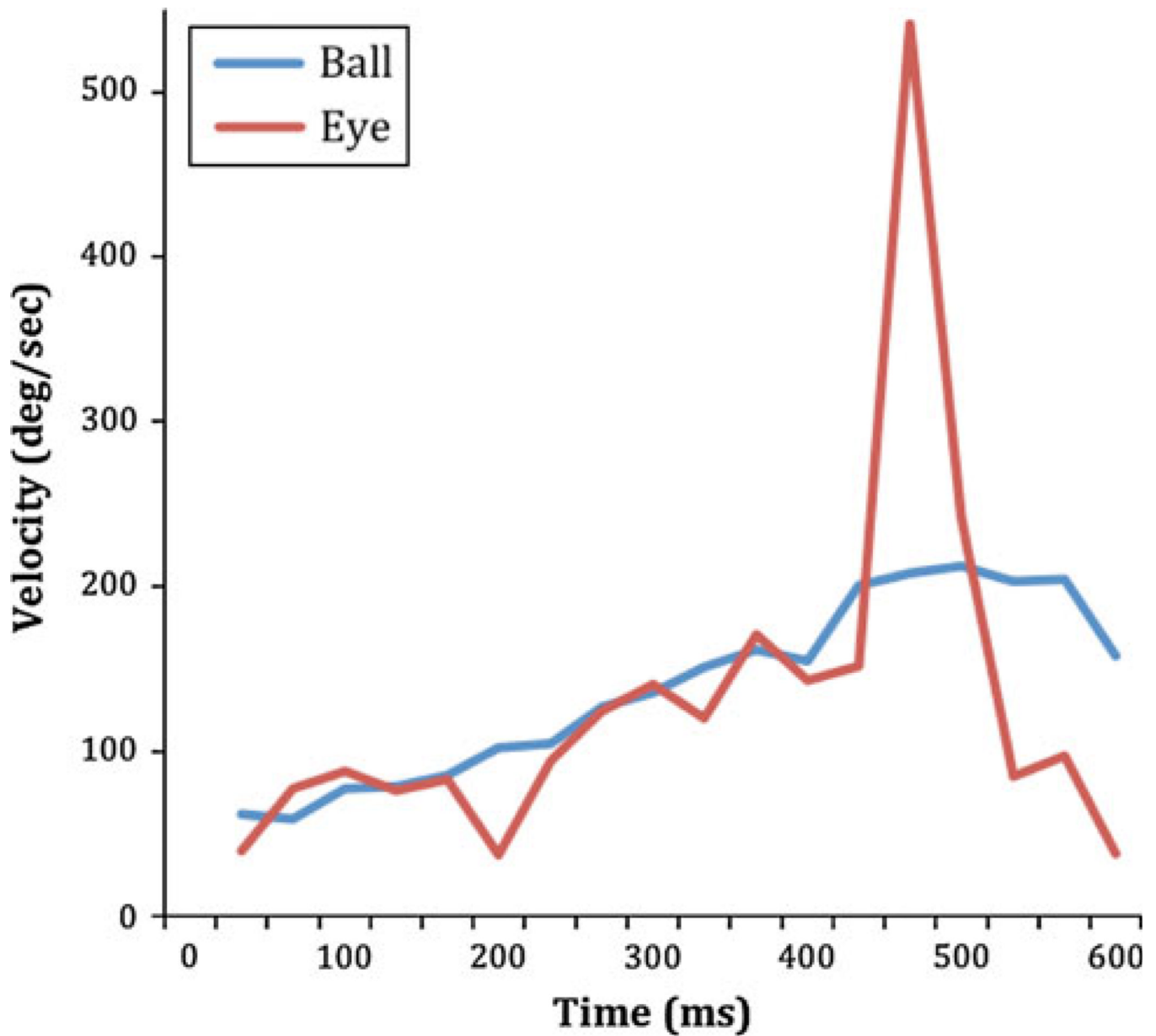
Fig. 4.
Angular velocity of ball relative to observer (*blue line*) during pursuit. The *red line* shows angular gaze velocity in space. Zero is the point at which the ball bounces off the front wall. The ball bounces on the floor 7,300 ms later. A saccade ahead of the ball is initiated at *t* = 430 ms
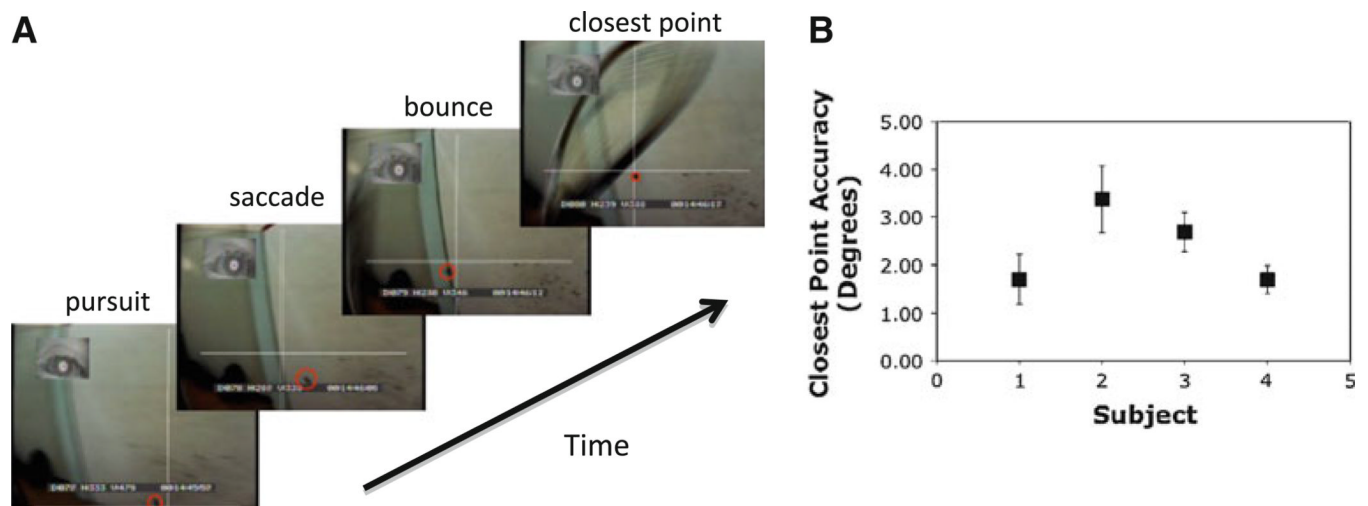
A

closest point

bounce

saccade

pursuit

Time

B

Fig. 5.
a Example of an anticipatory saccade prior to reflection off the wall. Pursuit of the ball precedes a saccade ahead of the ball. Gaze is held approximately stationary in space and the ball subsequently passes through this location. The ball is marked with a *red circle* for clarity. In the last frame, the racquet can be seen prior to the hit. **b** Accuracy of such saccades for 4 subjects. Minimum distance between gaze and ball is plotted. *Error bars* are ±1 SEM
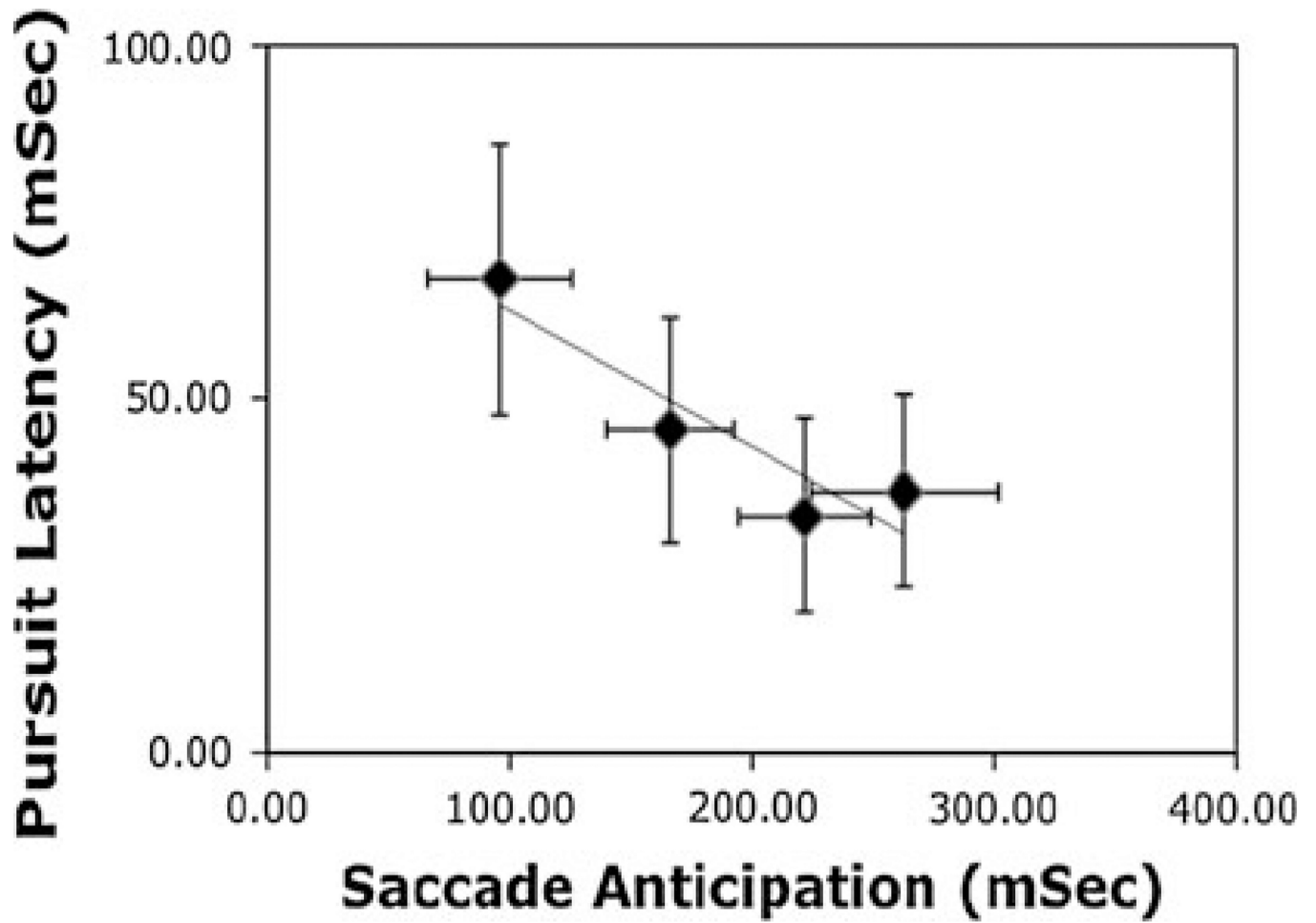
Fig. 6.
Latency for pursuit relative to the time when the ball passes closest to the center of gaze plotted against time the saccade was initiated relative to this time. Data points are for 4 subjects. *Error bars* are ±1 SEM
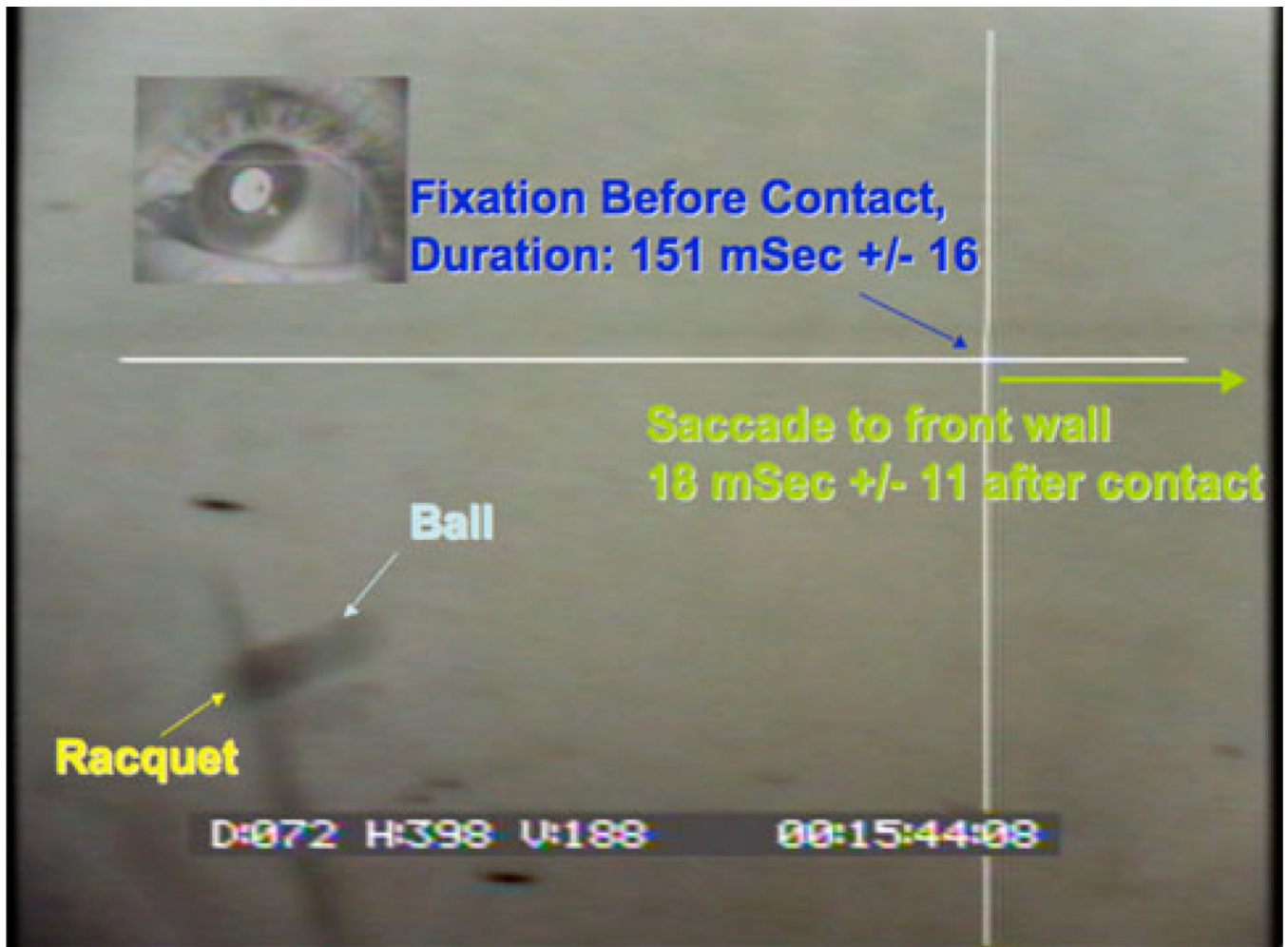
Fig. 7.
An example of fixation prior to contact of the ball with the racquet. Players stop pursuit 150 ms before contact and then saccade in the direction of the ball almost simultaneously with contact
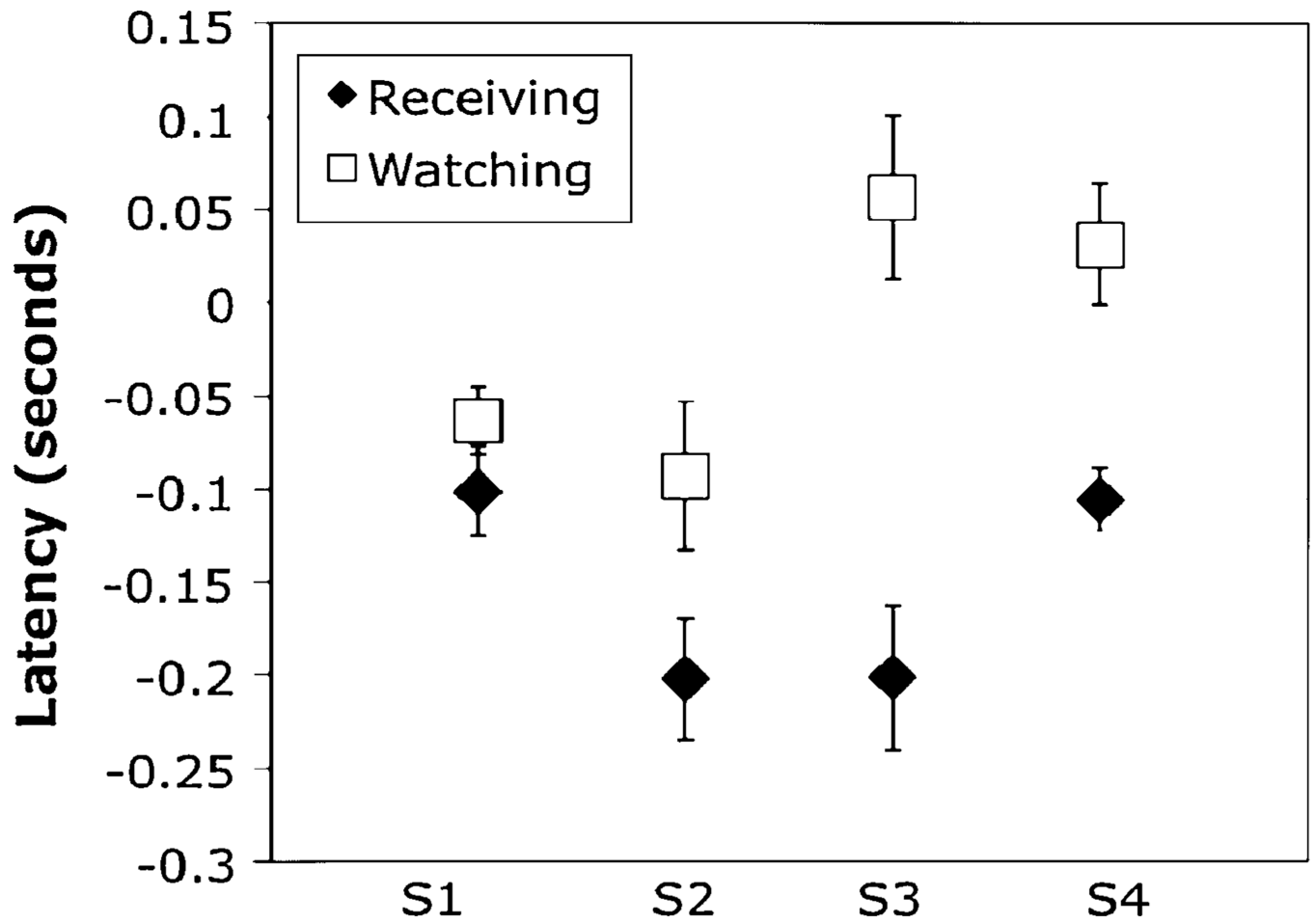
Fig. 8.
Latency of saccade to the front wall when the player must hit the ball (Receiving) compared with instances when the opponent must hit (Watching). *Negative values* indicate that the eye arrived prior to the ball. Data are for 4 subjects. *Error bars* are ±1 SEM