Ball Catching and the Role of Prediction

Keith Gorgos, 2006
Advised by Mary Hayhoe, Ph.D.
Department of Brain and Cognitive Sciences

How do we use our eyes on a daily basis for normal vision? Most people never even stop to think about how their eyes work to guide their actions during their daily routines. If they did, they may discover that vision is a highly specialized process. The eyes are constantly moving from location to location, seeking to acquire selective information about the world. Very little was known about how the eyes accomplished this until recently, when eye-tracking devices robust enough to allow unrestrained body movements were introduced. It has been shown in a number of everyday activities such as driving, table tennis, cricket, and reading that eye movements are highly specialized for the task at hand. This specialization suggests that eye movement patterns are not generated reflexively by the nervous system, but must be learned. The current paper investigates this aspect of eye movements. It has been previously shown that retinal motion, stereo information, and extra-retinal information from pursuit eye movements are important in the task of catching a ball. Stereo information is the result of a disparity produced through the integration of the left and right eye images and enables the perception of depth. Extra-retinal information includes knowledge of where the eye is positioned in the head (i.e., what direction the eye is pointed). Pursuit eye movements can be thought of those eye movements that occur when athletes “keep their eyes on the ball”. To further explore how our eyes are used to aid the accomplishment of daily living, we examined subjects’ eye, hand, and head movements as they caught balls thrown with a bounce. Furthermore, we examined the subjects’ head and eye movements as they threw balls to others and also as they observed two other individuals throwing and catching. Our investigation was an exploratory study in which we hoped to compare to previous studies and findings. We were looking to determine how we learn to make the predictions that are so necessary to our normal vision on a daily basis.

Our Approach

We examined this question by observing where the visual fixations occur, what visual information is needed in the activity, how specialized the eye movements were, whether or not there were any random or exploratory eye movements, and how the hands and head movements were coordinated.

Eye, head, and hand movements were recorded while subjects caught, threw and observed balls thrown with a bounce. Subjects’ eye movements and the view of the scene ahead were recorded with a head mounted eye-tracking camcorder. The image of the eye as it moved was recorded and embedded into the corner of the live scene image, which displayed the playing field from the subject’s point of view. This was advantageous to our analysis as it allowed us to see what the eye was doing at the same time we could examine the action in the field. The foveal direction of the eye (the fovea is a small area on the retina where fixations fall) was detected by the tracker and superimposed as a crosshair onto both video images of the eye and scene ahead. This direct mapping of the crosshair onto the image was very beneficial, as it allowed us to easily see complicated events as they took place, as opposed to a simpler eye-tracking device which would have just produced a list of coordinates. The image from the head mounted camera can be seen in figure 1a. Subjects were also recorded with a fixed scene camcorder, so that any activity in the field that occurred off the head camera could be obtained by matching the two video time stamps.

Hand movements were detected and recorded by two coordinate-tracking gloves worn by the subject. One glove was particularly specialized so as to measure the changes in the position of the subject’s fingers, such as when the fingers closed to make a catch.

Head movements were detected by using a stationary reference point in the field of view and observing when the field in the camera frame shifted. By measuring the distance the frame moved in regards to the reference point, the direction and magnitude of head movements were recorded.

Experimental Task, Condition & Setup

Each subject began the experiment bouncing a tennis ball off a wall with throws that bounced once, hit the wall, and returned to the subject. To some extent, this familiarization period enabled each subject to adapt and become accustomed to the properties of the tennis ball as well as the tasks of throwing and catching. A bouncy ball was not used during the familiarization period. This was done so that when it was introduced for the second half of the experiment, participants would be unfamiliar with its properties. This enabled us to see how the ability to pursue a ball depends on the experience with
a ball's dynamic properties.

Following the familiarization period, the subject and two other non-subject players were positioned in the field in the shape of a triangle, with each person as a "point" of the triangle, as seen in figure 1b. Players threw the ball with one bounce to the other players in a clockwise direction for the duration of the experiment. A tennis ball was used for the familiarization period (against the wall) and for the first half of the triangular sequence. For the second half, a bouncy ball that was slightly smaller and faster was used. In the triangular arrangement, three conditions were examined as the ball went around the triangle, Catching, Throwing and Observing (a throw and catch).

Results & Discussion
Catching

The first six tennis ball trials and the first six bouncy ball trials were analyzed for all five subjects. As found in cricket batsmen preparing to hit a ball, subjects' eyes fixate on the hands of the thrower preceding a catch, make an initial saccade (a rapid eye movement) to the region of the anticipated bounce point, and then pursue (follow closely) the ball until just before it reaches their hands.4

In order to examine the nature of the eye movements as the subjects caught the balls, we closely compared the time at which important events occurred with the time of the subject's eye movements. Our results showed that all of the initial saccades (towards the predicted bounce point area) occurred at a latency of 133 ms or less, in comparison to the release time from the hand. (A latency is the amount of time between a stimulus and a response.) It is known by vision scientists that it takes 200 ms to initiate a saccade based upon external stimuli. Thus, if an actual saccade occurs sooner than 200 ms after an event, that saccade was most likely made in anticipation of that event that required the saccade. Therefore, the saccades we observed were anticipatory in that they were based upon a prediction of where the ball would go. This is further emphasized by the fact that a number of the saccades even occurred before the release, some as much as 67 ms before (a "latency" of -67 ms). The average latency between \( t_r \) (time of the release) and \( t_s \) (start time of the saccade) for all five subjects in both tennis and bouncy trials is plotted in figure 2.0. While it may seem simple to say subjects "predict" or make "anticipatory" eye movements, it is a process of greater complexity than one may first realize. People do not usually realize the complexity of making predictions, because they are not even aware that predictions are being made: the brain does them almost automatically. Before someone can make an anticipatory saccade, they must combine the informative visual stimuli of the trial with predictions of the ball's speed, and other predictions of the ball's properties. As people gain more experience, they "automatically" become better, but how?

In the tennis trials, there is no decrease in the \( t_r - t_s \) latency as the number of trials increased, which is likely due to "familiarization period" in which participants became familiar with the properties of the tennis ball. However, the drastic decrease in the latency during the bouncy condition (saccades occurring much sooner over the trials) was rather striking. This suggests that subjects are adapting and learning with experience to the new ball's properties. This was shown by their increasing ability to anticipate and predict events, enabling them to become more successful in the task in terms of pursuing the ball. It seems as if the eye is trying to get itself into position at the bounce point sooner so that it can react to the ball off the bounce sooner, or have more time to adjust its gaze correctly. So how exactly are the subjects able to "get better" and constantly refine their actions? Our observations

![Figure 2.0: Earlier departure relative to release](image-url)
suggest that people maintain an internal model of the dynamic properties of the world, and rapidly update this model when errors occur. After experiencing errors on the first few trials (especially with the new bouncy ball), the subjects appear to have updated their internal model of the ball's dynamic properties, based on their recent experience with the newer ball.

The concept of an internal model is based on the idea that people have internal, perceptual representations of the physical world around them. When people receive visual and other sensory information regarding the properties of the objects in the physical world around them, the central nervous system processes this information and rapidly updates their current perceptual representations of the world. When errors occur in the present task, it is a signal to them that their current representation is "outdated", and revisions need to be made.

We also compared the end time of the saccade towards the bounce point ($t_B$) to the time the ball bounced ($t_b$), producing the latency of ($t_B - t_b$), as shown in figure 3.0. The end of the saccade can also be thought of as the start of the fixation near the bounce point. As shown by the graph, however, there is no real "latency", as all the latencies are negative. In other words, saccades reached the region near the predicted bounce before the bounce occurred. For both conditions, as the trials continue, there is a small downward trend in $t_B - t_b$, suggesting that experience with catching causes a slightly greater anticipation of the bounce event (the latency becomes more negative). This is especially true for the bounce condition, as the anticipation becomes greater at a much faster rate; the latency for the tennis trials decreases by an average of 33 ms, while the latency for the bouncy trials decreases by an average of 80 ms. That is, the saccade to the bounce point was arriving at its end position sooner for the bouncy trials.

The suggested conclusion is that as subjects gain more experience, they learn to anticipate faster. Due to errors they have made in tracking the ball, subjects feel a need to improve their anticipation of the bounce point. We believe they do this by subconsciously updating their perceptual models to the standards necessary to eliminate errors, based on the integration of visual and temporal cues with their prior experiences. The demand for this anticipation is increased when a newer, faster (bouncy) ball is introduced that travels 19 visual degrees faster. Thus, with the bouncy trial, because the participant realizes this drastic change in properties the anticipation becomes quicker, and at a faster rate. Due to prior experience with the tennis ball, the "adjustments" made in the tennis trials involve more "fine-tuning" as opposed to the rapid adjustments made for the bouncy trials. The greater anticipations made in the bouncy trials enable subjects to have more time fixated near the bounce point, gathering information from the approaching ball. In addition, while the gaze may be fixated above the anticipated bounce point, the subject's attention may be directed towards the ball in motion or on the actual bounce point. Thus, subjects may be gathering information through attention, while preparing for the next pursuit movement with the position of fixation.

The observations of anticipatory saccades that reached the bounce point region prior to the bounce, the saccades that anticipate the release from the hand, and the ability to predict the actual location of the bounce point demonstrate how essential prediction is to this ball catching activity. By arriving at the bounce point ahead of the ball, the subject has time to prepare for the change in direction and speeds caused by the floor at the bounce. If the gaze were to try and pursue the ball before the bounce point, it would probably overshoot the ball when the ball bounces, and have to correct for the error. By saccading ahead to a position in the ball's path beyond the bounce, it avoids the possibility of being thrown off by the bounce.

Once a subject's saccade brought his gaze to the bounce point region, fixation remained there for a length of time after the bounce, after which he began pursuing the ball with his gaze, a time referred to as $t_p$. This delay was referred to as the dwell time at the bounce point, or $t_B - t_p$, and is depicted in figure 4.0. The dwell time for the tennis trials seemed to stay fairly constant throughout the trials, while the bouncy trials demonstrated a more interesting pattern. The dwell time of the first bouncy trial is very close to the preceding tennis trials, but as the bouncy trials continue, the dwell time increases further, eventually becoming 65ms greater than the last tennis trial.

Although the increase was not statistically significant, due to large SEMs, the dwell time at the bounce point for the bounce ball increased as subjects had more experience with the ball's dynamic properties. The additional dwell time could be partially attributed to the faster anticipation to the bounce, allowing a longer dwell time. Additionally, it is possible some of the additional dwell time was a result of the pursuit movement (which occurs after the dwell time) starting later. Either way, it
seems the oculomotor system may require more time to respond to and start its attempts to pursue the bouncy ball than it did for the slower tennis ball, to which it had already adapted. However, to determine if there is actually an increase that can be considered statistically significant, additional participants in these tasks should be examined.

The subject's fixation direction at the time of the bounce point \( t(t) \) was recorded for all subjects in both tennis and bouncy trials, for which the results were very similar. Only 4 out of the total 36 fixations for the tennis trials were below the bounce point. Bouncy trials displayed similar results, with only 3 of 36 trials below the bounce point.

For 90 percent of all trials, the fixations during the bounce were above the bounce point—gaze was directed a mean of 5.9 visual degrees above the bounce point for all trials. Our results were similar to table tennis players whose saccades are "[...] clearly aimed just above the bounce point", as found by Land and Furneaux in 1997. Table tennis players directed their gaze an average of 5.3 degrees above the bounce point, with at least 90% of the fixations occurring above the bounce point as well. A possible reasoning for positioning the foveal direction just above the anticipated bounce point is that it reduces the distance the gaze will be "out of position", as compared to the foveal direction if it were directed under the bounce point. In other words, by positioning the foveal gaze just above the bounce point, the eye can always wait for the ball to rise up to its position to resume tracking. On the other hand, if the subjects tried to position their gaze at the bounce point itself, or under the bounce point, they would be increasing the distance across which they would need to make a saccadic 'catch-up' movement in the event of a positioning or speed prediction error. By keeping the gaze just above the bounce point, a positioning error will never put the gaze behind the ball at the bounce. Gaze positioned above the bounce always has a chance to resume tracking the ball as it rises up to its position and does not have extra ground to cover due to a fixation below the bounce.

Following the dwell time at the bounce point, subjects attempted to track the ball using smooth pursuit gaze movements. (These are a special class of eye movements driven by specialized neural circuitry that is thought to respond to image motion on the retina by an attended object.) However, accuracy depended on the experience with the properties of the ball. In order to compare the accuracy of these pursuit movements, we examined the position of the subject's gaze frame by frame and found the percentage of frames at which the subject was successfully pursuing the ball. That is, the frames when their gaze was directed on the ball (within a region of about 1.5 visual degrees).

As shown in figure 5.0, the subjects' accuracy with the tennis ball, is fairly stable. However, in the first bouncy ball trial, as shown in the figure, there is a drastic drop in accuracy of the pursuit in comparison to the sixth tennis trial. \( t(4) = 3.14, p<0.05 \). As the number of bouncy trials increase, however, there is a great deal of improvement, with the accuracy with the bouncy improving to be nearly equal to the accuracy of the early tennis trials. The accuracy on the last bouncy trial was found to be statistically greater than the first bouncy trial \( t(4) = 2.83, p < 0.05 \).

In general, during the pursuit period following the bounce, the majority of tennis trials involved no saccades. Following the tennis bounce, there was usually a short period of 'catch-up' pursuit in which a small gap between the ball and the gaze decreased, which was followed by a period of smooth pursuit in which the gaze accurately tracked the ball until it was close to the hands. In the later trials of the bouncy condition, this pattern was also very common. During the early bouncy trials, however, there was usually a large 'catch-up' saccade, followed by a number of smaller saccades, until the gaze was repositioned on the ball. Once the subject 'caught-up' (near the end of the ball's flight), they generally displayed some form of smooth pursuit movements (which tended to be 'shakier' and shorter than the tennis trials or later bouncy trials). The first saccade on the first bouncy trial was generally the largest: one subject's first saccade reached 11.7 visual degrees. Additionally, some subjects were never able to begin a period of smooth pursuit on the first trial—all their attempts to track the ball consisted of saccadic behavior. However, by the sixth bouncy trial, subjects were tracking the ball in a pattern similar to the tennis trials.

If subjects had experience with the properties of a ball, as they did with the tennis ball, and as they did with the later bouncy trials, they used smooth pursuit gaze movements to track the ball. However, they were forced to make 'catch-up' saccades on some of the trials when errors in their tracking occurred. The largest of saccades occurred on the first bouncy trials, when subjects had the least amount of experience and adaptation to the faster ball. This is parallel to the results found by Land and McLeod when examining cricket, who stated, “'catch-up' saccadic behavior is expected of someone who has not played cricket.” As trials and experience increased, the number and magnitude of saccades decreased, signifying a reduction in tracking errors. At the same time, smooth pursuit movements increased and the percent accuracy of these movements became more accurate. These changes in eye movements suggest that the subjects were rapidly updating their internal models of the world based on the ball's dynamic properties.

Throughout our analysis of the aspects of the catching trials, we saw a continuing pattern. Few errors and few adjustments were made to the eye movements during the tennis trials, while continuous adjustments were made as the result of errors during the bouncy trials. We discovered that success in the task of catching a ball is more complex than simply responding to visual cues. We learned that you can not keep your eye on the ball unless you have an accurate representation of the ball's properties, and can integrate this information with the
current visual stimuli. We believe this may be done through an internal model.

Observing

While the main focus of the experiment was on the catching aspect, eye movements were also recorded as subjects watched the throws between the two other players. The gaze and anticipatory patterns observed were task specific to the observing trials, as the patterns differed from the catching trials.

Similar to catching, when subjects watched another thrower and catcher, they made a saccade somewhere near the anticipated bounce point. The anticipation of the bounce point was much less accurate for observing trials, as shown in Figure 6.0, than it had been for the catching trials. As with catching, fixations were generally above the bounce point itself: 25 of 30 fixations (83%) were above the bounce point. However, fixations were spread across the horizontal axis much more than they had been in the catching. This is most is likely because the observer is not required to be as accurate, as the observer has nothing to do with making the catch successful. Thus, his eye movements do not gather as much information as they do in the catching trials. This is also probably why there was very little pursuit movement in the tennis trials. The observer is not making many errors pursuing the tennis ball, and thus is fairly "relaxed" in terms of gathering information. In other words, the observer already has knowledge of the ball's dynamic properties, and is having little trouble pursuing the tennis ball in the catching trials. Thus, it is possible that the observer does not feel a need to gather more information regarding the properties of the tennis ball as it goes between the two other individuals. Conversely, subjects were much more likely to pursue the bouncy ball when observing others.

Subjects in the bouncy trials pursued the ball somewhat after the bounce in about 83% of the bouncy trials compared to 22% in the tennis trials. This is likely because the subject was not familiar with the dynamic properties of the bouncy ball and thus, was trying to use the observation trials to learn the new properties of the bouncy ball. This gives more support to the concept that subjects may have an internal model that they are trying to update their when errors occur.

As in the catching trials, subjects made anticipatory saccades relative to important events in the balls path. Head movements also predicted these events. In general, anticipations on the observing trials were much greater than those on the catching trials. Again, this may be because subjects are not involved in the task themselves and thus are more willing to make "higher risk" predictions, as their "success" in observing is not that important.

Throwing

Analysis of throwing trials was not as extensive as the catching or observing trials, although interesting results were found nonetheless. The gaze and anticipatory patterns observed were much different than those seen in the observing and catching trials. When subjects threw the ball, they displayed a different pattern of movements, first making a saccade to a point on the floor beyond of the bounce point, and then to the catcher's hands, with both saccades preceding the ball by several hundred milliseconds. The fixation made just beyond the bounce point may serve to gather two pieces of information that allow the subject to throw the ball accurately to the other player. It is possible the fixation is positioned between the catcher's feet and the anticipated bounce point so that the thrower can easily switch attention back and forth between the two points and make an accurate throw. Attention to the feet allows the thrower to see the total distance the ball needs to go, while attention to the anticipated bounce point allows the user to see where he aims to throw the ball. The anticipatory saccades that precede the events in the throwing trials additionally demonstrate the importance of prediction for visual activities.

Overall

Throughout our observations of catching, observing, and throwing, we saw just how important predictions are to the current task. As stated earlier, it takes 200 ms to initiate a saccade based upon external stimuli. Thus, without predictions of the anticipated ball movement, the eye would not be able to track, or 'catch-up' to the ball's movement, as solely reactive behavior would leave the gaze behind the position of the ball, not on the ball as we saw with the tennis, and later bouncy trials. We also saw that when the properties of the physical world change, people are more likely to make errors. Errors indicate that there is a need to update one's internal representations of the physical world, which allows us to make necessary predictions for visual tasks. This was demonstrated by the reoccurring pattern of adjustments being made to the eye movements for the bouncy trials. While the current investigation did not focus on the effect of the changes in latencies and gaze durations in performing successfully physically (i.e. making a successful catch as opposed to dropping it), it seems the proactive nature of our visual system prefers to be as efficient as possible. Rather than constantly having to 'catch up' to a moving object and recalculate saccadic movements based on errors, individuals prefer to be more efficient. That is, they prefer to establish smooth pursuit movement early on. By increasing the dwell time and decreasing the $t_b - t_i$ and $t_b - t_f$ latencies, it seems as if the participants desire to arrive to the fixation point near the bounce sooner. Unlike a saccade, a fixated eye allows new information to be gathered from the world. Thus, getting to the fixation point sooner allows the eye more time to be prepared for the approaching ball and observe its position and
speed. It seems likely that this would enable the eye to begin smooth pursuit more accurately than a fixation that arrived later and did not allow for much information gathering. It seems likely that accurate smooth pursuit is preferred to multiple 'catch-up' saccades, as the saccades involve multiple calculations that try to estimate the next position of the ball's path at which the ball can be intercepted to resume smooth pursuit. Smooth pursuit, on the other hand, seems fairly easy to maintain, once it has been established, as it simply involves maintaining the current speed of movement. Thus, while it has not been proven by our experimental results, it is a possibility that changes are made in the latencies and dwell time so that more time is allowed in a state of fixation, when information can be gathered. This seems as if it would help to establish smooth pursuit earlier so less attentional resources are used in tracking the ball, as constant 'catch-up' calculations would not be necessary. On the other hand, during the tennis trials, few errors or adjustments were made and it seems likely that subjects may have established an accurate internal model from prior experience in the familiarization period.

Conclusion

It has been shown in that past that a number of everyday activities require highly specialized eye movements for the task at hand. This specialization suggests that eye movement patterns are not generated reflexively by the nervous system; rather, they must be learned. We investigated this aspect of eye movements and our findings demonstrate that vision cannot be a solely reactive system, but instead, must involve complex, proactive processes, that acquire information from the world essential to daily activity by anticipating events rather than reacting to them. From our analysis of subjects catching, observing others throw, and throwing balls with a bounce, we see that eye movements are highly specific to the task at hand, supporting a number of experiments done previously regarding such activities as cricket, table tennis, driving and reading music. While all task variations involved the similar condition of a ball thrown to someone with one bounce, the eye movements still changed when the actual task (catching, throwing, or observing) itself changed. One thing, however, was apparent throughout the trials: the eye movements were constantly proactive, attempting to gather information from the environment rather than reacting to it after the event occurred. Our analysis suggests that by anticipating events, observers position their bodies in ways that allow them to gather essential information to the current task. The notion of proactive eye movements has also been supported by other studies, including those by Land and Furneaux in 1997 and Flanagan and Johnson in 2003.

The patterns of eye activity we observed were similar to those found in cricket by Land and MacLeod, who showed that anticipating the position of the bounce point is important for intercepting the ball with the bat and similar to those found in table tennis by Land and Furneaux. Our results show that prediction is a major factor in the task of catching a ball. In addition to the anticipatory saccadic eye movements, the head movements and pursuit movements we observed reveal that visual information is gathered so as to plan for a predicted state of the world.

The findings of our research also suggest that people maintain a sophisticated internal model of the dynamic properties of the world (which seems necessary for continuity of behavior) and rapidly update this model by constantly comparing their internal model to the errors or accuracies that occur in their responses to stimuli. As we gather visual and non-visual sensory information from the world regarding the present task, we refine the model so that we are able to plan ahead and act appropriately in future situations. This was shown by our results when we switched from the tennis ball to the bouncy ball during the catching trials. We saw that the ability to pursue the ball depends on experience with the ball's dynamic properties—when people have little experience with these properties, errors are made. After experience with the properties of the new ball, people were able to update their internal models based on the learned properties of the world. This enabled them to respond in a predictive, and thus more successful, manner that involved more accurate pursuit. Therefore, our results also support preceding studies that suggest prediction, which is dependent on proactive eye movements, plays an important role in our everyday activities—even if we may fail to realize it.


Keith Gorgan will graduate with a B.S. in Brain & Cognitive Science in May, 2006. He worked with Professor Mary M. Hayhoe, Ph.D. for the duration of his sophomore year. After graduation, he is considering careers related to BCS, as well as the possibility of attending both Law and Business school, to further expand on the diversity of his education.