Cues for predictive eye movements in naturalistic scenes

Alexander Goettker

Justus Liebig Universität Giessen, Giessen, Germany
Center for Mind, Brain and Behavior,
University of Marburg and Justus Liebig University,
Giessen, Germany ✉

Nils Borgerding

Justus Liebig Universität Giessen, Giessen, Germany ✉

Linus Leeske

Justus Liebig Universität Giessen, Giessen, Germany ✉

Karl R. Gegenfurtner

Justus Liebig Universität Giessen, Giessen, Germany
Center for Mind, Brain and Behavior,
University of Marburg and Justus Liebig University,
Giessen, Germany ✉

We previously compared following of the same trajectories with eye movements, but either as an isolated targets or embedded in a naturalistic scene—in this case, the movement of a puck in an ice hockey game. We observed that the oculomotor system was able to leverage the contextual cues available in the naturalistic scene to produce predictive eye movements. In this study, we wanted to assess which factors are critical for achieving this predictive advantage by manipulating four factors: the expertise of the viewers, the amount of available peripheral information, and positional and kinematic cues. The more peripheral information became available (by manipulating the area of the video that was visible), the better the predictions of all observers. However, expert ice hockey fans were consistently better at predicting than novices and used peripheral information more effectively for predictive saccades. Artificial cues about player positions did not lead to a predictive advantage, whereas impairing the causal structure of kinematic cues by playing the video in reverse led to a severe impairment. When videos were flipped vertically to introduce more difficult kinematic cues, predictive behavior was comparable to watching the original videos. Together, these results demonstrate that, when contextual information is available in naturalistic scenes, the oculomotor system is successfully integrating them and is not relying only on low-level information about the target trajectory. Critical factors for successful prediction seem to be the amount of available information, experience with the stimuli, and the availability of intact kinematic cues for player movements.

Introduction

Over the last decades eye movements became widely used as a window into the brain, mind, and cognitive processes (König et al., 2016; Shaikh & Zee, 2018). The success of eye movement research is related to results that demonstrate a direct link between basic oculomotor paradigms, such as the antisaccade task (Hallett, 1978; for reviews, see Hutton & Ettinger, 2006; Munoz & Everling, 2004) or saccadic adaptation (McLaughlin, 1967; for a review see Pélisson, Alahyane, Panouillères, & Tilikete, 2010), and different cognitive metrics or learning mechanisms. In addition, the continuous nature of eye movement responses even allows one to map the dynamics of neuronal representations (Pack & Born, 2001). However, one thing that most of these influential paradigms have in common is the use of highly controlled but very simplistic stimuli (e.g., a simple target disk). Classically, oculomotor control is assumed to be mainly driven by early stages of visual processing (Lisberger, 2010; Lisberger, 2015) manipulated in such paradigms. However, multiple studies across varying settings revealed that higher level factors such as the task (Chen et al., 2021; Rothkopf, Ballard, & Hayhoe, 2016; Tatler, Hayhoe, Land, & Ballard, 2011; Yarbus, 1967) or expectations (Diaz, Cooper, Rothkopf, & Hayhoe, 2013; Jörges & López-Moliner, 2019; Kowler, Rubinstein, Santos, & Wang, 2019; Vo & Henderson, 2009) also have a strong impact on how we move our eyes. Therefore, a critical question remains unanswered: Can effects and mechanisms identified with these simple stimuli

be generalized to eye movement behavior with more naturalistic stimuli and tasks?

Although a major line of research mainly used mainly artificial stimuli, another line of work has looked at eye movements under naturalistic settings. The most influential early work was performed when using mobile eye trackers to record eye movements during everyday tasks (Hayhoe & Ballard, 2005; Land & Hayhoe, 2001). In these studies, eye movements were recorded while observers prepared a tea or making a peanut-butter and jelly sandwich. One of the key results of these studies was that observes mostly fixated on objects they interacted with, but they already moved their gaze predictively to the next relevant object before they completed the preceding action. This suggests that an important goal of the oculomotor system is to bring gaze to relevant locations at the right time to gather information. However, this is not a trivial task, as there are substantial internal processing delays (Nowak & Bullier, 1997; Schlag & Schlag-Rey, 2002). For simple stimuli, we know that it is possible to make accurate saccades to moving targets despite these processing delays (Fleuriet, Hugues, Perrinet, & Goffart, 2011; Goettker, Braun, & Gegenfurtner, 2019; Ron, Vieville, & Droulez, 1989; Schreiber, Missal, & Lefèvre, 2006). It is also possible to keep track of a predictably moving target without lag (Terry Bahill & McDonald, 1983). Also, for naturalistic stimuli, close to zero delay tracking has been observed when watching videos of outdoor scenes (Vig, Dorr, Martinetz, & Barth, 2011). But, in contrast to the behavior observed for simple stimuli, predictive eye movements seem to occur more frequently under more naturalistic contexts. This can be driven by knowledge about the physics of the world (Diaz et al., 2013), monitoring and planning a task (Sullivan, Ludwig, Damen, Mayol-Cuevas, & Gilchrist, 2021), or the semantic information provided by the scene (Henderson, 2017). However, these results so far came from different lines of work in vastly different paradigms, and until recently a direct comparison of oculomotor behavior for simple and naturalistic stimuli was missing.

In our recent study (Goettker, Pidaparthy, Braun, Elder, & Gegenfurtner, 2021) we tried to fill this gap. We made use of a hand-labeled dataset of ice hockey videos (Pidaparthy & Elder, 2019), which allowed us to present the movement of the puck (represented by a disk) in front of a gray background. We compared this condition with the original videos. For both conditions, the task of the observers was to follow the puck movements with their eyes, and, due to the identical trajectories, the low-level motion characteristics were comparable. When we computed the delay that observers had when tracking the motion trajectory, we observed a striking difference: With only the target movement in front of a gray background, the estimated delay was around 180 ms, which is expected with purely reactive behavior. In contrast, when seeing the original videos, even our non-expert observers had an average delay close to zero. This suggests that observers were able to make use of complex naturalistic input and integrated it to significantly reduce their tracking delay. However, because the conditions in our previous study were the two extreme cases of either no additional information present (when just seeing the disk) or all additional information available (original video), it is unclear which cues and information observers used to achieve this benefit in tracking performance. The goal of the current study was to answer that question by manipulating four factors: the level of expertise of the observers, the amount of peripheral information available, and the availability and reliability of positional and kinematic cues.

In our original study, we had studied non-expert observers, and even they reached almost zero-delay tracking when watching hockey videos. Therefore, we wanted to investigate how experts perform in our task, as it is known that across many different aspects (Brams et al., 2019), and especially in sports (Memmert, 2009; Vickers, 2009), expertise leads to differences in behavior and performance. We combined the difference in expertise with the investigation of the role of the amount of peripheral information available. We achieved this by using manipulated versions of the videos, where only a certain area around the puck position was visible. This is of interest due to two reasons: First, can experts make better use of additional information by better integration them for oculomotor control (Brams et al., 2019; Casanova, Oliveira, Williams, & Garganta, 2009)? Second, additional information might also be related to more crowding and limits in peripheral processing (Rosenholtz, 2016; Rosenholtz, Li, & Nakano, 2007; Whitney & Levi, 2011); therefore, additional peripheral information could also lead to no difference or even worse performance. For the other factor of interest, we used a second experiment in which we manipulated information along two axes that are known to be helpful in anticipation: positional and kinematic cues (Loffing & Cañal-Bruland, 2017). For positional cues, we added information about the player positions (comparable to the classical Heider & Simmel, 1944), but did not show the kinematic cues of player movements. To manipulate kinematic cues, we kept the video intact but played it in reverse or flipped it vertically to reduce their reliability (Pavlova, Krägeloh-Mann, Birbaumer, & Sokolov, 2002; Pavlova & Sokolov, 2000). Together, the combination of these different manipulations should give us a good starting point to understand which information observers used to gain the predictive advantage in the video condition.

Experiment 1. The role of expertise and peripheral information

Methods

Observers

Twenty-six observers took part in the experiment. Half of the participants were expert viewers, defined by reporting a good understanding of tactical knowledge in ice hockey games and the self report of regularly watching ice hockey games (on average 1.88 ± 1.29 games per week) and being a fan of the sport for at least a year. In contrast, the novice group had never watched an ice hockey game before. With respect to other characteristics such as gender (in both groups there were four persons who identified as female, nine who identified as male) or age ($M = 23.46 \pm 1.71$ for the experts, $M = 23.31 \pm 1.75$) the groups were matched. All participants were naïve with respect to the purpose of the study and gave informed consent at the beginning of the experiment. The experiment adhered to the tenets of the Declaration of Helsinki and was approved by the local ethics committee. Participants were paid for their effort with 8 euros/h.

Setup

Participants sat at a table in a dimly illuminated room with their head positioned on a chin rest. In this position, their eyes were roughly at the height of the center of a monitor (3840×2160 pixels, 30-Hz refresh rate, Philips 288P6LJEB/00; Philips, Amsterdam, Netherlands) at a distance of 70 cm. In this setup, the monitor spanned approximately 49×26 degrees of visual angle. The experiment was programmed and controlled with MATLAB (MathWorks, Natick, MA) using Psychtoolbox (Kleiner, Brainard, & Pelli, 2007). Gaze was recorded from one eye with a desk-mounted eye tracker (EyeLink 1000 Plus; SR Research, Kanata, ON, Canada) at a sampling frequency of 1000 Hz. Before each block a nine-point calibration was used, and an additional drift check was performed at the start of each trial.

Stimuli and conditions

We selected 12 of the original 18 videos (each lasting for 10 seconds) from our previous paper (Goettker et al., 2021) for the follow-up experiments. In this study, we wanted to investigate the factors modulating the prediction effect we observed earlier. Therefore, we selected the videos that showed the strongest differences between the video and the disk condition in the original study. We selected eight videos that contained the labeled passes and four of the other videos. The

average effect on the delay for the selected videos we estimated for the selected videos was a significant decrease of the lag in the disk condition by –176.426 ms in the video condition. Please note that all of the videos were filmed with a static camera; therefore, no artifacts due to camera movements were present, and observers had to keep track of the puck just by moving their eyes. In addition to the expert/novice group, we created four different conditions for each of the videos that systematically varied the amount of peripheral information that was available (see Figure 1). Two conditions were identical to our previous experiment: The *disk condition* was just a small black disk (diameter = 0.3 deg, contrast = 1) moving along the labeled puck position (see original paper for more details) in front of uniform gray background with no peripheral information available. In the *video condition*, the original, unaltered hockey clip was shown; therefore, all available peripheral information was presented. We created two intermediate conditions, where we showed only a cutout of the video around the labeled puck position. We created the cutouts by using a two-dimensional (2D) Gaussian window centered on the puck position with the original video presented in the center that then was turning into a grayscale image toward the periphery. We used a standard deviation of 3 deg for the *small context condition* and a standard deviation of 8 deg for the *large context condition*. We showed in our earlier paper (Goettker et al., 2021) that tracking behavior in the disk condition is independent of contrast, at least down to 10%. This renders it quite unlikely that the effects can be explained by contrast differences between the context conditions.

Importantly, participants were explicitly told that their task is to follow the target (a disk in the disk condition or the puck in the rest of the conditions). Each video was shown as a single trial. Each trial started with a central fixation cross, which was used for a drift check. Observers started the trial by looking at the cross and pressing the space bar at their own pace, then the video played for 10 seconds. After the video ended, a gray screen appeared, and a message appeared that indicated that the next trial was loading. The message stayed on the screen for roughly 5 seconds, during which the next video was loaded into Psychtoolbox to prepare for the next trial.

All observers completed all four conditions in a random order. The order was matched between the expert and novice groups such that each recruited expert saw the conditions in the same order as one of the recruited novices. In each condition, observers completed three blocks; in each block, all of the videos were shown in a random order. Each participant completed 144 trials (4 conditions $\times$ 3 blocks $\times$ 12 videos). One block took roughly 5 to 7 minutes, and participants completed all 12 blocks in one 1.5-hour
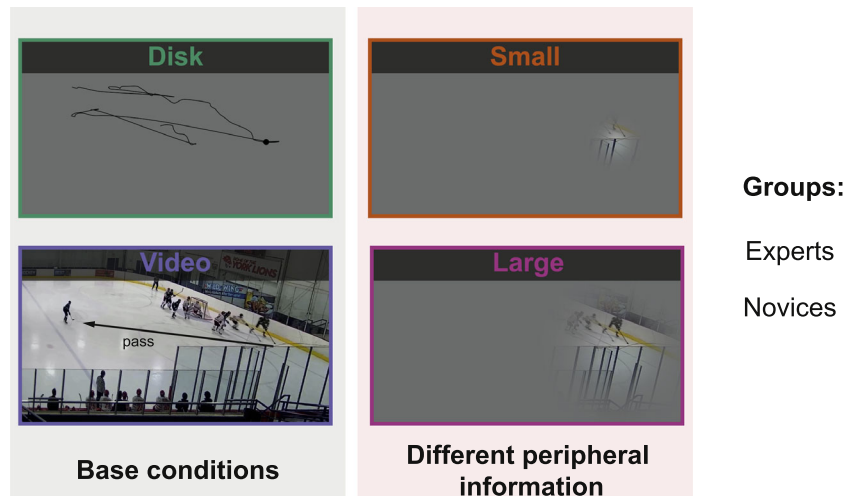
Figure 1. Conditions of Experiment 1. Depiction of the four different conditions. Disk and video conditions are the conditions previously tested in Goettker et al. (2021). In the disk condition, the whole trajectory of the puck is illustrated as an example, but please note that this was never visible to the observer. For the video condition, a pass between two players is illustrated, as these situations are a focus of the analysis. The small and large context condition manipulated the available peripheral information by showing only certain cutouts around the pass position. All four conditions were performed by each observer, but observers were either experts or novices with respect to ice hockey.

session. Across all participants, this led to a total of 3744 trials (2 groups × 13 observers × 144 trials).

Data analysis

Raw gaze positions were saved during each trial and analyzed offline with custom analysis scripts in MATLAB). Blinks in the data were linearly interpolated. Gaze position data were filtered with a second-order Butterworth filter with a cutoff frequency of 30 Hz. Afterward, eye velocity was computed based on the difference in horizontal and vertical eye positions in consecutive samples and was then converted into the absolute 2D eye velocity in deg/s by taking the square root of the squared differences and multiplying it by the sampling rate. Because the hand-leveled puck position (see Pidaparthy & Elder, 2019) was only available at the frame rate of the video (30 Hz), it was upsampled via linear interpolation to match the sampling rate of eye position. Saccades were detected based on the EyeLink criteria (velocity >30 deg/s and acceleration >4000 deg/s^2). Following that, for each trial we extracted different metrics related to different aspects of tracking performance.

For overall performance we computed the tracking error as the median Euclidean distance between gaze and puck that was computed at each sample between the current gaze position and the upsampled target position. We removed the first 500 ms from the analysis, because in the beginning of each trial participants needed to fixate in the center of the screen and initially

always had to search for the target. We also computed the number of saccades, and the proportion of the video spent in pursuit (defined as eye velocity > 3 deg/s for at least 100 ms with no saccade) as general measures of tracking performance.

To estimate the delay between eye and target movement we used a cross-correlation approach. We used a sliding window with a width of 500 ms. For each window, we computed the correlation between eye and target position for varying delays (−200 to +400 ms in steps of 10 ms) separately for horizontal and vertical eye positions. Then the window was shifted by 250 ms, and the procedure was repeated until the end of the video was reached. At the end of the video, the median correlation across the different sliding windows for each delay was computed. To account for autocorrelation within the data, we estimated the correlation within each clip by performing 1000 samples for 500-ms windows for eyes and pucks at random time points (for example, the eye vector after 1 second and the target vector after 8 seconds) again for horizontal and vertical eye position. We then subtracted those from the median. Finally, we took the average of the correlations obtained for horizontal and vertical movements to obtain one correlation value for each of the different delays between gaze and puck/disk movement for a given trial. This approach allowed us to quantify the delay between eye and target movement independent of the specific oculomotor response.

We also estimated the delay separately for saccadic and pursuit eye movements. For pursuit eye movements,

we also took a cross-correlation approach. But, instead of using a sliding window across the whole video, we calculated the cross-correlation for each pursuit segment found in the data. Pursuit segments were identified as time windows of at least 100 ms, where the eye velocity was above 3 deg/s and no saccade occurred. We did not estimate a pursuit delay when the overall proportion of pursuit for a respective observer and condition was below 0.2. To estimate the delay for saccades, we computed the relative position of each saccade endpoint with respect to the target position. For this, it was important whether the eyes were ahead of the target or lagged behind it, which cannot be captured by a simple distance metric. We therefore projected the saccade endpoint on a straight line connecting the target position 100 ms before saccade onset (the saccadic dead time) and at saccade end. This captures how far ahead or behind the eyes were with respect to the puck. To make different saccades more comparable across varying puck speeds, we converted these spatial errors into temporal errors. These specify at which time the puck would either reach the endpoint or when the puck had reached that endpoint. Across all conditions and subjects, eye movement data consisted of 9% saccades and 32% pursuit, with the remainder being considered as fixations, including small drift movements.

As a special example to visualize predictions, we analyzed 11 hand-labeled passes between players in more detail. As a measurement of prediction, we estimated the time of gaze arrival at the targeted player. We searched for the first instance around the time of the pass when gaze was closer than 3 deg to the final puck position. To control for the residual time that the puck still required toward the target location, we also measured the time that the puck arrived within 3 deg of the final location and subtracted this value from the estimated time of gaze arrival. For an exploratory analysis to investigate the influence of the available peripheral information on the arrival time at the pass target, we computed the arrival time for each pass separately and fitted a linear regression comparing the pass distance and the arrival time at the pass target. Here, the slope and intercept of this metric showed how the behavior at a pass was modulated by the interaction of pass distance and peripheral information.

To assess the accuracy of the prediction, we computed the error after the pass. We calculated the median gaze position between 200 and 300 ms after the pass offset and projected it onto the pass trajectory. The comparison of the actual endpoint and the projected endpoint gave us the error and the direction of the error was reflected in an overshoot or undershoot of the target position after the end of the pass.

Exclusion criteria and statistical analysis

Trials were excluded from the analysis based on two criteria: (1) if during the trial data were missing for more than 500 ms in a row, and (2) if the computed average error was larger than 10 deg. Based on these criteria, 3575 of 3744 trials (96%) were included in the analysis.

For each measure the average across all 12 videos across all three blocks was computed for each subject, so that there was one value per subject per condition. The obtained values were based on averages across values from each video (e.g., position error), each pass (e.g., time to arrive at pass target), or all saccades and pursuit segments that were present across all videos from one condition (estimate of saccade and pursuit delay). The averages were computed and then statistical analyses were performed in JASP. Participants were grouped by their expertise, and repeated-measure analyses of variance (ANOVAs) with the factor peripheral information (disk, small, large, video) and the intersubject factor expertise (novice, expert) were computed. If the sphericity assumption was violated, Greenhouse–Geisser correction was used to correct the degrees of freedom, and the corrected values are reported. Post hoc t-tests were used to assess differences between the conditions, and to correct for multiple testing the adjusted p value according to Holm was reported. The default option of pooled error terms for repeated-measures (RM) factors was kept active.

Results

The goal of Experiment 1 was to investigate the influence of expertise as well as the amount of available peripheral information. To quantify oculomotor behavior, we used different metrics; we start with reporting measurements quantifying changes in overall eye movement behavior, followed by computations of the delay between eye and target movements. In the last step, we take a closer look at eye movement behavior in pass situations, as they are a prime example to visualize predictive behavior.

Description of overall behavior

To quantify overall tracking behavior, we looked at three metrics: position error, number of saccades, and proportion of pursuit eye movements (see Methods for details). We assessed the influence of our manipulated variables with a repeated-measures ANOVA with the factor peripheral information (disk, small, large, video) and the between-subject factor expertise (novice vs. expert). For position error we observed a significant main effect of the available peripheral information, $F(1.583, 37.995) = 38.908$, $p < 0.001$, as well a main
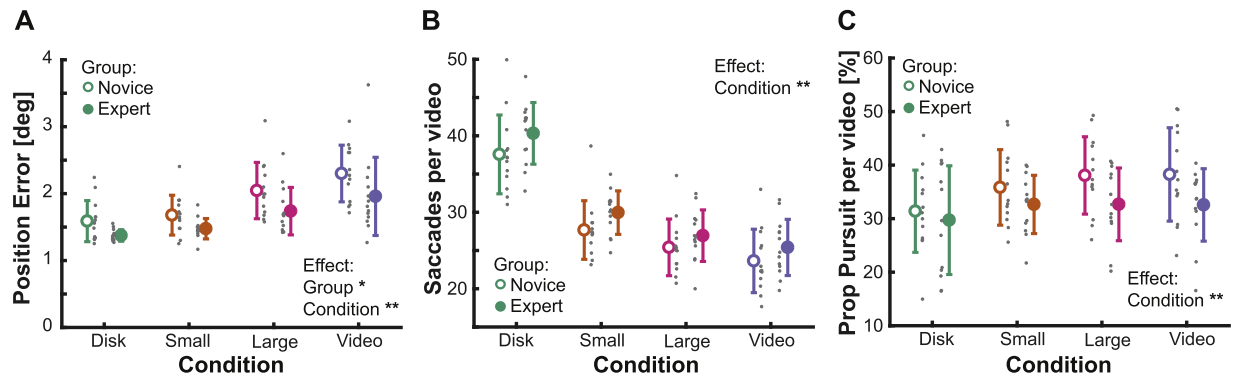
Figure 2. Overall eye movement behavior based on expertise and available peripheral information. (**A**) Average position error across the whole video is shown for the different conditions (different colors), separated based on the level of expertise (open vs. filled symbols). Circles show the average and standard deviation, and gray circles show the individual data points. Note here that individual data points are slightly jittered horizontally for illustration purposes, but sometimes can still overlap. Because expertise was a between-subject factor, the data points cannot be matched to each other. (**B**) Average number of saccades per 10-second video, depicted as in (**A**). (**C**) Proportion of time spend with pursuit eye movements per 10-second video, depicted as in (**A**).

effect of expertise, $F(1, 24) = 5.405$, $p = 0.029$, but no significant interaction, $F(1.583, 37.995) = 0.520$, $p = 0.557$). Experts had significantly lower error across all conditions (see Figure 2A), but the position error significantly increased with increasing peripheral information (for all comparisons, all $t > 3.531$ and all $p < 0.001$, except for disk vs. small, where $t = 1.414$, $p = 0.162$) with the highest position error when tracking the puck in the full video. This suggests a slight increase in position error when more complex visual information was available, which might have made it more difficult to find the puck or could suggest a different tracking strategy. However, please note here that just the distance between eye and puck only tells a part of the story. The distance alone does not reflect where observers were looking with respect to the puck movement. A similar average distance can be observed with a consistent lag behind the trajectory, as well as a consistent lead. We look at the tracking delay in more detail in the next section to assess this problem.

When investigating how many saccades and how much pursuit eye movements were used to track the puck, we observed an influence of the available peripheral information. The number of saccades, $F(1.911, 45.863) = 419.478$, $p < 0.001$ and the proportion of pursuit, $F(1.581, 37.954) = 8.153$, $p = 0.002$, were significantly affected. There was no significant influence of expertise: number of saccades, $F(1, 24) = 2.125$, $p = 0.158$; proportion of pursuit, $F(1, 24) = 2.287$, $p = 0.144$. Also, there was no interaction (both $F < 1.415$, both $p > 0.253$). The number of saccades decreased with additional peripheral information (for all comparisons across conditions, all $t > 3.711$, all $p < 0.001$) (see Figure 2B). This replicated our previous finding where we also observed roughly 30% fewer saccades from the disk to the video

condition. At the same time, the proportion of the video spent with pursuit eye movements increased (see Figure 2C), as the disk condition led to significantly less pursuit than all other conditions (all $t > 3.252$, all $p < 0.007$), but no other comparison reached significance (all $t < 1.026$, all $p > 0.926$). This suggests that additional, more complex peripheral information leads to smoother tracking, with fewer saccadic eye movements and more time spent with pursuit.

Delay in tracking behavior

After characterizing general oculomotor behavior, we computed what we consider the most interesting metric: the estimated delay between eye and target movements. We estimated the delay via a cross-correlation (see Methods for details). Figure 3 shows the average cross-correlation across all observer for each of the respective delays, separately for novices (Figure 3A) and experts (Figure 3B). Overall, the pattern looks comparable: The disk condition led to the latest peak, whereas the peak shifted toward zero with an increasing amount of peripheral information. To quantify the effects, we estimated the peak of the cross-correlation for each observer and again computed an ANOVA. We observed a significant effect of peripheral information, $F(2.184, 52.425) = 81.384$, $p < 0.001$, and expertise, $F(1, 24) = 9.180$, $p = 0.006$, but no significant interaction, $F(2.184, 52.425) = 0.861$, $p = 0.437$. Across all conditions, experts showed a significantly smaller delay when tracking the puck. For both groups, the delay reduced from around 160 ms in the disk condition to close to zero delay in the video condition (see Figure 3C). The differences between all levels of peripheral information were significant (all $t > 2.819$, all $p <$
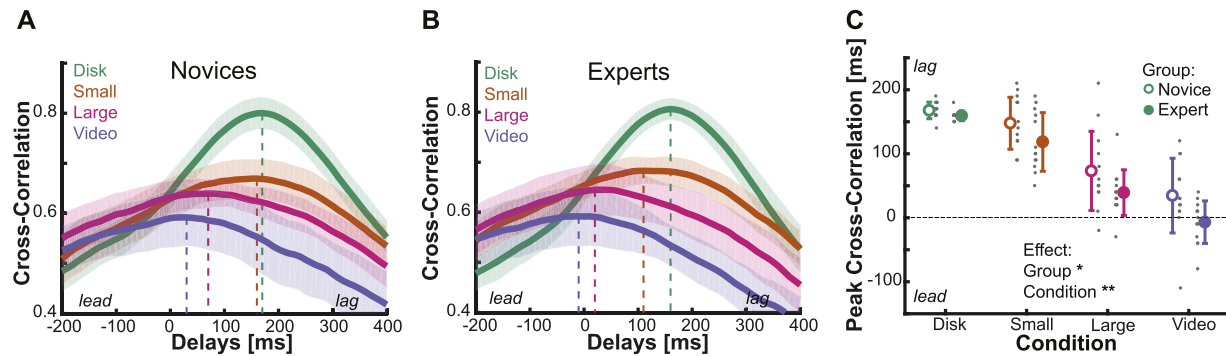
Figure 3. Estimated delay between eye and target movement. (**A**, **B**) Magnitude of the cross-correlation first averaged across all videos per observer and then across observers for different delays between eye and target movement, with data for novices (**A**) and for experts (**B**). The vertical dashed lines indicate the peak of the cross-correlation, which serves as an estimate of the average delay between eye and target movements. The different colors indicate the different conditions. The shaded area shows the standard deviation across observers. (**C**) The peak of the cross-correlation for each observer averaged across the different conditions (color-coded) and level of expertise (open vs. filled symbols). Circles show the mean and standard deviation, and gray dots indicate individual data. Note here that individual data points are slightly jittered horizontally for illustration purposes, but sometimes can still overlap.

0.006), suggesting an efficient use and integration of additional peripheral information and contextual cues that are available. Even only the little additional information that was present in the small condition led to a significant improvement over the disk condition, suggesting that there were some cues around the puck (e.g., the skate movements of a player) that allowed better predictions. There was a lack of significant interaction between expertise and the available peripheral information, but experts reached the minimum delay that novices showed when watching the full video when watching only the video with the large cutout (see Figure 3C), and then kept improving further when having access to all available information.

Delay for saccadic and pursuit eye movements

We observed a reduction of the average delay when observers had access to more information. Such an improvement could come from more predictive saccadic eye movements or could be achieved by more accurate pursuit. To address this point, in the next analysis we tried to estimate the contributions of saccadic and pursuit eye movements separately (see Methods for more details). For both saccadic and pursuit eye movements (see Figure 4), we again found a significant influence of the available peripheral information, $F(1.854, 44.502) = 167.527, p < 0.001$ for saccades; $F(2.713, 51.553) = 28.556, p < 0.001$ for pursuit). For both eye movements we observed the same pattern: With an increasing amount of peripheral information, the estimated delay decreased. However, although the metrics are not directly comparable (for saccades, it

represents the relative position of the saccade landing position to the target; for pursuit, the estimate is based on cross-correlation over a longer time period), the overall difference in scale is interesting. It seems that even in the disk condition, saccades only landed slightly behind the target, but landed more and more ahead of the target with increasing information (see Figure 4A). In contrast, pursuit eye movements show the expected lag of around 170 ms in the disk condition, but then this delay decreased close to zero in the video condition (see Figure 4B). For the separate estimates of pursuit eye movements, we did not observe a significant effect of expertise, $F(1, 24) = 2.360, p = 0.141$, nor an interaction, $F(2.713, 51.553) = 1.093, p = 0.357$, with peripheral, although the pattern of results looks similar to the overall estimates (compare to Figure 3C). For saccadic delays we did observer a significant effect of expertise, $F(1, 24) = 5.378, p = 0.029$, with experts showing lower delays across the conditions. In addition, there was also a significant interaction, $F(1.854, 44.502) = 4.986, p = 0.013$, which was driven by comparable estimated delays for the disk condition, but then an increasingly larger difference between experts and novices the more peripheral information became available (see Figure 4A), suggesting that experts benefited more from the additional peripheral information. The estimated delays for saccadic and pursuit eye movements were highly correlated across the different peripheral conditions, $r(49) = 0.62$, $p < 0.001$ for novices; $r(46) = 0.67, p < 0.001$ for experts (see Figure 4C), thus suggesting similar use of additional information for saccadic and pursuit control, although saccadic eye movements seemed to be more strategically placed by expert observers.
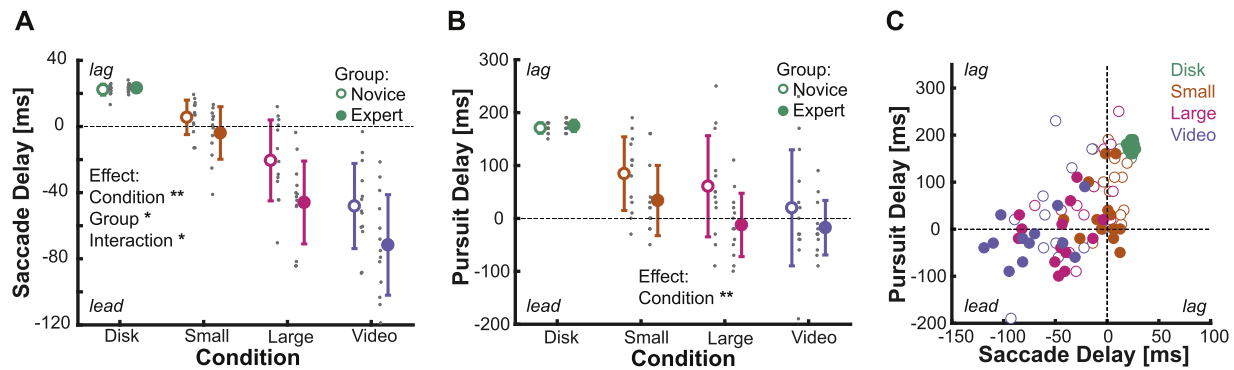
Figure 4. Separate estimate of delay for saccadic and pursuit eye movements. (**A**) The delay with respect to the target estimated for saccadic eye movements. Averages across subjects are shown for the different conditions (color-coded) and level of expertise (open vs. filled symbols). Gray data points show individual data. Note here that individual data points are slightly jittered horizontally for illustration purposes, but sometimes can still overlap. Error bars depict the standard deviation. (**B**) Same depiction as in (**A**), but this time for the estimate of pursuit delay. (**C**) Estimated delay for saccadic and pursuit eye movements for each observer plotted against each other. Open symbols indicate data from novices; filled symbols, data from experts. The color of the data points reflects the respective condition.

Passes as examples for prediction

To visualize predictive behavior better, we took a closer look at one specific situation in ice hockey games: passes between players. In a pass situation, a successful prediction allows one to already look at the receiving player before the puck arrives, which helps to analyze what is going to happen next. To quantify that, we computed when gaze arrived at the receiving player (see Figure 5A). We again found a significant impact of peripheral information, $F(1.842, 44.204) = 102.481$, $p < 0.001$, as well as a significant influence of expertise, $F(1, 24) = 5.807$, $p = 0.024$, with no interaction, $F(1.842, 44.204) = 0.298$, $p = 0.726$. The effects were again in the expected directions. Gaze arrived earlier at the targeted player for experts as well as with additional peripheral information (significant differences across all conditions, with all $t > 4.365$ and all $p < 0.001$).

These results suggest that observers were able to use the very little additional context in the small context condition to make some predictions about where the pass will go. However, this little context should only help when the targeted player becomes visible. To check for that, we computed the average arrival time across observers for each of the labeled passes based on the distance of the pass on the screen (see Figure 5B for the novice data; expert data look similar). We then fitted linear regressions between the distance of the pass and the arrival time for each individual observer and condition. We computed repeated-measures ANOVAs with the factor peripheral information (disk, small, large, video) and the between-subject factor expertise (novice vs. expert) for the slope and intercept of the regressions. For the slope, we observed a significant

effect of peripheral information, $F(1.653, 39.679) = 17.682$, $p < 0.001$, and no influence of expertise, $F(1, 24) = 1.765$, $p = 0.197$, nor an interaction, $F(1.653, 39.679) = 0.685$, $p = 0.483$. Post hoc t-tests showed that the effect was driven by the significantly higher slopes for the small (both $t > 5.31$, both $p < 0.001$) and the large (both $t > 3.53$, both $p < 0.002$) condition compared to the disk or the video condition. There was no difference between the video and disk condition, $t(25) = 1.180$, $p = 0.279$, nor between the small and the large condition, $t(25) = 1.768$, $p = 0.163$. For the intercept we again observed a significant impact of peripheral information, $F(2.293, 55.025) = 48.379$, $p < 0.001$, but no influence of expertise, $F(1, 24) = 1.053$, $p = 0.315$, nor an interaction, $F(2.293, 55.025) = 0.443$, $p = 0.671$. The effect was driven by a significantly higher intercept in the disk condition (all $t > 8.404$, all $p < 0.001$), compared to all other conditions, but there were no more differences across the rest of the conditions (all $t < 2.343$, all $p > 0.066$). Together, these results show that the disk and video condition mainly differed in the intercept but had on average very little slope (see Figure 5C). This is an interesting finding for the video conditions, as it suggests that even information in the far periphery can be integrated successfully for predictive eye movement. The two intermediate conditions clearly showed an influence of the pass distance. Both conditions showed positive slopes, indicating that the larger the distance of the pass the lesser the predictive advantage. This was to be expected, because for long pass distances the targeted players simply only became visible quite late.

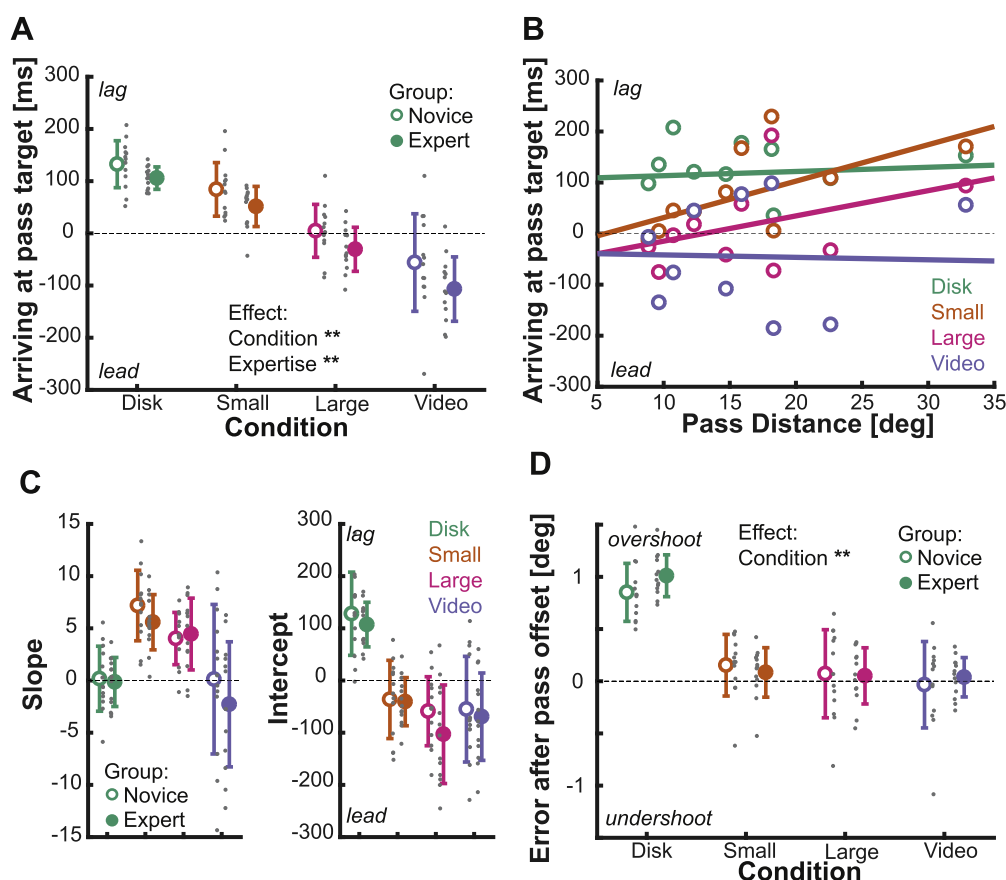Not only did the observers use the contextual cues available to arrive at the targeted player earlier, but

Figure 5. Eye movement behavior in pass situations. (**A**) The average times observers arrived at the pass target are shown for the different conditions (color-coded) and level of expertise (open vs. filled symbols). Gray data points show individual data. Note here that individual data points are slightly jittered horizontally for illustration purposes, but sometimes can still overlap. Error bars depict the standard deviation. (**B**) Average arrival time across novice observer for each of the labeled passes and their respective distances. Solid lines show linear regression to the average data points. The different colors indicate the different peripheral conditions. (**C**) Average slope (left) and intercept (right) of the linear regression for pass distance and arriving time, otherwise depicted as in (**A**). (**D**) Directional error after pass offset. Negative values reflect an undershoot of the pass distance and positive values an overshoot, depicted the same as in (**A**).

they also arrived there more accurately (see Figure 5D). We computed the directional error 200 ms after pass offset (see Methods for details), which has the benefit of allowing us to assess whether there was a systematic overshoot of the targeted player. For the error after the pass, the effect of expertise disappeared, $F(1, 24) = 0.160$, $p = 0.693$, but we still found a significant influence of peripheral information, $F(2.754, 66.096) = 95.395$, $p < 0.001$. This suggests that novices and experts were able to identify the targeted player at this point and fixated at them, but this was only possible for conditions with naturalistic information (disk vs. other conditions, all $t > 12.816$, all $p < 0.001$). When seeing only the disk, observers tried to extrapolate the target trajectory, but they could not predict when the pass would stop, so they systematically overshot the targeted player. However, even the small context provided enough available information (for all other

comparisons, all $t < 1.840$ and all $p > 0.210$) for an accurate fixation after pass offset.

Experiment 2. The role of positional and kinematic cues

Methods

The setup and data analyses were as in Experiment 1. The depiction of conditions can be seen in Figure 6.

Observer

Fifteen new observers (12 identified as female, three as male) took part in Experiment 2. All observers in
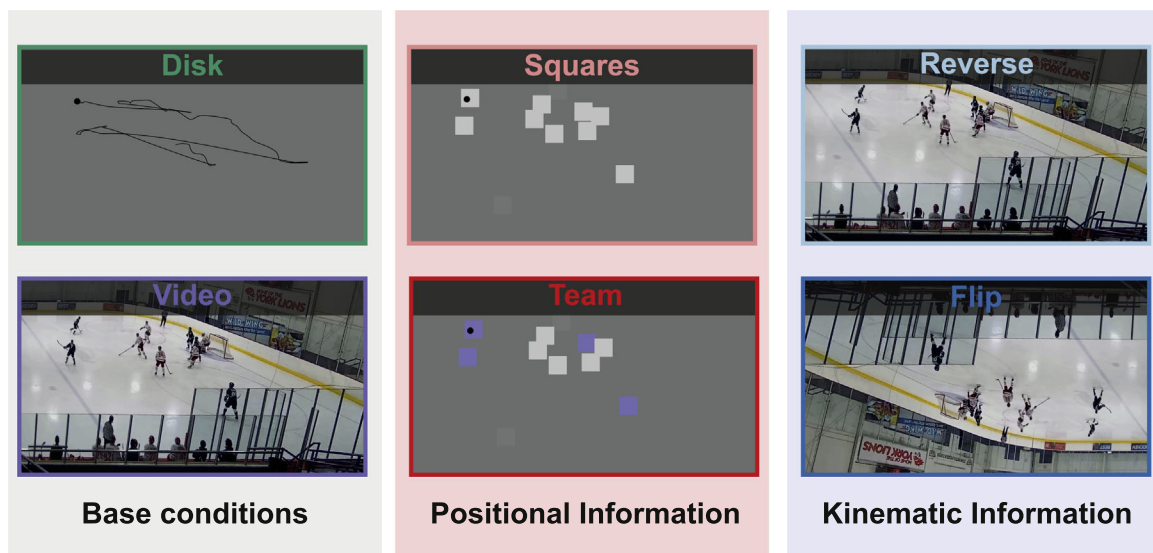
Figure 6. Conditions of Experiment 2. Depiction of the six different conditions. Disk and video conditions were identical to Experiment 1 and our previous study. As an example, the whole trajectory of the puck is illustrated in the disk condition, but note that this was never visible to the observer. For the positional conditions, the player positions were replaced by squares to provide positional information about the players. In the team condition, the same information was presented, but additional information about the player identity was available. For the kinematic condition, the original video was shown, but this time it was either temporally reversed or flipped vertically. This manipulated the reliability of kinematic information about player movement and, especially in the reverse condition, even changed the causal structure of movements. All six conditions were performed by each observer in random order.

Experiment 2 were novices. The age of the observers ranged from 19 to 54 years with a mean of 26.20 years. All participants were naïve to the purpose of the study and gave informed consent at the beginning of the experiment. The experiment adhered to the tenets of the Declaration of Helsinki and was approved by the local ethics committee. Participants were paid 8 euro/h for their participation.

Stimuli and conditions

As in Experiment 1 we again used 12 different 10-second videos of ice hockey games across the different conditions. We kept the *disk condition* and the *video condition*, but, because the focus of Experiment 2 was the role of positional and kinematic cues, four additional conditions were added.

To manipulate the positional cues, we made use of additional labels of the position of players in the videos (Pidaparthy & Elder, 2019). In one condition (*squares*) we replaced each player with a white square (contrast = 0.6, size = 1 deg), so that positional information about the rough position of the players was available but no information about their team or their kinematic cues like body movements. To increase the positional cues, we also had a condition (*teams*) with squares, but this time they were colored according to the team of their respective identity (players with dark jerseys were

shown in blue, the white team in white, and referees in light gray).

To manipulate kinematic cues, we showed observers the original hockey videos but one time flipped vertically (*flip*) or in reverse (*reverse*). Both conditions made it more difficult to process the kinematic cues of player movements, but the reverse condition was especially difficult, as it also changed the causal structure of puck movements; for example, when a pass between players started, the player who seemed to start the pass did not perform a physically correct movement at this time.

Again, across all conditions observers were explicitly told that their task was to follow the target (disk in the disk or positional conditions or puck in the video and kinematic conditions). Participants started each trial by pressing a space bar when looking at a central fixation cross, which was then followed by a video for 10 seconds. At the end of the video, a small break occurred where the next video was prepared.

All observers completed all six conditions in random order. In each condition observers finished three blocks and within each block all 12 sequences were randomly presented. Each participant completed 216 trials (6 conditions × 3 blocks × 12 videos). One block took roughly 5 to 7 minutes, and participants completed all 18 blocks in two 1-hour sessions, typically separated by multiple days. Across all participants this led to a total of 3240 trials (15 observers × 216 trials).
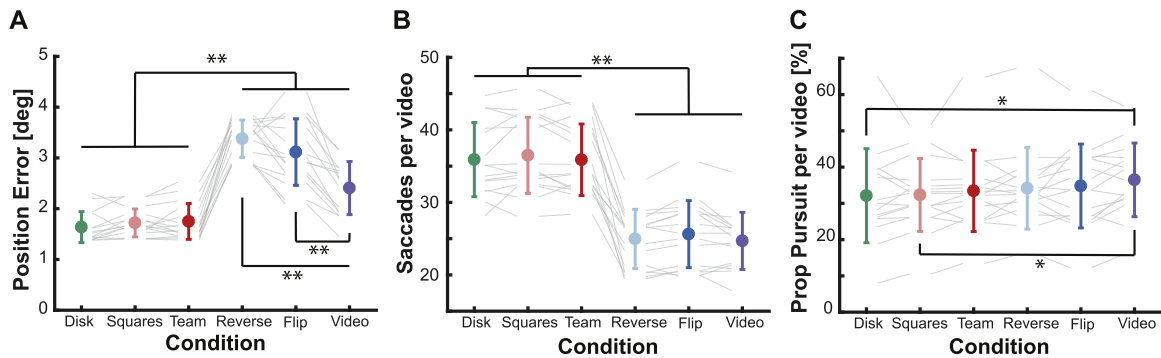
Figure 7. Overall eye movement behavior for different positional and kinematic cues. (**A**) Average position error across observers shown across the different conditions (color-coded). Circles show the mean and standard deviation, and gray lines show the individual subjects across all conditions. (**B**) Average number of saccades per video across the different conditions, depicted as in (**A**). (**C**) Average proportion of pursuit eye movements across the different conditions. Black lines indicate significant comparisons. **$p < 0.001$, *$p < 0.05$.

Exclusion criteria and statistical analysis

Exclusion criteria were identical to the first experiment: (1) if, during the trial, data were missing for more than 500 ms in a row, and (2) if the computed average error was larger than 10 deg. Based on these criteria, 3141 of 3240 trials (92%) were included in the analysis.

We computed the same metrics and averages as for Experiment 1 and then statistical analyses were performed in JASP. Repeated-measures ANOVAs with the one factor of condition (disk, squares, team, flip, reverse, video) were computed. If the sphericity assumption was violated, degrees of freedom were Greenhouse–Geisser corrected, and corrected values are reported. To assess differences between the conditions, post hoc t-tests with pooled error terms for the repeated measure factor were used. Adjusted p values according to the Bonferroni–Holm method are reported to correct for multiple comparisons.

Results

The goal of Experiment 2 (see Figure 6) was to investigate the influence of positional cues and kinematic information on predictive eye movements. We again begin by reporting metrics related to overall eye movement behavior and then look at the critical variable of the estimated delay between eye and target movements. To visualize the predictive behavior, we end with a look at eye movement behavior in pass situations.

Description of overall behavior

Overall, our conditions significantly influenced the overall tracking error, $F(2.259, 31.626) = 98.905$, $p < 0.001$. Observers were more accurate in conditions when they could track a disk target (see Figure 7A) and

had a gray background (all $t > 6.058$, all $p < 0.001$). The additional squares representing the player positions did not lead to a difference in the position error (all $t < 1.03$, all $p > 0.916$). Although the kinematic conditions and the video condition overall had the higher position error, the unaltered video condition still had a significantly lower error than the conditions where we manipulated the video (both $t > 6.536$, both $p < 0.001$). Again, please note here that the distance error by itself does not reflect where the eyes were with respect to the target trajectory; to assess this we again take a look at the estimates of delay in the next section.

A similar pattern to the results for the distance error was present for the number of saccades across the conditions; for main effect, $F(1.290, 18.066) = 74,397$, $p < 0.001$. The disk and positional conditions had comparable number of saccades (all $t < 0.627$) and were significantly different from the kinematic conditions and the original videos (all $t > 10.359$, all $p < 0.001$) (see Figure 7B). However, the latter ones also had comparable numbers of saccades (all $t < 0.939$). For proportion of pursuit, there was substantial variability across observers with averages ranging from 10% to 60%. There was a significant main effect across all conditions, $F(3.070, 42.980) = 3.255$, $p = 0.030$, but only the comparison of the disk and the square condition with the video condition reached significance (both $t > 3.235$, both $p < 0.026$; for all other comparisons, $t < 2.363$ and $p > 0.272$). There was a similar trend as in the first experiment: The video condition seems to led to a larger proportion of pursuit ($M = 36.48\% \pm 10.17\%$) than the disk condition ($M = 32.11\% \pm 12.97\%$).

Delay in tracking behavior

After looking at some broader eye movement characteristics, we again looked at the most interesting variable: the estimated tracking delay. As in
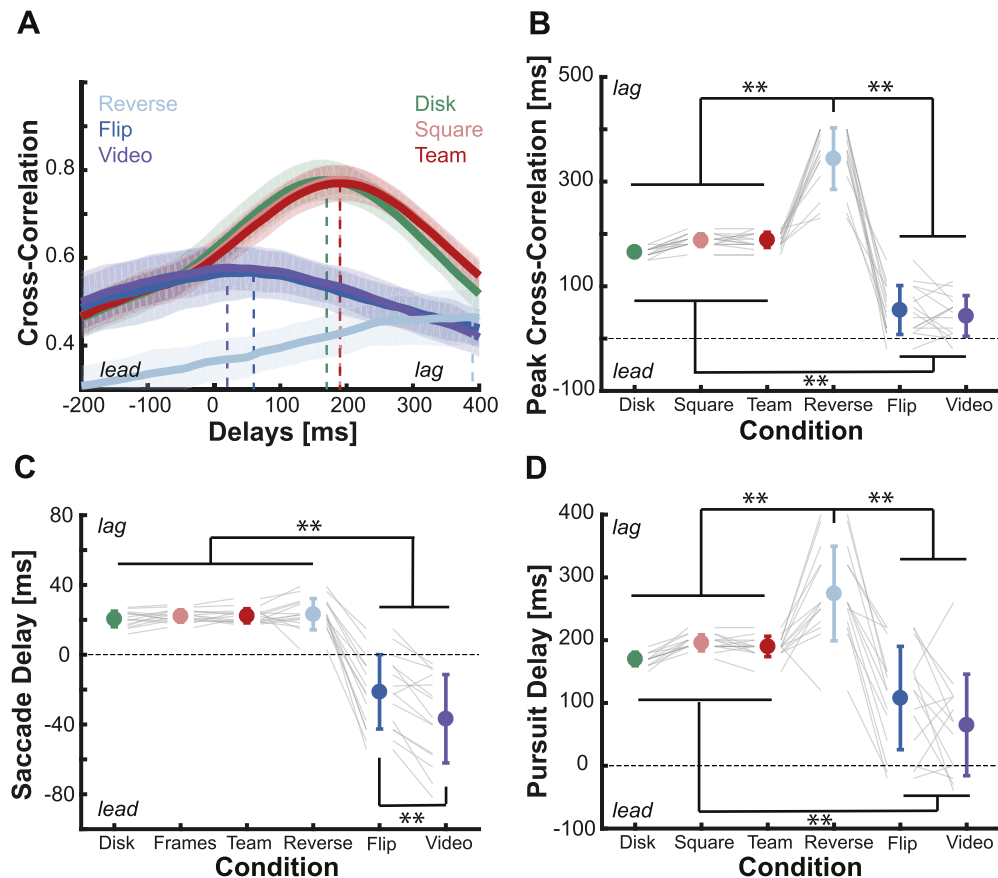
Figure 8. Tracking delay for different positional and kinematic cues. (**A**) Cross-correlation for different delays between eye and target movement averaged across observers for each condition (color-coded). The vertical dashed lines indicate the peak of the cross-correlation, which serves as an estimate of the average delay between eye and target movement. The shaded area shows the standard deviation across observers. (**B**) The peak of the cross-correlation across the whole video for each observer and each condition (color-coded). Circles and error bars show the means and standard deviations. Gray lines depict the individual values of observes. (**C**, **D**) Estimates of the delay for either saccadic (**C**) or pursuit (**D**) eye movements, depicted as in (**B**). **$p < 0.01$.

Experiment 1, we computed cross-correlations across varying delays across all conditions (see Figure 8A). To quantify the effect, we computed the peak of the cross-correlation for each observer and compared them across the conditions. There was a separation of the conditions into three different groups (see Figure 8B). The positional conditions (squares and team) led to a delay comparable to that of the disk condition (all $t < 1.853$, $p > 0.272$). The reverse condition was significantly worse than all other conditions with an average delay of 344 ms ($SD = 58.895$, all $t > 12.337$, all $p < 0.001$). The flip condition led to a similar delay as the regular video condition, which was significantly lower than the disk and positional conditions (all $t > 8.790$, all $p < 0.001$). Therefore, the positional cues available in the positional conditions did not seem to help to produce predictive eye movement behavior. If the causal structure of kinematic cues was impaired, the tracking delay became even higher than in the disk condition, where no cues were available. A vertical flip

of the video, with intact causal structure but more difficult kinematic cues, still led to the significant prediction advantage we observed for the original videos.

As in Experiment 1, we also tried to estimate the delay separately for saccadic (see Figure 8C) and pursuit eye movements (see Figure 8D) to investigate their respective contributions. We found comparable patterns for both eye movements; the positional conditions led to a similar lag as the disk condition (all $t < 0.428$ for saccades; all $t < 1.346$ for pursuit). The flip condition showed a similar predictive advantage as the video condition (although it was slightly lower than the video condition for saccades, with $t = 3.622$ and $p = 0.004$). Interestingly, the reverse condition showed similar estimates of saccadic delay to the positional conditions (all $t < 0.645$), suggesting that saccades in this condition show a similar reactive behavior as for the disk. However, for the pursuit estimates the reverse condition again showed the highest lag (all $t > 3.775$,
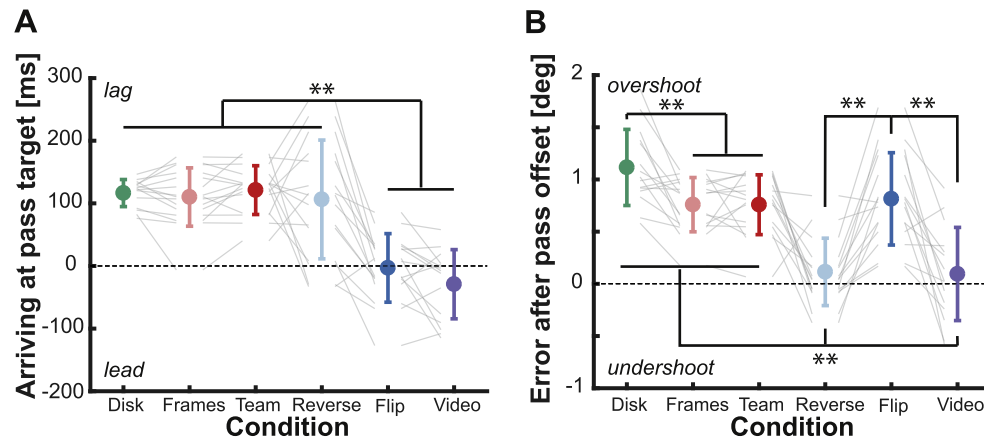
Figure 9. Eye movements in pass situations. (**A**) Average time observers arrived at the pass target across the different conditions. Circles and error bars show the means and standard deviations, and gray lines indicate the individual values. (**B**) Directional error 200 ms after pass offset. Positive values reflect an overshoot, negative values an undershoot. Depiction is the same as in (**A**). **$p < 0.01$.

all $p < 0.002$), suggesting that the impaired structure of kinematic cues also led to an impairment of pursuit performance. It might be that, to gain a predictive benefit for pursuit eye movements, it is critical to be able to make longer predictions into the future, which is difficult for the reverse condition.

Passes as examples for prediction

To look at the predictions in more detail, we again analyzed pass situations. Comparable to the results for saccadic delays, we saw that observers looked at the pass target roughly 100 ms after the puck arrived (see Figure 9A), suggesting a reactive gaze movement. However, for the flip and original video condition, we again observed the predictive advantage (all $t > 6.218$, all $p < 0.001$). Similar to Experiment 1, we observed that, when comparing the disk and video condition, observers used the contextual cues to arrive earlier at the target with greater accuracy (see Figure 9B). The positional conditions, led to significantly lower overshoot than the disk condition (both $t > 2.780$, both $p < 0.048$), suggesting that in pass situations observers made use of the available positional cues. The flip condition produced a similar error as the positional conditions (all $t < 2.347$, all $p > 0.109$). The reverse condition led to no systematic over- or undershoot, potentially suggesting that the motion cue of a pass and the pass situation in general helped observers to find the puck and look and fixate at the targeted player.

Discussion

A critical assumption in oculomotor research is that it is possible to generalize results obtained with simple, artificial stimuli to eye movement control in naturalistic, complex scenes. However, when we directly compared tracking performance for the same trajectories with simple stimuli and in complex naturalistic scenes (Goettker, Agtzidis, Braun, Dorr, & Gegenfurtner, 2020; Goettker et al., 2021), we found that observers leverage additional information present in naturalistic scenes to achieve less tracking delay in their eye movements. But which kind of information is used by the oculomotor system to gain that predictive advantage? Here, we studied four potential influence factors: expertise, the amount of available peripheral information, and positional and kinematic cues. Across most of our measurements (e.g., position error or predictive behavior in pass situations) (see Figure 2A and Figure 5A), experts showed significantly better performance than novices. This suggests a role of experience with the complex stimuli that allowed a better integration of the available information. In general, the more peripheral information that became available, the better the predictions of all observers (see Figure 3). Additionally, when it comes to additional peripheral information, experts seemed to be able to use them more efficiently for their saccade placement (see Figure 4A; see also the next section for a more detailed discussion of expertise). When we manipulated the reliability of positional and kinematic cues in the video, the most critical component seems to be access to intact kinematic cues from players. An artificial depiction of the player positions in the video did not lead to a predictive advantage (Figure 8) and was of only limited help in pass situations (Figure 9B). When the causal structure of kinematic cues was lost by playing the video in reverse, the tracking behavior was severely impaired, and the lag became even higher than without any contextual cues (Figure 8). Only the

condition where kinematic cues were more difficult to process but still intact due to vertical flip of the video led to similar predictive behavior that was observed when watching the original videos (see Figure 8). Together, these results demonstrate that, when complex, naturalistic information is available, the oculomotor system is successfully integrating additional cues and is not relying only on low-level information about the target trajectory.

The role of expertise

Across a broad range of research areas, the comparison of experts and novices has provided interesting insights into how experience changes behavior (Brams et al., 2019; Gegenfurtner, Lehtinen, & Säljö, 2011; Memmert, 2009; Vickers, 2009). Especially in sports, there has been a focus on differences in eye movement behavior between experts and novices. For example, differences have been overserved in cricket (Land & McLeod, 2000), badminton (Abernethy & Russell, 1987), chess (Sheridan & Reingold, 2014; Sheridan & Reingold, 2017), soccer (Casanova et al., 2009), table tennis (Ripoll, Fleurance, & Cazeneuve, 1987), and many more situations (for a review, see Brams et al., 2019). The general result seems to be that experts are better at identifying and predicting important information and therefore can allocate their gaze quicker to important locations (Brams et al., 2019). Although most of these findings were based on comparisons of active players, similar effects have also been seen when simply observing a game passively, which is more similar to the experiments we conducted (Smuc, Mayr, & Windhager, 2010).

Our results match well the current understanding of expertise effects. In our experiment, the experts showed lower tracking errors and more predictive behaviors than novices, revealing a more efficient tracking strategy. For example, in pass situations experts are better in predicting where the pass will go. This allows them to fixate the receiving player to analyze what is going to happen next. Similar observations have been made for other sports, as well (Vansteenkiste, Vaeyens, Zeuwts, Philippaerts, & Lenoir, 2014; Williams, Davids, Burwitz, & Williams, 1994). Although for most measurements there was no interaction between the available peripheral information and expertise, for the placement of saccadic eye movements experts benefited more than novices from additional peripheral information (see Figure 4A). Therefore, next to a general improvement of tracking performance in the expert group, this suggests a superior ability to make sense of and recognize potential peripheral cues, such as the position of player or formations (Gorman, Abernethy, & Farrow, 2013), and use those for better predictive saccadic eye movements.

Although the expert group outperformed the novice group, the novices also benefited significantly from additional complex peripheral information and did quite well in the task. Across both experiments presented here, and in our previous study (Goettker et al., 2021), novices were able to track the puck with on average close to zero delay when viewing the original ice hockey clips. This clearly demonstrates that novices can successfully integrate complex peripheral information to guide predictive eye movement behavior (Vig et al., 2011). The expertise effect in our task (see next section for the potential role of task constraints) did not lead to a fundamental difference in tracking behavior; it merely reflected domain-specific training of this predictive behavior that shows up with many years of experience in difficult situations (Tenenbaum, Sar-El, & Bar-Eli, 2000). In naturalistic tasks and contexts, it seems to be a common feature of the oculomotor system to use knowledge about the world and task to perform predictive eye movements (Diaz et al., 2013; Goettker et al., 2021; Land & Furneaux, 1997; Land & Hayhoe, 2001; Sullivan et al., 2021).

Integration of peripheral information

One interesting aspect of our experiment is that, due to the task constraint of tracking the puck, experts might have even been at a disadvantage. A free-viewing study of soccer games showed that experts in general seem to fixate more on the position away from the ball, such as the position of other players, whereas novices fixated on the ball more frequently (Williams et al., 1994). Therefore, peripheral information became even more important in our task than it typically might be to predict what is going to happen next. Peripheral processing is a complex topic (Rosenholtz, 2016), but in general resolution drastically decays with distance to the fovea (Grünert & Martin, 2020; Strasburger, Rentschler, & Jüttner, 2011) with additional effects of clutter and crowding (Rosenholtz et al., 2007).

Our manipulations of the viewing window size revealed that cues for guiding predictive eye movements are available at multiple different distances. We observed a small predictive advantage when approximately 3 deg of the video around the puck position was available. A distance of 3 deg to the fovea has been suggested as being the critical area based on well-known research on the quiet eye (Vickers, 2016). Quiet eye is defined as the ability to quickly bring and keep the gaze close to the target before an action toward it, and it has been related to performance in different tasks. From a vision science perspective, the importance of the central 3 deg distance seems rather arbitrary, as 3 deg away from the fovea resolution is already less than half of the foveal resolution and the eye is never really quiet (see Spering & Schütz, 2016). However, we think our results showed that giving observers access to 8 deg around the target

allowed for a significantly better prediction. This result provides direct evidence that information farther out in the periphery is also used and is important for guiding behavior. Because quiet eye is focused mostly on manual interception toward the fixated object, the use of such eccentric information to guide future eye movements to critical locations might not be directly comparable. However, making correct predictions about what is happening next is certainly highly important for our interactions with the environment. Additionally, we even observed an improvement in predictions when we showed the original video. Therefore, even information farther out than 8 deg seems to be useful. Thus, it seems that the oculomotor system is leveraging all available cues to improve tracking performance.

The role of positional and kinematic cues

Research on predictions in complex, naturalistic environments, such as sports, has mainly focused on two areas: kinematic cues and positional factors (Loffing & Cañal-Bruland, 2017). Kinematic cues are related to information about the body movement of other players (which we manipulated with the flip and reverse condition), and positional cues are related to knowledge about the position (which we manipulated in the squares and team condition) of an actor or their preferences.

Detecting and analyzing biological motion are fundamental skills of human observers (Johansson, 1973; Neri, Morrone, & Burr, 1998). The high sensitivity of human observers to biological motion became visible when results showed that even only a few visible light points attached to the joints of a walking human (so-called point-light displays) allow immediate recognition of a walking person. These limited displays even allow making judgments about the sex (Kozlowski & Cutting, 1977) or emotional state (Atkinson, Dittrich, Gemmell, & Young, 2004) of the walker. The ability to use biological motion cues to make correct predictions about future behavior has been extensively studied in the sport context (Casanova et al., 2009; Loffing & Cañal-Bruland, 2017; Müller & Abernethy, 2012). For example, tennis players can use the movement of the opponent to make predictions about where they will hit the ball (Shim, Carlton, Chow, & Chae, 2005). In our study, we chose to manipulate the quality of available kinematic information. Research with point-light walkers have shown that it is more difficult to recognize movements that are either played in reverse (Pavlova et al., 2002) or inverted (Pavlova & Sokolov, 2000). In the flip condition it was more difficult to analyze the biological motion cues, but the structure of the scene was otherwise intact. For the flip condition, even naïve observers in Experiment 2 could leverage the kinematic cues to gain a comparable predictive advantage as for the original video (see Figure 8). Interestingly,

when we played the video in reverse, not only were individual kinematic cues more difficult to interpret, but also the causal structure of the movements was lost. For example, in a pass situation, the hitting movement of the stick did not start a pass but actually was now only visible after the player received a pass with no consequences. This impairment of causality really seemed to confuse the observations reflected in only reactive saccades (Figure 8C and Figure 9A), and resulted in an overall much higher tracking lag than in the disk condition without any additional cues. Unexpected and surprising events are known to increase fixation durations (Vo & Henderson, 2009), which in dynamic scenes could be related to higher tracking delays. Together, such observations demonstrate that even naïve observers can successfully use kinematic cues to guide their eye movements. The underlying biological motion processing can deal with more difficult situations (e.g., an unnatural orientation), but the use of these cues breaks down when there are conflicts with our expectations of causal structure and our understanding of intuitive physics (McCloseky, 1983).

As mentioned earlier, next to kinematic cues a second helpful source of information can be positional cues (Loffing & Cañal-Bruland, 2017) or other situational information (Schläppi-Lienhard & Hossner, 2015). Previous research has demonstrated in squash that experts are able to predict above chance where an opponent will hit the ball, even before the opponent actually starts to move (Abernethy, Gill, Parks, & Packer, 2001). Such predictions can be based on the position of the opponent (Loffing & Hagemann, 2014), domain-specific and situation-specific knowledge (Schläppi-Lienhard & Hossner, 2015), or information about the structure and organization of other players on the field (Casanova et al., 2009; Williams, 1993; Williams Hodges, North, & Barton, 2006). We aimed at providing such positional cues by showing information about the player position as simple squares. However, in both positional conditions and whether or not the team identity of the players was visible, the cues led to no predictive advantage in eye movement behavior. Across all metrics, performance was very similar to the disk condition, where no additional cues were present. Only in the error directly after a pass was there a slight improvement (see Figure 9B). This could mean that observers were able to use the positional cues to fixate more accurately on the player position after the pass. Why the positional cues did not lead to a predictive advantage in our experiment could be based on three reasons.

First, because the player positions were detected automatically (Pidaparthy & Elder, 2019) and were not manually labeled, they were not perfect. Especially when players were close together, occasionally one of the players could disappear. Additionally, because

we made the 2D sizes of all squares constant, the three-dimensional sizes of the players at different distances were not accurately represented. This might have reduced the reliability of the perceived player positions and led to a lower weighting of this potential cue. They did not seem to serve as a distractor, as observers still achieved the same performance as in the disk-only condition. Second, most of the studies reporting positional effects investigated experts in the respective sport. Because in our second experiment observers were again novices, it could have been that they simply were not able to make sense of certain player configurations and use those to their advantage. Third, as the result for the kinematic cues already shows, most of the predictive advantage seems to be based on the analysis of player movements to predict where they will go next. Also, the stick of the players, which was especially critical to predict the occurrence and direction of a pass, was never visible. Thus, it seems that kinematic cues were just much more relevant for more precise predictive eye movements in our task, as they included the information about the future direction of a player movement or the direction of a pass.

Oculomotor control in naturalistic scenes

Across all conditions, the most striking differences in eye movement behavior were how tracking delays changed depending on expertise, available peripheral information, and other available contextual cues. However, in addition to the tracking delay, there were substantial other differences. For example, the number of saccades decreased significantly when the puck was embedded in a naturalistic context. In our previous study, we interpreted this effect as the need for fewer small corrective saccades due to better predictions (Goettker et al., 2021). However, based on these new results, we must revise that statement, because in the reverse condition (see Figure 7), where tracking delay was the highest across all tested conditions, the number of saccades was also significantly reduced. Therefore, the lower number of saccades is presumably more related to the complex visual input and in turn the less clearly defined position targets (Heinen, Badler, & Watamaniuk, 2018). The reduction in number of saccades was often accompanied by more active pursuit eye movements (see Figure 2). Also, the delay estimates separately for saccadic and pursuit eye movements were correlated (see Figure 4C), which suggests a tight coupling between the two movements (Goettker & Gegenfurtner, 2021; Krauzlis, 2004; Orban De Xivry & Lefèvre, 2007). However, saccadic eye movements seem to be used in particular to bring the gaze to locations that are anticipated to become important in the near future (see Figure 4A and Figure 8A).

Our results indicate that eye movements leverage all available information around the target, experience,

and kinematic cues to reduce tracking delay. Instead of using only low-level sensory input to explain oculomotor control (Lisberger, 2010; Lisberger, 2015) or gaze selection (Itti, Koch, & Niebur, 1998), our results demonstrate the importance of higher level information being available in naturalistic contexts. We want to emphasize that, although high-level information becomes more important in naturalistic and complex tasks, this does not render low-level information useless. A successful model must incorporate both low- and high-level factors (Torralba, Oliva, Castelhano, & Henderson, 2006), and most predictions are based on a combination on both types of information (Badler & Heinen, 2006). For example, to make predictions about the future movement direction of a player, the correct analysis of biological motion is important, in the same way the extraction of the direction of a pass is critical to determine where it will go to anticipate the receiving player.

Our results provide empirical evidence for a suggestion made by Henderson (2017): In naturalistic scenes and tasks, gaze control is mainly based on predictions. When the eyes are sometimes on a longer leash and the deviation from the puck position is even slightly higher (see Figure 2 and Figure 7; although note here that this error is also driven by predictive eye movements), the higher level information is used to guide the gaze to crucial positions at the right time or even ahead of time (e.g., a player receiving a pass). This allows a better analysis of critical information to predict what is going to happen next. Such a strong influence from higher level factors can be easily seen when looking at the selection of fixation points: During free-viewing of a scene, predictions based on low-level image properties have been highly successful (Itti et al., 1998; Kümmerer, Bethge, & Wallis, 2022), but a simple task can render these predictions useless (Einhaeuser, Rutishauser, & Koch, 2008; Rothkopf et al., 2016; Tatler et al., 2011; Yarbus, 1967). Gaze control in natural tasks is more focused on fulfilling certain task demands, such as guiding an action or locating an object (Diaz et al., 2013; Hayhoe & Ballard, 2014; Land & Hayhoe, 2001; Sullivan et al., 2021); therefore, guidance by low-level factors becomes less important.

Keywords: natural videos, eye movements, expertise, periphery

Data availability: Raw and aggregated data, videos of the stimuli with overlaid gaze data and stimulus information can be found at https://osf.io/kqup8/. Analysis is available at https://github.com/AlexanderGoettker/CuesforNaturalisticVideos.

Commercial relationships: none.
Corresponding author: Alexander Goettker.
Email: alexander.goettker@psy.jlug.de.
Address: Justus Liebig Universität Giessen, Giessen, Germany.

References

Abernethy, B., Gill, D. P., Parks, S. L., & Packer, S. T. (2001). Expertise and the perception of kinematic and situational probability information. *Perception, 30*(2), 233–252, https://doi.org/10.1068/p2872.

Abernethy, B., & Russell, D. G. (1987). The relationship between expertise and visual search strategy in a racquet sport. *Human Movement Science, 6*(4), 283–319, https://doi.org/10.1016/0167-9457(87)90001-7.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717–746, https://doi.org/10.1068/p5096.

Badler, J. B., & Heinen, S. J. (2006). Anticipatory movement timing using prediction and external cues. *Journal of Neuroscience, 26*(17), 4519–4525, https://doi.org/10.1523/JNEUROSCI.3739-05.2006.

Brams, S., Ziv, G., Levin, O., Spitz, J., Wagemans, J., Williams, A. M., . . . Helsen, W. F. (2019). The relationship between gaze behavior, expertise, and performance: A systematic review. *Psychological Bulletin, 145*(10), 980–1027, https://doi.org/10.1037/bul0000207.

Casanova, F., Oliveira, J., Williams, M., & Garganta, J. (2009). Expertise and perceptual-cognitive performance in soccer: A review. *Revista Portuguesa de Ciências Do Desporto, 9*(1), 115–122.

Chen, Y., Yang, Z., Ahn, S., Samaras, D., Hoai, M., & Zelinsky, G. (2021). COCO-Search18 fixation dataset for predicting goal-directed attention control. *Scientific Reports, 11*(1), 8776, https://doi.org/10.1038/s41598-021-87715-9.

Diaz, G., Cooper, J., Rothkopf, C., & Hayhoe, M. (2013). Saccades to future ball location reveal memory-based prediction in a virtual-reality interception task. *Journal of Vision, 13*(1):20, 1–14, https://doi.org/10.1167/13.1.20.

Einhaeuser, W., Rutishauser, U., & Koch, C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision, 8*(2):2, 1–19, https://doi.org/10.1167/8.2.2.

Fleuriet, J., Hugues, S., Perrinet, L., & Goffart, L. (2011). Saccadic foveation of a moving visual target in the rhesus monkey. *Journal of Neurophysiology, 105*(2), 883–895, https://doi.org/10.1152/jn.00622.2010.

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise differences in the comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23*(4), 523–552, https://doi.org/10.1007/s10648-011-9174-7.

Goettker, A., Agtzidis, I., Braun, D. I., Dorr, M., & Gegenfurtner, K. R. (2020). From Gaussian blobs to naturalistic videos: Comparison of oculomotor behavior across different stimulus complexities. *Journal of Vision, 20*(8):26 1–16, https://doi.org/10.1167/JOV.20.8.26.

Goettker, A., Braun, D. I., & Gegenfurtner, K. R. (2019). Dynamic combination of position and motion information when tracking moving targets. *Journal of Vision, 19*(7):2 1–22, https://doi.org/10.1167/19.7.2.

Goettker, A., & Gegenfurtner, K. R. (2021). A change in perspective: The interaction of saccadic and pursuit eye movements in oculomotor control and perception. *Vision Research, 188*, 283–296, https://doi.org/10.1016/j.visres.2021.08.004.

Goettker, A., Pidaparthy, H., Braun, D. I., Elder, J. H., & Gegenfurtner, K. R. (2021). Ice hockey spectators use contextual cues to guide predictive eye movements. *Current Biology, 31*(16), R991–R992, https://doi.org/10.1016/j.cub.2021.06.087.

Gorman, A. D., Abernethy, B., & Farrow, D. (2013). The expert advantage in dynamic pattern recall persists across both attended and unattended display elements. *Attention, Perception, & Psychophysics, 75*(5), 835–844, https://doi.org/10.3758/s13414-013-0423-3.

Grünert, U., & Martin, P. R. (2020). Cell types and cell circuits in human and non-human primate retina. *Progress in Retinal and Eye Research, 78*, 100844, https://doi.org/10.1016/j.preteyeres.2020.100844.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18*(10), 1279–1296, https://doi.org/10.1016/0042-6989(78)90218-3.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194, https://doi.org/10.1016/j.tics.2005.02.009.

Hayhoe, M., & Ballard, D. (2014). Modeling task control of eye movements. *Current Biology, 24*(13), R622–R628, https://doi.org/10.1016/j.cub.2014.05.020.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57*(2), 243, https://doi.org/10.2307/1416950.

Heinen, S. J., Badler, J. B., & Watamaniuk, S. N. J. (2018). Choosing a foveal goal recruits the saccadic system during smooth pursuit. *Journal of Neurophysiology, 120*(2), 489–496, https://doi.org/10.1152/jn.00418.2017.

Henderson, J. M. (2017). Gaze control as prediction. *Trends in Cognitive Sciences, 21*(1), 15–23, https://doi.org/10.1016/j.tics.2016.11.003.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology, 43*(3), 302–313, https://doi.org/10.1111/j.1469-8986.2006.00403.x.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254–1259.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14*(2), 201–211, https://doi.org/10.3758/BF03212378.

Jörges, B., & López-Moliner, J. (2019). Earth-gravity congruent motion facilitates ocular control for pursuit of parabolic trajectories. *Scientific Reports, 9*(1), 14094, https://doi.org/10.1038/s41598-019-50512-6.

Kleiner, M., Brainard, D., & Pelli, D. (2007). *What's new in Psychtoolbox-3?. Perception, 36*, 1–16, https://doi.org/10.1068/v070821.

König, P., Wilming, N., Kietzmann, T. C., Ossandón, J. P., Onat, S., Ehinger, B. v., . . . Kaspar, K. (2016). Eye movements as a window to cognitive processes. *Journal of Eye Movement Research, 9*(5), 1–16, https://doi.org/10.16910/jemr.9.5.3.

Kowler, E., Rubinstein, J. F., Santos, E. M., & Wang, J. (2019). Predictive smooth pursuit eye movements. *Annual Review of Vision Science, 5*, 223–246, https://doi.org/10.1146/annurev-vision-091718-014901.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics, 21*(6), 575–580, https://doi.org/10.3758/BF03198740.

Krauzlis, R. J. (2004). Recasting the smooth pursuit eye movement system. *Journal of Neurophysiology, 91*(2), 591–603, https://doi.org/10.1152/jn.00801.2003.

Kümmerer, M., Bethge, M., & Wallis, T. S. A. (2022). DeepGaze III: Modeling free-viewing human scanpaths with deep learning. *Journal of Vision, 22*(5):7, 1–27, https://doi.org/10.1167/jov.22.5.7.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 352*(1358), 1231–1239, https://doi.org/10.1098/rstb.1997.0105.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25–26), 3559–3565, https://doi.org/10.1016/s0042-6989(01)00102-x.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball. *Nature Neuroscience, 3*(12), 1340–1345, https://doi.org/10.1038/81887.

Lisberger, S. G. (2010). Visual guidance of smooth-pursuit eye movements: Sensation, action, and what happens in between. *Neuron, 66*(4), 477–491, https://doi.org/10.1016/j.neuron.2010.03.027.

Lisberger, S. G. (2015). Visual guidance of smooth pursuit eye movements. *Annual Review of Vision Science, 1*(1), 447–468, https://doi.org/10.1146/annurev-vision-082114-035349.

Loffing, F., & Cañal-Bruland, R. (2017). Anticipation in sport. *Current Opinion in Psychology, 16*, 6–11, https://doi.org/10.1016/j.copsyc.2017.03.008.

Loffing, F., & Hagemann, N. (2014). On-court position influences skilled tennis players' anticipation of shot outcome. *Journal of Sport and Exercise Psychology, 36*(1), 14–26, https://doi.org/10.1123/jsep.2013-0082.

McCloseky, M. (1983). Intuitive physics. *Scientific American, 248*(4), 122–131.

McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception & Psychophysics, 2*(8), 359–362.

Memmert, D. (2009). Pay attention! A review of visual attentional expertise in sport. *International Review of Sport and Exercise Psychology, 2*(2), 119–138, https://doi.org/10.1080/17509840802641372.

Müller, S., & Abernethy, B. (2012). Expert anticipatory skill in striking sports. *Research Quarterly for Exercise and Sport, 83*(2), 175–187, https://doi.org/10.1080/02701367.2012.10599848.

Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience, 5*(3), 218–228, https://doi.org/10.1038/nrn1345.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395*(6705), 894–896, https://doi.org/10.1038/27661.

Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In: K. S. Rockland, J. H. Kass, & A. Paters (eds.),

Extrastriate Cortex in Primates: Cerebral Cortex, Vol. 12 (pp. 205–241). Boston, MA: Springer, https://doi.org/10.1007/978-1-4757-9625-4_5.

Orban De Xivry, J. J., & Lefèvre, P. (2007). Saccades and pursuit: Two outcomes of a single sensorimotor process. *Journal of Physiology, 584*(1), 11–23, https://doi.org/10.1113/jphysiol.2007.139881.

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature, 409*(6823), 1040–1042, https://doi.org/10.1038/35059085.

Pavlova, M., Krägeloh-Mann, I., Birbaumer, N., & Sokolov, A. (2002). Biological motion shown backwards: The apparent-facing effect. *Perception, 31*(4), 435–443, https://doi.org/10.1068/p3262.

Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics, 62*(5), 889–899, https://doi.org/10.3758/BF03212075.

Pélisson, D., Alahyane, N., Panouillères, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience & Biobehavioral Reviews, 34*(8), 1103–1120, https://doi.org/10.1016/j.neubiorev.2009.12.010.

Pidaparthy, H., & Elder, J. (2019). Keep your eye on the puck: Automatic hockey videography. In: *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1636–1644). Piscataway, NJ: Institute of Electrical and Electronics Engineers, https://doi.org/10.1109/WACV.2019.00179.

Ripoll, H., Fleurance, P. H., & Cazeneuve, D. (1987). Analysis of visual patterns of table tennis players. In: J. K. O'Regan, & A. Lévy (Eds.), *Eye Movements from Physiology to Cognition* (pp. 616–617). Amsterdam: Elsevier Science, https://doi.org/10.1016/B978-0-444-70113-8.50085-0.

Ron, S., Vieville, T., & Droulez, J. (1989). Target velocity based prediction in saccadic vector programming. *Vision Research, 29*(9), 1103–1114, https://doi.org/10.1016/0042-6989(89)90059-x.

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science, 2*(1), 437–457, https://doi.org/10.1146/annurev-vision-082114-035733.

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision, 7*(2):17, 1–22, https://doi.org/10.1167/7.2.17.

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2016). Task and context determine where you look. *Journal of Vision, 7*(14):16, 1–20, https://doi.org/10.1167/7.14.16.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: Delays and localization errors in the visual system. *Nature Reviews Neuroscience, 3*(3), 191–215, https://doi.org/10.1038/nrn750.

Schläppi-Lienhard, O., & Hossner, E.-J. (2015). Decision making in beach volleyball defense: Crucial factors derived from interviews with top-level experts. *Psychology of Sport and Exercise, 16*, 60–73, https://doi.org/10.1016/j.psychsport.2014.07.005.

Schreiber, C., Missal, M., & Lefèvre, P. (2006). Asynchrony between position and motion signals in the saccadic system. *Journal of Neurophysiology, 95*(2), 960–969, https://doi.org/10.1152/jn.00315.2005.

Shaikh, A. G., & Zee, D. S. (2018). Eye movement research in the twenty-first century—a window to the Brain, mind, and more. *The Cerebellum, 17*(3), 252–258, https://doi.org/10.1007/s12311-017-0910-5.

Sheridan, H., & Reingold, E. M. (2014). Expert vs. novice differences in the detection of relevant information during a chess game: Evidence from eye movements. *Frontiers in Psychology, 5*, 941, https://doi.org/10.3389/fpsyg.2014.00941.

Sheridan, H., & Reingold, E. M. (2017). Chess players' eye movements reveal rapid recognition of complex visual patterns: Evidence from a chess-related visual search task. *Journal of Vision, 17*(3):4, 1–12, https://doi.org/10.1167/17.3.4.

Shim, J., Carlton, L. G., Chow, J. W., & Chae, W.-S. (2005). The use of anticipatory visual cues by highly skilled tennis players. *Journal of Motor Behavior, 37*(2), 164–175, https://doi.org/10.3200/JMBR.37.2.164-175.

Smuc, M., Mayr, E., & Windhager, F. (2010). The game lies in the eye of the beholder: The influence of expertise on watching soccer. In: S. Ohlsson (Ed.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*. Wheat Ridge, CO: Cognitive Science Society.

Spering, M., & Schütz, A. C. (2016). How eye movements improve vision and action – comment on Vickers. *Current Issues in Sport Science (CISS), 1*, 114, https://doi.org/10.15203/CISS_2016.114.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5):13, 1–82, https://doi.org/10.1167/11.5.13.

Sullivan, B., Ludwig, C. J. H., Damen, D., Mayol-Cuevas, W., & Gilchrist, I. D. (2021). Look-ahead fixations during visuomotor behavior: Evidence from assembling a camping tent. *Journal of Vision, 21*(3):13, 1–17, https://doi.org/10.1167/jov.21.3.13.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision:

Reinterpreting salience. *Journal of Vision, 11*(5):5, 1–23, https://doi.org/10.1167/11.5.5.

Tenenbaum, G., Sar-El, T., & Bar-Eli, M. (2000). Anticipation of ball location in low and high-skill performers: A developmental perspective. *Psychology of Sport and Exercise, 1*(2), 117–128, https://doi.org/10.1016/S1469-0292(00)00008-X.

Terry Bahill, A., & McDonald, J. D. (1983). Smooth pursuit eye movements in response to predictable target motions. *Vision Research, 23*(12), 1573–1583, https://doi.org/10.1016/0042-6989(83)90171-2.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786, https://doi.org/10.1037/0033-295X.113.4.766.

Vansteenkiste, P., Vaeyens, R., Zeuwts, L., Philippaerts, R., & Lenoir, M. (2014). Cue usage in volleyball: A time course comparison of elite, intermediate and novice female players. *Biology of Sport, 31*(4), 295–302, https://doi.org/10.5604/20831862.1127288.

Vickers, J. (2016). Origins and current issues in Quiet Eye research. *Current Issues in Sport Science (CISS), 1*, 101, https://doi.org/10.15203/CISS-2016.101.

Vickers, J. N. (2009). Advances in coupling perception and action: The quiet eye as a bidirectional link between gaze, attention, and action. *Progress in Brain Research, 174*, 279–288, https://doi.org/10.1016/S0079-6123(09)01322-3.

Vig, E., Dorr, M., Martinetz, T., & Barth, E. (2011). Eye movements show optimal average anticipation with natural dynamic scenes. *Cognitive Computation, 3*(1), 79–88, https://doi.org/10.1007/s12559-010-9061-4.

Vo, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision, 9*(3):24, 1–15, https://doi.org/10.1167/9.3.24.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168, https://doi.org/10.1016/j.tics.2011.02.005.

Williams, A. M. (1993). Advance cue utilization in soccer. *Science and Football, 11*, 284–290.

Williams, A. M., Davids, K., Burwitz, L., & Williams, J. G. (1994). Visual search strategies in experienced and inexperienced soccer players. *Research Quarterly for Exercise and Sport, 65*(2), 127–135, https://doi.org/10.1080/02701367.1994.10607607.

Williams, A. M., Hodges, N. J., North, J. S., & Barton, G. (2006). Perceiving patterns of play in dynamic sport tasks: Investigating the essential information underlying skilled performance. *Perception, 35*(3), 317–332, https://doi.org/10.1068/p5310.

Yarbus, A. L. (1967). *Eye Movements and Vision* (pp. 171–211). Heidelberg: Springer.