Correspondence
Ice hockey spectators use contextual cues to guide predictive eye movements

Alexander Goettker¹,²,⋆, Hemanth Pidaparthy¹, Doris I. Braun¹,², James H. Elder³, and Karl R. Gegenfurtner¹,²

Eye movements are an integral part of human visual perception. They allow us to have a small foveal region with exquisite acuity and at the same time a large visual field. For a long time, eye movements were regarded as machine-like behaviors in response to visual stimulation, but over the past few decades it has been convincingly shown that expectations, intended actions, rewards and many other cognitive factors can have profound effects on the way we move our eyes. In order to be useful, our oculomotor system must minimize delay with respect to the dynamic events in the visual scene. The ability to do so has been demonstrated in situations where we are in control of these events, for example when we are making a sandwich or tea, and when we are active participants, for example when hitting a cricket ball. But what about scenes with complex dynamics that we do not control or directly take part in, like a hockey game we are watching as a spectator? A semantic influence on gaze fixation location during viewing of tennis videos has been suggested before. Here we use carefully annotated hockey videos to show that the brain is indeed able to exploit the semantic context of the game to anticipate the continuous motion of the puck, leading to eye movements that are fundamentally different than when following exactly the same motion without any context.

To directly measure the influence of context, and dissociate the modulation of real-time prediction by complex visual context from the kinematics of the target, we made use of a large collection of ice hockey videos. Unlike most sport videos, they were filmed with a static camera and thus avoid any potential artifacts based on camera movement. Furthermore, we went through the time-consuming process of hand-labelling the positions of the puck throughout the videos. We selected a subset of 18 different 10 s clips from these hockey videos and presented them to 15 naïve participants with no expertise in ice hockey while tracking their eye movements. Participants were asked to follow the puck with their eyes. Having access to the puck positions allowed us to show participants’ targets moving along the same trajectories as the puck in the videos but without the context information. We created three conditions (Figure 1A): first, a black disk moving in front of a uniform gray background (disk+gray); second, a black disk moving in front of a static presentation of the first frame of each clip (disk+static); and third, the original hockey clips (full video). Each participant saw all 18 clips in all three context conditions. This allowed a direct comparison of oculomotor behavior for the same target trajectories, but different levels of context complexities. (See supplemental information for detailed methods.)

We found striking differences in the estimated delay between gaze and puck position across the three conditions. Based on average correlations of horizontal and vertical eye and target positions with varying delays (see Supplemental experimental procedures), we estimated a delay of roughly 170 ms for the two conditions without the dynamic context. This kind of a delay can be expected based on inherent processing delays when tracking an unpredictable moving puck.

Figure 1. Gaze when tracking the puck.
(A) Illustration of the paradigm and example gaze positions for the three context conditions. The large central image is a still frame taken from one of the 18 clips, showing a pass situation. The colors of the small squares indicate gaze positions of representative observers for all three conditions: green for disk presentations on a gray background, orange for disk presentations on a static background and violet for the full video. The two insets to the right illustrate the puck movement trajectories shown in the two context-deprived conditions. The black contour represents the trace of the puck position over the 10 s example clip. (B) Magnitude of the cross-correlation between gaze and puck position. Without the dynamic context the time delay was around 170 ms, roughly corresponding to neural processing delays expected for unpredictable motion. (C) Average number of saccades per 10 s hockey clip across the conditions. (D) Arrival time of gaze at targeted player. (E) Gaze error projected onto the trajectory of the pass 200 ms after the pass ended. All error bars and shaded areas depict the standard error of the mean. Light gray lines depict the individual values.
target. In contrast, the estimated delay for the same trajectory of the puck was close to 0 ms in the full video condition (Figure 1B), demonstrating that with additional context information the visual oculomotor system can overcome internal processing delays to track complex target movements without significant time lags.

We also observed interesting differences in saccadic eye movements for the three conditions. The average number of saccades executed during each 10 s clip was just over 30% lower in the full video condition (25.31) compared with disk+gray (37.12) and disk+static (37.78) conditions (Figure 1C; both $t_{14} > 13.63$, $p < 0.001$). Interestingly, this reduction was especially prominent for small saccades with amplitudes below 5 deg (Figure S1A), suggesting that fewer corrective saccades were needed to follow the puck in the full video condition.

While the availability of the full video context greatly reduced the average tracking delay, overall tracking errors were in fact somewhat higher for the full video condition (2.24 deg) relative to the disk+gray (1.71 deg) and disk+static (1.83 deg) conditions (both $t_{14} > 3.77$, $p < 0.002$). We conjectured that this may in part be due to anticipatory eye movements made in the full video condition. To explore this possibility, we examined a specific class of event that might invite anticipatory saccades: passes between players. In this situation, without any context, the puck makes sudden and unpredictable accelerations, both at the beginning and at the end of the pass. On average, participants arrived at the endpoint of passes around 125 ms earlier (both $t_{14} > 9.80$, $p < 0.001$; Figure 1D) when seeing the full video in contrast to the disk conditions. This demonstrates predictive and even anticipatory behavior of the oculomotor system based on the movement and position of the players. The predictive behavior during passes contributes to the above zero-delay estimates but is not the only explanation (Figure S1C,D). In addition to this difference in timing, there was also a significant difference in the accuracy of the following fixation. Participants fixated very close to the end position of the pass when watching the video clips, but overshot the targeted position when only seeing the movement trajectory (both $t_{14} > 6.05$, $p < 0.001$; Figure 1E). We speculate that without information about the position of the player receiving the pass, the oculomotor system relied more on the kinematics of the puck. Given internal processing delays, this led to overshoot.

Our results reveal fundamentally different eye movement strategies when viewing complex but meaningful action sequences compared to strategies seen when viewing simple target motion sequences predominantly used in prior eye movement studies. Based on a direct comparison of the same target trajectories with and without context information, we experimentally show a direct and causal influence of contextual information during passive viewing of complex naturalistic scenes. The contextual information allows for knowledge-driven predictions, to ensure that important information can be processed with the highest visual acuity without any delay.

While near-zero lag between gaze and target has previously been demonstrated when correlating gaze and a saliency map metric based on low level image characteristics, our findings provide direct evidence that, in real-world scenarios, semantic knowledge about the dynamic scene is directly integrated into our oculomotor response. This makes sure the eyes are at the right point at the right time during crucial moments, like when a pass is received by a player. At other times, the eyes may be on a longer leash, so to speak. Small corrective saccades are skipped, introducing larger position errors, in favor of larger saccades right to the anticipated recipient of a pass. The advantage is improved fixation accuracy at these critical moments, which, we speculate, might lead to improved performance on perceptual tasks, for example to judge whether the receiving player will immediately shoot the puck at the goal. Improved performance on such tasks would surely benefit athletes but might also be of more general importance for anticipating probable outcomes of diverse dynamic scenarios involving multiple active agents.

SUPPLEMENTAL INFORMATION

Supplemental information includes one figure, a detailed description of the methods, a link to the publicly available dataset, supplemental results regarding the influence of target contrast, as well as the author contributions and can be found with this article online at https://doi.org/10.1016/j.cub.2021.06.087.

ACKNOWLEDGEMENTS

The authors thank Stefanie Ackermann for her careful help with data collection. A.G., D.I.B., and K.R.G. were supported by the Deutsche Forschungsgemeinschaft (DFG; project number 222641018–SFB/TRR 135 Project A1). H.F and J.H.E. wish to acknowledge the support of the NSERC CREATE Training program in Data Analytics and Visualization (NSERC CREATE DAV), the York Vision: Science to Applications (VISTA) program and the York Research Chair program.

DECLARATION OF INTERESTS

Two authors of this paper are founders and shareholders of the company AttentiveVision (www.attentivevision.com), which owns the datasets from which the data used for this paper were drawn.

REFERENCES

1. Lisberger, S.G. (2015). Visual guidance of smooth pursuit eye movements. Annu. Rev. Vis. Sci. 1, 447–468.
2. Yarbus, A.L. (1967). Eye Movements and Vision (Boston, MA: Springer).
3. Tatler, B.W., Hayhoe, M.M., Land, M.F., and Ballard, D.H. (2011). Eye guidance in natural vision: Reinterpreting salience. J. Vis. 11, 5.
4. Kowler, E., Rubinstein, J.F., Santos, E.M., and Wang, J. (2019). Predictive smooth pursuit eye movements. Annu. Rev. Vis. Sci. 5, 223–246.
5. Land, M.F., and Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? Vis. Res. 41, 3559–3565.
6. Land, M.F., and McLeod, P. (2000). From eye movements to actions: How batsmen hit the ball, Nat. Neurosci. 3, 1340–1345.
7. Henderson, J.M. (2017). Gaze control as prediction. Trends Cogn. Sci. 21, 15–23.
8. Pidaparthi, H., and Elder, J. (2018). Keep your eye on the puck: Automatic hockey videography. IEEE Wint. Conf. Appl. Comp. Vis. 2019, 1636–1644.
9. Vg, E., Dorr, M., Martinez, T., and Barth, E. (2011). Eye movements show optimal average anticipation with natural dynamic scenes. Cog. Comp. 3, 79–98.
10. Vickers, J.M. (2016). Origins and current issues in Quiet Eye research. Curr. Iss. Sport Sci. 1, 101.

Justus Liebig University Giessen, Giessen, Germany. *Center for Mind, Brain and Behavior (CMBB), Justus Liebig University Giessen, Giessen, Germany. *York University, Toronto, Canada. *Lead Contact.
*E-mail: Alexander.Goettker@psy.jlug.de