Accurately tracking visual moving stimuli with gaze requires a higher fidelity motion representation than perceptual discrimination

Ilja Wagner, T Scott Murdison, Karl R Gegenfurtner, Alexander Goettker

PNAS NEXUS

PNAS Nexus, 2025, **4**, pgaf167
https://doi.org/10.1093/pnasnexus/pgaf167
Advance access publication 23 May 2025
Brief Report

Accurately tracking visual moving stimuli with gaze requires a higher fidelity motion representation than perceptual discrimination

Ilja Wagner[a], T. Scott Murdison[b], Karl R. Gegenfurtner[a,c] and Alexander Goettker[a,c,*]

[a]Allgemeine Psychologie, Justus Liebig University Giessen, 35396 Giessen, Germany
[b]Reality Labs, Meta Platforms Inc., Redmond, WA 98052, USA
[c]Center for Mind, Brain and Behavior, Universities of Marburg, Giessen, and Darmstadt, Germany
*To whom correspondence should be addressed: Email: Alexander.Goettker@psy.jlug.de
Edited By Stephen Fleming

Abstract

The human visual system evolved to process and interpret continuous visual input. However, visual signals in modern societies increasingly emanate from digital displays, which emulate continuous motion by presenting a series of discrete frames at some fixed temporal frequency. How does noncontinuous, sampled motion influence visual perception and action? We asked participants to complete a combined oculomotor and speed-discrimination task, where they tracked the motion of a stimulus via smooth pursuit eye movements and compared the velocity of the pursuit stimulus with a reference velocity. Motion quality was manipulated by presenting discrete frames at different temporal frequencies (120, 60, 30, and 15 Hz), covering a range of values corresponding to different motion qualia, from quasi-continuous (120 Hz) to highly discontinuous (15 Hz). We found that participants could reliably discriminate velocities, even at low refresh rates. However, measures of pursuit quality declined with decreasing refresh rate and increasing target velocity: a perception–action dissociation. The decline in pursuit accuracy can be explained by a model estimating a proxy for motion quality from the spatiotemporal energy profile of the stimuli. Using this estimate, our model predicts pursuit accuracy for arbitrary combinations of target velocity and refresh rate. Our results highlight the dissociation between perception and action in digital displays with sampled motion; while we may be able to accurately discriminate motion at reduced frame rates, acting on moving stimuli requires a higher fidelity motion representation.

Keywords: smooth pursuit eye movements, motion perception, applied perception

Introduction

Accurately perceiving and interpreting motion signals are fundamental tasks for the human visual system. Although our visual system has evolved to analyze a continuous stream of visual information in natural environments, in modern societies, visual motion is increasingly provided discretely as a series of frames on digital displays (sampled motion). Understanding the effects of sampled motion on visual perception is of great interest in applied research for the design of future display technologies (1). At the same time, sampled motion can be leveraged as a tool for understanding the basic processing of visual information. Previous research, for example, used sampled motion to demonstrate that visual perception (2) and oculomotor tracking (3–6) of sampled motion stimuli are both influenced, among other things, by the temporal characteristics of a display, i.e. its refresh rate or how often the display content is updated per time interval. Here, we provide a systematic comparison of the influence of sampled motion on visual perception and oculomotor behavior. Our results demonstrate a dissociation between the perceptual and oculomotor responses to sampled motion: we observed a systematic influence of sampled motion on oculomotor behavior. This was accounted for by a model that, first, computes an estimate for motion quality, and second, uses this estimate to predict pursuit gain. In contrast, perceptual discrimination thresholds were unaffected by the quality of sampled motion. These results have important implications for the comparability of results across different lab settings, the integration of sensory information for motor control and perception, and applied questions in cinematic and head-mounted display design.

Results

We measured smooth pursuit eye movements and visual perception of human observers ($n = 24$), who gave informed consent, and the study was approved by the local ethics committee of the Justus Liebig University Giessen, Department of Psychology (proposal Giessen LEK 2020-0015). Participants tracked the motion of a Gaussian blob that moved in a step ramp (7). At the end of each trial, participants reported whether the blob moved slower or faster than a reference stimulus (Fig. 1A). We manipulated target velocity (7, 8.50, 10, 11.50, and 13 degrees of visual angle per

Competing Interest: T.S.M. is employed by Meta Platforms. This study did not evaluate any products or services related to the company.
Received: July 16, 2024. **Accepted:** May 2, 2025

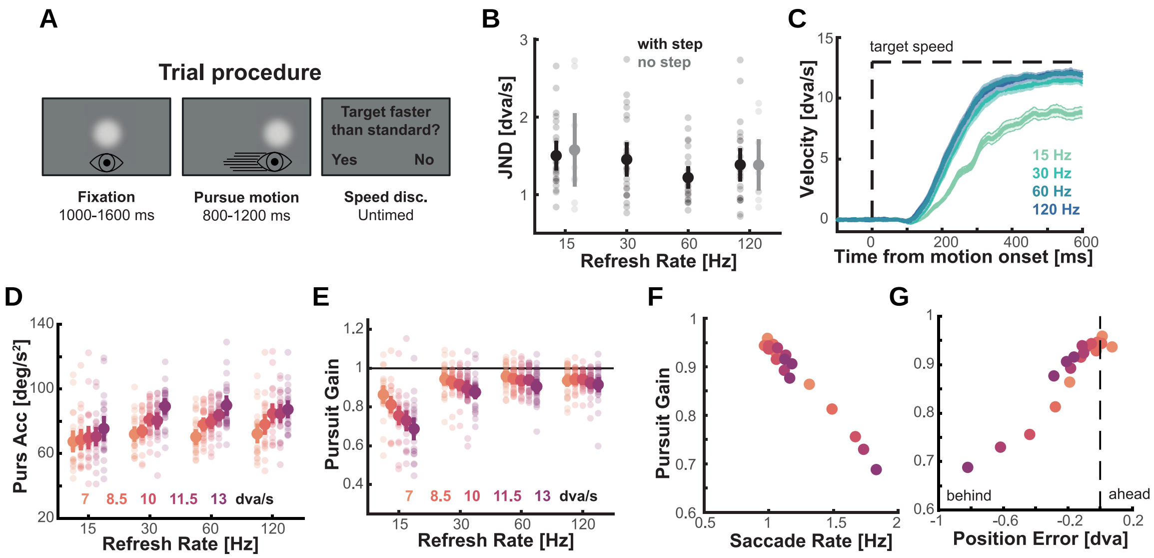

Fig. 1. Perceptual performance and oculomotor response to sampled motion. A) Participants had to visually track the motion of a Gaussian blob, and, at the end of each trial, judge whether the blob moved faster than a comparison stimulus (shown at various points throughout the experiment). B) Degrading the motion stimulus by reducing the monitor's refresh rate had no effect on participant's speed-discrimination performance. Black data points show the average results of the main experiment. Gray data points show the average results from a control experiment, where the moving blob was presented without an initial step before motion onset. Small grayish dots show the individual data. C) The average velocity profiles for the speed of 13 dva/s for the different refresh rates aligned on motion onset. The shaded areas depict SEM. D and E) The accuracy of the oculomotor tracking response gradually decreased with decreasing refresh rate. This was visible in the initial pursuit acceleration (D) as well as pursuit gain (E). The impact was most pronounced at high target velocities. Small data points are data from individual participants, while large data points are means over participants. Different colors are different target velocities. Error bars are 95% CIs. F and G) Correlations between pursuit gain and the rate of corrective saccades (F) and the tracking error (G) during the trial. The data points represent the mean across observers for each of the conditions.

second [dva/s]) and the display's refresh rate (15, 30, 60, and 120 Hz), covering a range of parameter values that are typically used to study motion perception and oculomotor behavior (see Supplementary Material for details). Our results show a dissociation between perception and oculomotor responses.

Varying the monitor's refresh rate had no discernible effect on participants' ability to discriminate between different speeds (120 Hz: M [mean just noticeable difference {JND}] = 1.38 dva/s, $CI_{95\%}$ = [1.17, 1.60 dva/s]; 60 Hz: M = 1.22 dva/s, $CI_{95\%}$ = [1.08, 1.36 dva/s]; 30 Hz: M = 1.45 dva/s, $CI_{95\%}$ = [1.23, 1.67 dva/s]; 15 Hz: M = 1.50 dva/s, $CI_{95\%}$ = [1.31, 1.69 dva/s]), $F(3, 45) = 1.55$, $P = 0.214$ (Fig. 1B). The lack of perceptual effect did not depend on the presence of an initial step that varied in size depending on the velocity and could have served as a cue for the perceptual task (120 Hz: M [mean JND] = 1.38 dva/s, $CI_{95\%}$ = [1.05, 1.71 dva/s]; 15 Hz: M = 1.58 dva/s, $CI_{95\%}$ = [1.10, 2.05 dva/s]), $t(9) = 1.41$, $P = 0.193$.

In contrast, manipulations of target velocity and refresh rate clearly affected participants' oculomotor response (see Fig. 1C for the velocity traces for the medium speed of 10 dva/s). We compared the initial open-loop acceleration of pursuit and the closed-loop gain during steady-state pursuit (Fig. 1D and E). Pursuit acceleration increased with target velocity and with refresh rate. Pursuit gain decreased with increasing target velocity and lower refresh rate. We quantified the modulation by velocity by fitting linear functions to the acceleration and gain data of individual participants, separately for each refresh rate, and by inspecting the corresponding slopes. A robust ANOVA with slopes as the dependent variable and refresh rate as the independent variable revealed that the effect of target velocity on pursuit acceleration was smaller for lower refresh rates (120 Hz: M (mean slope) = 2.778 dva/s^2, $CI_{95\%}$ = [2.131, 3.426 dva/s^2]; 60 Hz: M = 3.168 dva/s^2, $CI_{95\%}$ = [2.698, 3.638 dva/s^2]; 30 Hz: M = 2.394 dva/s^2, $CI_{95\%}$ = [1.574, 3.215 dva/s^2]; 15 Hz: M = 1.586 dva/s^2, $CI_{95\%}$ = [0.843, 2.328 dva/s^2]), $F(2.85, 42.73) = 4.72$, $P = 0.007$. For pursuit gain, the influence of velocity was more pronounced for lower refresh rates (120 Hz: M [mean slope] = –0.004, $CI_{95\%}$ = [–0.008, 0]; 60 Hz: M = –0.007, $CI_{95\%}$ = [–0.012, –0.003]; 30 Hz: M = –0.011, $CI_{95\%}$ = [–0.015, –0.008]; 15 Hz: M = –0.029, $CI_{95\%}$ = [–0.037, –0.021]), $F(1.47, 22.04) = 26.18$, $P < 0.001$. The lower pursuit gain at high velocities and lower refresh rates was compensated by a higher frequency of catch-up saccades ($r(18) = –0.90$, $P < 0.001$; see Fig. 1F). However, even with more corrective saccades, position error during tracking (quantified as the difference between eye and target position) was still higher for the conditions with reduced pursuit gain ($r(18) = 0.92$, $P < 0.001$; see Fig. 1G).

The effect of refresh rate on oculomotor tracking performance was accounted for by a model, which received simulations of stimulus motion across refresh rates and target velocities as input (Fig. 2A) and used those to compute a proxy for motion quality. This proxy was based on the spatiotemporal energy profile at the stimulus speed compared to the total energy (see Supplementary Material for details). This metric is closely related to the spatial separation between successive position updates ($r = 0.99$), since the magnitude of these steps increases the difference to the natural scenario of a continuous visual input. We chose to use a motion quality metric based on spatiotemporal energy for generalizability to the field of motion perception (8) and applied display design (9). The motion quality estimate was then related to pursuit gain via a nonlinear transformation (Fig. 2B). Overall, our model was able to accurately predict the average oculomotor tracking performance across different refresh rates and stimulus velocities, $r(18) = 0.98$, $P < 0.001$, and generalized well to the data of a new group of participants, who tracked targets moving with a set of velocities that were not tested in our main experiment, $r(7) = 0.97$, $P < 0.001$ (Fig. 2C). To quantify the successful generalization to the new data set, we compared the root-mean-square error (RMSE) of a

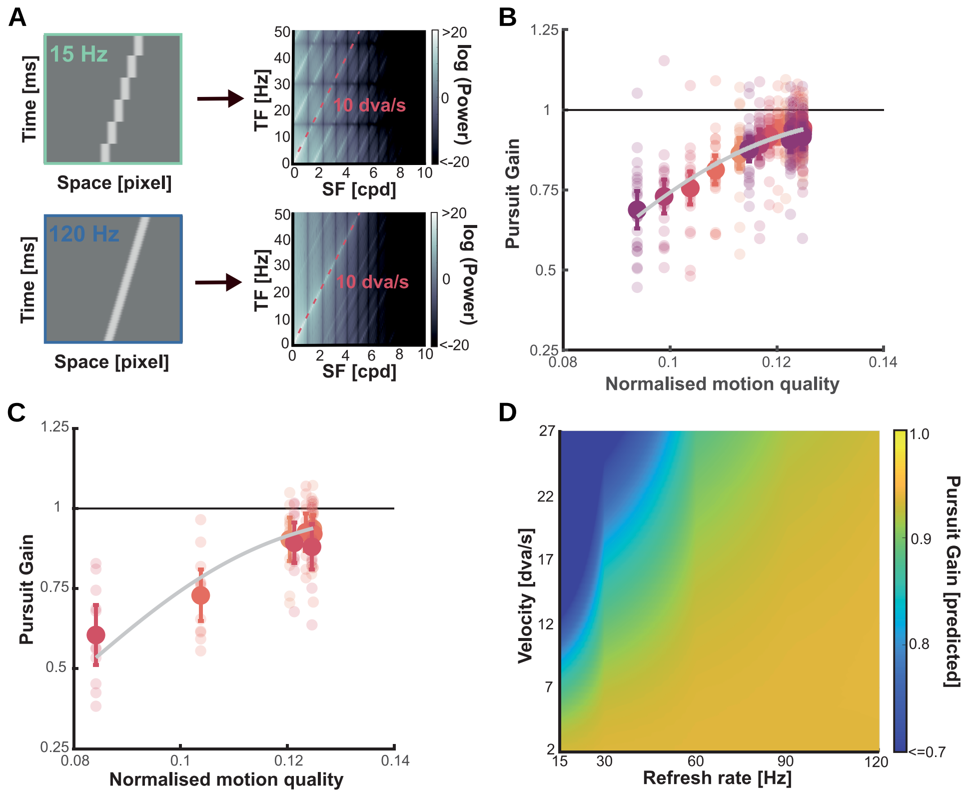

Fig. 2. Predicting oculomotor performance from motion quality. A) Depending on the refresh rate of the monitor and the target's velocity, the spatiotemporal profile of the Gaussian blob differed (here are two examples for 10 dva/s) and the stimulus either transitioned smoothly from its start to its end position (e.g. 120 Hz condition) or it appeared to jump from one location to another (e.g. 15 Hz). We took this stimulus as input and then computed the spatiotemporal energy of the stimulus. Here, most of the energy should lie on the diagonal, reflecting the original target speed of 10 dva/s. Please note that the power spreads further for the 15 Hz condition, and there is the appearance of multiple aliases of the velocity diagonal. We then computed a motion quality estimate as the power that was along the velocity diagonal divided by the overall power in the spectrum. B) The accuracy of the oculomotor tracking response in the main experiment depends in a nonlinear manner on the motion quality of the moving stimulus (estimated from its spatiotemporal energy profile). Data points are the same as in Fig. 1C. For each participant, we assume the same motion quality estimate, since the stimulus, presented in the experiment, did not differ across participants. The gray line corresponds to a Gaussian cumulative distribution function, fitted to the oculomotor data from the main experiment. The mean and SD of the function were treated as free parameters. C) Our model generalized to a new group of participants and a new set of target velocities (4 and 16 dva/s), without the need to refit model parameters. D) Our model can predict pursuit gain for arbitrary combinations of refresh rate and target velocity, again, without the need to refit model parameters. C and D) We used the function parameter from the main experiment to predict the relationship between gain and motion quality in the validation condition and generate predictions for pursuit gain across arbitrary combinations of refresh rate and velocity.

model that used the fitted parameters from the main experiment to predict data in the validation experiment and a model that was re-fitted to the new data and observed virtually identical results (RMSE old parameter: $M = 0.01$, $CI_{95\%} = [0.07, 0.13]$; RMSE new parameter: $M = 0.01$, $CI_{95\%} = [0.07, 0.12]$, $t(11) = 0.761$, $P = 0.463$). Since we can simulate stimulus motion under arbitrary combinations of refresh rate and target velocity, our model also allows us to predict the impact of motion quality on pursuit behavior for any combination of these two factors (Fig. 2D).

Discussion

Our results have four key implications for all fields that use display devices to study visual perception and oculomotor behavior, as well as for cinema, graphics, and displays industries more generally.

First, motion perception is often explained by using spatiotemporal energy models (8). Those models can, for example, explain why apparent motion leads to a veridical motion percept, even at relatively low refresh rates, and under what conditions observers will perceive motion artifacts (9). Our results extend those classical models and allow inferences about oculomotor behavior, drawn directly from a proxy measure for the motion quality of a stimulus. The fact that motion discrimination performance in our paradigm, unlike oculomotor responses, did not scale with the observed differences in motion quality suggests that perceptual judgments were at least not exclusively based on the spatiotemporal energy of a stimulus and information is integrated differently for oculomotor control and perceptual decisions (10). The consistent perceptual accuracy might be related to a longer time window, within which visual information about motion energy, motion duration, and movement distance can be integrated into perceptual decisions. For example, for the binary perceptual judgment at the end of the trial, the velocity of the target can be estimated by relating the perceived distance the target traveled to the perceived motion duration in the current trial. Such an estimate is independent of the presented motion quality but cannot be used by the oculomotor system, which needs to produce a continuous output and therefore is limited to an instantaneous readout of visual information from a moving stimulus. Alternatively, performance differences between oculomotor tracking and motion discrimination tasks might also be related to differences in the sensitivity to stimulus features (11) or differential processing of the same motion signal for visual perception and smooth pursuit eye movements (12). Overall, our results provide an example of the different sensitivity of discrete perceptual judgments and continuous oculomotor responses. This highlights the qualitative as well as quantitative differences between the processing of sensory signals for perception and oculomotor control (13), which seem to be fine-tuned to the different goals of the system: the oculomotor system needs to provide a quick and continuous output to bring and keep objects of interest on the retina, whereas the perceptual system can integrate information over longer times to create the most reliable visual percept (14).

Second, these results provide interesting insights into the processing and consequences of sampled motion. A previous study by Churchland and Lisberger found that sampled motion for a specific temporal window (between 20 and 30 Hz) can lead to an increase in the initial pursuit acceleration, and those trials were also judged to be faster by human observers (4). We do not have the data for perceived speed across refresh rates, but our oculomotor data do not show this effect for pursuit initiation. There are several potential reasons for these diverging results. (i) The motion stimulus in our experiment was a single target that was staying visible on the screen between position updates, whereas the motion stimulus in the work by Churchland and Lisberger was a dot cloud with multiple moving targets that only were briefly flashed for each position update. The intermittently presented dot cloud requires integration of the different individual movements into a coherent motion and leads to a different representation of the motion in the spatiotemporal frequency spectrum. This might have led to differences in the processing of the information. Studies with motion stimuli more similar to ours also observed results similar to ours. Initial pursuit acceleration was lower for the lower refresh rates (5), independently of whether the stimulus was only briefly flashed or kept on the screen between position updates. Other recent work (6) also showed that with increasing refresh rates, the behavior went from a more discrete tracking with higher rate of corrective saccades to a more continuous response with higher pursuit gain and less corrective saccades. (ii) The velocities tested in the experiment of Churchland and Lisberger were higher (16 or 32 dva/s) than in our experiment. This might be critical, as the explanation for the increased acceleration was that while apparent motion reduces the firing rate of MT neurons, it does so asymmetrically with a larger decrease for slower velocities, which then biases the average response of the population to faster speeds. Thus, the asymmetry might not have been as large for our slower velocities, and therefore, we did not see the effect. Independently of the potential differences in the pursuit acceleration, our results highlight that also the steady-state phase of the pursuit response is affected by sampled motion. While the pursuit initiation and steady-state phase have been related to different motion mechanisms (15), our results suggest that the position-based motion mechanism that tries to keep the eye on the target during steady-state pursuit is also affected by the refresh rate of the display.

Third, our results emphasize the importance of considering properties and limitations of stimuli and display devices when comparing findings from different laboratories and in the ergonomic design of consumer displays. For example, when using video recordings as stimuli, the native frame capture rate is often only 30 Hz, which, for fast velocities, should lead to a reduction in pursuit gain, independent of the content that is presented. In addition, different labs rely on different display devices when presenting stimuli, which might lead to different oculomotor responses, depending on the frame rate at which moving stimuli are presented and the velocities at which stimuli move. In those studies, a reduction in the accuracy of the oculomotor response to moving stimuli might not necessarily reflect the limitations of the oculomotor system but limitations in the hardware on which stimuli were presented. This matches previous behavioral observations that show a very strong influence of the refresh rate of the sampled motion for faster velocities (5).

Fourth, one of the main goals in (commercial) display design is to create metrics to formally quantify and evaluate the experience of observers when viewing dynamic content (16) across a range of different capture and display pipelines, from cinema to mobile phones to head-mounted virtual and augmented reality displays. Using subjective measures, such as questionnaires, some studies demonstrated that observers prefer high over low frame rates (17) when viewing noncinematic content (e.g. video games), while other studies report preferences for lower refresh rates when viewing cinematic content, likely due to habituation to the idiosyncratic aesthetic of cinematic content (soap opera effect; (16)). Those diverse results highlight the influence of factors beyond low-level properties of a stimulus on the subjective visual experience of observers and the limitations of using purely subjective measures when attempting to quantify observer's viewing experience. Thus, while measurements of the experience with a display are often purely based on subjective perceptual metrics, using additional behavioral markers, such as the oculomotor response to some dynamic content, might help disambiguate between different factors that influence the experience when judging the fidelity by which digital displays reproduce continuous motion. Pursuit quality is closely linked with the ability to successfully interact with moving objects (18, 19), and therefore, a reduction in it can have severe consequences for immersive, interactive virtual environments. Such an effect would not be captured by simple questionnaires asking about perceived display quality or the visibility of artifacts but requires detailed behavioral studies. This highlights the additional power of implicit, objective behavioral markers and demonstrates the relevance of using continuous oculomotor markers in digital display design (20).

Acknowledgments

The authors thank Benedikt Bohn and Nils Borgerding for helping with data collection.

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This study was partially funded by Meta Platforms, Inc. I.W., K.R.G., and A.G. were also supported by the Deutsche Forschungsgemeinschaft (DFG), project number 222641018, SFB/TRR135 Project A1.

Author Contributions

Experimental design, data analysis, and modeling: I.W. and A.G.; data collection, data visualization, and writing the first draft: I.W.; revising the manuscript: I.W., T.S.M., K.R.G., and A.G.; funding acquisition: K.R.G. and A.G.

Data Availability

Data are publicly available at osf: https://osf.io/2agtk/.

References

1 Mantiuk RK, Denes G. 2021. 34-5: Adaptive refresh rate and resolution: exploiting the limitations of spatio-temporal vision. *SID Int Symp Dig Tech Pap*. 52(1):465–468.

2 Castet E. 1995. Apparent speed of sampled motion. *Vision Res*. 35(10):1375–1384.

3 Boström KJ, Warzecha A-K. 2009. Ocular following response to sampled motion. *Vision Res*. 49(13):1693–1701.

4 Churchland MM, Lisberger SG. 2001. Shifts in the population response in the middle temporal visual area parallel perceptual and motor illusions produced by apparent motion. *J Neurosci*. 21(23):9387–9402.

5 Liston DB, Stone LS. 2014. Oculometric assessment of dynamic visual processing. *J Vis*. 14(14):12.

6 Coudiere A, Danion FR. 2024. Eye-hand coordination all the way: from discrete to continuous hand movements. *J Neurophysiol*. 131(4):652–667.

7 Rashbass C. 1961. The relationship between saccadic and smooth tracking eye movements. *J Physiol*. 159(2):326–338.

8 Adelson EH, Bergen JR. 1985. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*. 2(2):284.

9 Watson AB, Ahumada AJ, Farrell JE. 1986. Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. *J Opt Soc Am A*. 3(3):300.

10 Lisi M, Morgan MJ, Solomon JA. 2022. Perceptual decisions and oculomotor responses rely on temporally distinct streams of evidence. *Commun Biol*. 5(1):189.

11 Simoncini C, Perrinet LU, Montagnini A, Mamassian P, Masson GS. 2012. More is not always better: adaptive gain control explains dissociation between perception and action. *Nat Neurosci*. 15(11):1596–1603.

12 Tavassoli A, Ringach DL. 2010. When your eyes see more than you do. *Curr Biol*. 20(3):R93–R94.

13 Spering M, Carrasco M. 2015. Acting without seeing: eye movements reveal visual processing without awareness. *Trends Neurosci*. 38(4):247–258.

14 Goettker A, Stewart EEM. 2022. Serial dependence for oculomotor control depends on early sensory signals. *Curr Biol*. 32(13): 2956–2961.e3.

15 Wilmer JB, Nakayama K. 2007. Two distinct visual motion mechanisms for smooth pursuit: evidence from individual differences. *Neuron*. 54(6):987–1000.

16 Chapiro A, Atkins R, Daly S. 2019. A luminance-aware model of judder perception. *ACM Trans Graph*. 38(5):1–10.

17 Wilcox LM, Allison RS, Helliker J, Dunk B, Anthony RC. 2015. Evidence that viewers prefer higher frame-rate film. *ACM Trans Appl Percept*. 12(4):1–12.

18 Schroeger A, Goettker A, Braun DI, Gegenfurtner KR. 2024. Keeping your eye, head, and hand on the ball: rapidly orchestrated visuomotor behavior in a continuous action task. bioRxiv 628893. https://doi.org/10.1101/2024.12.17.628893, preprint: not peer reviewed.

19 Fooken J, Kreyenmeier P, Spering M. 2021. The role of eye movements in manual interception: a mini-review. *Vision Res*. 183: 81–90.

20 Goettker A, MacKenzie KJ, Murdison TS. 2020. Differences between oculomotor and perceptual artifacts for temporally limited head mounted displays. *J Soc Inf Disp*. 28(6):509–519.