# The Impact of Visually Simulated Self-Motion on Predicting Object Motion – A Registered Report Protocol

Björn Jörges¹\*, Laurence R. Harris¹

¹ Center for Vision Research, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada

\* Corresponding Author

# Abstract

To interact successfully with moving objects in our environment we need to be able to predict their behavior. Predicting the position of a moving object requires an estimate of its velocity. When flow parsing during self-motion is incomplete – that is, when some of the retinal motion created by self-motion is incorrectly attributed to object motion – object velocity estimates become biased. Further, the process of flow parsing should add noise and lead to object velocity judgements being more variable during self-motion. Biases and lowered precision in velocity estimation should then translate to biases and lowered precision in motion extrapolation. We investigate this relationship between self-motion, velocity estimation and motion extrapolation with two tasks performed in a realistic virtual reality (VR) environment: first, participants are shown a ball moving laterally which disappears after a certain time. They then indicate by button press when they think the ball would have hit a target rectangle positioned in the environment. While the ball is visible, participants sometimes experience simultaneous visual lateral self-motion in either the same or in the opposite direction of the ball. The second task is a two-interval forced choice task in which participants judge which of two motions is faster: in one interval they see the same ball they observed in the first task while in the other they see a ball cloud whose speed is controlled by a PEST staircase. While observing the single ball, they are again moved visually either in the same or opposite direction as the ball or they remain static. We expect participants to overestimate the speed of a ball that moves opposite to their simulated self-motion (speed estimation task), which should then lead them to underestimate the time it takes the ball to reach the target rectangle (prediction task). Seeing the ball during visually simulated self-motion should increase variability in both tasks. We expect to find performance in both tasks to be correlated, both in accuracy and precision.

# Introduction

We are constantly immersed in complex, dynamic environments that require us to interact with moving objects, for example when passing, setting, or hitting a volleyball, or when deciding whether we can make a safe left turn before another car reaches the intersection. In such situations, it can often be important to predict how objects in our environment will behave over the next few moments. Predictions allow us, for example, to time our actions accurately despite neural delays [1–3] in perceiving moving objects and issuing and executing motor commands [4,5]. Delays of between 100ms and 400ms between visual stimulation and motor response are generally assumed [6]. Without predicting or anticipating motion, we would thus always be acting on outdated positional information and be doomed to fail when attempting to intercept moving objects, particularly when they are relatively fast. Further, and perhaps more obviously, predictions are also important when moving objects are occluded during parts of their trajectory [7–9], or when the observer has to avert their eyes or turn away from the target.

When an observer is moving while attempting to interact with moving objects in their environment, further difficulties arise. Even in the simplest case, when the observer has a good view of the target and predictions are only necessary to compensate for neural delays, the visual system needs to separate retinal motion created by observer motion from retinal motion due to object motion in order to judge an object’s trajectory accurately. A prominent hypothesis on how this is achieved is the Flow Parsing hypothesis [10–16]. This hypothesis states that to solve this problem humans parse optic flow information and subtract this visual stimulation attributed to self-motion from global retinal motion. The remaining retinal motion can then be scaled with an estimate of the distance between the observer and the moving object [17] to obtain the world-relative velocity of the object. While this process was originally posited as a purely visual phenomenon [18,19], more recent studies have shed some light on the multisensory nature of this process: Dokka and her colleagues [20] found, for example, that compensation for self-motion in speed judgements was more complete when both visual and vestibular information indicated self-motion than when either visual or vestibular cues indicated that the observer was at rest. They further found precision to be lower in the presence of self-motion, which they attributed to self-motion information being noisier than optic flow information. Subtracting noisy self-motion information [21] from less noisy optic flow [20] would then make the estimate of object motion noisier in a moving observer than in a static observer. It is important to note that such a flow parsing mechanism should be active even while the participant at rest. The assumption here is that the noise added during flow parsing is proportional to the self-motion speed (in a Weber’s Law-like fashion), that is, when the observer is not moving at all the added noise is minimal, whereas higher self-motion speeds lead to more noise.

Several studies have shown that flow parsing is often incomplete [9,22–27]. A candidate explanation for this incompleteness is an inaccurate estimation of self-motion: typically, studies have used stimuli where only some cues indicated that the observer was moving (usually visual and/or vestibular cues), while efferent copy cues indicated that their bodies were at rest. If all self-motion cues are integrated after weighting them (e.g., according to their relative reliabilities [21,28,29]), this would then lead to an underestimation of self-motion, which is consistent with the biases found in the studies cited above. Further, while we did not find evidence for an effect of self-motion on precision in a recent study [30], we believe it is likely due to the fact that the effect was noisier than anticipated, resulting in a lack of statistical power.

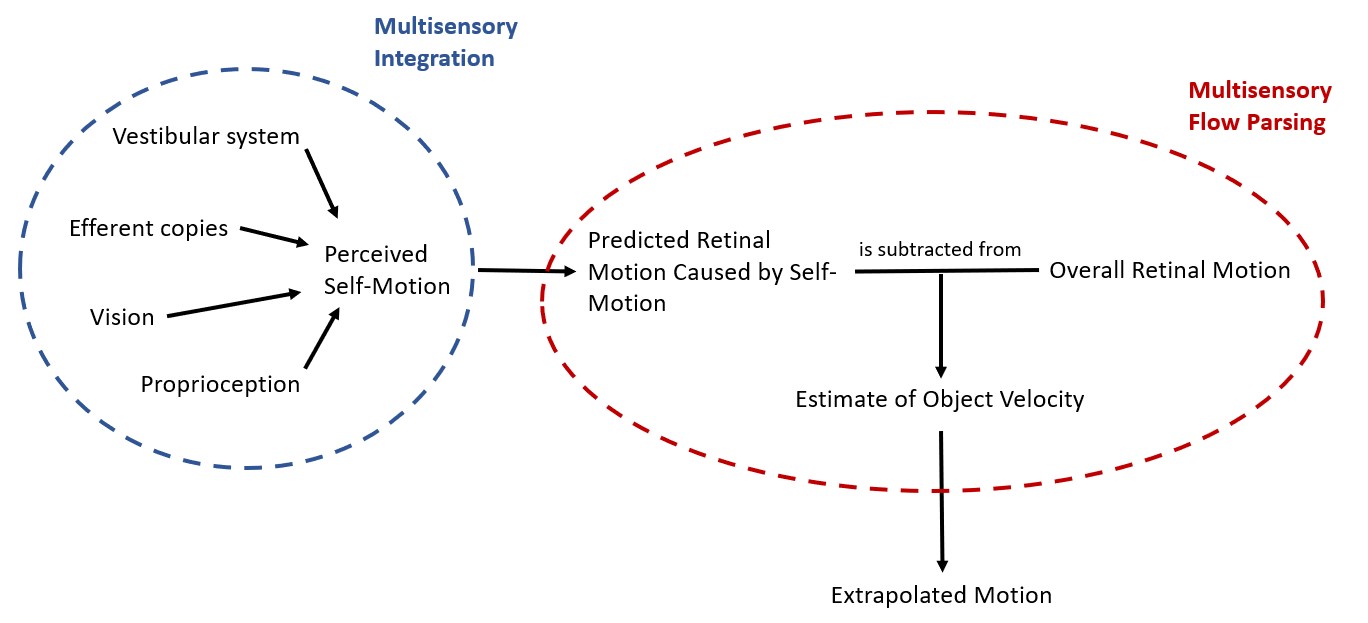


Figure 1: Schematic of the processes at play when predicting motion during self-motion. First, an estimate of the self-motion velocity is obtained by integrating the available cues from different sensory sources. This velocity estimate is used to predict the retinal motion that would be caused by self-motion. Finally, an estimate of the physical object velocity is obtained by subtracting the predicted retinal motion due to self-motion from the global retinal motion and the further trajectory of the object is extrapolated based on this estimate.

Figure 1 shows a simple schematic of the processes we assume to be at play when predicting object motion during self-motion: the organism first estimates its own motion in the environment from the various sources of information available to it. Based on this self-motion estimate, the organism then makes a prediction about the retinal motion that this self-motion would be expected to create. The predicted retinal motion is then subtracted from the total observed retinal motion and any remaining motion is attributed to the object, a process we call “multisensory flow parsing” to distinguish it from the purely visual conceptualization of flow parsing brought forward by authors like Wertheim [15,16] or, more recently, Dupin and Wexler [10]. This step adds noise to the velocity estimate because the self-motion estimate is noisier than the retinal input [20]. The further trajectory of the object would then be extrapolated based on this estimate of its velocity.

This schematic provides an overview over the general mechanism that might be at play when predicting future motion of an object observed during self-motion. Depending on the motion profiles of observer and object further complications may arise. For example, retinal speeds depend not only on the physical speeds and directions of the observer and object but also on the distance between them and can therefore change systematically even when observer and object move at constant physical speed without changing direction. To obtain a veridical representation of the physical velocity of the object, the observer thus has to perform additional computations, including estimating their distance to the object, the direction of the object in relation to their own direction of motion, and the necessary transformations to obtain the physical object speed from these values [17].

Some studies suggest that biases incurred while estimating motion, e.g., due to the Aubert-Fleischl effect which lowers the perceived speed of a target when an observer tracks it with their gaze [31], or due to low contrast in the stimulus [32], might transfer to biases in motion extrapolation based on these speed estimates. It seems straight-forward that any biases and precision differences observed in the perception of speed would correlate perfectly with errors and precision differences in time-to-contact judgements. However, there are two complications: first, participants might integrate biased and less precise speed information obtained during self-motion with prior information they have formed in response to previous exposure to the stimulus. They might thus extrapolate motion biased on a combination of prior information and (biased and more variable) online information. Further, it has been reported that under certain circumstances perceptual judgements and action-related tasks can be based on separate cues (see, e.g., [33–35]). While it remains an appealing hypothesis, it should thus not be assumed that biases and variability differences in time-to-contact judgements reflect only biases and variability differences in speed estimation. Studying to what extent biases and variability differences in online information acquired while viewing a target influence the way we extrapolate its further motion will help us better understand the predictive mechanisms at play not only when the target is occluded or the observer averts their gaze from the target but also when timing interceptive actions accurately despite unavoidable neural delays [36]. In the present study, we therefore investigate how biases in speed estimation elicited by visual self-motion impact the prediction of object motion.

More specifically, we will test three interconnected hypotheses:

* Predictions of where a moving object will be in the future will be biased (Hypothesis 1a) and more variable (Hypothesis 1b) in the presence of visually simulated self-motion
* Object speed judgements will be biased (Hypothesis 2a, which constitutes a replication of an earlier result of Jörges & Harris, 2021) and more variable (Hypothesis 2b, a more highly powered follow-up to a hypothesis for which Jörges & Harris, 2021 did not find significant support) in the presence of visually simulated self-motion
* The effect of visually simulated self-motion on motion extrapolation can be predicted from its effect on speed estimation, both in terms of bias (Hypothesis 3a) and its variability (Hypothesis 3b)

# Methods

## Apparatus

We programmed all stimuli in Unity 2020.3.30f1. Given the on-going COVID-19 pandemic, some participants who are owners of head-mounted VR devices (HMDs) are tested in the safety of their home. Our experiment is compatible with all major modern HMDs. To minimize variability introduced by the use of different HMDs, each program we use to present stimuli limits both the horizontal and the vertical field of view to 80°. We don’t expect there to be relevant differences in frame rate, as Unity caps the frame rate at 60 Hz. Our programs are, further, much less demanding in processing power than any VR application remote participants are likely to run; that is, frame rate dropping below 60 Hz should occur almost never. If it is safely possible and an ethics approval is granted, we might also test some participants in person in our laboratory. For these participants, we will use an VIVE Pro Eye.

## Participants and Recruitment

We recruit participants with HMDs in their possession online for them to perform the experiment in their homes. Recruitment occurs through social media (such as Twitter, Facebook, and Reddit). Some remote participants might also be recruited through the professional recruitment service XpertVR. Since recruiting participants with VR equipment at home is not trivial, we might also rely on York University participant pools to recruit participants for in-person testing. In this case, all applicable guidelines for safe face-to-face testing are fulfilled and exceeded. Participants receive a monetary compensation of 45 CAD for participation in the experiment; participants recruited through the York University participant pools may receive course credit instead of a monetary compensation. All participants are screened for stereo-blindness with a custom Unity program (downloadable on GitHub: https://github.com/b-jorges/Stereotest) in which participants have to distinguish the relative depth of two objects that are matched in retinal size. Participants are only included if they answer correctly on 16 out of 20 trials. The simulated disparity is 200 arcsec. While this allows only for a coarse assessment of the participants’ stereovision, our experiment is not critically dependent on a high stereoacuity. We test 20 men and 20 women (see Power Analysis). The experiment was approved by the local ethics committee and is conducted in accordance with the Declaration of Helsinki.

## Stimulus

Each participant performs two main tasks in an immersive VR environment: a prediction task and a speed estimation task. Every participant completes both tasks, and we counterbalance the order in which they complete them such that 20 participants (10 women and 10 men) start with the prediction task, and 20 participants (10 women and 10 men) start with the speed estimation task. All programs we used to present the stimuli are available on Open Science Foundation (https://osf.io/gakp5/), and the Unity projects can be downloaded on Open Science Foundation as well (https://osf.io/6mz4w/).

For both tasks, we display a circle in the middle of the observer’s field of view that moves with their head rotation in front of the Earth-stationary simulated world. Participants are instructed to keep this circle surrounding the fixation cross. When the center of the circle is with 2.5° (vertically and horizontally) of the fixation cross, it disappears to indicate to the participant that their head is positioned correctly. We further record head position whenever the ball or the ball cloud (see below) are visible. Since recording head position on each frame might slow down the program on older systems, we opt to record the mean rotation (horizontally and vertically) over bins of five frames.

**Prediction** – In the prediction task (see Figure 2A, and see also this video on OSF (https://osf.io/rkg23/), we first show participants a ball of 0.4 m diameter moving laterally 8m in front of them at one out of three speeds (4, 5, 6 m/s). We have used this range of speeds in our previous study [30], while the size of the ball was diminished slightly in comparison to this study (see description of the speed estimation task for the rationale). The ball can travel to the left or to the right. It appears to the left of the observer when it travels to the right, and on the right of the observer when it travels to the left such that the visible part of the trajectory is centered in front of the observer. At the same time, a target rectangle is presented on the side towards which the ball is travelling. The ball disappears after 0.5s and participants press the space bar on their keyboard in the moment they think the ball would hit the target. In order to curtail biases in speed perception due to the Aubert-Fleischl phenomenon [37], participants are asked to keep their eyes on a fixation cross that is presented straight ahead of them and slightly below the stimulus (see Figure 2A) and moves with the observer when they experience visual self-motion. The target is presented at a distance that depends on the speed of the ball and the occlusion duration, which can be 0.5 s, 0.6 s, or 0.7 s. Speeds and occlusion durations are chosen such that, when the participant keeps their gaze on the fixation cross, the whole trajectory (including the invisible part) unfolds within a field of view of 60°, which is well within the effective field of view of any modern HMD. The distance between the point where the ball disappears (the “point of disappearance”; see Figure 2C) and the target rectangle is given by the following equation:

|  |  |
| --- | --- |
|  | [1] |

While the target is visible, participants experience lateral visual self-motion either in the same direction as the ball or in the opposite direction as the ball, or they remain stationary. The self-motion speed ramps up in a Gaussian fashion over the first 50 ms until it reaches 4 m/s, then remains constant for 400 ms, and finally ramps down again over the last 50 ms before the ball becomes invisible. Overall, the observer moves 1.8 m over 500 ms. Please note that the different motion profiles elicit very different retinal speeds: observer motion in the opposite direction of the ball elicits higher retinal speeds overall than for a static observer or for observer motion in the same direction as the ball. Table 1 displays the mean absolute retinal speeds across the trajectory for all conditions. While our previous results [30] suggest that the role of the retinal speeds for the overall precision in motion estimation is subordinate to other sources of variability, retinal speeds are highly correlated with the expected effect of our self-motion manipulation on variability.

Table 1: Mean absolute retinal speeds across the visible part of the trajectory for each combination of ball speed and observer motion profile. The script in which we derive these values can be found on GitHub (https://github.com/b-jorges/Predicting-while-Moving/blob/main/Geometry%20Prediction.R).

|  |  |  |  |
| --- | --- | --- | --- |
|  | Ball Speed | | |
|  | 4 m/s | 5 m/s | 6 m/s | |
| Observer Static | 0.4 °/s | 28.4 °/s | 34.0 °/s | |
| Same Direction | 22.2 °/s | 5.9 °/s | 11.6 °/s | |
| Opposite Directions | 44.4 °/s | 50.4 °/s | 55.8 °/s | |

We further add a range of occlusion durations (0.1s, 0.2s, 0.3s, 0.4s, 0.8s, 0.9s, 1s) while the observer is static to get an estimate of how variability changes in response to different occlusion durations. Overall, participants complete 225 trials (3 ball speeds \* 3 self-motion profiles \* 3 occlusion durations \* 5 repetitions + 3 ball speeds \* 6 occlusion durations \* 5 repetitions), which takes around 10 minutes.

Participants complete a brief training of 18 trials before starting the main experiment (see this video on OSF: https://osf.io/4js5w/). The ball travels at one of three speeds (2.5, 3.5, 4.5 m/s), going either left or right, and three occlusion durations (0.45, 0.55, 0.65 s). In the training, the ball reappears upon pressing the spacebar in the position it would have been in at that moment. No visual self-motion is simulated in the training. This allows participants to estimate their error (spatially) and helps them familiarize themselves with the task and the environment.

|  |  |
| --- | --- |
| **A** | **B** |
| **C** | **D** |

Figure 2: **A**. Screenshot from the prediction task while the ball was visible. **B**. Screenshot from the speed estimation task while the ball cloud was presented. **C**. Schematic of the prediction task. **D**. Schematic of the speed estimation task.

**Speed estimation –** In the speed estimation task (video on OSF: https://osf.io/xqkgy/), participants are presented with two motion intervals and have to judge which of them is faster. In one interval, they view a ball travelling to the left or to the right. As for the prediction task, this ball can have one of three speeds (4, 5, 6 m/s), and the participant can also experience visual self-motion in the same direction or in the opposite direction or remain static (see Figure 2D). Except for the self-motion intervals, the scene is static just like in the prediction experiment, and object motion occurs relative to the scene. The second motion consists of a ball cloud of 2.5 m width and 1 m height at the same distance to the observer (see Figure 2B). Each ball in this cloud has the same diameter as the main target (0.4m) and balls are generated at one side of the cloud and then move to the other side where they disappear. In our previous study [30], we used smaller balls for the ball cloud, which may have been a factor their judgements of speed being consistently overestimated relative to the single ball. We therefore decided to use the same ball size for both the single ball and the elements of the ball cloud. At any given moment, between 8 and 12 balls are visible. All the balls in the ball cloud move at the same speed and are visible either until they reach the opposite side of the cloud area or until the motion interval ends after 0.5s observation time. The speed of these balls is constant throughout each trial and is governed by PEST staircase rules [38]. For each condition, we employ two pests: one starts 30% above the speed of the single ball from the other motion interval, while the other starts 30% below. The initial step size is 0.6 m/s and each pest terminates either after 37 trials, or when the participant has completed at least 30 trials and the step size drops below 0.03 m/s. We modified the original PEST rules such that the step size is always twice the initial step size (that is, 1.2 m/s) for the first ten trials in order to spread out the values presented to the observer and allow for more robust JND estimates. We limit the range of speeds the ball cloud can take to between one third of the speed of the single ball and three times the speed of the single ball. Participants are asked to maintain fixation on the fixation cross that had the same characteristics as in the prediction task, that is, it is always presented slightly below the stimulus and it moves with the participant as they experience visual self-motion. To keep the visual input identical across both tasks, the target rectangle from the prediction task, while irrelevant for the speed estimation task itself, is present in this task as well. Overall, participants perform between 30 and 37 trials in 18 staircases (two start values, three speeds and three motion profiles) for a total of between 540 and 666 trials.

Before proceeding to the main tasks, participants complete a training session. This training session consists of one PEST of reduced length (between 20 and 27 trials) that starts 30% above the speed of the ball (3 m/s). Participants need to achieve a final step size of below 0.3; otherwise, they are asked to repeat the training. If they fail the training a second time, we exclude them from the analysis. The participants do not experience visually simulated self-motion in this training. This task – including the training – takes about 40 minutes to complete.

Participants can choose to receive the instructions as PDF (can be downloaded from GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Instructions%20Predicting%20while%20moving.pdf) or watch a video (which can be viewed on YouTube: <https://youtu.be/qHTWVyjn0QI> and <https://youtu.be/JyOZ-duRGmU>, respectively).

## Modelling the Predictions

To obtain specific predictions corresponding to each hypothesis, we built models of the underlying perceptual processes for both the prediction and the speed estimation task. The instantiation of the model for the prediction task can be found here (on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Analysis%20Prediction.R), and the instantiation of the speed estimation model can be found here (on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Analysis%20Speed%20Estimation.R). The implementation of the model that relates performance in both tasks can be found here (on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Predictions%20Correlations.R). While a detailed discussion of these models can be found in Appendix A, the most important assumptions are the following. The first assumptions reflect our hypotheses:

* The speed of the ball is overestimated by 20% of the presented self-motion speed when observer and ball move in opposite directions [30], with a between-participant standard deviation of 30%. In the prediction task, this overestimation of speed should lead to an underestimation of the time it takes for the ball to travel the occluded distance. No biases are assumed for the Same Directions and Observer Static motion profiles. This assumption reflects Hypothesis 2a.
* While we previously did not find evidence for an impact of visual self-motion on precision [30], we believe that the higher self-motion speeds in this study might enable us to uncover small effects that were not apparent at lower self-motion speeds. The variability in perceived speed is 20% higher for the Opposite Directions motion profile, with a between-participant standard deviation of 30%. No differences in variability are assumed for the Opposite Directions and Static motion profile. This assumption reflects Hypothesis 2b.
* The same effects of self-motion on accuracy and precision are also at play in the prediction task. This assumption reflects Hypotheses 1a and 1b.
* Participants display the same biases in perceived object speed in response to self-motion in the opposite direction of the ball in both the speed estimation and the prediction task. Similarly, variability is impacted equally in both tasks. This assumption reflects Hypotheses 3a and 3b.

We further need to make several assumptions about how the participants process the stimulus independently of the presence of visual self-motion:

* We neglect how exactly participants recover physical speed from angular speed and the perceived depth but to acknowledge the added complexity, we assume a Weber Fraction of 10% for the estimation of the ball speed, which is slightly higher than is generally reported for less complex speed discrimination tasks [39], with a between-participant standard deviation of 1.5%.
* *Prediction task only:* We assume that the computations executed by the visual system are approximated accurately with the physical equation for distance from speed and time (d = v\*t), such that the extrapolated time () can be estimated from the distance between the point of disappearance and the target () and the perceived speed of the ball ():

|  |  |
| --- | --- |
|  | (3) |

* *Prediction task only:* The distance between the point of disappearance of the ball and the target rectangle (that is, the occluded distance) is estimated with a Weber Fraction of 5%, as reported in the literature [40]. We further assume that this distance is estimated accurately or, if biases in depth perception impact accuracy, that these biases also impact the perceived speed of the ball in such a way that these biases cancel out. Out of these two scenarios, we believe the latter is more likely, as it is quite established in the literature that depth is underestimated in VR [41], and an underestimation of depth would lead to an underestimation of the physical distance between the point of disappearance and the target rectangle. However, this bias in depth perception should also lead the observer to underestimate the physical speed of the ball in the same way, causing both biases to cancel each other out. Between-participant variability is neglected here.

**Prediction** – For the prediction task, under the above assumptions, participants are expected to respond between 0.12 and 0.2s earlier during the Opposite Direction motion profile than during the Static motion profile (Figure 3A). Our model further predicts that visual self-motion during motion observation will lead to higher variability in responses. Measuring the relation between self-motion and variability is not straight-forward because self-motion should cause an underestimation of the occlusion duration. If most noise behaves according to Weber’s Law, noise should be proportional to the mean length of the extrapolated interval. A shorter predicted interval should thus in turn be related to lower variability (in absolute terms) even if self-motion has no direct effect on precision. Figure 3B illustrates the expected relationship between biases in prediction, the motion profile, and variability in responses.

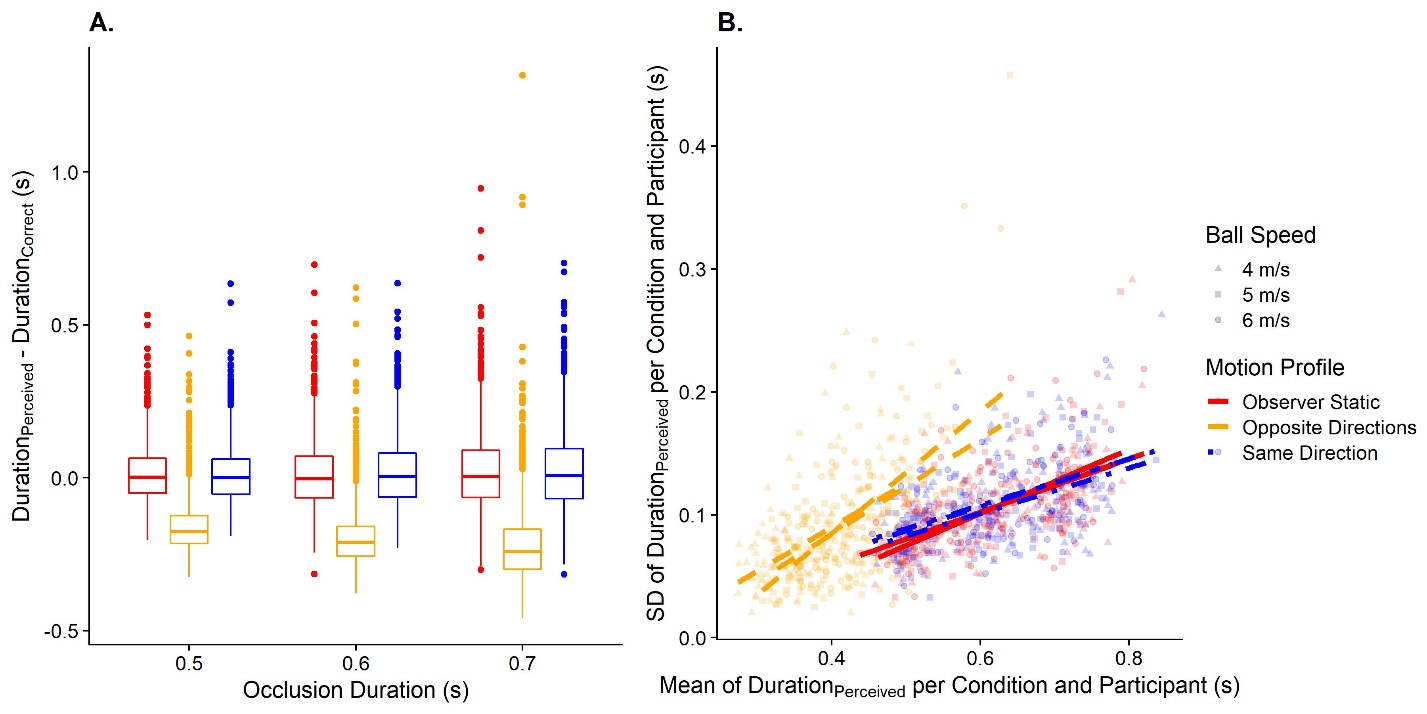


Figure 3: A. Predicted data for the timing error in the prediction task, divided up by occlusion durations (x axis) and motion profile (color-coded; left-most: “Observer Static”; in the middle: “Opposite Directions”; right-most “Same Directions”). B. Predicted data for variability in the prediction task. The y axis displays the standard deviation of the extrapolated duration per condition and participant, while the x axis corresponds to the mean of the extrapolated duration per condition and participant. The motion profile is coded with different colors and line types (red and continuous for “Observer Static”, yellow and dashed for “Opposite Directions” and blue and dashed-and-dotted for “Same Direction”). The lines are regression lines for their respective condition.

**Speed estimation** – Here, we expect to replicate the findings from our previous study [30]: There we found that participants largely estimated speed with the same degree of accuracy when they were static as when they were moving in the same direction as the target. In line with these results, visually simulated self-motion in the opposite direction to the ball should lead to an overestimation of ball speed (Hypothesis 2a; see Figure 4A). Since we use a higher self-motion speed than in our previous study, we also expect that precision will be lower for visual self-motion in the opposite direction to the ball (Hypothesis 2b, see Figure 4B).

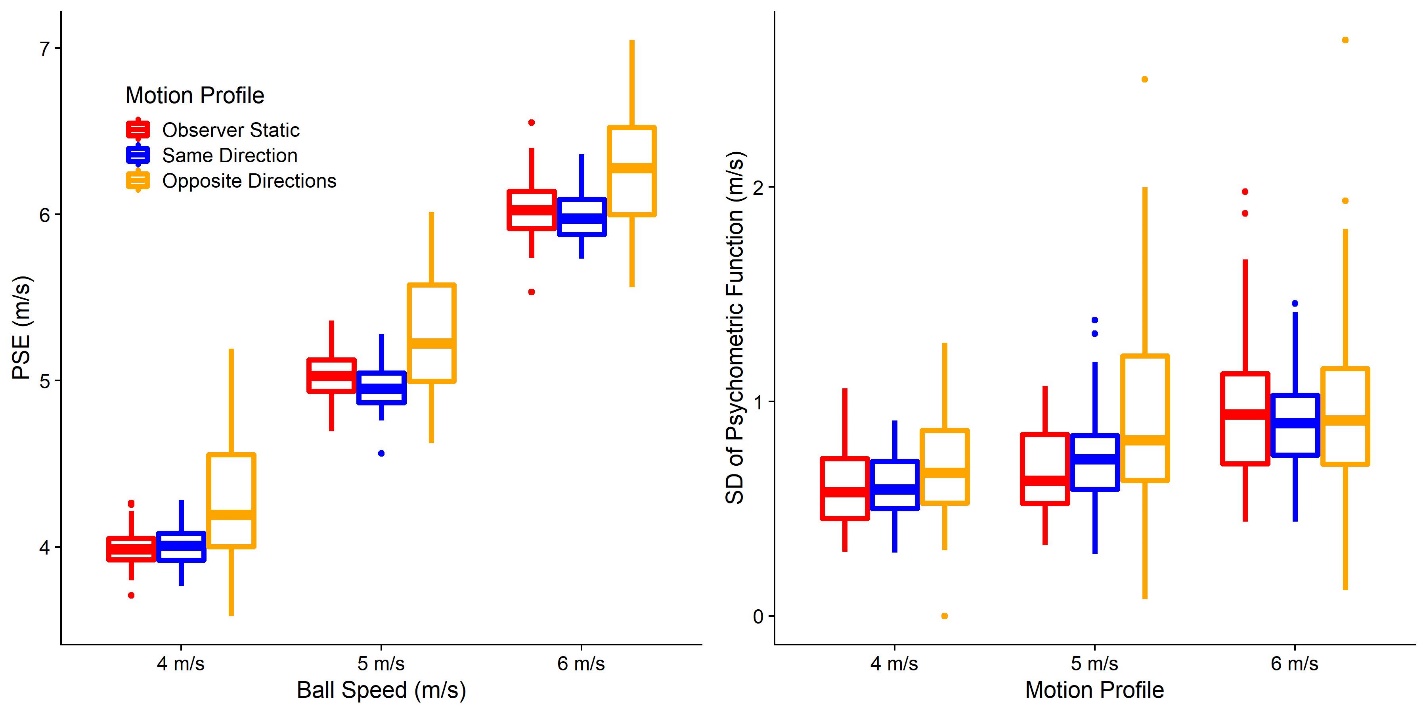


Figure 4. A. Predicted PSEs (y axis) for each ball speed (x axis) and motion profile (color-coded; left-most: “Observer Static”; in the middle: “Opposite Directions”; right-most “Same Directions”). B. As A. but for the predicted JNDs.

**A link between speed estimation and predicted time-to-contact** – We further expect the errors observed in the prediction task in response to self-motion to correlate with the errors in the speed estimation task in response to self-motion, indicating that performance in speed perception translate to errors in predicted time-to-contact, both in terms of accuracy (Figure 5A) and precision (Figure 5B).

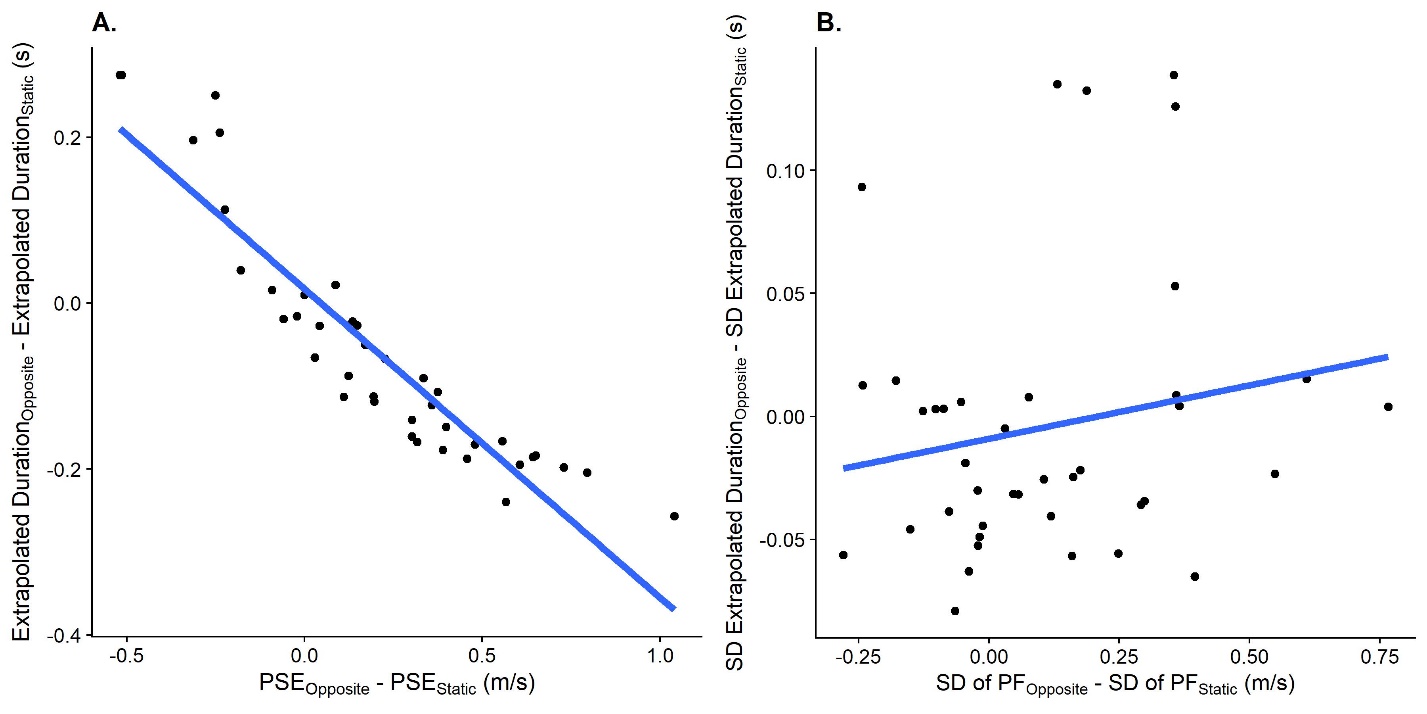


Figure 5: A. Relationship between the difference in PSEs between the Opposite Directions motion profile and the Observer Static motion profile in the speed estimation task (x axis) and the difference in predicted durations between these motion profiles (y axis). One data point corresponds to one participant. B. As A., but for the relation between the JND differences in the speed estimation task between the “Opposite Directions” motion profile and the “Observer Static” motion profile and the differences in standard deviations between these motion profiles.

## Data analysis plan

We first perform an outlier analysis. For the prediction task, we exclude all trials where the response timing was more than three times the occlusion duration, which indicates that the participant has not paid attention and missed the trial. For the speed estimation task, we exclude participants where more than 20% of presented ball cloud speeds were at the limits we set for the staircase (one third of the speed of the single ball and three times the speed of the single ball). For all analyses related to precision, we further exclude all conditions where we obtained a standard deviation of 0.01 or lower. According to our simulations, this should occur very rarely, and taking the log of such low values, as we do for the precision analyses to counteract the expected skew in these distributions, would lead to extremely small numbers that could bias results unduly. We also remove all trials where the head rotation was outside of the permitted range (+- 2.5°) for half or more of the recorded bins.

Unless noted otherwise, we compute bootstrapped 95% confidence intervals as implemented in the confint() function for base R [42] to determine statistical significance.

**Prediction** – To test Hypotheses 1a regarding accuracy, we use Linear Mixed Modelling as implemented in the lme4 package [43] for R. The corresponding script can be found here (on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Analysis%20Prediction.R). We fit a model with the temporal error as dependent variable, the motion profile (“Observer Static”, “Same Direction” and “Opposite Directions”) as fixed effect, and random intercepts and random slopes for the speed of the ball per participant, as well as random intercepts for the occlusion duration as random effects. In Wilkinson & Rogers notation (1973), this model reads as follows:

|  |  |
| --- | --- |
| *+ (1 | Occlusion Duration)* | (4) |

We expect the regression coefficient corresponding to the motion profile “Opposite Directions” to be negative and significantly different from zero.

For Hypothesis 1b regarding precision, we need to take into account one possible confound: differences in timing accuracy can impact variability: overestimating the time it takes the ball to hit the rectangle could be connected to a higher variability, while underestimating the time could lead to lower variability. For this reason, we first compute the means and standard deviations of extrapolated durations for each condition and participant. We then fit a test model with the standard deviations as dependent variable, the mean timing error and the motion profile as fixed effects, and random intercepts as well as random slopes for ball speeds per participant and random intercepts for the occlusion durations as random effects:

|  |  |
| --- | --- |
| *+ (1 | Occlusion Duration)* | (5) |

We further fit a null model without the motion profile as fixed effect:

|  |  |
| --- | --- |
| *+ (1 | Occlusion Duration)* | (6) |

We then compare both models by means of a Likelihood Ratio Test to determine whether the motion profile explains significantly more variability than the test model which already takes into account biases in extrapolated time. We will not interpret the regression coefficients as means and standard deviations are likely to be correlated, which may lead to biased regression coefficients.

**Speed Estimation** – To test Hypotheses 2a and 2b (script can be found here, on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Analysis%20Speed%20Estimation.R) regarding speed estimation, we first use the R package quickpsy [45] to fit psychometric functions to the speed estimation data, separately for each participant, speed and motion profile. Quickpsy fits cumulative Gaussian functions to the data by direct likelihood maximization. The means of the cumulative Gaussians correspond to the Points of Subjective Equality (PSEs) and their standard deviations correspond to the 84.1% Just Noticeable Differences (JNDs).

To assess whether the motion profile biased the perceived speed significantly, we fit a Linear Mixed Model with the PSEs as dependent variable, the self-motion profile as fixed effect, and random intercepts and random slopes for the ball speed per participant as random effects:

|  |  |
| --- | --- |
|  | (7) |

We expect that the regression coefficient for the motion profile “Opposite Directions” will be positive and significantly different from zero, indicating that speed is overestimated when observer and target move in opposite directions as compared to when the observer is static.

Regarding precision, the same considerations apply as for the prediction task: in addition to a direct effect of the self-motion profile, biases elicited by the different self-motion profiles can impact precision. For this reason, we use a model comparison-based approach similar to the one used above. Separately for the “Same Direction” and “Opposite Directions” motion profiles, we first fit a test model that contains the log JNDs as dependent variable, the self-motion profile and the PSEs as fixed effects, and random intercepts as well as random slopes for ball speed per participant as random effects.

|  |  |
| --- | --- |
|  | (8) |

We also fit a null model without the motion profile as fixed effect:

|  |  |
| --- | --- |
|  | (9) |

Finally, we compare both models with a Likelihood Ratio Test and we expect the test model (Equation 8) to be a significantly better fit than the null model (Equation 9). As for the prediction task, we will not interpret the regression coefficients obtained in this analysis.

**A link between speed estimation and prediction** – To test Hypotheses 3a and 3b (script can be found here, on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Analysis%20Correlation.R), we first prepare the prediction data by computing means and standard deviations of the extrapolated time for each participant. We then calculate the difference in performance (mean and standard deviations for the prediction task and PSEs and JNDs for the speed estimation task) between the “Opposite Directions” motion profile and the “Observer Static” motion profile for both tasks for each participant.

For accuracy, we then determine to what extent PSE differences between the Opposite Direction motion profile and the Observer Static motion profile obtained in the speed estimation task predict the mean extrapolated time in the prediction task. For this purpose, we fit a Linear Model with the difference in mean motion extrapolation errors between the motion profiles as the dependent variable and the difference in PSEs between the motion profiles as the independent variable:

|  |  |
| --- | --- |
|  | (9) |

We expect that the regression coefficient for the fixed effect “PSE Difference (Opposite – Static)” will be negative and significantly different from zero, indicating that a stronger effect of self-motion on PSEs is linked to a larger effect of self-motion on the estimated time-to-contact.

For precision, the same complication as for Hypothesis 1b applies: A correlation between the effect of visual self-motion on the precision of speed estimation and on the precision of the predicted times-to-contact could be due to biases introduced by visual self-motion. If visual self-motion in the opposite direction, for example, leads to too-early responses, the extrapolated intervals become shorter. A shorter interval, in turn, would lead to higher precision. Therefore, to test whether the difference in precision observed in the speed estimation task was significantly related to the variability in the prediction task even after accounting for biases, we need to determine whether the effect of visual self-motion on JNDs predicts any variability beyond the variability that is already explained by the bias in motion extrapolation. To test this hypothesis, we first fit a test model with the variability difference between the “Opposite Directions” motion profile and the “Observer Static” motion profile in the prediction task as the dependent variable and the mean difference between these motion profiles and the difference in JNDs in the speed estimation task as the independent variables (as a measure of bias introduced by visual self-motion):

|  |  |
| --- | --- |
|  | (10) |

We also fit a null model without the JND difference as independent variable:

|  |  |
| --- | --- |
|  | (11) |

Then, we use a Likelihood Ratio Test to determine whether the test model (with the JND difference as fixed effect) was significantly better than the null model. We expect the test model (Equation 10) to be a significantly better fit than the null model (Equation 11). As above, the regression coefficient will not be interpreted.

**Model fitting –** These analyses only serve to demonstrate that performance in both tasks is related, but they don’t provide insight into how strongly they are related. We will therefore use the models outlined in the section “Modelling the Predictions”, and described more in detail in Appendix A, to fit parameters that capture biases and precision differences in perceived speed due to self-motion (which we had set to 20% on average to model our predictions and conduct the power analyses) in both tasks separately.

We will use a two-step approach to fit these parameters: since we expect biases to affect variability in performance but do not expect variability to affect biases, we first set the variability parameter (capturing the impact of self-motion on the variability of perceived speed) to zero and fit the accuracy parameter (capturing the impact of self-motion on mean perceived speed). To do so, we minimize the root median squared error between the observed mean difference in timing errors between baseline and the self-motion condition and the difference in timing error in the simulated dataset across all speeds and occlusion durations. We perform this optimization by using the Brent method [46] as implemented in the optimize function in base R. In the second step, we set the accuracy parameter for each participant to the one fitted in the first step and fit the precision parameter. Here, we minimize the root median squared error between the difference between the standard deviation of the observed difference in timing errors in the baseline condition and the self-motion condition and the respective simulated values. We use the same approach to obtain these parameters for the speed estimation task as well but use the observed and simulated PSEs (for accuracy) and JNDs (for precision).

Once we have obtained one accuracy parameter and one precision parameter for each participant and task, we perform a simple linear regression between the parameters fitted for the prediction task and the speed estimation task (separately for accuracy, see Equation 11, and precision, see Equation 12) to determine to what extent performance in one task is indicative of performance in the other:

|  |  |
| --- | --- |
|  | (11)  (12) |

Since these parameters are scaled the same way for both tasks (the effect of self-motion on accuracy/precision as a fraction of presented self-motion speed), we expect regression coefficients of around 1 in both analyses. A value of above 1 would mean that the effect of self-motion is stronger in the prediction task than in the speed estimation task, and vice-versa. To test this prediction we compute 95% confidence intervals which we expect to contain a value of 1.

**An effect of visual self-motion in the same direction as the ball** – While our earlier results [30] suggest that visual self-motion in the same direction as the observer should not have any effect on perceived speed, we perform all analyses outlined in this section equivalently for the “Same Direction” motion profile as well.

## Power analysis

Since power for complex hierarchical designs cannot be computed analytically, we used Monte Carlo simulations to determine power for all statistical tests outlined in the previous section: we used our models for the prediction task and the speed estimation task to first simulate full datasets. Then, we performed the analyses detailed above over each of these simulated datasets and determined the results for each combination of number of participants and number of trials. To keep the computational cost manageable, we used the faster, but more bias-prone Satterthwaite approximation, as implemented in the lmerTest package [47] for R, to assess statistical significance rather than bootstrapped confidence intervals. The script used for the power analyses can be found here (on GitHub: https://github.com/b-jorges/Predicting-while-Moving/blob/main/Power%20Analysis.R).

We repeated this process 250 times for all combinations of 20, 30 and 40 participants, 5, 9 and 13 repetitions per condition in the prediction task, and 20 to 27, 30 to 37, and 40 to 47 trials per pest, which makes for an average of 50, 70 and 90 trials per condition, respectively, for the speed estimation task. The results are shown in Figure 6. For the precision in the prediction task, using 9 repetitions per condition appears to add a considerable amount more power than using only 5 repetitions, while the added benefit of another 4 repetitions for a total of 13 is small. However, the prediction task is very quick to do, taking only around 10 minutes even with 13 repetitions per condition. Similarly, 70 trials per condition increases the power to detect an effect on precision significantly more than using only 50 trials, while the added benefit of 90 trials is marginal. Since the speed estimation task takes much longer and is more fatiguing than the prediction task, we judge this marginal increase in power to be not worth the additional time spent by the participant. We thus opt for a combination of 40 participants, 13 repetitions per condition in the prediction task, and 70 trials per condition in the speed estimation task, which allows us to achieve a power of at least 0.85 for all statistical tests.

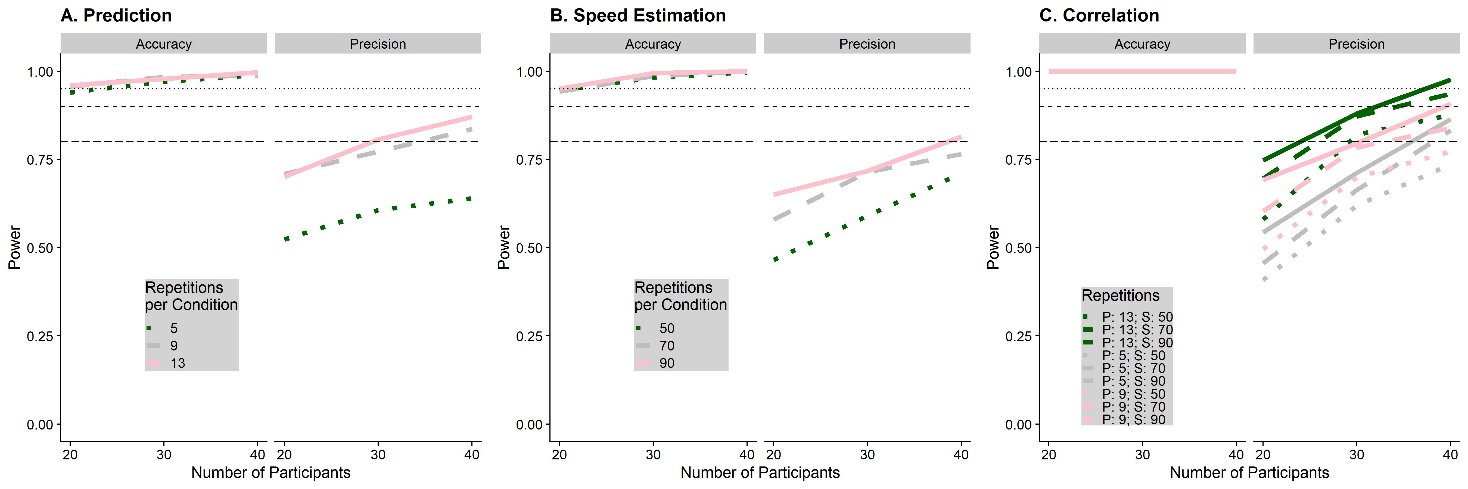


Figure 6: Simulated power for the prediction task (A), the speed estimation task (B) and the correlation between performance in speed estimation and speed prediction (C), separately for the statistical tests referring to biases (accuracy) and variability (precision). The number of participants for which we simulated power is on the x axis, while the number of trials for each task is coded with different shades of green and line types. The horizontal lines indicate a power level of 0.8, 0.9 and 0.95 respectively.

We also used our power analyses to determine that all of our statistical analyses led to an expected false positive rate of 0.05 in absence of a true effect.

# Acknowledgements

BJ and LRH are supported by the Canadian Space Agency (CSA) (CSA: 15ILSRA1-York). LRH is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada (NSERC: RGPIN-2020-06093). The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

# Competing Interests

The authors declare no competing interests.

# Data Availability and Open Science

All scripts and all data obtained with regards to this project will be made available in the project GitHub repository (https://github.com/b-jorges/Predicting-while-Moving/). Larger files such as the programs used to present stimuli and the respective Unity projects are hosted on OSF (https://osf.io/eayf7/).

# References

1. Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, et al. Signal timing access the macaque visual system. J Neurophysiol. 1998;79: 3272–3278. doi:10.1152/jn.1998.79.6.3272

2. Maunsell JHR, Gibson JR. Visual response latencies in striate cortex of the macaque monkey. J Neurophysiol. 1992;68: 1332–1344. doi:10.1152/jn.1992.68.4.1332

3. DiCarlo JJ, Maunsell JHR. Using neuronal latency to determine sensory-motor processing pathways in reaction time tasks. J Neurophysiol. 2005;93: 2974–2986. doi:10.1152/jn.00508.2004

4. de Rugy A, Marinovic W, Wallis G. Neural prediction of complex accelerations for object interception. J Neurophysiol. 2012;107: 766–771. doi:10.1152/jn.00854.2011

5. Zago M, McIntyre J, Senot P, Lacquaniti F. Internal models and prediction of visual gravitational motion. Vision Res. 2008;48: 1532–1538. doi:10.1016/j.visres.2008.04.005

6. Foxe JJ, Simpson G V. Flow of activation from V1 to frontal cortex in humans: A framework for defining “early” visual processing. Exp Brain Res. 2002;142: 139–150. doi:10.1007/s00221-001-0906-7

7. Jörges B, López-Moliner J. Earth-Gravity Congruent Motion Facilitates Ocular Control for Pursuit of Parabolic Trajectories. Sci Rep. 2019;9: 1–13. doi:10.1038/s41598-019-50512-6

8. Spering M, Schutz AC, Braun DI, Gegenfurtner KR. Keep your eyes on the ball: smooth pursuit eye movements enhance prediction of visual motion. J Neurophysiol. 2011;105: 1756–1767. doi:10.1152/jn.00344.2010

9. Kreyenmeier P, Fooken J, Spering M. Context effects on smooth pursuit and manual interception of a disappearing target. J Neurophysiol. 2017;118: 404–415. doi:10.1152/jn.00217.2017

10. Dupin L, Wexler M. Motion perception by a moving observer in a threedimensional environment. J Vis. 2013;13: 1–14. doi:10.1167/13.2.15

11. Warren PA, Rushton SK. Perception of object trajectory: Parsing retinal motion into self and object movement components. J Vis. 2007;7: 2. doi:10.1167/7.11.2

12. Rushton SK, Warren PA. Moving observers, relative retinal motion and the detection of object movement. Curr Biol. 2005;15: R542–R53. doi:10.1016/j.cub.2005.07.020

13. Warren PA, Rushton SK. Evidence for flow-parsing in radial flow displays. Vision Res. 2008;48: 655–663. doi:10.1016/j.visres.2007.10.023

14. Warren PA, Rushton SK. Perception of scene-relative object movement: Optic flow parsing and the contribution of monocular depth cues. Vision Res. 2009;49: 1406–1419. doi:10.1016/j.visres.2009.01.016

15. Wertheim AH. Perceiving motion: relativity, illusions and the nature of perception. Neth J Psychol. 2008;64: 119–125. doi:10.1007/bf03076414

16. Wertheim AH. Motion perception during self-motion: The direct versus inferential controversy revisited. BBS 17:293-355. Behav Brain Sci. 1999;22: 337–340. doi:10.1017/S0140525X99001831

17. Layton OW, Niehorster DC. A model of how depth facilitates scenerelative object motion perception. PLoS Comput Biol. 2019;15: 1–29. doi:10.1371/journal.pcbi.1007397

18. Dupin L, Wexler M. Motion perception by a moving observer in a threedimensional environment. J Vis. 2013;13: 1–14. doi:10.1167/13.2.15

19. Warren PA, Rushton SK. Perception of object trajectory: Parsing retinal motion into self and object movement components. J Vis. 2007;7: 1–11. doi:10.1167/7.11.2

20. Dokka K, MacNeilage PR, DeAngelis GC, Angelaki DE. Multisensory self-motion compensation during object trajectory judgments. Cereb Cortex. 2015;25: 619–630. doi:10.1093/cercor/bht247

21. Fetsch CR, Deangelis GC, Angelaki DE. Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory. Eur J Neurosci. 2010;31: 1721–1729. doi:10.1111/j.1460-9568.2010.07207.x

22. Garzorz IT, Freeman TCA, Ernst MO, MacNeilage PR. Insufficient compensation for self-motion during perception of object speed: The vestibular Aubert-Fleischl phenomenon. J Vis. 2018;18: 1–9. doi:10.1167/18.13.9

23. Probst T, Loose R, Niedeggen M, Wist ER. Processing of visual motion direction in the fronto-parallel plane in the stationary or moving observer. Behav Brain Res. 1995;70: 133–144. doi:10.1016/0166-4328(95)80003-4

24. Gray R, MacUga K, Regan D. Long range interactions between object-motion and self-motion in the perception of movement in depth. Vision Res. 2004;44: 179–195. doi:10.1016/j.visres.2003.09.001

25. Dyde RT, Harris LR. The influence of retinal and extra-retinal motion cues on perceived object motion during self-motion. J Vis. 2008;8. doi:10.1167/8.14.5

26. Hogendoorn H, Alais D, MacDougall H, Verstraten FAJ. Velocity perception in a moving observer. Vision Res. 2017;138: 12–17. doi:10.1016/j.visres.2017.06.001

27. Niehorster DC, Li L. Accuracy and tuning of flow parsing for visual perception of object motion during self-motion. Iperception. 2017;8: 1–18. doi:10.1177/2041669517708206

28. Butler JS, Campos JL, Bülthoff HH. Optimal visual–vestibular integration under conditions of conflicting intersensory motion profiles. Exp Brain Res. 2014;233: 587–597. doi:10.1007/s00221-014-4136-1

29. Campos JL, Butler JS, Bülthoff HH. Multisensory integration in the estimation of walked distances. Exp Brain Res. 2012;218: 551–565. doi:10.1007/s00221-012-3048-1

30. Jörges B, Harris LR. Object speed perception during lateral visual self-motion. Attention, Perception, Psychophys. 2021;84: 25–46. doi:10.3758/s13414-021-02372-4

31. Jörges B, López-Moliner J. Determining Mean and Standard Deviation of the Strong Gravity Prior through Simulations. 2020; 1–19. doi:10.31234/osf.io/exp93

32. de’Sperati C, Thornton IM. Motion prediction at low contrast. Vision Res. 2019;154: 85–96. doi:10.1016/j.visres.2018.11.004

33. Kammers MPM, van der Ham IJM, Dijkerman HC. Dissociating body representations in healthy individuals: Differential effects of a kinaesthetic illusion on perception and action. Neuropsychologia. 2006;44: 2430–2436. doi:10.1016/j.neuropsychologia.2006.04.009

34. Witt JK, Proffitt DR. Perceived slant: A dissociation between perception and action. Perception. 2007;36: 249–257. doi:10.1068/p5449

35. Bradshaw MF, Watt SJ. A dissociation of perception and action in normal human observers: The effect of temporal-delay. Neuropsychologia. 2002;40: 1766–1778. doi:10.1016/S0028-3932(02)00039-8

36. Zago M, McIntyre J, Senot P, Lacquaniti F. Visuo-motor coordination and internal models for object interception. Exp Brain Res. 2009;192: 571–604. doi:10.1007/s00221-008-1691-3

37. Dichgans J, Wist E, Diener HC, Brandt T. The Aubert-Fleischl phenomenon: A temporal frequency effect on perceived velocity in afferent motion perception. Exp Brain Res. 1975;23: 529–533. doi:10.1007/BF00234920

38. Taylor MM, Creelman CD. PEST: Efficient Estimates on Probability Functions. J Acoust Soc Am. 1967;41: 782–787. doi:10.1121/1.1910407

39. Kaiser MK. Angular velocity discrimination. Percept Psychophys. 1990;47: 149–156. doi:10.3758/BF03205979

40. Norman JF, Todd JT, Perotti VJ, Tittle JS. The Visual Perception of Three-Dimensional Length. J Exp Psychol Hum Percept Perform. 1996;22: 173–186. doi:10.1037/0096-1523.22.1.173

41. Jamiy F El, Marsh R. Distance estimation in virtual reality and augmented reality: A survey. IEEE Int Conf Electro Inf Technol. 2019;2019-May: 063–068. doi:10.1109/EIT.2019.8834182

42. R Core Team. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing,. Vienna, Austria; 2017. Available: http://www.r-project.org/.

43. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. J Stat Softw. 2015;67. doi:10.18637/jss.v067.i01

44. Wilkinson GN, Rogers CE. Symbolic Description of Factorial Models for Analysis of Variance. J R Stat Soc. 1973;22: 392–399.

45. Linares D, López-Moliner J. quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. R J. 2016;8: 122–131. Available: https://journal.r-project.org/archive/2016-1/linares-na.pdf

46. Brent RP. Algorithms for Minimization Without Derivatives. IEEE Trans Automat Contr. 1974;19: 632–633. doi:10.1109/TAC.1974.1100629

47. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models . J Stat Softw. 2017;82. doi:10.18637/jss.v082.i13

48. Jörges B. Data Analysis and Power Simulations with General Linear Mixed Modelling for Psychophysical Data – A Practical , R-Based Guide. PsyArXiv. 2021. doi:10.31234/osf.io/ack8u

## Appendix A

Since we assume, for the sake of visualizing our predictions and performing the power analysis, that self-motion impacts the perceived speed equally for both tasks, we first generate parameters indicating biases and variability differences for each participant: for each participant separately, we draw the effect of self-motion in the opposite direction of the target on accuracy from a normal distribution with a mean of 20% of the presented visual self-motion and a standard deviation of 30%. This reflects the rather noisy effect we found in our previous study [30]. For the impact of self-motion in the opposite direction of the target, we draw the values from a normal distribution with a mean of 20% and a standard deviation of 30%. While we did not find such effect previously, the higher self-motion speeds in this study help raise the power and should allow us to find effects for which our previous study did not have enough power. In line with the previous study, no effects were assumed for self-motion in the same direction as the object.

**Modelling the speed estimation task** – To model participant performance, we closely follow the procedure outlined by Jörges [48]: we first generate means and standard deviations for (cumulative Gaussian) psychometric functions for each condition and participant. We compute the means by taking the presented speed of the stimulus (4, 5, or 6 m/s) and adding the effect of self-motion on perceived speed that we have drawn before for each participant. For the standard deviations, we part from a Weber Fraction of 10% for speed judgements commonly reported in the literature [39], and draw the Weber Fractions for each participant from a normal distribution with a mean of 10% and a standard deviation of 1.5%, which corresponds to between-participant variability in their general ability to discern speeds, and convert these values into standard deviations. We then add the value chosen above for the effect of self-motion as a fraction of the presented speed. As a next step, we choose 70 speed values for each condition which our staircase procedure is likely to present for each participant. We achieve this by drawing values from Cauchy distributions with their location parameters at the presented speed and their scale parameter at 10% of the presented speed. We then compute the probability of each participant for each condition and each simulated presented stimulus speed to judge that the ball cloud is faster than the single ball, and use these probabilities to draw binary yes/no answers from the Bernoulli function. Finally, we apply the process outlined under Data analysis plan to obtain simulated PSEs and JNDs for each participant and condition.

**Modelling the prediction task** – For the prediction task, our model for time-to-contact is based on the physical equation for distance from time and speed; time-to-contact can then be estimated from the ball’s speed and the distance between its point of disappearance and the target rectangle.

|  |  |
| --- | --- |
|  | (3) |

We use the same simulated participants as for the speed estimation task (i.e., with the same PSEs and JNDs, and with the same parameters for the effect of self-motion on PSEs and JNDs). We further assume here that the distance between observer and stimulus is either estimated correctly, or it is misestimated consistently, such that the effect on perceived speed and perceived distance cancel out, and that this distance is estimated with a Weber Fraction of 5% [40]. We then draw a (simulated) perceived distance for each trial and participant from a normal distribution with its mean at the correct distance, and a standard deviation that corresponds to this Weber Fraction of 5%. For the perceived speed, we add each participants self-motion effect on their personal PSE and JND, respectively, and then draw a (simulated) perceived speed from a normal distribution that uses the PSE for each participant and trial as mean and the JND for each participant and trial as standard deviation. Finally, we obtain the simulated extrapolated duration by dividing the simulated distance by the simulated speed for each participant and trial and subtract the occlusion duration to obtain the timing error.