OBJECT SPEED PERCEPTION DURING LATERAL SELF-MOTION

Björn Jörges[1]*, Laurence R. Harris*[1]*

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

* Corresponding Author

Abstract

Judging the velocity of objects during observer self-motion requires disambiguating retinal stimulation caused by the observer's own movement and retinal stimulation caused by object movement. According to the Flow Parsing hypothesis, observers first estimate their own motion based on visual and other cues. They then subtract the retinal motion corresponding to their own movement from the total retinal stimulation and interpret the remaining stimulation as pertaining to object motion. While the phenomenon has been studied to some extent for motion-in-depth and rotational motion, lateral motion has been largely neglected. The Flow Parsing hypothesis yields predictions both for the precision and the accuracy of target speed estimation during self-motion: firstly, subtracting noisier self-motion information from retinal input should lead to a decrease in precision when the observer is moving during motion observation. Furthermore, when self-motion is only simulated visually, while other cues such as vestibular sensory information and data from efference copies are unavailable, self-motion is likely to be underestimated, which should yield a overestimation of target speed when target and observer move in opposite directions and an underestimation of target speed when target and observer move in the same direction.

Significance

When interacting with our inherently dynamic environment, humans are rarely standing still. It is thus crucial for us to obtain accurate and precise estimates of the behavior of objects in our environment even while we are moving ourselves. In this project, we investigate how humans judge the speed of other objects while moving themselves, which helps both understand everyday situations such as avoiding collisions while driving a car and constrain models of how humans time their interactions with a dynamic world.

Introduction

When observing a moving target while an observer is moving, the same retinal speeds can correspond to vastly different physical velocities. When an observer moves in the same direction as the target, the retinal speed of the object is partially cancelled out, and vice-versa. Observers must thus obtain an accurate estimate of their own velocity, and subtract or add the consequences of this movement it to the retinal motion of the target to obtain an accurate estimate of the object velocity. More specifically, the Flow Parsing Hypothesis (Dupin & Wexler, 2013; Warren & Rushton, 2008, 2009) posits that, to estimate object motion from ambiguous retinal input, observers first compute which components of retinal stimulation are caused by their own motion in the environment. Then, they subtract this self-motion information from the overall stimulation and attribute the remaining stimulation to object motion in the scene. When self-motion is experienced only visually while undergoing no physical motion, which entails a conflict between visual and vestibular input, self-motion is likely underestimated, leading to biases in judgments of object motion. This has been shown to some extent for vertical observer and object motion (Dyde & Harris, 2008), as well as for rotating observers (Probst, Loose, Niedeggen, & Wist, 1995) and motion in depth (Gray, MacUga, & Regan, 2004). Furthermore, it has been argued that self-motion information is noisier than retinal information concerning object motion, especially when observers have only visual information about their own movement at their disposal (Fetsch, Deangelis, & Angelaki, 2010). Subtracting noisy self-motion information from retinal motion in order to obtain an estimate of target velocity should thus decrease precision (Dokka, MacNeilage, DeAngelis, & Angelaki, 2015). This process is relatively straightforward for the consequences of angular self-motion, but for lateral motion the geometry for such a subtraction process requires additional computations involving, for example, an estimate of the distance of the object to the observer. More specifically, observers need to first estimate their own motion in an allocentric world frame by using retinal stimulation attributable to static object in the environment and other sensory and efferent information. Then, this estimate needs to be used to generate an estimate of the retinal stimulation caused by the observer's motion. The estimated retinal stimulation due to self-motion is subtracted from the total retinal stimulation, which allows to interpret the remaining retinal stimulation as object motion.

Remarkably, the literature is quite sparse with regards to lateral motion: (Dokka et al., 2015) investigated to what extent observer motion (visual cues only, vestibular cues only and both visual and vestibular cues) influenced the judged direction of vertical downwards motion with small lateral components. They found biases in line with insufficient flow parsing in all observer motion conditions, as well as increases in sensitivity. However, to our knowledge, no studies have investigated how lateral self-motion influences perceived lateral object speed. This is particularly relevant as the visual system has been shown to use velocity information to extrapolate object trajectories to compensate for noisy online information and neural delays (Aguado & López-Moliner, 2019; Aguilar-Lleyda, Tubau, & López-Moliner, 2018; Jörges & López-Moliner, 2019; López-Moliner, Brenner, Louw, & Smeets, 2010). The aim of this project is thus to verify the impact of self-motion on accuracy and precision for object velocity judgments in a lateral motion paradigm, which will further our understanding of Flow Parsing and help us understand the conditions under which Flow Parsing is incomplete. More specifically, our hypotheses are:

- When the observer is **static** during object motion observation, we expect the **highest accuracy** of velocity estimation.
- When the observer is **moving opposite to the object motion (e. g. observer moves to the right, object moves to the left)** during object motion observation, we expect them to **overestimate the observed velocity**.
- When the observer is **moving in the same direction as the target (e. g. both observer and target move to the right)** during object motion observation, we expect them to **underestimate the observed velocity**.
- Furthermore, we expect the **precision to be lower** when the subject experiences **self-motion during object motion observation** relative to when they are static.

Participants

We tested 16 participants from the population of PhD and undergrad students at York University with equal numbers of males and females. Due to the culturally independent nature of the phenomenon under study, we do believe our results are likely to be relevantly skewed by WEIRD people effects (Henrich, Heine, & Norenzayan, 2010). Participants will have normal or corrected-to-normal vision and have to achieve a stereoacuity of 63 arc seconds or below on the Fly Stereo Acuity Test. The project has received ethics approval from the Human Participant Ethics Review Sub-Committee at York University. Informed consent was obtained from all subjects and the experiment was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Apparatus

All the experiments were performed in virtual reality with participants remaining physically static and seated. We programmed the stimuli in Unity (2019.2.11f1), while object motion, self-motion and the psychophysical staircases were controlled in C# via its integration with Unity. The Unity project is available on Open Science Foundation (https://osf.io/m6ukw/). Stimuli were presented in an Oculus Rift. Participants responded by means of a finger mouse.

Setup

Our experiment consisted of a Two Interval Forced-Choice Task where participants were asked to indicate which of two intervals contained objects moving at the higher velocity. In one interval participants were presented a ball with a diameter of 0.33 m at a simulated distance of 8 m in front of them, travelling to the right or to the left with 6.6 or 8.0 m/s (four target motion profiles). During this interval, participants were either static or experienced simulated body movement to the left or to the right with a Gaussian velocity profile (three self-motion profiles), accelerating until reaching peak velocity after 0.25 s and then slowing down until coming to a halt at 0.5 s. The position in time x(t) was given by a cumulative Gaussian distribution with a mean of 0.25 s and a standard deviation of 0.08 s divided by 2, multiplied by -1 for trials with self-motion to the left. That is, participants moved 0.5 m over the course of 0.5 s, which amounts to a mean velocity of 1 m/s. In the other interval, participants were shown a cloud of smaller moving balls each with a diameter of 0.1 m as comparison. The balls appeared 1.25 m to the left of the observer (if the big target in the same trial moved to right) or to the right of the observer (if It moved to the left), then moved in the same direction as the big target and disappeared after having travelled 2.5 m. They were spread out vertically over a distance of 1 m, and 10 to 15 were visible at any given moment. The speed of these smaller balls was controlled by a PEST staircase. We employed two staircases for each combination of self-motion (left, right or static) and object motion (-8, -6.6, 6.6 and 8 m/s), one of which started 33 % above the target's speed, and the other one 33 % below target speed (two staircases for each combination of target motion and self-motion). Thus, there was a total of 4 target speeds x 3 motion conditions x 2 = 24 interleaved staircases. When participants answered that the ball cloud was faster, a lower velocity was displayed in the next trial of that PEST and vice-versa. The step sizes were governed by the following rules (Taylor &

Creelman, 1967): the initial step size was 1.2 m/s. For the first five trials for each PEST, the step size was maintained. Starting from the eleventh trial, after a reversal (subjects answered “PEST is slower” in the second-to-last trial and “PEST is faster” in the last trial or vice-versa), the step size was halved. After the second same answer, the step size was maintained. After the third same answer, the step size was either maintained, when the step size had been doubled before the last reversal, or doubled when the step size had not been doubled before the last reversal. After four same answers, the step size was always doubled. Each PEST ended when it converged (five consecutive trials with step sizes lower than 0.1) AND participants had judged at least 20 trials of the staircase. If the staircase did not converge, the PEST was terminated after 35 trials. The experiment ended when all 24 PESTs had terminated. This took about an hour. An short sequence of stimuli can be viewed under https://github.com/b-jorges/Motion-Perception-during-Self-Motion/blob/master/Figures/GIF%20of%20Stimulus.gif .

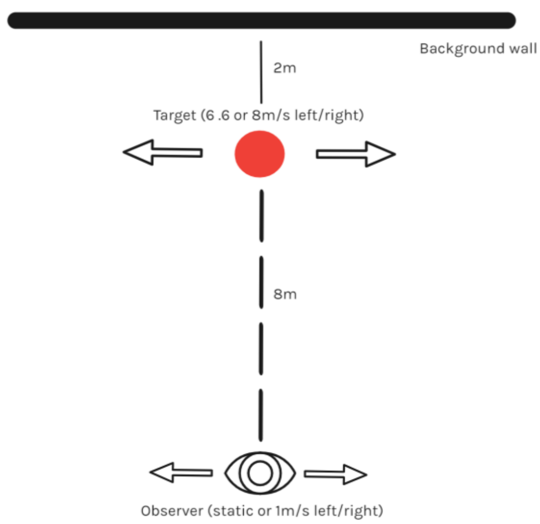

Figure 1: Top view of the stimulus scene in one of the test trials. The red circle represents the target, which moves laterally at 6.6 or 8 m/s for 0.5 s, that is, 3.3 or 4 m. The stylized eye indicates the position of the observer, who can be static or move to the left or to the right for 0.5 s with a Gaussian motion profile and a mean velocity of 1 m/s. The target is 8 m away from the observer, while the target is 2 m in front of the background wall.

Before starting the actual data collection, participants perform a training session with one PEST where the big target moves at 4 m/s. Subjects are asked to repeat the training if the step size in any of the last five trials is above 0.3 m/s. If they still fail to meet the criterion after a second repetition they are excluded from the experiment.

Analysis

To assess the **Just Noticeable Difference (JND)** as a measure of precision, we employ General Linear Mixed Modelling, implemented in the R package lme4, according to the recommendations in (Moscatelli, Mezzetti, & Lacquaniti, 2012). We first establish a Test Model, in which responses are fitted to a cumulative Gaussian, with subject ID (“Subject”) and horizontal velocity ($\text{Velocity}_{\text{horizontal}}$, with values -8, -6.6, 6.6 and 8 m/s) as random effects with random intercepts, and self-motion (binary variable “$Subject\ Motion$” with the values “Yes” and “No”) and difference in velocity between target and ball cloud (“$Difference$”) and their interaction as fixed effects. In lme4 syntax, this corresponds to:

$$Subject\ Motion * Difference + (1\ |Subject) + (1\ |\ Velocity_{Target}) \quad [1]$$

We furthermore establish a Null Model with subject and horizontal velocity as random effects with random intercepts, and subject motion profile and difference in velocity between target and ball cloud as fixed effects, but not their interaction:

$$Subject\ Motion + Difference + (1\ |Subject) + (1\ |\ Velocity_{Target}) \quad [2]$$

We then use an ANOVA to test whether the test model is significantly better than the null model. If the interaction term improves the model significantly, the subject motion profile has a relevant influence on the slope of the fitted cumulative Gaussian. We expect the interaction parameter to be lower for Motion = “Congruent” and Motion = “Incongruent”, thus putting into evidence that self-motion decreases precision in object velocity judgments during self-motion.

To assess the **Point of Subjective Equivalence (PSE)**, our Test Model contains the same random effects as above and the self-motion profile (ternary variable “$Motion\ Profile$” with the values “Congruent”, “No Motion” and “Incongruent”) and the velocity difference between target and ball cloud (“$Difference$”) as fixed effects (Moscatelli et al., 2012). The lme4 syntax is:

$$Motion\ Profile + Difference + (1\ |\ Subject) + (1\ |\ Velocity_{Target}) \quad [3]$$

The Null Model contains the same random effects, and only the difference in speed between target and ball cloud as fixed effect.

$$Difference + (1\ |\ Subject) + (1\ |\ Velocity_{Target}) \quad [4]$$

We compare both models with an ANOVA and expect the Test Model to be significantly better than the Null Model, indicating that self-motion has an impact on the PSE. Self-motion in the same direction as the target should decrease perceived target velocity, and self-motion in the opposite direction of the target should increase perceived target velocity.

Power Analysis

Based on the analysis plan above, we proceeded to a power analysis via simulation. The R code used for this power analysis is available online under https://github.com/b-jorges/Motion-Perception-during-Self-Motion/blob/master/PowerAnalysisMotionEstimation.R. We first created datasets that would roughly resemble the data we are expecting to collect. At the core of the simulation of these datasets is the assumptions that responses could be described by a cumulative Gaussian function (which approximates what is commonly known as “Psychometric Function”). The mean of the cumulative Gaussian corresponds to the PSE, and its standard deviation is proportional to the JND. We varied the means of the Gaussian according to the self-motion profile. Pilot data (see below) show consistently a bias to interpret the dot cloud as faster; when the observer is static, we thus assume a PSE of 2/3 of the presented velocity. When the observer moved opposite to the target, we expected the PSE to be higher than in the static condition, and when the observer moved with the target, we expect the PSE to be lower. We conducted the power analysis assuming a difference of 1/8 of the mean presented self-motion velocity; (Dokka et al., 2015) found biases up to 50 % of self-motion. Their task, directionality judgments about downward motion with a lateral left- or rightward component, bears some similarities to ours, but is different enough to warrant a more conservative estimate for the sake of the power analysis. For the standard deviation, we parted from a Weber fraction of 7 % for the static condition (McKee, 1981), which corresponds roughly to a standard deviation of 10 % of the PSE. Where the observer is moving, we expected increased JNDs and therefore an increased standard deviation. For the sake of this power analysis, we assume that the standard deviation in this case might be 1/4 higher than the standard deviation for a static observer. (Dokka et al., 2015) found increases of up to 200 % in thresholds from no self-motion to visually simulated self-motion. We choose a much more conservative value to account for task differences. Additionally, we varied the PSE and SD per subject by multiplying them with random values drawn from a normal distribution with a mean of 1 and a standard deviation of 0.1. To account for the fact that our staircase leads to a concentration of responses around the PSE, we drew the stimulus strengths from a Cauchy distribution with a location of 1 and a scale of 0.04. We drew 55 stimulus strengths for this distribution (per combination of target velocity and self-motion, we use two PESTs with about 27 trials each; see above) and fed them into the cumulative Gaussian we established per condition and subject. This yielded the answer probability per trial. We then used these probabilities to draw binary answers (PEST faster yes/no) from a Bernoulli distribution for each trial.

We simulated 500 of these data sets, conducted the analyses described above over each for 10, 12, 14, 16, 18 and 20 subjects. We report the percentage where the Test Model was significantly better than the Null Model in Table 1.

n	Power Accuracy	Power Precision
10	1	0.83
12	1	0.882
14	1	0.918
16	1	0.966
18	1	0.948
20	1	0.980

Table 1: Simulated power values for 10, 12, 14, 16, 18 and 20 participants.

While the effect should be easily detectable for the accuracy-based hypothesis, the precision hypothesis is somewhat harder to detect and requires at least 14 subjects (for a power above 0.9). As it is not very costly for us to add more subjects, **we aim for a power of 0.95, which should be achieved with roughly 16 subjects.** Note that, as the simulation process involves several sources of uncertainty, some variability is to be expected in the results, which explains why the simulated power for 18 subjects is lower than the power simulated for 16 subjects.

Pre-existing Data

We collected data from seven pilot participants. One (s07) was excluded because some of her PESTs did not converge. Two participants (s01 and s02) had previously done the task in 2D, but only their 3D data was included in the analysis. Pilot results are largely in line with our predictions: In terms of JNDs, we found that our Test Model was significantly better than the Null Model ($p = 0.02$), and effects trended in the direction of our hypothesis (regression coefficients of -0.078, SE = 0.034, for the interaction between self-motion present and the difference in velocity, which corresponds to a lower precision). For the PSEs, we found that our Test Model was significantly better than the Null Model ($p < 0.001$), and the effects go largely in the expected direction (regression coefficients of 0.072, SE = 0.05, for the main effect of congruent motion, and -0.25, SE = 0.053, for the main effect of incongruent motion; which corresponds to a lower perceived speed for congruent motion and self-motion, and a higher perceived speed for incongruent motion and self-motion). The code used for this analysis as well as the pilot data are available under https://github.com/b-jorges/Motion-Perception-during-Self-Motion/blob/master/AnalysisPilotData.R.

The pilot data were not included into the final analysis; we recruited 16 new subjects.

Open Practices

We will publish all raw data collected during this project in the GitHub repository https://github.com/b-jorges/Motion-Perception-during-Self-Motion/, as well as all the code used for analysis. Furthermore, the Unity project used to present the stimulus and collect data is available on OSF under https://osf.io/m6ukw/.

References

Aguado, B., & López-Moliner, J. (2019). Perceived speed of motion in depth modulates misjudgements of approaching trajectories consistently with a slow prior. *Vision Research*, *159*, 1–9. https://doi.org/10.1016/j.visres.2019.03.009

Aguilar-Lleyda, D., Tubau, E., & López-Moliner, J. (2018). An object-tracking model that combines position and speed explains spatial and temporal responses in a timing task. *Journal of Vision*, *18*(12), 12. https://doi.org/10.1167/18.12.12

Dokka, K., MacNeilage, P. R., DeAngelis, G. C., & Angelaki, D. E. (2015). Multisensory self-motion compensation during object trajectory judgments. *Cerebral Cortex*, *25*(3), 619–630. https://doi.org/10.1093/cercor/bht247

Dupin, L., & Wexler, M. (2013). Motion perception by a moving observer in a threedimensional environment. *Journal of Vision*, *13*(2), 1–14. https://doi.org/10.1167/13.2.15

Dyde, R. T., & Harris, L. R. (2008). The influence of retinal and extra-retinal motion cues on perceived object motion during self-motion. *Journal of Vision*, *8*(14), 1–10. https://doi.org/10.1167/8.14.5

Fetsch, C. R., Deangelis, G. C., & Angelaki, D. E. (2010). Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory. *European Journal of Neuroscience*, *31*(10), 1721–1729. https://doi.org/10.1111/j.1460-9568.2010.07207.x

Gray, R., MacUga, K., & Regan, D. (2004). Long range interactions between object-motion and self-motion in the perception of movement in depth. *Vision Research*, *44*(2), 179–195. https://doi.org/10.1016/j.visres.2003.09.001

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2–3), 61–83; discussion 83-135. https://doi.org/10.1017/S0140525X0999152X

Jörges, B., & López-Moliner, J. (2019). Earth-Gravity Congruent Motion Facilitates Ocular Control for Pursuit of Parabolic Trajectories. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-50512-6

López-Moliner, J., Brenner, E., Louw, S., & Smeets, J. B. J. (2010). Catching a gently thrown ball. *Experimental Brain Research*, *206*(4), 409–417. https://doi.org/10.1007/s00221-010-2421-1

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision Research*, *21*(4), 491–500. https://doi.org/10.1016/0042-6989(81)90095-X

Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, *12*(11), 1–17. https://doi.org/10.1167/12.11.26

Probst, T., Loose, R., Niedeggen, M., & Wist, E. R. (1995). Processing of visual motion direction in the fronto-parallel plane in the stationary or moving observer. *Behavioural Brain Research*, *70*(2), 133–144. https://doi.org/10.1016/0166-4328(95)80003-4

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient Estimates on Probability Functions. *The Journal of the Acoustical Society of America*, *41*(4A), 782–787. https://doi.org/10.1121/1.1910407

Warren, P. A., & Rushton, S. K. (2008). Evidence for flow-parsing in radial flow displays. *Vision Research*, *48*(5), 655–663. https://doi.org/10.1016/j.visres.2007.10.023

Warren, P. A., & Rushton, S. K. (2009). Perception of scene-relative object movement: Optic flow parsing and the contribution of monocular depth cues. *Vision Research*, *49*(11), 1406–1419. https://doi.org/10.1016/j.visres.2009.01.016