An object-tracking model that combines position and speed explains spatial and temporal responses in a timing task

David Aguilar-Lleyda
Vision and Control of Action (VISCA) Group, Department of Cognition, Development and Psychology of Education, Institut de Neurociències, Universitat de Barcelona, Barcelona, Catalonia, Spain
Present address: Centre d'Économie de la Sorbonne (CNRS & Université Paris), Paris, France

Elisabet Tubau
VISCA Group, Department of Cognition, Development and Psychology of Education, Institut de Neurociències, Universitat de Barcelona, Barcelona, Catalonia, Spain

Joan López-Moliner
VISCA Group, Department of Cognition, Development and Psychology of Education, Institut de Neurociències, Universitat de Barcelona, Barcelona, Catalonia, Spain

Many tasks require synchronizing our actions with particular moments along the path of moving targets. However, it is controversial whether we base these actions on spatial or temporal information, and whether using either can enhance our performance. We addressed these questions with a coincidence timing task. A target varying in speed and motion duration approached a goal. Participants stopped the target and were rewarded according to its proximity to the goal. Results showed larger reward for responses temporally (rather than spatially) equidistant to the goal across speeds, and this pattern was promoted by longer motion durations. We used a Kalman filter to simulate time and space-based responses, where modeled speed uncertainty depended on motion duration and positional uncertainty on target speed. The comparison between simulated and observed responses revealed that a single position-tracking mechanism could account for both spatial and temporal patterns, providing a unified computational explanation.

Introduction

In many daily life situations we need to interact with moving objects. Suppose we want to hit the skeet in a skeet-shooting game. A successful shot often involves syncing our actions (pulling the trigger) with particular positions or times of the moving object (skeet), so that we anticipate the future position where we want the target to be hit. But which information do we use to decide when to pull the trigger? And which information will maximize our performance? Traditionally, it has been hypothesized that we use optical variables (e.g., Lee, 1976, Bootsma & Oudejans, 1993) to time our actions, allowing our responses to be time-invariant across conditions (e.g., same timing across target speeds). However, the difficulty to reconceal various response patterns (i.e., including position-based responses) from a single framework has led to other proposals. These ones resort to using kinematic variables estimated from the target's movement (e.g., distance/position, speed, see Kwon & Knill, 2013; López-Moliner, Field, & Wann, 2007; Tresilian, 1999; Wann, 1996). Here, we propose a tracking model that can account for these different response patterns and that predicts performance across different conditions.

In an example like skeet shooting, we need to localize the (changing) position of the target, which does not only depend on retinotopic maps but also on other effects (Schlag & Schlag-Rey, 2002), like motion signals (De Valois & De Valois, 1991; Linares, López-Moliner, & Johnston, 2007; Maus, Fischer, & Whitney, 2013; Whitney, 2002). Position and motion interactions can

Citation: 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, 1–19, https://doi.org/10.1167/18.12.12.

https://doi.org/10.1167/18.12.12 Received May 28, 2018; published November 20, 2018 ISSN 1534-7362

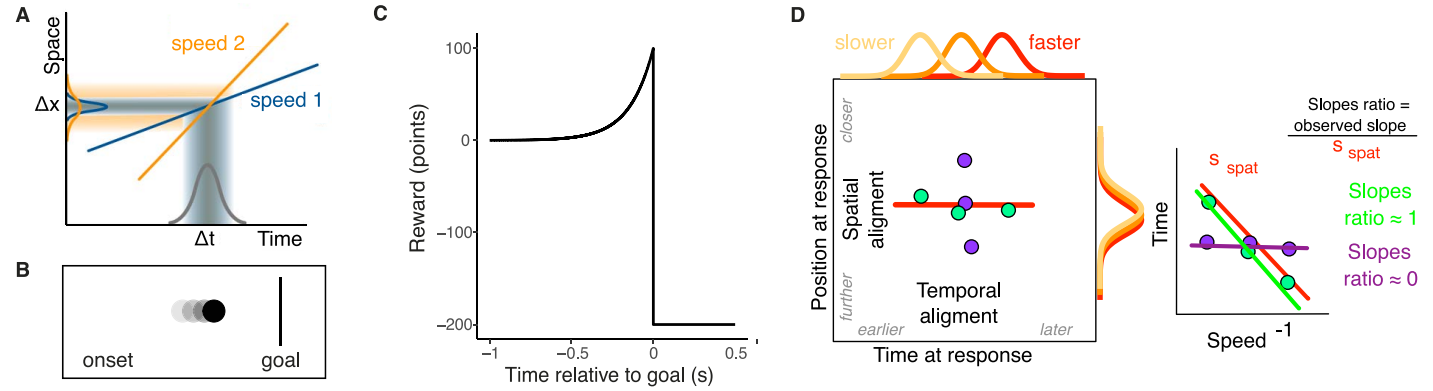

Figure 1. (A) Spatial variability as a function of the temporal resolution for two different speeds (denoted by the oriented line in the space-time plot). Speed 2 is faster than speed 1 (shallower slope). When sampling the position at a given temporal resolution (illustrated by the Gaussian width on the x axis), the spatial variability of the sampled position will depend on the speed of the moving target. As a consequence, variability increases for faster speeds. (B) Representation of the stimuli seen in every trial. A circle appears at a starting position and moves rightward toward a vertical line (goal). The circle stops when the participant presses a button. (C) Example of reward function used in the experiment, corresponding to the temporal reward condition, where reward is given as a function of the time remaining to the goal. (D) The schematics on the left presents two sets of hypothetical average response across speeds (each dot corresponding to an average for one speed). The purple responses are closely aligned in time, but not space, what would be an example of overt time-invariant responses. The green responses are aligned in space but not in time, and thus close to overt space invariance. The right part of the figure shows how the slopes ratio is calculated. First, for each speed the average time of the observed responses is plotted against the inverse of that speed. Then a slope is obtained from those three values (observed slope). Two different observed slopes can be seen as the purple and green lines. Then a slope corresponding to pure space-invariance (s_{spat}) is calculated, and the observed slope is divided by it, with the result being the slopes ratio. Slopes ratio closer to 1 denote overt space-invariance, while if they are close to 0 that denotes overt time-invariance.

be complex. Recently, Kwon, Tadin, and Knill (2015) put forward a model that tries to unify visual motion and position perception. The model implements an optimal object tracking system in which speed is integrated to update the position. The computational principle is based on a Kalman filter (Kalman, 1960) that couples position and motion information estimates. Measured sensory signals (of position or speed) and internal estimates based on past signals are optimally combined according to their respective uncertainties. If sensory signals are very reliable, they would be weighted more than the internal model, and vice versa. This unified tracking model successfully predicts many phenomena, such as perceptual biases like the motion-induced position shift (MIPS, De Valois & De Valois, 1991). Such a model is not only relevant to explain perceptual effects, but can also have consequences for sensoriomotor actions as well. Sensorimotor situations involving moving objects require that an action is made relative to the course of the object. Actions are planned to obtain a certain outcome, as defined by the situation's reward function. Due to the spatiotemporal nature of that course, though, the action may be planned by relying more on either temporal or spatial information. For example, when positional uncertainty is high, on a foggy day, a skeet shooter may rely more on motion information. A position-motion coupled model (Kwon et al., 2015) would make different predictions for conditions in which measured positional noise or speed noise are affected differently. For example, it would predict that measured position would be more variable for faster speeds (see Figure 1A), due to the limited visual temporal resolution (Linares, Holcombe, & White, 2009). In that case, the model would favor the use of motion information. On the other hand, motion duration could affect the reliability of speed measurements (Burr & Santoro, 2001; Neri, Morrone, & Burr, 1998), with shorter presentation times leading to noisier speed estimates. In such a case, the model would favor position measurements.

Up to date, it has not been examined how the use of positional or temporal cues determines performance when people interact with moving targets. This is what we aimed at in the present study. Specifically, we wanted to investigate whether responses (i.e., motor actions) based on one type of information enhance performance, compared to responses based on the other. The task we used in our study was similar to those commonly referred to as coincidence action (CA) timing task (López-Moliner et al., 2007; Tresilian, 1995). As seen in Figure 1B, in every trial a target underwent constant rightward motion toward a vertical line (goal), and participants had to press a button to stop the target. This task requires a timed response and therefore shares some similarities with an interceptive

action (IA). However there are important differences too that might affect the generalization of our results. We will address some of these limitations in the discussion. However, we allowed that participants did not have to merely synchronize their response with a visual event (e.g., alignment with the goal), but to decide how to respond based on a reward function. A numerical reward in points was given according to how close the target was to the goal when the button was pressed. In different conditions though, the reward was given as a function of the temporal proximity to alignment, or as a function of spatial proximity to the goal. Thus, we defined performance as the reward that participants accumulated in a series of trials where they received positive or negative feedback as a function of their responses. Reward increased exponentially as a function of the time or the spatial distance left for the target to align with the goal, with the maximum reward given if the participant pressed the button at the moment of alignment (see Figure 1C). However, any response after alignment was penalized. Participants were instructed to accumulate as many points as possible across trials. From trial to trial, the target could change its speed. We identified time-based or overt time-invariant response distributions as those where the target at the moment of response was at a similar remaining time to alignment across all speeds. Alternatively, position-based or overt space-invariant response distributions are those where responses across speeds are at a similar spatial distance from the goal. Then, unlike previous studies in which responses could only be planned by using timing information (Mamassian, 2008; Ota, Shinya, & Kudo, 2015) or in which performance depended on correctly hitting a series of spatially-defined reward or penalty areas (Gepshtein, Seydell, & Trommershäuser, 2007; Hudson, Wolfe, & Maloney, 2012; Neyedli & Welsh, 2013; O'Brien & Ahmed, 2013; Trommershäuser, Gepshtein, Maloney, Landy, & Banks, 2005; Trommershäuser, Landy, & Maloney, 2006; Trommershäuser, Maloney, & Landy, 2003), here we could distinguish whether participants were aiming at responding when the target was at a given temporal or spatial distance from the goal and whether any of these modes led to a higher overall reward than the other. Introducing trial-to-trial changes in initial time to alignment enabled us to determine whether the information used changed across the different motion durations. We also wanted to know whether impairing the processing of motion information would affect performance, as predicted by a position-motion coupled model. Thus, we included a condition in which the target moved under Second Order motion. Finally, to compare the behavioral performance in our different conditions with the model predictions, we implemented the Kalman filter model and conducted simulations with similar stimulus manipulations as in the experiments.

To anticipate, our results confirmed the predictions of the model: Longer motion durations promoted more overt time-invariant responses across speeds and enhanced performance, probably by refining speed estimates, while spatially-aligned responses were dominant for shorter motion durations and Second Order motion. In the latter cases, penalizations increased for faster speeds. The comparison with the model predictions shed further light into the response mechanism: a single one based on position tracking could account for the observed pattern across the different conditions, including both time and spatial invariant responses.

Model

The tracking model closely resembles the model presented in Kwon et al. (2015). However, it is simpler in the sense that we do not account for any pattern motion within the object (for instance, the changing texture within the retinal image of a ball as a product of rotational motion). The state vector $\mathbf{x_t}$ contains the position and speed of the object $[x_t, v_t]'$ at time t. The model assumes that the motion in the world is governed by the following process (or model) equation:

$$\mathbf{x_t} = A\mathbf{x_{t-1}} + w,\ w \sim \mathcal{N}(0, Q_t) \quad (1)$$

Where $A = [1\ \Delta t; 0\ 1]$. This equation denotes the internal model of the position and speed of the moving object and states that the next value of position is predicted as the previous positional value plus the time step multiplied by the previous value of speed. Q is the process covariance matrix and is defined as $Q = q \times q^{\mathrm{T}}$ with $q = [q^x q^v] = [0.9\ 1.5]$, where q^x and q^v are positional and speed process variance (Gaussian noise with zero mean). Model performance was unaffected by Δt across values smaller than 20 ms. We set $\Delta t = 2$ ms, which is the temporal resolution at which the object position was updated in the simulations. The position is updated based on first order information (target speed) without including an acceleration term in Equation 1. The reason for this is that our target stimuli moved at constant speed. The observer samples (or measures) the world's object position and speed, which are specified by the observation equations of a Kalman filter:

$$y_t^x = x_t + z_x,\ z_x \sim \mathcal{N}\left(0, r_t^x\right) \quad (2)$$

$$y_t^v = v_t + z_v,\ z_v \sim \mathcal{N}\left(0, r_t^v\right) \quad (3)$$

Where y_t^x is the measurement of world position at time t and z_x is the measurement Gaussian noise of the position which has zero mean and variance r^x. y_t^v is the corresponding speed measurement at time t and r^v is

the measurement variance of the object speed. In vector notation, like Q, the measurement noise covariance matrix R is defined as $R = r \times r^{\mathrm{T}}$, with $r = [r^x r^v]$. Importantly, the observer continuously updates the estimates of position and speed of the object in the world through the noise measurement of these two variables. We assume, like in Kwon et al. (2015), and as a consequence of using a Kalman filter, that the measurement at time t ($y_t = \left[y_t^x y_t^v\right]^{\mathrm{T}}$) is combined with the current internal estimate (prior) ($\widehat{\mathbf{x}}_\mathbf{t}^-$) to update the model to produce a posterior estimate ($\widehat{\mathbf{x}}_\mathbf{t}$) at time t.

$$\widehat{\mathbf{x}}_\mathbf{t} = \widehat{\mathbf{x}}_\mathbf{t}^- + \mathbf{K}\left(\mathbf{y_t} - \widehat{\mathbf{x}}_\mathbf{t}^-\right) \quad (4)$$

The difference $\mathbf{y_t} - \widehat{\mathbf{x}}_\mathbf{t}^-$ is called residual or innovation and reflects the discrepancy between the prediction from the prior ($\widehat{\mathbf{x}}_\mathbf{t}^-$) and the actual measurement. Usually a measurement matrix (e.g., H) is used that multiplies the prior. Since we are using the identity matrix we have ommited this term in Equation 4.

How much weight (K) is given to the current observation depends on its reliability. The higher the reliability of the observation, the larger the K (Kalman gain). As in the standard Kalman filter, the Kalman gain matrix was chosen so that it minimizes the posterior error covariance: $P = E[e \times e^{\mathrm{T}}]$, where $e = \mathbf{x} - \widehat{\mathbf{x}}$. A specific form that minimized the posterior error covariance is:

$$K = P^-(P^- + R)^{-1} \quad (5)$$

where P^- is the a priori estimate error covariance in which the error is $e = \mathbf{x} - \widehat{\mathbf{x}}^-$. The a priori uncertainty at time t is estimated from the a posteriori uncertainty P at time $t - 1$ by adding the process noise: $P^- = P + Q$.

The resulting posteriors (see Equation 4) of position and speed would correspond to the perceived position and speed, respectively.

Different experimental conditions should affect the measurement noise of position and speed differently. We can simulate these different conditions and see how well they can explain the observed human pattern of responses.

To conduct the simulations we make two assumptions based on previous literature. These concern measurement noise of position and speed, as well as the initial estimates of these two state variables. The first assumption concerns positional noise: We assume that uncertainty about the object position will scale with object speed. This has been reported in Linares et al. (2009) and in Brenner, van Beers, Rotman, and Smeets (2006). We modeled measured positional uncertainty the following way:

$$r^x = (dt \times v)^2 \quad (6)$$

Where v is the physical speed of the object and dt is the temporal resolution of the sampling, which was set to 10, 18, and 22 ms in different sets of simulations (Brenner et al., 2006). The positional variability cannot be disentangled from a timing variability (Brenner & Smeets, 2009) in our experiment. Additional timing cues, however, have been shown to increase temporal precision when available (Chang & Jazayeri, 2018). The second assumption refers to the speed measurement uncertainty. Based on the motion integration literature (Burr & Santoro, 2001; Neri et al., 1998) we assume that uncertainty decreases with longer presentation times:

$$r^v = (a/T)^2 \quad (7)$$

where T is the motion duration in a given condition ($T = 0.8, 1, 1.2$). In different sets of simulations we set a to 1.5, 2, and 3.

Methods

Participants

Twenty-one participants (11 women, all right-handed, age range 18–32) took part in the condition in which reward was defined in the temporal domain (temporal reward condition). Twenty-one participants of a completely new sample (nine women, all right-handed, age range 18–40) took part in the spatial reward condition. Twelve completely new participants (six women, all right-handed, age range 18–31) participated in the Second Order motion condition. All had normal or corrected-to-normal vision and were naïve about the aim of the experiment. In all cases, participants gave their informed consent. The study complied with the local ethics guidelines, in accordance with the Declaration of Helsinki, and was approved by the University of Barcelona's Bioethics Commission.

Apparatus and stimuli

Participants sat in a dimly lit room, approximately 50 cm in front of a Samsung SyncMaster 1100MB CRT monitor (Samsung, Cambridge, UK; 21-in., 1,024 × 768 resolution, 120 Hz refresh rate). They responded by pressing the button of an ancillary input device sampled at 120 Hz refresh rate with their dominant hand. The experiment was run on a Mac Pro 4.1 Quad-Core Intel Xeon at 2.66 GHz (Apple, Inc., Cupertino, CA).

Temporal and spatial reward conditions

The simplest way to give reward is as a function of only one variable: for instance, temporal proximity to alignment. However, if we found higher performance

for time-based responses, it could be argued that this was determined by the fact that the reward function was fostering responding as a function of time. Thus, we included both a condition where reward was given as a function of the temporal proximity to alignment, and one where it was given as a function of spatial proximity.

A trial started with both a white vertical line (goal) and a white circular target appearing on a black background (shown white in Figure 1B). The goal line (10 cm tall, 1 px width) was positioned 15 cm to the right of the center of the screen. The target had a radius of 0.3 cm (0.34°) and travelled from left to right. In each trial, the target constantly moved at one of three possible speeds (19.5 cm/s, 25 cm/s and 32 cm/s) (22°/s, 28°/s, and 35.5°/s). The chosen speeds followed a geometric progression to compensate for a constant Weber fraction, so that the discriminability between the slowest and mid speeds was presumably similar to the discriminability between the mid and fastest speeds. At every trial, the target also had one of three possible motion durations (0.8 s, 1 s, 1.2 s), where motion duration is defined as the time from motion onset until the target and the goal were aligned. The moment when the center of the target and the goal are at the same lateral position will henceforth be referred to as alignment. Motion durations were clearly discriminable from each other (Regan & Hamstra, 1993), while at the same time all were above the 500 ms over which basic temporal recruitment is built up (Krekelberg & Lappe, 1999, 2000). The combination of speeds and motion durations, which was random for each trial, resulted in nine different initial distances from the goal (15.6, 19.5, 23.4, 20, 25, 30, 25.6, 32, and 38.4 cm, in ascending order for speeds and motion durations). The target disappeared when the participant pressed the button or at a random point between 1.3 s and 1.4 s after movement onset, when the target had already completely crossed the goal. Visual feedback was given when required (see the following material). Note that interleaving different initial times to alignment with speeds made more difficult that participants could use a time-based response on a specific stimulus duration (e.g., 0.85 s) from motion onset rather than aiming at some temporal (or other type of) distance to alignment.
Temporal reward: The reward given after each response was a function of the response time t. Reward increased exponentially and the reward function $u(t)$ was such that pressing the button at alignment, centered at $t = 1$, was rewarded with 100 points. Responding after that was penalized with the subtraction of 200 points (Figure 1C). Thus, for time-defined rewards:

$$u(t) = \begin{cases} \beta \times \exp(\alpha t), & \text{if } t \leq 1 \\ -200, & \text{if } t > 1 \end{cases} \quad (8)$$

Where the linear parameter $\beta = 0.1008$ acted as a scaling parameter and the non-linear parameter $\alpha = 6.9$ defined the slope or rate of change of the reward as a function of time. Note that, for the motion durations of 0.8 s and 1.2 s, to make alignment be always at $t = 1$ we simply shifted the reward function so that the same reward was given across motion durations when the response was made at the same time to alignment.
Spatial reward: The reward given after each trial was based on the remaining spatial distance to alignment. The reward function was:

$$u(s) = \begin{cases} \beta \times \exp(\alpha s), & \text{if } s \leq 15 \\ -200, & \text{if } s > 15 \end{cases} \quad (9)$$

And the corresponding parameters were $\beta = 1.1109$ and $\alpha = 0.3$, so that the reward was also 100 points at alignment that was centered at position $s = 15$ (for the goal was positioned 15 cm right of the center of the screen). Later responses were penalized with −200 points. The value of these parameters was selected so that, for the medium speed, both the time and space condition had the same mapping between time/space and reward.

Second Order motion condition stimuli

In this condition the reward was temporally defined, with the only difference from the temporal reward condition being the visual stimuli presented to the participant. The target's lateral displacement was produced over a horizontal stripe (30 cm width and 3 cm height) (33.4° × 3.44°) consisting of a contrast-defined random dot texture (black and white dots) as in (Seiffert & Cavanagh, 1998). The size of the texture elements was 0.12 × 0.12 cm^2 (0.13 deg^2) and the dots were updated every 80 ms. The moving target (circle of 1 cm diameter; 1.15°) was also composed of the same texture elements and updated at the same rate. With such stimuli, participants had to obtain motion information from a change in contrast, not in luminance.

Procedure

Participants completed two blocks of practice trials (90 trials per block including all speeds and motion durations). In these practice trials, if their response time was within a temporal window of 200 ms centered on the goal, visual feedback was provided. The aim of the baseline session was to familiarize participants with the experimental paradigm, especially with the timing until alignment.

After the practice trials, the main part of the experiment started. Everything remained the same except for the feedback after each trial. Reward was introduced, so that each trial was rewarded according

to the response and condition (see the aforementioned material). After a nonpenalized trial, the reward won in that trial appeared on the screen, as well as the total reward won in that block until the moment. After a penalized trial, "−200" appeared next to the total reward. As before, blocks had 90 trials. Participants completed 12 blocks. They started each block with 5,000 points, and they were instructed to finish each block with as many points as possible. They were told that their final payoff would depend on the total points they accumulated throughout the experiment. When receiving the initial instructions, participants were shown a graphical depiction of the reward function, similar to Figure 1C. However, no reference was made to reward being related to time or space.

Data analysis

A trial was identified as an outlier if the moment of the response was above or below three standard deviations for that particular participant, block, motion duration, and target speed. These outliers were removed before proceeding with the analyses.

Temporal and spatial distributions of responses

We looked first at whether response times were temporally or spatially equidistant to the goal across speeds. We did this for each participant, block and motion duration. Figure 1D shows the main rationale based on how the response times and target position at the response time are distributed for the different target speeds.

A pure temporal invariance of responses would consist in, for all target speeds, aiming at responding when the time left to the goal is very similar (e.g., points vertically aligned in Figure 1D). On the contrary, a pure spatial invariance of responses would result in responding, for all speeds, in similar spatial distances to the goal (points aligned horizontally in Figure 1D). One can think of a tau-based response (Bootsma & Oudejans, 1993; Lee, 1976) when responses are time-invariant across speeds. Note, however, that we want to capture to which degree overt responses are temporally or spatially distributed across speeds. For example, suppose that one participant responds at a similar spatial distance to the goal across different speeds. We do not require this distance to be stable in absolute terms. For instance, this same participant may respond when the target is 1 cm away from the goal, and eventually shift to when it is 2 cm away, without her pattern being affected as long as this difference remains the same across speeds. Our experimental design allowed to determine the extent to which participants exploited more temporal or spatial cues in this way. This is described as follows.

To estimate the trend of the response distribution, we just fitted a linear model to the response times as a function of the reciprocal of the speed:

$$t_{resp} = s \times v^{-1} + b \quad (10)$$

Where the slope s would provide an indicator of a time-invariant response pattern (s very close to zero). b would account for a speed-independent temporal constant and would include the sensorimotor delays. We used the reciprocal of the speed because then the slope has a meaningful interpretation in spatial terms (s would correspond to the remaining distance to the goal at the time of the response).

We calculated s (the observed slope, see also Figure 1D) for each participant, block and motion duration. Although values of s closer to 0 informed us time-based responses, we did not know which range of s values could be related to mainly temporal information being used to plan responses—and the same for spatial information. That is why, for each participant and motion duration, and using her own data, we obtained the expected slope, as a function of response time, corresponding to a pure spatial arrangement of responses (i.e., as if the same spatial distance to the goal was used across speeds), that is the slope s_{spat} in Figure 1D. For this purpose, we took the mean of the responses in space (i.e., spatial position of the response) for the whole experiment, and divided it by each speed, thus obtaining the three expected response times that would correspond to a pure spatial alignment. Those were fitted to Equation 10 to estimate the slope s_{spat} that would correspond to an entirely positional response pattern. This slope (s_{spat}) was compared to s, the slope obtained from the observed data. More specifically we used the slopes ratio s/s_{spat} (shown in Figure 1D). Then, a perfect spatial arrangement of responses across speeds would correspond to a ratio of 1, while a ratio of 0 would denote an invariant temporal distance to the goal. This single value allows us to summarize the response behavior of each individual participant within a continuum capturing temporal and spatial invariant responses across speeds.

Some of the analyses relied on grouping participants depending on whether they exploited more temporal or spatial cues. We used a cutoff point. Our compromise choice was the value halfway (0.5) between a purely spatial (ratio = 1) and a purely temporal (ratio = 0) pattern. Thus, if the ratio was smaller than 0.5, the participant was considered to favor a time-based response. Otherwise, her pattern was categorized as space-based. In addition, by calculating a ratio between s and s_{spat}, we also obtained a value that was used as an indicator of how time or space-based responses were along a continuum. Note the fact that each partici-

pant's responses were split by motion duration. Thus, a participant's data could fall within both the temporal and spatial groups, depending on the main cue used to plan responses for a particular motion duration. We followed this grouping procedure for every group analysis detailed in the results section.

Comparison between conditions

Statistical testing was conducted with the behavioral and simulated response time distributions to see whether the different stimuli conditions (e.g., motion duration and target speed) produced similar effects in both types of distributions. One of the hypotheses that we had concerning response time distributions or the associated spatial position is the predicted larger spatial variability for faster speeds, that in our task would lead to a higher number of penalized trials. Concerning the slopes ratios, we expected them to reflect more spatial invariance across speeds in the Second Order motion condition or in conditions where speed estimations become more difficult (e.g., short motion durations).

Finally, we also looked at history effects. On a longer timescale, we examined possible learning effects across blocks. On a shorter timescale, we analyzed whether the time of the response was affected by the previous target speed, or by whether the previous trial had been penalized (response after the goal).

To make the comparisons, we used Linear Mixed Models (LMM) and ran analyses of variance (ANOVAs) on the output of the LMM. Since every condition was run with a different sample, participants are nested in conditions and LMMs allow us to specify the participant random effects in the nested design very easily. Target speed and motion duration were treated as fixed effects. We conducted the LMM and corresponding ANOVAS on response times and the spatial standard deviation. We used the lmer function implemented in the lme4 R package (v.1.0-6) (Bates, Maechler, Bolker, & Walker, 2015). Similarly we used a Generalized LMM with the same structure for the analysis of the fraction of penalized trials. To compare slopes ratios distributions, we used a different approach (see the following material).

Simulations

To test whether human performance could be predicted by the perceptual or posterior estimates of position and speed generated by the Kalman filter, we simulated data under different conditions of target speed and motion duration. The position and speed measurement noise was set according to the simulated conditions following Equations 6 and 7. The measurement noise was added to the physical position and speed in each simulated trial. The process noise was set to 0.9 and 1.5, respectively, for position and speed. We ran 1,000 trials for each combination of speed, motion duration, the scaling factor of the speed's measurement noise (a in Equation 7) and the scaling factor of the position's measurement noise (dt in Equation 6). The initial state of the position estimate was the actual physical initial position, and the initial state for the speed estimate was set to 0 cm/s. The initial variances for these initial estimates were 2 and 7, respectively, for position and speed. Performance was not affected by these initial values of state uncertainty.

Once all these parameters were set, and we obtained the input signal with the corresponding measurement noise, we ran the Kalman filter and, for each simulated trial, we obtained the posterior estimate for both position and speed. The filtered signal (posterior) was then used to simulate time-invariant and space-invariant responses in each trial, with its corresponding reward. For example, for time-invariant responses, we computed the time to the goal based on the ratio between posterior position and speed estimates and used a constant temporal distance to the goal across all simulated conditions. We added a zero mean Gaussian noise ($SD = 10$ ms) to simulate execution noise.

We then computed the slopes ratio (as explained in the data analysis section) and compared the respective distributions between simulated and behavioral data.

To compare both simulated and behavioral distributions we used a two-sample Kolmogorov-Smirnov (KS) test. The KS allowed to test whether two histograms are samples from the same (null hypothesis) or different distributions. We used the ks.test function implemented in R. In addition we also compared the join slopes ratios-reward bivariate distributions between behavioral and simulated data. We performed a multivariate two-sample E test for equal (bivariate) distributions (Székely & Rizzo, 2004) based on the distance between behavioral and simulated distributions. To calculate the distances between the two distributions and conduct the test we used the R package energy and p values were computed from bootstrap replicates (Székely & Rizzo, 2013). By comparing the behavioral bivariate distributions against simulated ones based on time- or space-invariant responses we can know which response mechanism explains the behavioral data better.

Results

The raw data and the code for generating the posterior postion and posterior speed using the Kalman filter simulations are available at the following link: https://osf.io/vmjne/?view_only=8e30e81e8e4f4c4c9fe5b257ca65f503

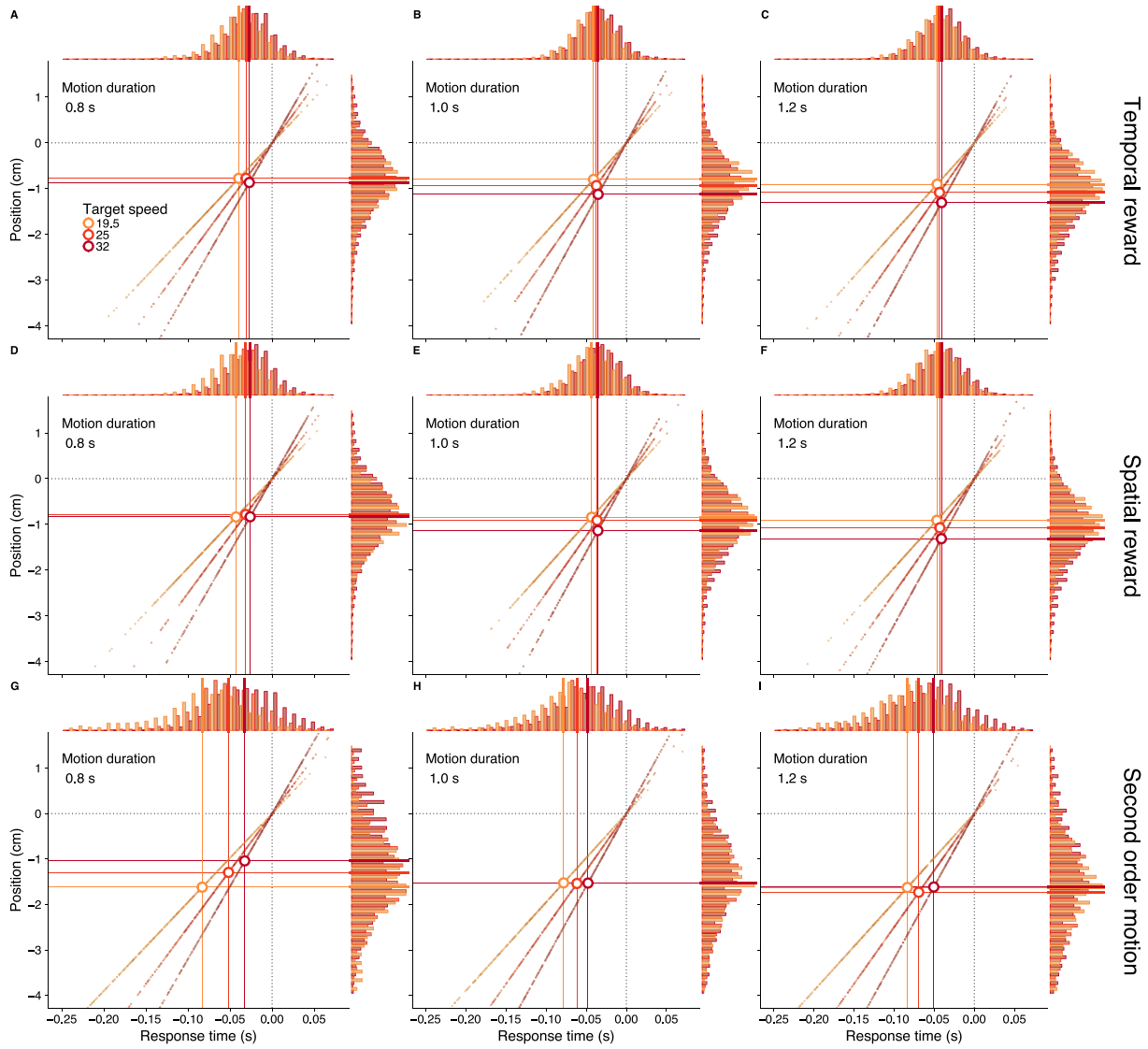

Figure 2. Position of the target at the moment of response, as a function of the response time, both expressed relative to the moment of alignment with the goal. Each panel shows, for a particular condition (rows) and motion time (columns), the full dataset of responses for all participants. Each data point, depicting a trial, is colored according to the target speed. Bigger hollow dots represent the mean of the responses for that speed, with vertical and horizontal lines helping visualize the average response time and position, respectively. Histograms of both response time and positions are displayed at the top and right borders, respectively, and color-coded by target speed.

Response times and position

All values referring to temporal responses reported henceforth are scaled so that alignment occurs at 0. Negative values indicate responses before alignment, and positive values indicate responses after alignment.

Outliers accounted for 0.348% of behavioral trials. These were removed from subsequent analyses. We first show results based on raw response times. Figure 2 shows the distributions of response times and positions of the target at the time of the response. The panels show the raw data points oriented in space (y-axis) and time (x-axis). The different target speeds result in three different orientations (speed is color-coded). How time and position are distributed in each oriented line is shown by the marginal histograms for each speed. By eye analysis, important differences can be seen between the different motion durations within each condition. For example, responses tend to be spatially aligned (same distance to the goal across speeds) for motion durations of 0.8 in the temporal and spatial reward conditons. Therefore, the resulting positional histograms for each speed overlap to a larger extent (closer means) than for longer motion

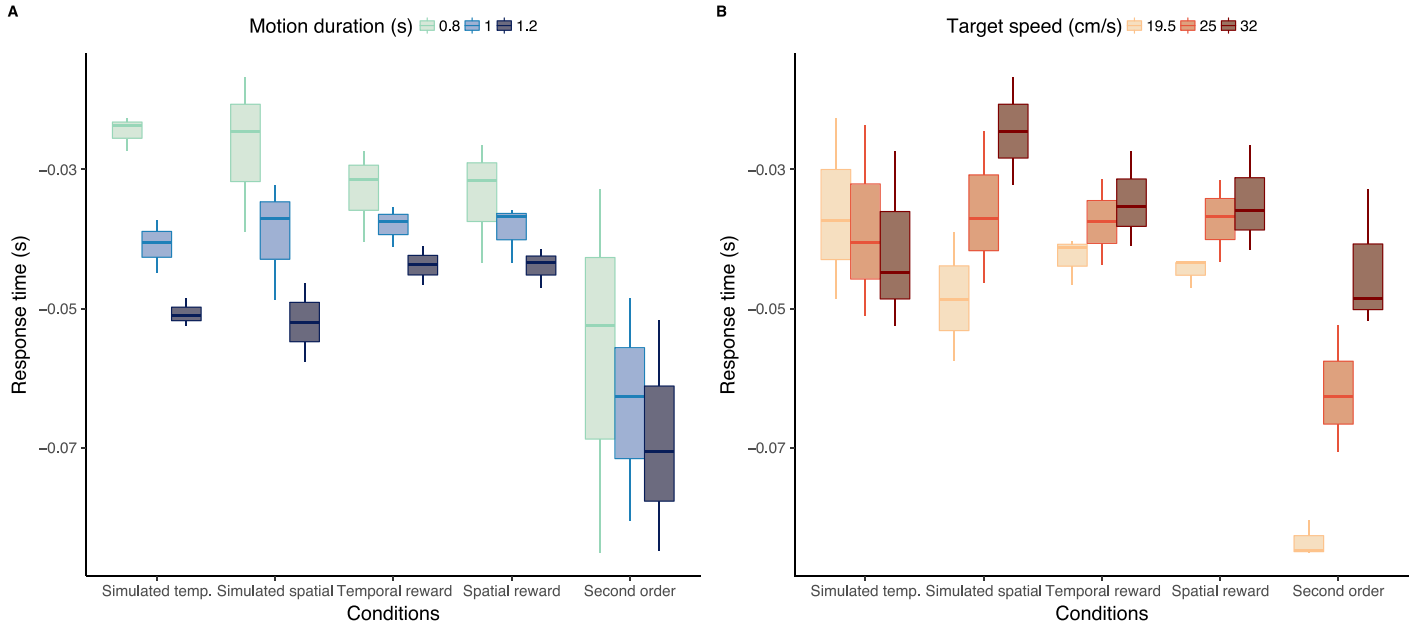

Figure 3. Boxplots of simulated (based on temporal and spatial invariance across speeds) and behavioral response times for the different conditions (0 corresponds to perfect alignment of target and goal at the moment of response; negative times denote responses before alignment). Boxes are split per motion duration (A) and target speed (B), color-coded accordingly. The central line, lower bound, and upper bound of the box represent the median, Q1 and Q3, respectively. The lower whisker corresponds to the smallest observation greater than or equal to the lower bound minus 1.5 times the interquartile range (IQR). The upper whisker corresponds to the largest observation less than or equal to the upper bound plus 1.5 times the IQR.

durations (1 and 1.2 s), for which the positional histograms per each speed are further away from each other. Following the rationale shown in Figure 1D and explained in the Methods section, this trend will result in slopes ratios close to 1. Alternatively, longer motion durations (1 and 1.2 s) elicit more temporally aligned responses across speeds (marginal temporal histograms at the top overlap more). For these cases, the slopes ratio will be close to 0. Note that a different pattern arises for the Second Order motion. In this condition, the responses are spatially aligned for the two longer motion durations.

Figure 3A shows the average data (in the form of boxplots) across participants (or repetitions for the simulated data) split by motion duration for the different conditions, while Figure 3B shows it split by target speed. As can be noted, response times were on average before alignment (temporal reward: −40 ms, spatial reward: −43 ms, Second Order motion: −87 ms), as expected if participants wanted to avoid penalizations (a more thorough analysis of reward will be presented as follows). In all behavioral conditions there was the same pattern for motion duration and target speed. Longer durations led to earlier responses and faster speeds to later responses. These main effects were significant for both variables (duration: $F = 452$, $p < 0.001$; speed: $F = 1{,}509$, $p < 0.001$), as yielded by an ANOVA on the output of the LMM (as described in the Methods section). Motion duration interacted with target speed ($F = 76.30$, $p < 0.001$): the later responses for faster speeds were modulated by motion duration (shorter durations produced later responses than longer durations). When exposed to Second Order motion, participants responded earlier on average (47 ms) than in the temporal reward condition), but differences between conditions were not significant ($F = 0.13$, $p = 0.88$).

More importantly, simulated data based on an spatial responses were able to reproduce the main trends of the behavioral data. The ANOVA conducted on a linear model on the simulated response times yielded the same significant effects of motion duration ($F = 1{,}620$, $p < 0.001$), target speed ($F = 2{,}030$, $p < 0.001$), and their interaction ($F = 29$, $p < 0.001$). This pattern (Figure 3) is based on simulating a space-invariant response across target speeds on the output of the Kalman filter (posterior estimate of position). However, the same pattern (i.e., same statistical differences among conditions) was obtained by simulating time-invariant responses, but only for the different motion durations. Importantly, the simulation of time-invariant resposnes could not capture the trend of the mean differences across target speed. Only the space-invariant responses could do so, as shown in Figure 3B. Therefore, simulations based on a single spatial response strategy were able

to reproduce the same trend as the behavioral data, both across motion durations and target speeds. This ability to capture the response patterns across all the different stimulus conditions initially favors a space-based mechanism.

History effects

We looked first at the effect of block number to test whether there were learning effects. We included block number and condition as fixed-effects in the linear model and participant as random effects. Although there were changes in response time across blocks (less than 1 ms/block in all conditions), the effect of block failed to reach significance ($F < 1$, $p = 0.77$), with the interaction between block and condition also being nonsignificant ($F < 1$, $p = 0.66$). Only condition had a significant effect in this model ($F = 4.47$, $p = 0.016$) in which target speed and motion duration were not included. Previous work have shown effects of target properties of previous trials (e.g., target speed in de Lussanet, Smeets, & Brenner, 2001). We set up another linear model with previous speed and condition as fixed effects and participants as random effects. Both previous target speed ($F = 26.6$, $p < 0.001$) and condition ($F = 14.7$, $p < 0.001$) had significant effects on response times as well as their interaction ($F = 4.9$, $p < 0.001$). Interestingly, however, the effect of the previous speed was opposite to the effect reported already for the current target speed: that is, for larger previous speeds the response in the next trial was earlier (e.g., about 8 ms earlier for 32 cm/s than for 19.5 cm/s). Since this may denote some correction after penalized trials for faster speeds (as reported later in the Results section), we had a closer look and introduced in the model whether the previous trial was penalized. Although reduced, the effect of the previous speed remained significant ($F = 3.23$, $p = 0.038$). Importantly, this effect was mainly caused by the correction (about 16 ms earlier) after the faster speed (32 cm/s) when the previous trial was penalized. This was the parameter that contributed most in the model. The penalization factor as a main effect did not reach significance ($F < 1$, $p = 0.33$).

We will next look at the reward and slopes ratio in order to shed light on the underlying mechanism.

Response pattern across speeds

As explained in the Methods section and Figure 1D, we obtained the slope of response times against the reciprocal of the target speed (slope s) and also the slope that, for each participant, motion duration and block, would correspond to a pure spatial arrangement of responses (slope s_{spat}). The ratio of these two slopes would denote whether a set of response times are overt time-invariant (slopes ratio = 0) or overt space-invariant (slopes ratio = 1) across the different target speeds. Figure 4 shows the histograms of the resulting slopes ratio for each motion duration and condition, together with the density of the slopes ratio obtained from the simulations based on spatial (solid green) and temporal (dotted purple) responses. The slopes ratio values contained in the histograms corresponding to experimental conditions are obtained per each participant and block. As can be seen, the slopes ratio obtained from simulating time-invariant responses fails to capture the change of location of the distributions obtained in the different motion durations and conditions. However, when the response is based on invariant positions across speeds, the resulting distributions are shifted in the same direction as the experimental distributions: leftward shifts (toward zero or more temporal responses) for longer motion durations. Densities obtained from the simulated spatial responses were not different from the histograms in the different experimental conditions (p values of the KS statistic shown in Figure 4).

The means across blocks of the slopes ratio are also shown in Figure 5 for the different conditions (different panels). Note that target speeds are used to compute the slopes ratios and cannot be used as grouping variable. Now it can be more clearly seen that different motion durations produced different kind of responses (equivalent to the shifts of the histograms in Figure 4). Larger durations promoted more time-invariant responses across speeds. Confidence intervals of the experimental slopes ratios included the zero for 1.2 and 1.0 s of motion duration in the temporal and spatial reward conditions, whereas the shortest duration (0.8) in these two conditions elicited responses that are consistent with constant spatial distance to the goal across speeds (slopes ratio very close to 1). As expected, the values of the slopes ratio are shifted rightwards in the Second Order motion condition, where luminance-based motion is impaired. The two longer durations now yield a perfect space-invariant pattern, while the ratio is very close to two for the 0.8 s duration. These larger ratios denote closer spatial positions to the goal for the fastest speed (32 cm/s) and farther distances to the goal for the slowest speed (19.5 cm/s).

Interestingly, the average simulated ratios based on the spatial response also depended on the motion duration (0.8 s: 1.054, 1 s: 0.583, and 1.2 s: 0.363). These means lie along the green solid lines in Figure 5. Note that this was only a consequence of lowering the noise in the measurement of the speed with longer motion durations, resulting in producing more similar

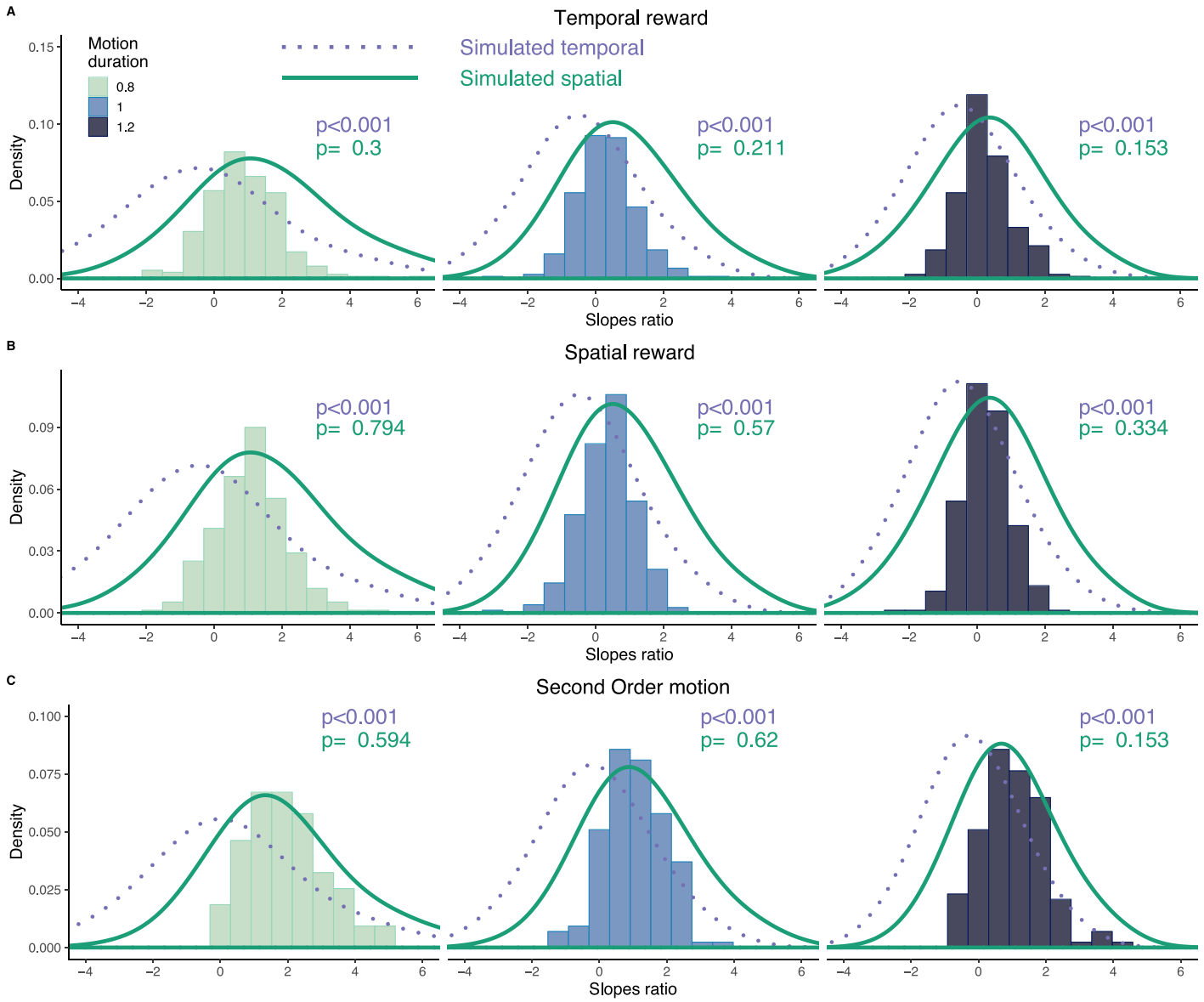

Figure 4. Histograms of observed slopes ratios, split per condition (rows) and motion times (columns and color-coded). Histograms are built from the slopes ratio for each participant and block. The densities of the corresponding distributions of simulated time-invariant (dotted blue) and space-invariant (solid green) slopes ratios are superimposed. *P*-values on each panel correspond to a two-sample Kolmogorov-Smirnov (KS) test, where the null hypothesis is that the distribution of temporal (purple) and spatial (green) simulated slopes ratios is no different from the histograms of the observed slopes ratios (i.e., both density and histogram belong to the same distribution).

response times across speeds without changing the response mechanism. The same pattern was obtained irrespective of whether the reward is temporally or spatially defined. Thus, the domain of the reward does not seem to affect the structure of the response distributions (i.e., temporal or spatial invariance across speeds). However, motion duration seems to be an important factor. We next need to know how the type of response (time or space-invariant) affected the accumulated reward and whether the model based on the Kalman filter can predict the reward across the different conditions and motion durations.

Reward and loss fraction

Figure 5 also shows that longer motion durations produced larger average rewards. The effect of motion duration on the reward was very significant ($F = 163.9$, $p < 0.001$), but the differences between conditions failed to reach significance ($F = 2.47$, $p = 0.094$). There was a significant interaction ($F = 6.63$, $p < 0.001$) as a consequence of the differential reward between the two longer motion durations in the different conditions (i.e., very similar for the Second Order motion). The

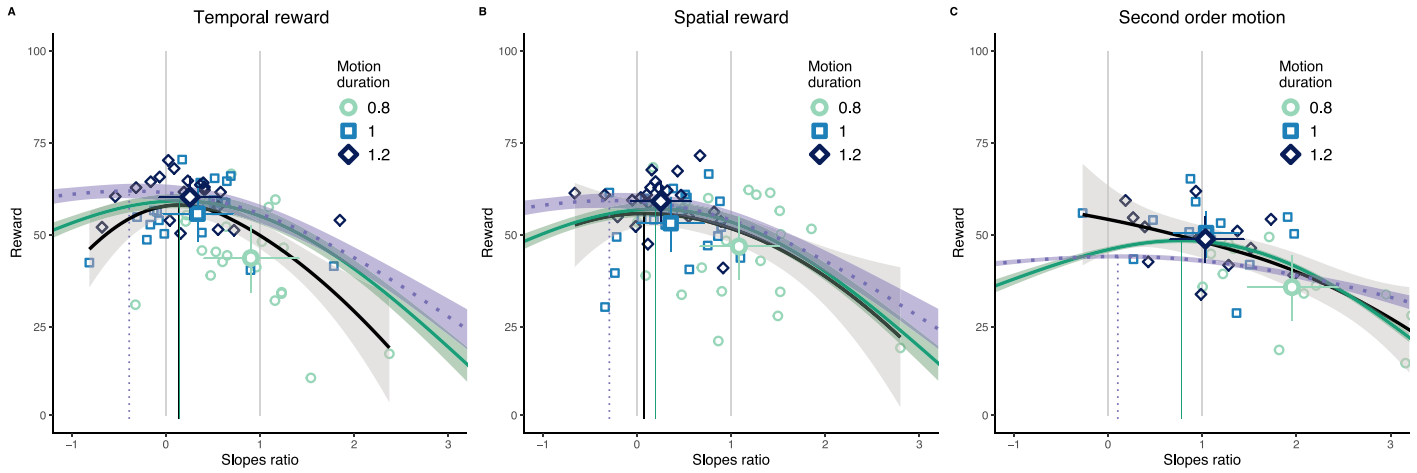

Figure 5. Average reward per trial (in points) as a function of the slopes ratio (0 = perfect temporal invariance across speeds; 1 = perfect spatial invariance across speeds; both denoted by gray vertical lines). Each participant is included in each panel with three data points; each point corresponding to her responses within a certain motion duration (shape- and color-coded). Bigger points represent the mean of each motion duration's distribution, with associated error bars that show 95% confidence intervals. The black solid lines denote a (Gaussian) smoothed version of the data points. Green and blue lines denote the pattern from the simulated data and what the model would predict based on space-invariant (green solid) or time-invariant (blue dotted) responses. Since we simulated different factors of speed measurement noise, we plot the predictions from the simulated condition that was closer to the behavioral data, according to the KS statistic value (see Methods).

black solid line in Figure 5 shows the overall relation between slopes ratio and reward. The average reward has a maximum, which is very close to a slopes ratio of zero in the temporal and spatial reward conditions. This is also so for the Second Order motion, in which reward is on average larger for slopes ratios close to zero. However, due to the overall rightwards shift of the slopes ratio we cannot observe the same non-monotonic pattern of the reward.

To explain why more overt time-invariant responses across speeds enhanced performance, we explored a hypothesis already mentioned in the introduction: that participants whose response distribution was consistent with time invariance dealt better with faster target speeds. The precision by which we acknowledge the change of position of a moving object is limited by the temporal resolution of the visual system, with spatial uncertainty being larger for faster speeds. Thus, spatial variability increases for faster speeds, as depicted in Figure 1A. The amount of penalized trials should thus be larger for faster speeds when the response pattern is consistent with spatial invariance. To test this, we grouped the whole dataset, split by the slopes ratio: slopes ratios smaller than 0.5 were considered to have more time-invariant patterns across speeds (i.e., more constant temporal distance to the goal), while slopes ratios larger than 0.5 included more space-invariant patterns (i.e., similar positions to the goal across speeds). Note, however, that negative slopes ratios do not correspond to temporal invariance across speeds but to earlier responses for faster target speeds.

We calculated the loss fraction for the different target speeds (Figure 6). For each speed, the loss fraction was simply the proportion of trials for which responses were made after alignment, and thus were penalized. We conducted a general linear mixed model (binomial responses) with the loss fraction as dependent variable, target speed and response type (spatial or temporal invariance) as dependent variables (fixed effects) and participants as random effects (we allowed both intercept and slope to vary). Temporal responses did not lead to a significant rise of penalized trials with increasing speed (average slope = 0.008, $p = 0.27$), while spatial responses did result in a significant increase of penalized trials (slope = 0.062, $p < 0.001$). These slopes correspond to the average values of the slopes across conditions, since the different conditions did not have a significant effect ($F < 1$).

Response mechanism

As mentioned already, the same data pattern across motion durations revealed by the behavioral response can be reproduced by both types of invariances across speeds. However, only the space-invariant response mechanism could capture the trend across target speeds. Furthermore, the same participants could produce responses that were consistent with either pattern, depending on the motion duration (e.g., longer duration generally leading to more time-invariant responses across speeds). As shown in Figure 5, the average reward seems to have a maximum at about a slopes ratio of

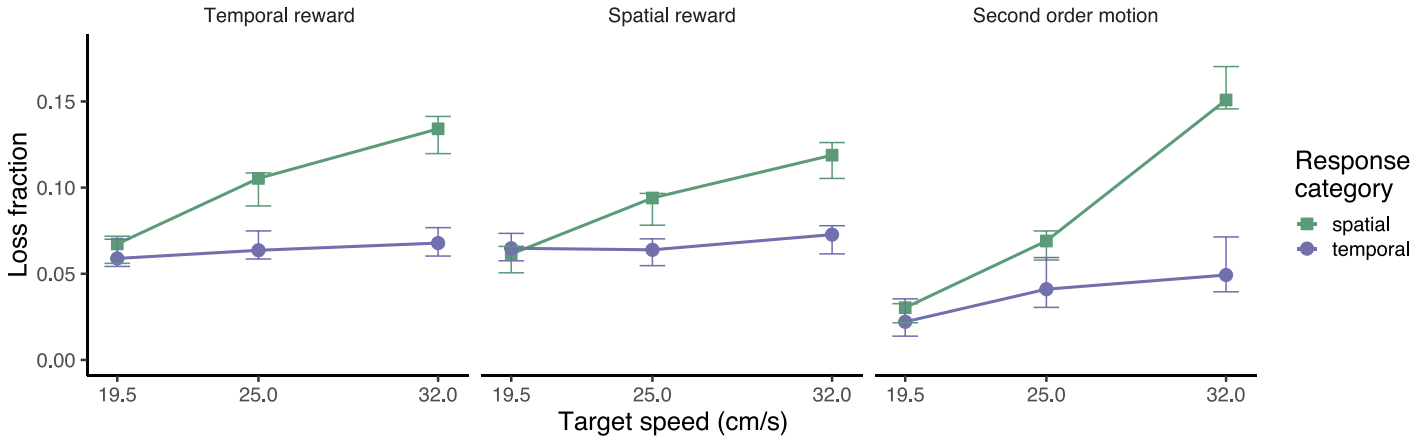

Figure 6. Average loss fractions between response pattern groups (green: mostly space-invariant overt responses; blue: mostly time-invariant overt responses) as a function of target speed, with a different for each condition. Error bars represent binomial 95% confidence intervals.

zero, and it drops as the slopes ratio deviates from this value. This trend is very well captured by simulating space-invariant responses and their corresponding reward (purple solid line in Figure 5). The maximum predicted reward for the two response mechanisms is denoted in Figure 5 by a vertical line. For the temporal and spatial reward conditions, the predicted maximum reward is very close to the observed maximum reward (undistinguishable in the temporal reward condition). The maximum reward predicted by simulated time-invariant responses deviates from the observed reward and is expected to peak at a lower (smaller then zero) slopes ratio. For the three conditions, the closest simulated distribution (minimum distance) to the behavioral data was obtained by simulating a space-invariant response (dist = 1.72, $p = 0.93$; dist = 3.28, $p = 0.55$; dist = 2.5, $p = 0.81$, respectively for the temporal reward, spatial reward and Second Order motion conditions). The best predictions obtained by simulating a time-invariant response on the output of the Kalman filter deviated more from the behavioral data (dist = 10.64, $p = 0.16$; dist = 16.33, $p = 0.03$; dist = 6.1, $p = 0.38$ respectively for the temporal reward, spatial reward and Second Order motion conditions). Despite the larger deviation between the simulated time-invariant and the behavioral distributions, we cannot reject the null hypothesis (i.e., not different from a time-invariant pattern) in the temporal reward and Second Order motion conditions.

The fact that aiming at similar (perceived) positions across speeds can also generate overt time-invariant responses is a new finding. Note that the mechanism to simulate the response is based on the posterior (perceived) position obtained by the Kalman filter, with the only difference between the critical conditions that yield different patterns (motion duration) being the measurement noise variance of the target position and speed. To further strengthen the bigger plausibility of a spatial mechanism, we simply took the slopes ratios obtained in the simulated trials with both response mechanisms (temporal and spatial invariance) and plotted them as a gradient as a function of the different simulated positional and speed measurement noise variances (Figure 7A). This figure shows that there is a clear dependency of the slopes ratio on both types of measurement noise variance. However, more importantly, it also shows that aiming at similar perceived positions across speeds (i.e., spatial invariance) covers a larger spectrum of the slopes ratio (left panel of Figure 7A), which is consistent with the observed slopes ratios in the behavioral data. As a consequence, an optimal balance between positional and speed uncertainty is observed when the slopes ratio is about 0. This is the situation with the least temporal variability across target speeds. This, and not a different mechanism, is why in our experiment overt time-invariant responses (slopes ratio close to 0) lead to the highest reward. Finally, if positional uncertainty increases too much, position would be further extrapolated for faster speeds, increasing response variability across speeds, and affecting the reward.

Discussion

Due to the spatiotemporal nature of interacting with moving objects, people can plan their actions by basing responses on temporal and/or spatial information. The specific information that we may use in different situations has been a long-standing debate in the visual control of timed actions. Our main finding suggests that a single response mechanism based on position tracking can account for both overt spatial and temporal response distributions. That is, temporal and spatial physical arrangements of targets at the time of the

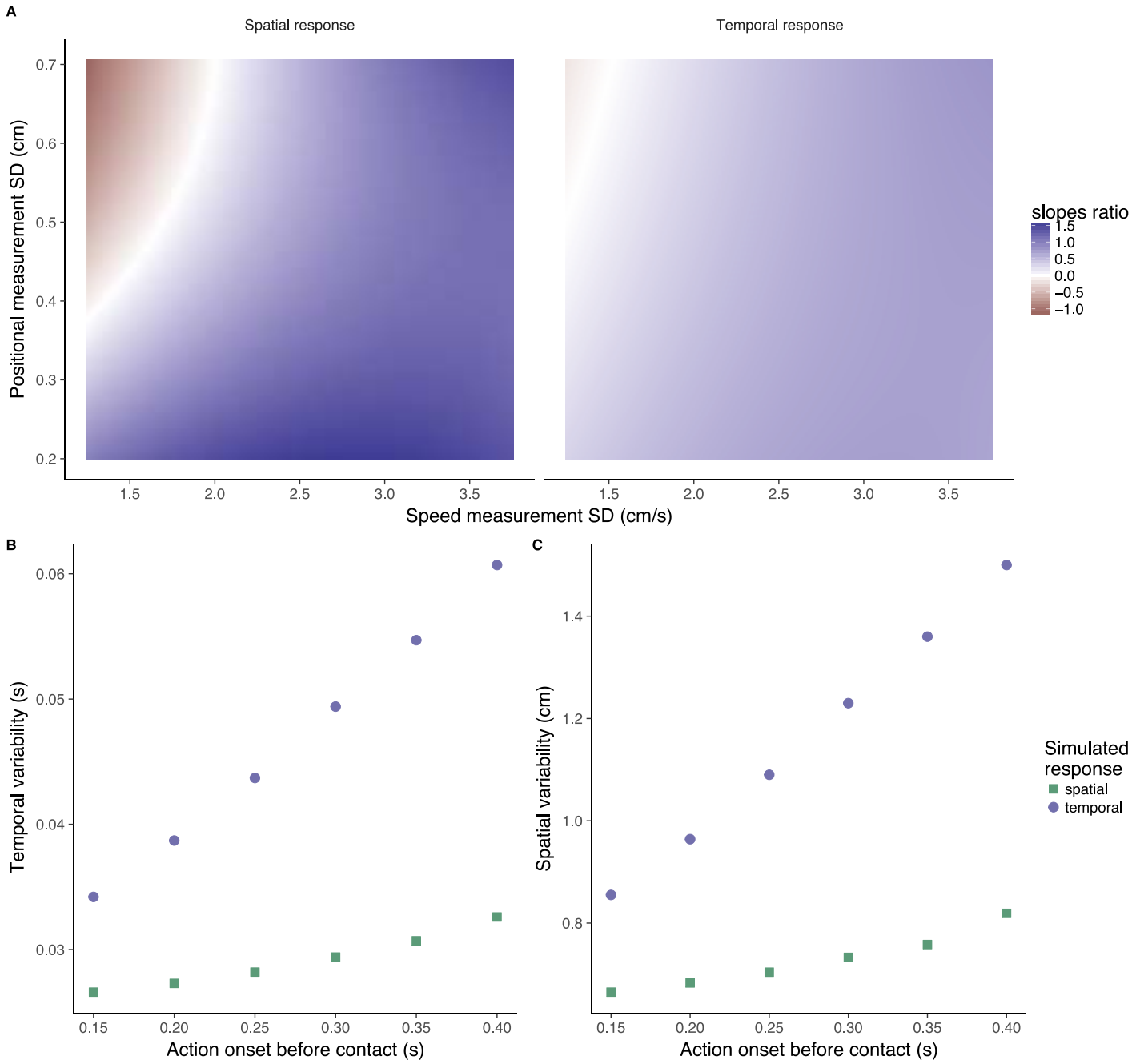

Figure 7. (A) Mean of slopes ratio obtained from the simulated data (color gradient) as a function of the speed measurement *SD* (x-axis) and the positional measurement *SD* (y-axis). The two panels denote the simulated response mechanism: spatial invariance (aiming at the same position across speeds) and temporal invariance (aiming at the same temporal distance from the goal across speeds). The white color in the gradient corresponds to the slopes ratio of zero (observed time-invariant distribution across speeds), which would correspond to an optimal balance between positional and speed uncertainties. Slopes ratios for values of measurement variability that were not included in the simulations were linearly interpolated. (B) Simulated temporal variability of response initiation as a function of the time of action initiation relative to the alignment between target and goal (earlier responses correspond to larger values). The color denotes the simulated response (spatial or temporal). (C) The same as in B, but with the spatial variability at response initiation.

response would originate from the same spatial mechanism. This can be concluded from simulated space-invariant responses on the output of the Kalman filter model being able to reproduce the different arrangement of overt responses (i.e., spatial and temporal alignments to the goal) across different motion durations. Simulated time-invariant responses failed to reproduce the same response pattern across motion durations.

Response mechanism and response criteria

Responses consistent with a constant temporal criterion to the region of interest have been traditionally explained through people using specific combinations of optical variables (e.g., tau), which can predict an invariant response pattern across different target speeds (Bootsma & Oudejans, 1993; Lee, 1976; López-Moliner & Bonnet, 2002; Smith, Flach, Dittman, & Stanard, 2001; Marinovic, Plooy, & Tresilian, 2010). In these situations, using some informational variable that carries temporal structure (e.g., rate of expansion in computing tau) would be necessary to generate invariant responses across speeds. However, depending on task constraints (e.g., range of target speeds), people can use variables of spatial nature, like visual angle, to solve a collision task while ignoring other variables that increase temporal precision. This would result in spatial response patterns (Smith et al., 2001). Similarly, response control based on spatial information has also been proposed in other past studies (Brenner et al., 2006; López-Moliner & Keil, 2012; Wann, 1996). This diversity of response patterns makes it difficult to reconcile most empirical findings into a general model of interceptive timing. This is mainly, but not the only reason, why timing has been regarded functionally separable from positioning (Tresilian, 1999b; Brenner & Smeets, 2015). In this sense, the Kalman filter model provides an implementation in which positional and temporal information (with the latter provided by speed) are coupled and the use of either of them alone is meaningless.

In addition to the nature of the relevant informational variable (or variables) that characterizes the mechanism, there is the issue of the response criterion or value along the relevant variable the triggers the response. Admittedly, we have simulated an invariant criterion (spatial and temporal). However, rather than assuming an invariant criterion, our simulations aim at capturing the main trends by focusing on either type of information (position or temporal). In this sense the use of an invariant criterion should be regarded as a simplification in our implementation but not as a fundamental feature of the model. It has been reported that different criteria of an informational variable can be used in different task conditions or target properties (e.g., different thresholds of rate of expansion for different target sizes, e.g., López-Moliner et al., 2007; Smith et al., 2001). Here, we have shown that the reward function can affect the response criteria. As reported in the Results section, earlier responses were recoded after a penalization irrespective of stimulus condition. Also, Figure 3B may give some evidence that participants are not using a unique criterion. A comparison between the boxplots of observed and simulated responses shows that, although the pattern and the average across speeds are the same, the boxplots of the simulated responses are more spread out. This may be a hint that participants are actually changing their spatial criterion to respond earlier for the fastest speed and later for the slowest one.

Also for the sake of simplicity we have conducted the simulation based on metrical units. However, in our task, for instance, participants could have used optical variables like the rate of contraction of the visual angle between the target and the goal (Bootsma & Oudejans, 1993; Tresilian, 1999b). The Kalman filter model could operate with these variables as well. The reason we have not used optical information is because models based on optical variables do not normally make predictions on the measured uncertainty of these optical variables: the trade-off of the sampling noise with the internal uncertainty since neural processing, for example, is less emphasized. This means that the differences across motion durations cannot easily be accommodated by using optical variables alone without resorting to their sampling variability. Models that currently only combine optical variables, then, should be reformulated, either by including uncertainty measures or an active role of internal models in interceptive tasks (Kwon & Knill, 2013). In line with our findings, this has already been proposed in optical models for catching within the context of the outfielder problem (Belousov, Neumann, Rothkopf, & Peters, 2016) or estimating collision avoidance (Keil & López-Moliner, 2012).

Positional and speed uncertainty balance

The Kalman filter model finds an optimal balance between position and speed measurement uncertainty. In addition, these can be weighted differently depending on the stimulus' conditions or task constraints, something not addressed in our study. In one of our conditions we hampered the motion system by removing luminance-based motion as a main source for speed estimation. By doing this we supposedly increased the speed measurement noise, which in turn made the tracking favor positional information. When this happened, the distribution of target positions at the time of response became more position- (or space) invariant even for the longest motion durations. However, we claim that this distributional change is not caused by a different response mechanism, but it arises as a consequence of the internal balance of uncertainties in the system. More specifically, we propose that while the balance between position and speed uncertainties may change the overt distribution of responses, the mechanism would still be based on continuously sampling the target's position in space until it reaches a spatial distance that triggers the response. The simi-

larity in performance between simulations and experimental responses supports this claim, since the similarity of the shifts observed in Figure 4 is well predicted by the expected uncertainties of position and speed.

The sampling error of position increases with faster speeds, but this error can be ameliorated by prediction and temporal integration from speed. We found that the more overtly time-invariant responses were, the larger the average reward in our task. We claim that this enhanced performance is a consequence of an optimal ratio between position and speed uncertainties, with lower speed measurement noise favoring the predictive and integrative properties of speed. As a consequence, a more stable performance across speeds is achieved. This is shown in Figure 6. In contrast, more overt spatial invariance increases the number of penalized trials as target speed becomes faster. This pattern is predicted by the model due to larger speed measurement noise or smaller positional noise. Our results can be consistent with the recent finding that additional cues to timing (e.g., auditory signal) increase precision in a time-to-contact task (Chang & Jazayeri, 2018). It has been shown that people integrate estimates from timing and time-to-contact based on speed. In our model, we suggest that by adding further timing cues the positional measurement variance would be reduced.

When more emphasis is put on the predictive side (relatively more reliance on speed to predict future positions), it is then more likely to circumvent the constraints of the visual system (i.e., temporal resolution). This results in monitoring high speeds not leading to an increase in spatial variability and error (Brenner et al., 2006). This explanation is also consistent with previous findings on motion detection. Spatial displacements or speed thresholds have been both revealed to drive discrimination performance depending on the conditions. Nakayama and Tyler (1981) reported that minimum displacement was used when the object could be easily tracked, while for gratings it depended on the temporal frequency (TF), with TFs smaller than one being consistent with using a speed threshold. This was replicated by de la Malla and López-Moliner (2010), in addition to revealing the use of distance-based thresholds in expanding/contracting stimuli. Seiffert and Cavanagh (1998) also identified position displacement as the crucial variable for Second Order motion, consistent with our findings.

We used the slopes ratio to characterize how space- or time-invariant overt responses were, where a slopes ratio of 1 would correspond to responses purely aligned in space, while 0 would correspond to complete alignment in time. However, it needs to be explained why we identified slopes ratios smaller than 0 in our data. This can be linked to a response pattern where participants would respond earlier (both in time and in space) for faster speeds. This is consistent with motion extrapolation (Linares et al., 2007; Nijhawan, 1994), and would be predicted by smaller values of target speed measurement noise (or, equivalently, larger values of measured positional noise).This finding is in accordance with the account that the position tracking model of Kwon et al. (2015) gives of the motion-induced position shift (MIPS).

Pursuit and saccadic eye movements in interceptive timing

Although we have not recorded eye movements, we think that eye movements can play an important role in determining the uncertainties of target speed and position of the moving target. Not only do eye movements affect perceived speed (Turano & Heidenreich, 1999; Goettker, Braun, Schutz, & Gegenfurtner, 2018), but also, and more importantly, eye movements help to avoid systematic biases caused by texture motion within the target in hitting (de la Malla, Smeets, & Brenner, 2017). In interception tasks people tend to pursuit the target until about 100 to 150 ms before the hitting (period the would correspond to sensorimotor delays) and make a saccade to the hitting position (de la Malla et al., 2017, Cámara, de la Malla, Lopez-Moliner, & Brenner, 2018). Smooth pursuit might help to refine target speed estimates based on first-order information for time-to-contact, while neglecting possible accelerations (Benguigui & Bennett, 2010). However, improving speed estimates might not be the only benefit of smoothly pursuing the target. Consistently with our model, positional input appears to be also relevant for the pursuit system (Blohm, Missal, & Lefevre, 2005), showing a coupling between the saccadic (based on position error) and smooth pursuit systems that parallels the coupling between position and speed in our Kalman filter model.

Task limitations

The model has been compared with results obtained in a coincidence action (CA) timing task, which differs in some aspects from a fast interception action (IA) task (see Tresilian, 1995, for a comparison of different type of tasks). One important difference is related to the viewing time. Usually, targets are visible for shorter durations in fast interception than in our experiments. This might have promoted a stronger reliance on planning with no chance to exert online control (i.e., movement time virtually null in CA). Since we try to characterize the initiation of responses, we conducted additional simulations with earlier action onset times (from 150–400 ms prior to contact with the goal) or

equivalently with further spatial positions that would elicit the same average onset times. We then looked at the response variability (both temporal and spatial). The results are plotted in Figure 7B and 7C, respectively, for temporal and spatial variability at action onset. It can be seen that only the simulated spatial responses allow to keep both variabilities at reasonable values with earlier action onsets. This denotes that a spatial response mechanism could at least be applied in faster actions without incurring in large behavioral variability, specially if errors at action onset cannot be corrected online. Importantly, it has been repeatedly observed that actions are initiated earlier for targets moving more slowly, and later (but executed more quickly) for faster targets (e.g., Bootsma & van Wieringen, 1990; Laurent, Montagner, & Sevelsbergh, 1995; Wallace, Stevenson, Weeks, & Kelso, 1992). Our simulations, but only those based on spatial responses, are able to reproduce the action initiation pattern.

Another important difference between IA and CA tasks is the variability of response times, with the temporal standard deviation being about five times larger in CA than in IA tasks under the same conditions (McLeod et al., 1985; Tresilian, 1994). This is probably due to the more specific feedback available in fast IA that may allow the ability to refine temporal precision. If so, this component is missing in our modelling of response initiation. Finally, related to the type of task, fast IA are not affected by cognitive factors. In this sense, we avoided using occlusions which, when longer than sensorimotor delays, tend to make the task permeable to cognitive processes (Tresilian, 1995).

Trial-to-trial corrections and penalized trials

Our results suggest that reward might not only depend on the participants' slopes ratio. An obvious factor is where exactly people aim at responding. Aiming at earlier positions would lead to less reward when the response was before the goal, but also to reducing the chances of being penalized. An element directly related to this is how participants corrected their responses on a trial-to-trial basis. For instance, after a penalty (response after the goal), some participants may have decided to shift their aimpoint earlier so as to avoid future penalizations. This particular behavior was proven to be significant for trials where the last one had featured the fastest target speed. It could also happen that, after a series of nonpenalized responses, participants may be slowly responding later, to win more. These different dynamics are not analyzed here or incorporated in the model, but are important to characterize the nature of the responses and could represent an interesting future direction.

Conclusion

In a coincidence timing task with a moving object, overt responses across speeds can be aligned in time or in space, with the former leading to enhanced performance. We present a model that samples and updates the target's position by combining positional and speed information optimally according to their uncertainties. This model shows how a common spatial mechanism could account for both temporal and spatial patterns responses, with one or the other pattern arising from different balances between position and speed uncertainties. Thus, performance would be better for time-invariant overt responses due to an optimal balance of the uncertainties, not because of a different underlying mechanism.

Keywords: decision-making, position-tracking, visual motion, Kalman filter, coincidence timing, timed actions

Acknowledgments

We thank Cristina de la Malla for useful comments on the manuscript. Funding was provided by the Catalan government (2017SGR-48) and grant PSI2017-83493-R (AEI/Feder, UE).

Commercial relationships: none.
Corresponding author: Joan López-Moliner.
Email: j.lopezmoliner@ub.edu.
Address: Vision and Control of Action (VISCA) Group, Department of Cognition, Development and Psychology of Education, Institut de Neurociències, Universitat de Barcelona, Barcelona, Catalonia.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, *1*: 48.

Belousov, B., Neumann, G., Rothkopf, C. A., & Peters, J. R. (2016). Catching heuristics are optimal control policies. In *Advances in neural information processing systems* (pp. 1426–1434).

Benguigui, N., & Bennett, S. J. (2010). Ocular pursuit and the estimation of time-to-contact with accelerating objects in prediction motion are controlled independently based on first-order estimates. *Experimental Brain Research*, *202*(2): 327–339.

Blohm, G., Missal, M., & Lefevre, P. (2005). Direct

evidence for a position input to the smooth pursuit system. *Journal of Neurophysiology*, *94*(1): 712–721.

Bootsma, R., & Oudejans, R. (1993). Visual information about time-to-collision between two objects. *Journal of Experimental Psychology: Human Perception & Performance*, *19*, 1041–1052.

Bootsma, R. J. & van Wieringen, P. C. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1): 21–29.

Brenner, E. & Smeets, J. B. J. (2009). Sources of variability in interceptive movements. *Experimental Brain Research*, *195*(1): 117–133.

Brenner, E. & Smeets, J. B. J. (2015). How people achieve their amazing temporal precision in interception. *Journal of Vision*, *15*(3):8, 1–21, https://doi.org/10.1167/15.3.8. [PubMed] [Article]

Brenner, E., van Beers, R. J., Rotman, G., & Smeets, J. B. J. (2006). The role of uncertainty in the systematic spatial mislocalization of moving objects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 811–825.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research*, *41*, 1891–1899.

Cámara, C., de la Malla, C., López-Moliner, J., & Brenner, E. (2018). Eye movements in interception with delayed visual feedback. *Experimental Brain Research*, *236*(7): 1837–1847.

Chang, C.-J. & Jazayeri, M. (2018). Integration of speed and time for estimating time to contact. *Proceedings of the National Academy of Sciences*, *115*(12): E2879–E2887.

de la Malla, C. & López-Moliner, J. (2010). Detection of radial motion depends on spatial displacement. *Vision Research*, *50*, 1035–1040.

de la Malla, C., Smeets, J. B., & Brenner, E. (2017). Potential systematic interception errors are avoided when tracking the target with one's eyes. *Scientific Reports*, 7.

de Lussanet, M. H., Smeets, J. B., & Brenner, E. (2001). The effect of expectations on hitting moving targets: Influence of the preceding target's speed. *Experimental Brain Research*, *137*(2): 246–248.

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving gabors. *Vision Research*, *31*(9), 1619–1626.

Gepshtein, S., Seydell, A., & Trommershäuser, J. (2007). Optimality of human movement under natural variations of visual–motor uncertainty. *Journal of Vision*, *7*(5):13, 1–18, https://doi.org/10.1167/7.5.13. [PubMed] [Article]

Goettker, A., Braun, D. I., Schütz, A. C., & Gegenfurtner, K. R. (2018). Execution of saccadic eye movements affects speed perception. *Proceedings of the National Academy of Sciences*, 201704799.

Hudson, T. E., Wolfe, U., & Maloney, L. T. (2012). Speeded reaching movements around invisible obstacles. *PLoS Computational Biology*, *8*(9), e1002676.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, *82*(1), 35–45.

Keil, M. S. & López-Moliner, J. (2012). Unifying time to contact estimation and collision avoidance across species. *PLoS Computational Biology*, *8*(8): e1002625.

Krekelberg, B., & Lappe, M. (1999). Temporal recruitment along the trajectory of moving objects and the perception of position. *Vision Research*, *39*, 2669–2679.

Krekelberg, B., & Lappe, M. (2000). The position of moving objects. *Science*, *289*, 1107a.

Kwon, O.-S. & Knill, D. C. (2013). The brain uses adaptive internal models of scene statistics for sensori- motor estimation and planning. *Proceedings of the National Academy of Sciences*, *110*(11): E1064–E1073.

Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences*, *112*(26): 8142–8147.

Laurent, M., Montagne, G., & Savelsbergh, G. J. (1994). The control and coordination of one-handed catching: The effect of temporal constraints. *Experimental Brain Research*, *101*(2): 314–322.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*, 437–459.

Linares, D., Holcombe, A. O., & White, A. L. (2009). Where is the moving object now? Judgments of instantaneous position show poor temporal precision (sd = 70 ms). *Journal of Vision*, *9*(13):9, 1–14, https://doi.org/10.1167/9.13.9. [PubMed] [Article]

Linares, D., López-Moliner, J., & Johnston, A. (2007). Motion signal and the perceived positions of moving objects. *Journal of Vision*, *7*(7):1, 1–7, https://doi.org/10.1167/7.7.1. [PubMed] [Article]

López-Moliner, J., & Bonnet, C. (2002). Speed of response initiation in a time-to-contact discrimina-

tion task reflects the use of η. *Vision Research*, *42*, 2419–2430.

López-Moliner, J., Field, D. T., & Wann, J. P. (2007). Interceptive timing: Prior knowledge matters. *Journal of Vision*, *7*(13):11, 1–8, https://doi.org/10.1167/7.13.11. [PubMed] [Article]

López-Moliner, J., & Keil, M. (2012). People favour imperfect catching by assuming a stable world. *PLoS One*, *7*(4), 1–8.

Mamassian, P. (2008). Overconfidence in an objective anticipatory motor task. *Psychological Science*, *19*(6), 601–606.

Marinovic, W., Plooy, A. M., & Tresilian, J. R. (2010). The effect of priming on interceptive actions. *Acta Psychologica*, *135*: 30–37.

Maus, G. W., Fischer, J., & Whitney, D. (2013). Motion-dependent representation of space in area mt+. *Neuron*, *78*(3), 554–562.

McLeod, P., McLaughlin, C., & Nimmo-Smith, I. (1985). Information encapsulation and automaticity: Evidence from the visual control of finely timed actions. *Attention and Performance XI,* (pp. 391–406).

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familar position cues. *Vision Research*, *21*, 427–433.

Neri, P., Morrone, M. C., & Burr, D. C. (1998, October 29). Seeing biological motion. *Nature*, *395*(6705), 894–896.

Neyedli, H. F., & Welsh, T. N. (2013). People are better at maximizing expected gain in a manual aiming task with rapidly changing probabilities than with rapidly changing payoffs. *Journal of Neurophysiology*, *111*(5), 1016–1026.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*, 256–257.

Ota, K., Shinya, M., & Kudo, K. (2015). Motor planning under temporal uncertainty is suboptimal when the gain function is asymmetric. *Frontiers in Computational Neuroscience*, *9*, 88.

O'Brien, M. K., & Ahmed, A. A. (2013). Does risk-sensitivity transfer across movements? *Journal of Neurophysiology*, *109*(7), 1866–1875.

Regan, D., & Hamstra, S. J. (1993). Dissociation of discrimination thresholds for time to contact and for rate of angular expansion. *Vision Research*, *33*, 447–462.

Schlag, J., & Schlag-Rey, M. (2002). Through the eye, slowly: Delays and localization errors in the visual system. *Nature Reviews Neuroscience*, *3*, 191–200.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research*, *38*, 3569–3582.

Smith, M., Flach, J., Dittman, S., & Stanard, T. (2001). Monocular optical constraints on collision control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 395–410.

Székely, G. J. & Rizzo, M. L. (2004). Testing for equal distributions in high dimension. *InterStat*, *5*(16.10).

Székely, G. J. & Rizzo, M. L. (2013). Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, *143*(8): 1249–1272.

Tresilian, J. R. (1994). Approximate information sources and perceptual variables in interceptive timing. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 154–173.

Tresilian, J. R. (1995). Perceptual and cognitive processes in time-to-contact estimation: Analysis of prediction-motion and relative judgement tasks. *Perception & Psychophysics*, *57*, 231–245.

Tresilian, J. R. (1999). Visually timed action: Time-out for 'tau'? *Trends in Cognitive Science*, *3*, 301–309.

Tresilian, J. R. (1999b). Analysis of recent empirical challenges to an account of interceptive timing. *Perception & Psychophysics*, *61*(3): 515–528.

Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *Journal of Neuroscience*, *25*(31), 7169–7178.

Trommershäuser, J., Landy, M. S., & Maloney, L. T. (2006). Humans rapidly estimate expected gain in movement planning. *Psychological Science*, *17*, 981–988.

Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America A*, *20*(7), 1419–1433.

Turano, K. A. & Heidenreich, S. M. (1999). Eye movements affect the perceived speed of visual motion. *Vision Research*, *39*(6): 1177–1187.

Wallace, S. A., Stevenson, E., Weeks, D. L., & Kelso, J. S. (1992). The perceptual guidance of grasping a moving object. *Human Movement Science*, *11*(6): 691–715.

Wann, J. (1996). Anticipating arrival: Is the tau-margin a specious theory? *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1031–1048.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in Cognitive Science*, *6*, 211–216.