



Flexible prediction of opponent motion with internal representation in interception behavior

Kazushi Tsutsui^{1,2} · Keisuke Fujii^{1,3,4} · Kazutoshi Kudo^{2,5} · Kazuya Takeda⁶

Received: 29 April 2021 / Accepted: 30 July 2021 / Published online: 11 August 2021
© The Author(s) 2021

Abstract

Skilled interception behavior often relies on accurate predictions of external objects because of a large delay in our sensorimotor systems. To deal with the sensorimotor delay, the brain predicts future states of the target based on the current state available, but it is still debated whether internal representations acquired from prior experience are used as well. Here we estimated the predictive manner by analyzing the response behavior of a pursuer to a sudden directional change of the evasive target, providing strong evidence that prediction of target motion by the pursuer was incompatible with a linear extrapolation based solely on the current state of the target. Moreover, using neural network models, we validated that nonlinear extrapolation as estimated was computationally feasible and useful even against unknown opponents. These results support the use of internal representations in predicting target motion, suggesting the usefulness and versatility of predicting external object motion through internal representations.

Keywords Prediction · Extrapolation · Interception · Pursuit · Navigation · Internal representation

1 Introduction

Skilled motor behavior often relies on accurate predictions of external objects and environments (Wolpert and Flanagan 2001; Yarrow et al. 2009). For example, to hit a moving ball with a bat or capture an evasive opponent, it is essential to accurately predict target motion as well as own motion (Land and McLeod 2000; Brault et al. 2012; Fujii et al.

2014). In such cases, it would be easy to intercept a stationary or slow moving target, but may be difficult to intercept a fast moving target because there is a large delay in our sensorimotor systems. Sensorimotor delay, which is associated with receptor transduction, neural conduction, central processing and muscle activation, is inevitable in animals, and can be several hundreds of milliseconds in human interception behavior (Smeets and Brenner 1994; Franklin and Wolpert 2011). As a result, for the successful interception of fast moving targets, we would need to compensate for this sensorimotor delay by a prediction of future states (e.g., position and velocity) based on the current state available. The predictive mechanism to deal with sensorimotor delay is a prominent problem in interception behavior, and numerous studies have been done on predictions of target motion (Hayhoe 2017; Brenner and Smeets 2018; De la Malla et al. 2019; Fiehler et al. 2019). However, the understanding of target motion prediction is rather limited because these have predominantly studied target objects that move in a predictable manner, such as at a constant velocity (Brenner and Smeets 1996, 2007, 2009, 2015a, b; Brenner et al. 1998, 2013; Brouwer et al. 2000, 2002; De Lussanet et al. 2001) or accelerated by gravity (Lacquaniti and Maioli 1989; Zago et al. 2004, 2009, 2010; Senot et al. 2005, 2012; Zago and Lacquaniti 2005; López-Moliner et al. 2010; López-Moliner

Communicated by Benjamin Lindner.

✉ Kazushi Tsutsui
k.tsutsui6@gmail.com

¹ Graduate School of Informatics, Nagoya University, Nagoya, Japan

² Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan

³ RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

⁴ PRESTO, Japan Science and Technology Agency, Tokyo, Japan

⁵ Graduate School of Interdisciplinary Information Studies, The University of Tokyo, Tokyo, Japan

⁶ Institutes of Innovation for Future Society, Nagoya University, Nagoya, Japan

and Brenner 2016); it remains an open question how we predict interactive targets such as evasive opponents, which appear to be less predictable.

When we try to intercept an opponent, there are two basic manners that the brain may use to predict the future state of the opponent. One is linear extrapolation, and the other is nonlinear extrapolation. The first predictive manner has generally been assumed to estimate a target motion (or trajectory) based only on the current sensory information, namely position and velocity, of the target. That is, in this manner, the pursuer predicts that the target would move straight ahead regardless of the situation, in each instant. This linear extrapolation, which assumes a simple mechanism to compensate for the sensorimotor delay, is consistent with experimental observations in both humans (Engel et al. 1999; Engel and Soechting 2000) and non-humans (Borghuis and Leonardo 2015). The alternative manner has been assumed to estimate target motion based on internal representations in addition to the current information. That is, target motion is estimated through a mapping between the current state and the future state. Such transformations (or representations) are termed internal models and are thought to be acquired through prior experience. This manner allows us to make nonlinear extrapolation with our perceptible information and is supported by the fact that we can successfully catch a falling ball (Zago et al. 2004; López-Moliner and Brenner 2016). Specifically, even though the human visual system is poor at perceiving acceleration, the brain can accurately predict the motion of a ball accelerated by gravity. Although the use of internal models in predicting target motion is still controversial (Baurès et al. 2007; Zago et al. 2008), the idea is attractive in that it has the potential to predict target motion with greater accuracy by nonlinear extrapolation. However, it is unclear whether nonlinear extrapolation with an internal model is used for predicting the motion of opponents, who seem to have less stable rules of motion than free-falling objects, and, if so, whether it can accurately predict the opponent motion.

To address these questions, we conducted an experiment in which participants played a one-on-one chase and escape task on a screen with joystick controllers. Three experimental conditions for the width of the pitch (narrow, square, and wide) were studied to examine the situational dependence of the predictive manners. We analyzed the response behavior of the pursuer (defender) to a sudden directional change of the target (attacker) to estimate the predictive manner adopted by the pursuer, providing strong evidence that the pursuer would make a nonlinear extrapolation of the opponent motion. Moreover, we validated the feasibility and effectiveness of nonlinear extrapolations using neural network models which learn the mapping between the current state and the future state from the experimental data. Our results suggest the usefulness and versatility of the

prediction of external objects through internal representations, and provide an insight into the predictability of others' behavior.

2 Methods

2.1 Participants

Twelve males participated in the experiment (aged 22–31, mean = 25.9, s.d. = 3.0). All participants were right-handed, had normal or corrected-to-normal vision, had some experience in amateur sports, and were naïve to the purpose of the study. This study was approved by the Ethics Committee of the University of Tokyo of Arts and Sciences. Informed consent was obtained from each participant before the experiments. Participants were recruited in pairs and every member of each pair took in turn the roles of both attacker and defender. They each received 1,000 yen per hour as a reward.

2.2 Apparatus and stimuli

Participants were seated in a chair, and they operated the joystick of an Xbox One controller that could tilt freely in any direction to control a disk on the screen. The stimuli were presented on a 27-inch monitor (ASUS SWIFT PG278Q) at a refresh rate of 120 Hz. A black rectangle surrounding the disks was defined as the play area, or "pitch." The width of the pitch was 7.5, 15.0, 30.0 cm in the narrow, square, and wide condition, respectively, with a consistent height of 15.0 cm. The velocity of each disk on the screen was determined by the degree of inclination of the joystick on their respective controllers. The maximum speed of both the attacker and the defender was set to 5.5 cm per second. The diameter of each disk was 1.0 cm, and the central position of each disk on the screen during the trials was recorded at 120 Hz on a computer (MacBook Pro) with Psychtoolbox version 3.0. The viewing distance of the participants was about 50 cm, and a partition prevented direct viewing of the hands or controller of the other player.

2.3 Procedure and design

Each participant controlled either a red disk representing an attacker or a blue disk representing a defender on the screen (Fig. 1). The participant controlling the attacker was asked to get past the defender and reach the end line (yellow line) behind the defender (Fig. 1, lower left panel), whereas the participant controlling the defender was asked to catch the attacker without him reaching the end line (Fig. 1, lower right panel). "Catch" was defined as contact between the outer disk edges. If the attacker left the boundaries of the pitch (black rectangle), the trial was deemed a successful

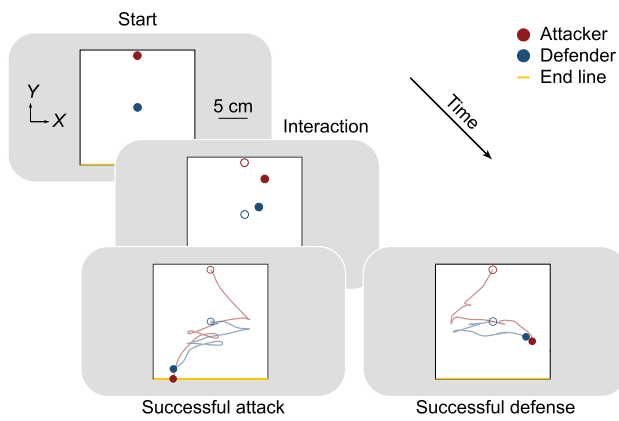


Fig. 1 Experimental setup. Illustration of the experimental task. Participants ($n = 12$) controlled either an attacker (red disk) or a defender (blue disk) on a screen using the joystick of a controller. The initial location of the attacker was the upper middle (red circle) and that of the defender was the middle (blue circle) of the pitch (upper panel). The participant controlling the attacker was asked to move past the defender and reach the end line (lower left panel), whereas the participant controlling the defender was asked to “catch” (contact) the attacker without him reaching the end line (lower right panel). If the attacker moved out of pitch boundaries (black rectangle), the trial was deemed a successful defense

defense. The start trial position of the attacker was in the upper middle of the pitch (red circle), while that of the defender was in the center of the pitch (blue circle). The experimental task began with a start cue. A high-pitched beep sounded as feedback to a successful attack. Conversely, a low-pitched beep sounded after a successful defense. The number of successful attacks was indicated at the end of each block. The experimental block consisted of 50 trials, with a warm-up of 10 trials to get used to the task. There were three experimental conditions (narrow, square, and wide), and each participant played one block in turn on both the attacker and the defender under each experimental condition. In total, there were 60 warm-up trials and 300 experimental trials for each participant (or each pair). The order of the experimental conditions was counterbalanced across pairs.

2.4 Behavioral analysis

We recorded the onscreen X and Y positions of the attacker and defender. All behavioral analyses, except for the evaluation of unpredictability of target motion (Fig. 2b), were performed using data recorded at 120 Hz. Because we thought it would be reasonable to use data with a temporal resolution closer to human perception in evaluating the unpredictability of the target motion, we used the downsampled data in this analysis. Specifically, we first downsampled the recorded data to 20 Hz based on the previous studies (Pöppel 1997; Mrotek and Soechting 2007a, b) and then, using the downsampled data, we calculated the difference in the target

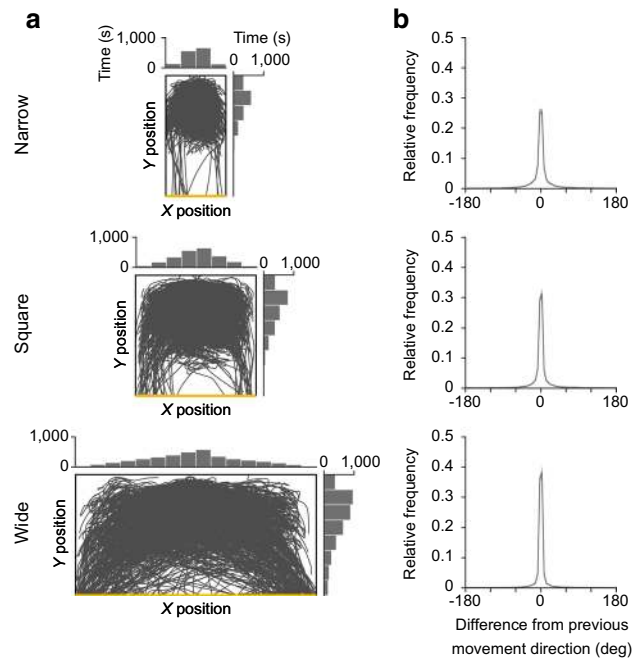


Fig. 2 Characteristics of target motion. **a** Attacker paths with histograms of X and Y positions, in all trials ($n = 600$) for each of the experimental conditions (narrow, square, and wide). Bin width was set to divide the pitch into 4, 8, and 16 columns for the three conditions, respectively. **b** Relative frequency distribution of the difference in attacker moving direction between at time t and $t + 1$ ($\Delta t = 50$ ms) for each of the three experimental conditions. Bin width was set to 5 degrees. To quantify the unpredictability of attacker motion, we calculated the entropy. One-way repeated-measures ANOVA with the Holm–Bonferroni method was used ($F_{(2, 22)} = 91.152, P < 0.001, \eta^2 = 0.800$; narrow vs. square: $t_{11} = 5.835, P < 0.001$; square versus wide: $t_{11} = 8.745, P < 0.001$; narrow versus wide: $t_{11} = 11.533, P < 0.001$)

(attacker) moving direction between time t and time $t + 1$ ($\Delta t = 50$ ms) and the entropy (see Supplementary Fig. 1). The following behavioral analyses are performed using data recorded at 120 Hz. Directional change was defined as velocity in the X direction crossing zero, and response time as the temporal difference between the directional bin changes of the attacker and the defender, distinguishing between positive and negative X velocities. The response time might be affected by the movement directions of both attacker and defender, but since our interest was in the approximate value, we only focused on the time difference for simplicity. We limited the range of response times from 0 to 500 ms, and removed any response times longer than 500 ms from the analyses to exclude responses where the defender had given up trying to catch the attacker. A short latency response was defined as the response less than 150 ms based on the results of the simple reaction task (see Supplementary Figs. 2 and 3). To calculate the values of each variable within each horizontal position on the pitch, we divided the pitch into 4, 8 and 16 columns for the three experimental conditions,

respectively. For the column that containing missing values, such as because the participants did not go to that position, the mean and s.e.m. were calculated excluding the missing values, which are shown by dotted circles and lines in each figure. The frequency of directional change was defined as the average number of occurrences per second. Specifically, the frequency was calculated by dividing the number of directional changes by the time spent in each X column. Note that we focused mainly on the direction of movement in the behavioral analysis, because both attackers and defenders were moving at almost maximum speed most of the time; the proportion of movement speed that exceeded 90% of the maximum speed was more than 90% in all conditions for both attackers and defenders.

2.5 Computational model

Because we thought it would be reasonable to use data with a temporal resolution closer to human perception in modeling human prediction, we also used the data downsampled to 20 Hz in the analyses of the computational model. At every time t , the neural network models receives an input vector \mathbf{x}_t and emits a hidden state vector from the last layer \mathbf{h}_t^l that parameterizes a predictive distribution of the target position $\hat{\mathbf{y}}_t$ at next time-step $t + 1$ ($\Delta t = 50$ ms) relative to the current position. To correspond to the available sensory information of the pursuer (or defender) as assumed in previous research, the input vector $\mathbf{x}_t \in \mathbb{R}^6$ in this study is composed of position and velocity information. Specifically, this consisted of the velocity vectors of the attacker $\mathbf{v}_t^a = (v_X^a, v_Y^a)_t$ and defender $\mathbf{v}_t^d = (v_X^d, v_Y^d)_t$ and a range vector $\mathbf{r}_t = (r_X, r_Y)_t$, which is defined as a vector pointing from the position of the pursuer to that of the target. Our three neural network models contained three weight layers. The dimensions of the hidden state vectors of the first layer $\mathbf{h}_t^1 \in \mathbb{R}^{64}$, second layer $\mathbf{h}_t^2 \in \mathbb{R}^{128}$, and last layer $\mathbf{h}_t^l \in \mathbb{R}^5$ were determined according to previous research (Alahi et al. 2016). The hidden state vector of the last layer \mathbf{h}_t^l was used to generate a bivariate Gaussian distribution parameterized by the mean $\boldsymbol{\mu}_t = (\mu_X, \mu_Y)_t$, standard deviation $\boldsymbol{\sigma}_t = (\sigma_X, \sigma_Y)_t$, and correlation coefficient ρ_t following the previous researches (Graves 2013; Alahi et al. 2016). The predicted distribution $\hat{\mathbf{y}}_t$ at time t is given by $\hat{\mathbf{y}}_t \sim \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\sigma}_t, \rho_t)$.

In our neural network models, the input and hidden layers differed among the networks, while the output layer was common. In the LN model, all layers were composed of the fully connected layer without nonlinearity,

$$\mathbf{h}_t = \mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h$$

where \mathbf{W}_{xh} , and \mathbf{b}_h denote the input-to-hidden weight matrix and the bias vector, respectively. In the NN model, only the output layer is the fully connected layer without nonlinearity,

and the other layers are composed of the fully connected layers with nonlinearity,

$$\mathbf{h}_t = \varphi(\mathbf{W}_{xh}\mathbf{x}_t + \mathbf{b}_h)$$

where $\varphi(x) = \max(0, x)$ is the rectified linear unit (ReLU) for nonlinearity (Glorot et al. 2011). In the recurrent neural network (RNN) model, the input, output, and hidden layers are, respectively, the fully connected layer without nonlinearity, that with nonlinearity, and a recurrently connected layer,

$$\mathbf{h}_t = \tanh(\mathbf{W}_{xh}\mathbf{x}_t + \mathbf{W}_{hh}\mathbf{h}_{t-1} + \mathbf{b}_h)$$

where \mathbf{W}_{hh} is the hidden-to-hidden (or recurrent) weight matrix and \mathbf{h}_{t-1} is the hidden state vector at the previous time-step $t - 1$.

The LSTM model was designed to be better at storing and accessing information than standard RNNs, and the hidden layer of the RNN model is replaced with an LSTM layer below,

$$\begin{aligned} \mathbf{i}_t &= \sigma(\mathbf{W}_{xi}\mathbf{x}_t + \mathbf{W}_{hi}\mathbf{h}_{t-1} + \mathbf{b}_i) \\ \mathbf{f}_t &= \sigma(\mathbf{W}_{xf}\mathbf{x}_t + \mathbf{W}_{hf}\mathbf{h}_{t-1} + \mathbf{b}_f) \\ \mathbf{o}_t &= \sigma(\mathbf{W}_{xo}\mathbf{x}_t + \mathbf{W}_{ho}\mathbf{h}_{t-1} + \mathbf{b}_o) \\ \mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tanh(\mathbf{W}_{xc}\mathbf{x}_t + \mathbf{W}_{hc}\mathbf{h}_{t-1} + \mathbf{b}_c) \\ \mathbf{h}_t &= \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \end{aligned}$$

where $\sigma(x) = 1/(1 + \exp(-x))$ is the logistic sigmoid function, \mathbf{i} , \mathbf{f} , \mathbf{o} , \mathbf{c} , and \mathbf{h} are the input gate, forget gate, output gate, memory cell, and hidden state activation vectors, respectively, at time-step t . $\mathbf{h}_0 = \mathbf{c}_0 = 0$. The \mathbf{W} terms denote weight matrices, the \mathbf{b} terms are biases, and \odot is the Hadamard (element-wise) product. The deep neural network models (DNN, DRNN, and DLSTM) had two hidden layers each.

The neural network models were trained to minimize the loss $\mathcal{L} = -\sum_{t=1}^T \log \mathbb{P}(\mathbf{y}_t | \mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\sigma}_t, \rho_t))$ where $\mathbf{y}_t (= \mathbf{x}_{t+1})$ denotes the actual target position at next time-step $t + 1$. Network parameters were iteratively optimized via stochastic gradient descent with the Adam optimizer (Kingma and Ba 2015). The learning rate and batch size was 0.0003 and 16, respectively, in all neural network models and experimental conditions. These hyper-parameters were selected using a grid search on pre-experimental data (Supplementary Table 1) to make full use of the experimental data. The networks were trained for the experimental data of 11 participants (550 trials) and tested on the experimental data of the other one participant (50 trials) in each model and condition; that is, model performance was evaluated by “leave-one participant-out cross-validation.” The average number of training data (time-steps) was 27,428 (range: 26,465–28,551), and that of the test data was 2493 (range: 1370–3456).

To evaluate model performance using the test data, μ_t was used for the predicted coordinates \hat{y}_t in the one time-step prediction and compared with the actual coordinates y_t . The predicted coordinates \hat{y}_t , and the displacement to reach them, were used as model input for position and velocity of the attacker at the next time-step $t + 1$ in the sequential (or trajectory) prediction. In this case, we assumed that own state (position and velocity of the defender) could be used up to 250 ms ahead based on estimations with the internal model of own motion (Wolpert et al. 1998, 2011; Kawato 1999; Imamizu et al. 2000). In addition, in the sequential prediction, for RNN, LSTM, DRNN, and DLSTM models, a 2.5 s observational period was provided to “warm” the hidden state. The prediction and the observation for it were kept within the same trial, and never crossed between trials.

The linear (L) model predicts that the target continued to move in the same direction at a constant speed. We first calculated the target displacement from time $t - 1$ to the current time t , and added the displacement to the current position to predict the position at time $t + 1$. The curvilinear (C) model predicts that the target continues to move at the same speed and angular velocity along a circular arc. We thus calculated the target displacements from time $t - 2$ to time $t - 1$ and that from time $t - 1$ to the current time t , and then computed the angular change $\Delta\theta$ per a time-step using the displacements. When predicting target position at time $t + 1$, movement speed is equal to the magnitude of the latest displacement and movement direction is the angle of the latest displacement plus $\Delta\theta$. Consequently, for $\Delta\theta = 0$, the predictions of the linear and curvilinear models are equal.

2.6 Statistical analysis

No statistical methods were used to predetermine sample sizes, but our sample sizes were chosen based on standards in the field. All quantitative data are reported as mean \pm s.e.m. across participants. The data were analyzed using one-, two- or three-way repeated-measures analysis of variance (ANOVA), as appropriate. For these tests, Mauchly's test was used to test sphericity; if the sphericity assumption was violated, degrees of freedom were adjusted by the Greenhouse–Geisser correction. P values were adjusted by the Holm–Bonferroni method for multiple comparisons. The column containing missing values was excluded from statistical analyses (Figs. 3e, 6a). The data distribution was assumed to be normal for multiple comparisons, but this was not formally tested. Two-tailed statistical tests were used for all applicable analyses. The significance level was set at an alpha value of 0.05. The method of Holm was used to adjust the P values in multiple testing (Holm 1979). When reporting K P values for K distinct tests, the Holm method is to compare the r th smallest P value (for $r = 1, \dots, K$) among the K P values with $0.05/(K - r + 1)$, and the test result is

considered statistically significant after adjustment for the multiple tests if the r th smallest P value is less than $0.05/(K - r + 1)$. However, if the r th smallest P value is the first that exceeds $0.05/(K - r + 1)$, then the test results associated with the $(K - r + 1)$ largest P values are considered statistically nonsignificant according to the Holm method. To make the presentation simpler, we let the adjusted P value be $(K - r + 1)$ times the original P value and simply compare the adjusted P value with 0.05 to determine whether a particular test result is statistically significant after adjustment. Specific test statistics, P values, and effect sizes for the analyses are detailed in the corresponding figure legends and in Supplementary Table 2. All statistical analyses were performed using R version 4.0.2 (The R Foundation for Statistical Computing).

3 Results

Our task required participants to control either an attacker (target; red disk) or a defender (pursuer; blue disk) on a screen using the joystick of a controller (Fig. 1). The participant controlling the attacker was asked to move past the defender and reach the end line (Fig. 1, lower left panel). On the other hand, the participant controlling the defender was asked to catch the attacker before the attacker reached the end line. A “catch” was regarded as a case where the outer edges of the disks were in contact (Fig. 1, lower right panel). If the attacker moved out of the pitch bounds (black rectangle), the trial was deemed a successful defense. The velocity of each disk on screen was determined by the degree of joystick inclination on the respective controllers, and the disks had equal maximum speed (magnitude of velocity). Three experimental conditions (narrow, square, and wide) were set to examine whether the predictive manner changed for targets with different rules of motion; based on the previous research (Tsutsui et al. 2019a), we reasoned that targets would change direction more frequently in the narrow pitch condition, whereas move more linearly in the wide pitch condition. The proportion of successful defenses were 0.97, 0.91, and 0.59, respectively, with mean trial durations of 2.47, 3.99, and 6.38 s, for each experimental condition.

3.1 Characteristics of target motion

We first examined the characteristics of the target motion. The target (attacker) paths show highly varied motion (Fig. 2a). To quantify the unpredictability of target motion (or the effectiveness of a linear extrapolation) from the perspective of the pursuer (defender), we calculated the difference in movement direction of the target between at a time t and time $t + 1$ ($\Delta t = 50$ ms) for each experimental condition and calculated the entropy (Fig. 2b). As expected, the

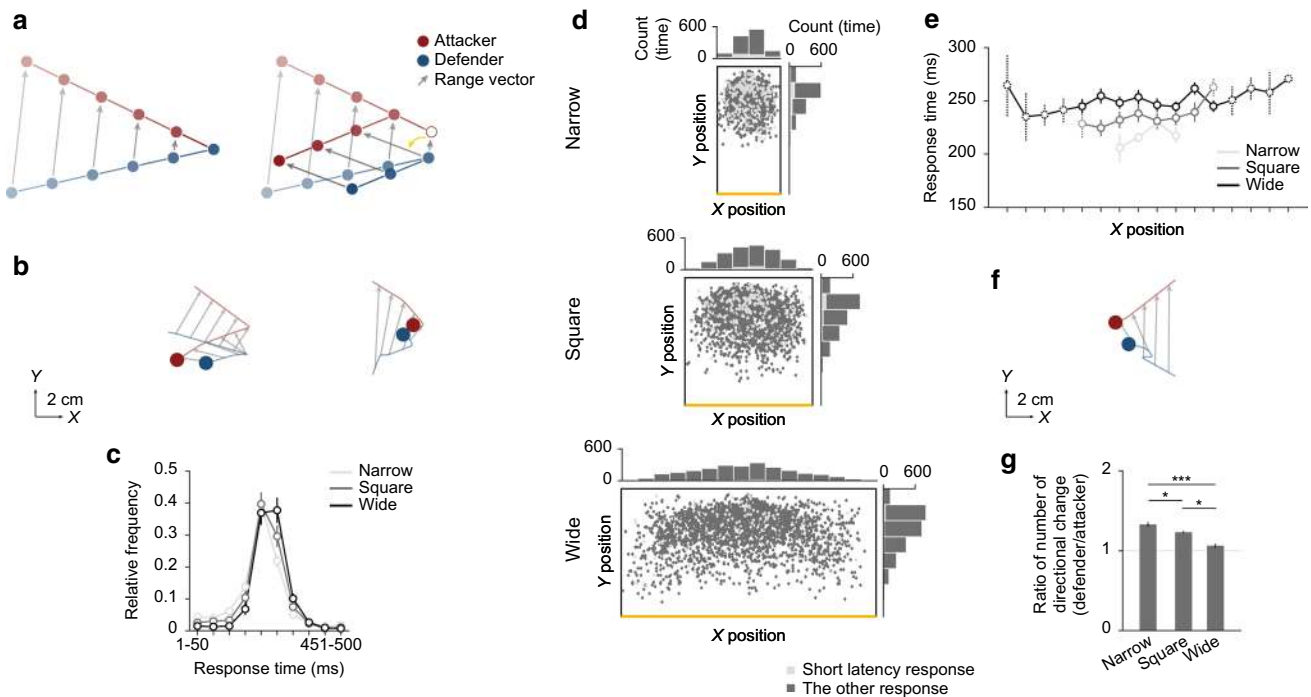


Fig. 3 Anticipatory response to directional change of target movement by pursuer. **a** Traditional description of pursuit of a target that moves straight (left panel), and a target that changes movement direction (right panel). The pursuer (blue disk) often moves along the time-optimal (shortest) path to intercept a target (red disk). The gray arrow denotes the range vector from pursuer to target at each instant. Assuming that the pursuer predicts the near-future position of the target using a linear extrapolation (red circle), the directional change of the pursuer is necessarily delayed by one step from that of the target due to sensorimotor delay of the pursuer. **b** Example trajectories of passive (left panel) and anticipatory (right panel) responses. In many cases, the pursuer reactively changed its moving direction with respect to that of the target, while the pursuer, in some case, changed its moving direction anticipatory. **c** Relative frequency distribution of the response times. Response time was defined as the temporal difference between the target and pursuer in the zero-crossing of their horizontal velocities. Bin width was set to 50 ms. We compared the proportion of the short latency response (less than 150 ms). One-way repeated-measures ANOVA with the Holm–Bonferroni method was used ($F_{(2, 22)}=17.386$, $P<0.001$, $\eta^2=0.427$; narrow versus square: $t_{11}=3.012$, $P=0.024$; square versus wide: $t_{11}=3.014$, $P=0.024$; narrow versus wide: $t_{11}=5.685$, $P<0.001$). **d** Spatial distribution and histogram of response for each of the experimental conditions. Light and dark gray dots denote the short and other latency

responses, respectively. Bin width was set to divide the pitch (horizontal position) into 4, 8 and 16 columns for the respective experimental conditions. **e** Mean response times within each horizontal position on the pitch. Bin width was set to divide the pitch into 4, 8, and 16 columns for the respective experimental conditions. Dashed circles and error bars denote bins containing missing values and that were excluded from statistical analysis. Because the number of bins was different across the conditions and we were interested in differences of response times within the pitch, we used one-way repeated-measures ANOVA with the Holm–Bonferroni method for each experimental condition ($F_{\text{narrow}(3, 33)}=1.958$, $P=0.140$, $\eta^2=0.075$; $F_{\text{square}(2.16, 23.72)}=1.064$, $P=0.365$, $\eta^2=0.033$; $F_{\text{wide}(3.72, 40.95)}=2.715$, $P=0.046$, $\eta^2=0.076$). For detailed statistics, see Supplementary Table 2. **f** Example trajectory of anticipatory response failure. In this case, the pursuer probably changed its moving direction in an incorrect anticipation of a directional change by the target. **g** Ratio of directional changes in the horizontal (X) position of the defender versus that of the attacker. One-way repeated-measures ANOVA with the Holm–Bonferroni method was used ($F_{(2, 22)}=66.279$, $P<0.001$, $\eta^2=0.738$; narrow vs. square: $t_{11}=4.124$, $P=0.017$; square vs. wide: $t_{11}=8.545$, $P<0.001$; narrow vs. wide: $t_{11}=9.966$, $P<0.001$). For all panels, quantitative data represent the mean \pm s.e.m across participants. * $P<0.05$; *** $P<0.001$

proportion of linear movement of the target decreased as the pitch narrowed. This indicates that linear extrapolation would not work effectively as the pitch narrowed.

3.2 Anticipatory response to target movement by pursuer

Then, to determine the predictive manner used by the pursuer to extrapolate the target motion, we analyzed the response behavior of the pursuer to sudden directional changes of the

target. Pursuit behaviors, from insects to mammals, have often been described as movements toward the estimated future position of the target based on its current position and velocity (Olberg et al. 2000; Fajen and Warren 2004; Ghose et al. 2006; Olberg 2012; Kane et al. 2015; Tsutsui et al. 2019b) (Fig. 3a, left panel). In other words, this description (or model) assumes a linear extrapolation of target motion by pursuers in each moment. Accordingly, the directional change of the pursuer would necessarily be one step behind that of the target owing to the sensorimotor

delay (Fig. 3a, right panel). Conversely, the temporal difference of directional changes between the target and pursuer allows us to estimate the predictive manner of the pursuer. Specifically, if a pursuer adopts linear extrapolation to estimate the future target position (or motion), the response of the pursuer to a directional change of the target should be purely reactive, whereas this would not necessarily be the case when adopting nonlinear extrapolation, as the response may include anticipatory components (e.g., extremely short latency response).

Thus, we examined the temporal differences in directional changes in the horizontal (X) position between the target and pursuer. The directional change of the pursuer was basically reactive, occurring after perceiving that of the target (Fig. 3b, left panel), but, in some case, was anticipatory, occurring before perceiving that of the target (Fig. 3b, right panel). The frequency distribution of the response times, defined as the temporal difference between the target and pursuer in the zero-crossing of their horizontal velocities, included extremely short latency responses (Fig. 3c), and the proportion of short latency responses (less than 150 ms) increased as the pitch width narrowed (0.14, 0.09, and 0.04, respectively). These results strongly indicate that a linear extrapolation of target motion was insufficient to explain the predictive manner of the pursuer. It also indicates that the pursuers may flexibly change their predictions depending on the target motion or situation.

To examine the spatial factors that result in a short latency response, we next focused on its spatial distribution. In studies on eye movement, explicit barriers are known to promote anticipatory eye movements (Kowler 1989; Kowler et al. 2014, 2019). If the pursuer's short latency responses would be distributed near the edges of the pitch, an explicit barrier may have been employed. However, the short latency responses were distributed at the middle as well as the edges of the pitch (Fig. 3d), and the mean response times for the horizontal position on the pitch were almost uniform (Fig. 3e). These results indicate that the defender made a short latency response even in situations where the information on explicit barriers would be difficult to use for prediction, suggesting that a short latency response, that is, a nonlinear extrapolation of target motion, may involve using clues from rules of target motion rather than explicit barriers. In addition, we found the cases in which the directional change of the pursuer failed in anticipation of that of the target (Fig. 3f). The ratio of the directional changes of the pursuer versus the target was greater than 1 for each experimental condition (Fig. 3g). If the defender uses a linear extrapolation (i.e., behave passively), the ratio of the directional change of defender to that of attacker should be equal to (or less than) 1. On the other hand, if the defender makes an incorrect nonlinear extrapolation (as shown in Fig. 3f), the ratio could exceed 1. This result therefore suggests that

anticipatory responses with nonlinear extrapolations were attempted under all experimental conditions, but that such nonlinear extrapolations were not always spatiotemporally appropriate.

3.3 Verification of predictability of target motion with neural network models

We thus sought to determine whether short latency responses were coincidental by examining the predictability of target motion with neural network models which predict the future position of the target through a mapping between the current state and the future state in a learning-based method (Fig. 4a). The input of the neural network models corresponded to the sensory information often used in chase (Ghose et al. 2006; Kane et al. 2015) or escape (Domenici 2002) models, namely the velocity vectors of pursuer and target, and the range vector, which is defined as a vector pointing from the position of the pursuer to that of the target. The output was the estimated position of the target, as represented by a bivariate Gaussian distribution. Note that inputs of the neural network models did not include accelerations or information on explicit barriers such as distance to the edge of the pitch. The models were trained using the error-based algorithm, a comparison between the predicted and actual consequences of the target position, using data from all but one participant, whose data were used to test the fidelity of the model prediction. To obtain insights into the important information for accurate prediction of the target motion, we computed three neural network models. The linear neural network (LN) model was composed only of linear transformations, while the nonlinear neural network (NN) model included a widely-used nonlinear transformation (Glorot et al. 2011; Lecun et al. 2015). The recurrent (nonlinear) neural network (RNN) model had a recurrent structure in addition to the nonlinear transformation. Thus, if nonlinearity is necessary for accurate prediction, the performance of the NN model should outperform that of the LN model, and if recurrence (time-series information) is necessary for accurate prediction, the performance of the RNN model should outperform that of the NN model. For comparison, we also computed two extrapolation models, linear (L) and curvilinear (C), as proposed in previous studies (Mrotek and Soechting 2007a, b; Borghuis and Leonardo 2015).

We first examined the model accuracy in predicting the movement direction of the target at the next time-step. To visualize the association between the predicted and actual angles, we showed the two-dimensional relative frequency distribution (Fig. 4b). The angles in this figure represent the movement direction of the target with respect to the pursuer. This relative movement direction was defined between -180 and 180 degrees, with 0 degree

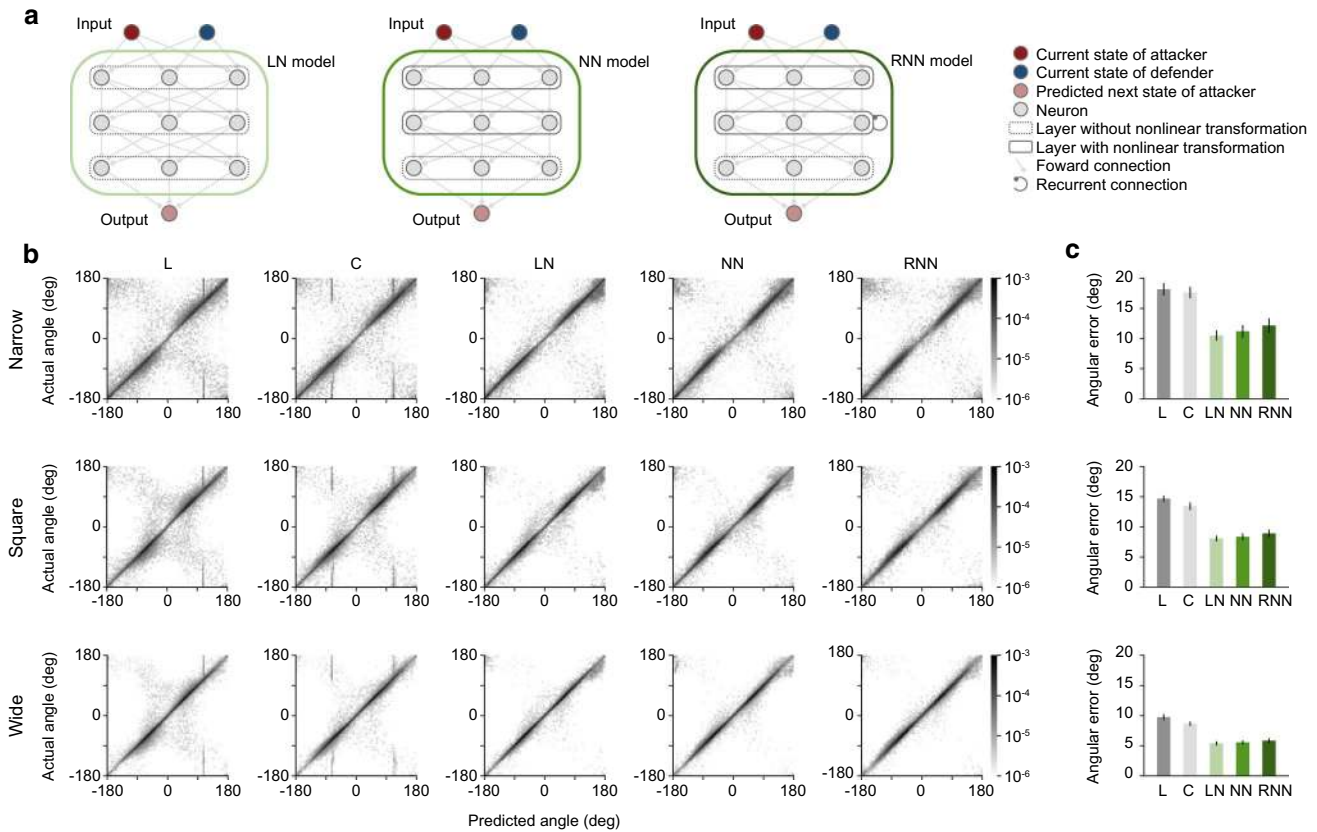


Fig. 4 Prediction of target movement direction at the next time-step with models. **a** Illustration of neural network models. The models predicted a next state of attacker (target) using the current states of attacker and defender (pursuer). The linear neural network (LN) model was composed only of linear transformations (left panel). The nonlinear neural network (NN) model included nonlinear transformations (middle panel). The recurrent neural network (RNN) model had a recurrent structure in addition to the nonlinear transformation (right panel). **b** Two-dimensional relative frequency distribution between predicted and actual directions of target movement for each of the experimental conditions (Pooled data for all partici-

pants). L, C, LN, NN, and RNN denote the linear, curvilinear, linear neural network, nonlinear neural network, and recurrent neural network model, respectively. Bin width was set to 5 degrees. **c** Angular error of the model prediction for each of the experimental conditions. Two-way repeated-measures ANOVA with the Holm–Bonferroni method was used ($F_{\text{condition}(1,29, 14,18)} = 68.184$, $P < 0.001$, $\eta^2 = 0.393$; $F_{\text{model}(1,60, 17,63)} = 151.485$, $P < 0.001$, $\eta^2 = 0.341$; $F_{\text{condition} \times \text{model}(1,91, 21,02)} = 10.900$, $P < 0.001$, $\eta^2 = 0.020$). For detailed statistics, see Supplementary Table 2. For all panels, quantitative data represent the mean \pm s.e.m across participants

indicating movement directly toward the pursuer, and positive and negative values indicating movement to the left and right sides, respectively, with respect to the pursuer. As shown in this figure, the predicted and actual angles were roughly matched in all models, while some deviation was found especially where the angles had different signs (Fig. 4b, second and fourth quadrants). Note that the deviations were lower in the neural network models (LN, NN, and RNN) than in the conventional ones (L and C). On average, the neural network models showed better agreement between the angles than the conventional ones under all conditions (Fig. 4c). These results suggest that the neural network models can more accurately predict the target motion for various situations including sudden directional change (e.g., left to right, or vice versa) (see also Supplementary Fig. 4).

3.4 Flexible and accurate longer-term prediction by neural network models

Given that the sensorimotor delay during the pursuit was about 250 ms (see Fig. 3c), it would be desirable to be able to predict 250 ms ahead to spatiotemporally match our own motion with the target motion accurately for successful interception. Thus, we then examined the model performance of sequential prediction for the target motion by testing the prediction accuracy up to 250 ms ahead in each model (Fig. 5a). In this analysis, the estimated target state (position and velocity) at time t was used sequentially as input for the prediction at next time-step $t + 1$, up to 250 ms (5 time-steps) ahead. Representative examples show that the neural network models were able to accurately predict a variety of trajectories, including the

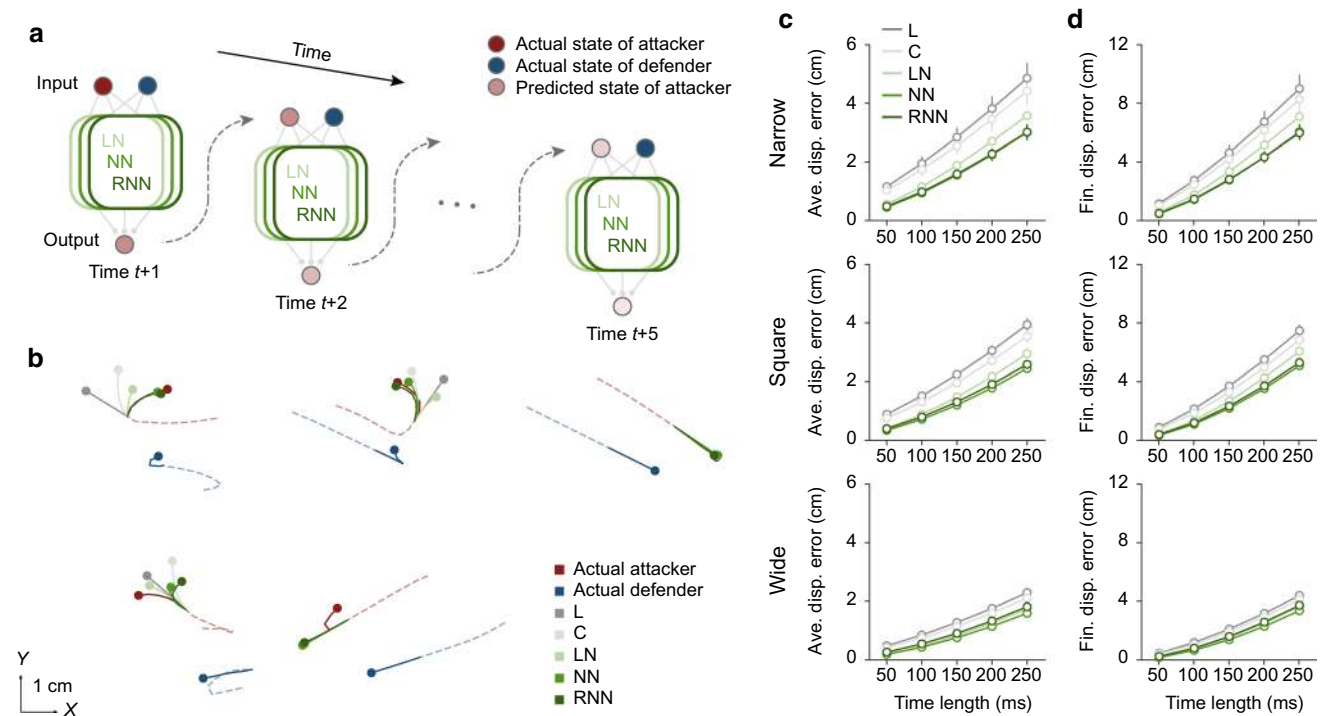


Fig. 5 Sequential prediction of target motion with models. **a** Illustration of sequential prediction by neural network models. The predicted state of the attacker (target) was used as a part of model input at the next time-step, and the prediction was made sequentially up to 5 steps (250 ms) ahead. Assuming that the own state of the defender (pursuer) could be accurately estimated by the internal model of own motion, we used the actual state of the defender as a part of the model input for sequential prediction. The same procedure was used for all neural network models. **b** Examples of predicted and actual trajectories. L, C, LN, NN, and RNN denote the linear, curvilinear, linear neural network, nonlinear neural network, and recurrent nonlinear neural network models, respectively. Red and blue lines show the actual trajectory of attacker and defender, respectively. The disks denote the end points of the predicted and actual trajectories. For ease of visibility, the trajectories of attacker and defender from the 10 time-steps (500 ms) before, to the time of prediction start, are shown by dashed red and blue lines. **c** Average displacement error of the sequential model prediction up to 5 time-steps (250 ms) ahead for each of the experimental conditions. Three-way repeated-measures ANOVA with the

Holm–Bonferroni method was used ($F_{\text{condition}(1.11, 12.26)}=21.006, P<0.001, \eta^2=0.144; F_{\text{model}(1.18, 12.93)}=76.106, P<0.001, \eta^2=0.0952; F_{\text{time length}(1.00, 11.04)}=471.038, P<0.001, \eta^2=0.511; F_{\text{condition}\times\text{model}(1.30, 14.29)}=8.486, P=0.008, \eta^2=0.015; F_{\text{model}\times\text{time length}(1.39, 15.33)}=44.454, P<0.001, \eta^2=0.009; F_{\text{condition}\times\text{time length}(1.12, 12.35)}=20.470, P<0.001, \eta^2=0.032; F_{\text{condition}\times\text{model}\times\text{time length}(1.48, 16.25)}=6.433, P=0.014, \eta^2=0.002$). For detailed statistics, see Supplementary Table 2. **(d)** Final displacement error of the sequential model prediction up to 5 time-steps (250 ms) ahead. Three-way repeated-measures ANOVA with the Holm–Bonferroni method was used ($F_{\text{condition}(1.12, 12.28)}=20.660, P<0.001, \eta^2=0.105; F_{\text{model}(1.24, 13.61)}=62.994, P<0.0001, \eta^2=0.049; F_{\text{time length}(1.01, 11.07)}=505.531, P<0.001, \eta^2=0.637; F_{\text{condition}\times\text{model}(1.34, 14.70)}=7.689, P=0.010, \eta^2=0.008; F_{\text{model}\times\text{time length}(1.44, 15.87)}=38.561, P<0.001, \eta^2=0.009; F_{\text{condition}\times\text{time length}(1.13, 12.39)}=19.900, P<0.001, \eta^2=0.036; F_{\text{condition}\times\text{model}\times\text{time length}(1.48, 16.33)}=5.915, P=0.017, \eta^2=0.002$). For detailed statistics, see Supplementary Table 2. For all panels, quantitative data represent the mean \pm s.e.m across participants

straight, gentle curve, and sharp curve phases (Fig. 5b, upper panels). Even though the predictions were occasionally incorrect (Fig. 5b, lower panels), on average, the neural network models made more accurate predictions than the conventional ones for both the average and final displacement errors (Fig. 5c, d). These results indicate that the neural network models also worked well in predicting target motion over a longer period of time and therefore would be of practical usefulness.

3.5 Ineffectiveness of the recurrent structure

Although the RNN model appeared to contain richer information due to its recurrent structure, its predictive performance was similar to or slightly lower than that of the NN model. While we also examined whether predictive performance improved using the long short-term memory (LSTM) model, which can hold information for longer time periods, the prediction accuracy was almost the same

(Supplementary Figs. 5, 6, and 7). To clarify why the recurrent structure did not lead to a performance improvement in the sequential prediction, we investigated the properties of target motion in terms of directional changes in the horizontal position. Based on the findings of the previous research (Tsutsui et al. 2019a), we focused on frequency distributions in spatial and temporal aspects regarding the change in the horizontal direction of the target. First, we investigated the spatial bias in the frequency per time of the directional change of the target and found that it was almost uniform within the pitch under all experimental conditions (Fig. 6a). Next, we investigated the relative frequency distribution of the time interval between directional changes and found that it decayed exponentially over time, particularly after a second peak at approximately 500 ms (Fig. 6b). These results indicate that directional changes in the horizontal position of the target showed the Poisson-like property known as “memorylessness” (see also Supplementary Fig. 8). In other words, whether the target changes movement direction in any moment may be little influenced by the prior process, and it suggests that this spatiotemporal property of target motion may be a reason why the recurrent structure did not lead to improvement of performance in the sequential prediction.

4 Discussion

Traditionally, pursuit behavior, from insects to mammals, has been described as movement along a local shortest path toward the estimated future position of the target based on its current position and velocity (Land and Collett 1974; Olberg et al. 2000; Fajen and Warren 2004; Ghose et al. 2006; Olberg 2012; Kane et al. 2015; Tsutsui et al. 2019b). In these studies, it is often (implicitly) assumed that the motion (or trajectory) of a target is predicted by linear

extrapolation, and under such an assumption, the pursuer (or defender) should be purely reactive to a sudden directional change of the target (or attacker). Here we have shown that, in striking contrast to these traditional descriptions, pursuers sometimes change their movement direction before perceiving (or even without occurring) a directional change of the target. Our results are consistent with a recent finding that pursuit behavior relies on predictions through target models (Mischiati et al. 2014), presenting the possibility that the predictive mechanisms that humans (or animals) use to compensate for sensorimotor delays during pursuit are more sophisticated than previously thought.

Previous studies on human interception behavior have predominantly used target objects moving in a predictable manner—at a constant velocity (Brenner and Smeets 1996, 2007, 2009, 2015a, b; Brenner et al. 1998, 2013; Brouwer et al. 2000, 2002; De Lussanet et al. 2001) or accelerated by gravity (Lacquaniti and Maioli 1989; Zago et al. 2004, 2009, 2010; Senot et al. 2005, 2012; Zago and Lacquaniti 2005; López-Moliner et al. 2010; López-Moliner and Brenner 2016). Presumably, the reason that these experimental paradigms have dominated, despite the fact that we often encounter less predictable situations in daily life or sports, is that these paradigms allow investigation under strict experimental controls. However, due to their simplicity, these paradigms may occasionally allow multiple interpretations for experimental observations. For example, in catching a falling ball, some researchers have proposed that an internal model, which allows us to extrapolate a nonlinear trajectory, is used to predict the target motion because we can catch a ball accelerated by gravity in spite of being poor at perceiving accelerations, while some others have questioned this proposal (Baurès et al. 2007). The question results from considerations that the capture of a falling ball can also be performed by continuous prediction using a linear extrapolation in each

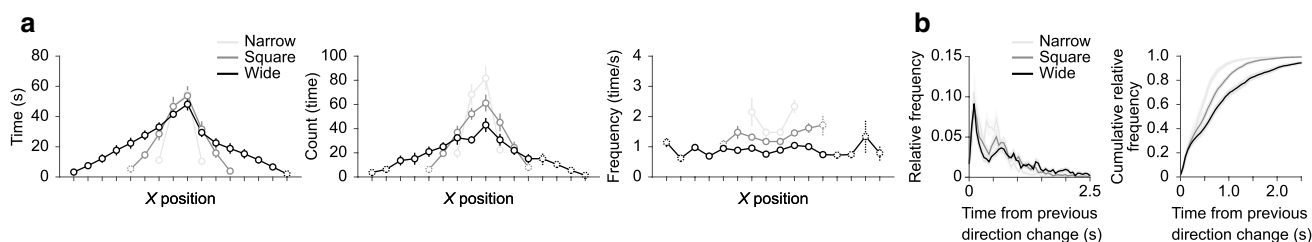


Fig. 6 Spatiotemporal property of change in the horizontal direction of the target. (a) Frequency distribution of time spent of the attacker (left), frequency distribution of directional changes in horizontal position of the attacker (middle), and frequency distribution per time of directional changes (right) within each horizontal position on the pitch. Bin width was set to divide the pitch into 4, 8, and 16 columns for the respective conditions. Dashed circles and error bars denote that the bin contained missing values. Because the number of bins was different across the conditions and we were

interested in differences of the frequency per time within the pitch, we used one-way repeated-measures ANOVA with Holm–Bonferroni method in each experimental condition ($F_{\text{narrow}(1.32, 14.52)} = 2.870$, $P = 0.104$, $\eta^2 = 0.150$; $F_{\text{square}(2.16, 23.78)} = 2.115$, $P = 0.140$, $\eta^2 = 0.142$; $F_{\text{wide}(3.66, 40.25)} = 2.159$, $P = 0.096$, $\eta^2 = 0.125$). **b** Frequency distribution (left) and cumulative frequency distribution (right) of time from previous directional change. For all panels, quantitative data represent the mean \pm s.e.m across participants

instance. Indeed, in this case, the predictions of the two predictive manners are not much different (Baurès et al. 2007). However, our experiment allowed us to distinguish between the two manners. Our results that pursuers anticipatorily changed their movement directions obviously cannot be explained by continuous linear extrapolation, and support the idea that an internal model is used to predict target motion.

Predictions of target motion through internal representations have an ability to comprehensively describe the experimental observations in various situations. We found that pursuers flexibly change their frequency of anticipating directional change of the target in response to the expectation that the target would go straight in each instance. This result implies that in situations where the target is always straight ahead, the pursuer will predict that the target moves in a straight line. In such situations, the prediction of nonlinear extrapolation is equivalent to that of linear extrapolation. Following this reasoning, a linear extrapolation conventionally considered as the basis or default (Mrotek and Soechting 2007a, b) in predicting target motion may instead be considered as the prediction in a special situation when a target moves at a constant velocity. This novel perspective may explain contradictions such as situational and individual differences in the prediction of target motion (Mrotek and Soechting 2007a, b).

On the basis of computational neuroscience studies, the existence of an internal model in the central nervous system has been established (Wolpert et al. 1998, 2011; Kawato 1999; Imamizu et al. 2000). In general, internal models are associated with predicting the motion of one's own body (e.g., arm) and tools, and can be used to maintain stability in the presence of feedback (or sensorimotor) delays when trying to make rapid movements under feedback control. On the other hand, some researchers have proposed that the notion of internal models can be extended to predicting the behavior of other persons (Wolpert et al. 2003), but there is little experimental evidence. Here, we have shown that neural network models were able to learn a flexible and accurate predictions that could be useful against unknown opponents, and these results suggest the feasibility of acquiring the “internal model of opponent motion,” especially in this type of interaction.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00422-021-00891-9>.

Acknowledgements We thank Leon Omura for useful discussion and comments on this manuscript.

Funding This work was supported by the JSPS (Grant Numbers 25242059, 17J10922, 19H04941, 20H04075, and 21H04892) and the JST (Grant Number JPMJMI18C7).

Data availability The data supporting the findings of this study are available in figshare at <https://doi.org/10.6084/m9.figshare.14405354>.

Code availability The codes are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Ethics approval This study was approved by the Ethics Committee of the University of Tokyo of Arts and Sciences.

Consent to participate Informed consent was obtained from each participant before the experiments.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alahi A, Goel K, Ramanathan V, et al (2016) Social LSTM: human trajectory prediction in crowded spaces. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, pp 961–971
- Baurès R, Benguigui N, Amorim MA, Siegler IA (2007) Intercepting free falling objects: better use Occam's razor than internalize Newton's law. *Vis Res* 47:2982–2991. <https://doi.org/10.1016/j.visres.2007.07.024>
- Borghuis BG, Leonardo A (2015) The role of motion extrapolation in amphibian prey capture. *J Neurosci* 35:15430–15441. <https://doi.org/10.1523/JNEUROSCI.3189-15.2015>
- Brault S, Bideau B, Kulpa R, Craig CM (2012) Detecting deception in movement: the case of the side-step in rugby. *PLoS ONE* 7:1–13. <https://doi.org/10.1371/journal.pone.0037494>
- Brenner E, Smeets JBJ (1996) Perception and action are based on the same visual processing. *Investig Ophthalmol Vis Sci* 37:19–31
- Brenner E, Smeets JBJ (2007) Flexibility in intercepting moving objects. *J Vis* 7:1–17. <https://doi.org/10.1167/7.5.14>
- Brenner E, Smeets JBJ (2009) Modifying one's hand's trajectory when a moving target's orientation changes. *Exp Brain Res* 196:375–383. <https://doi.org/10.1007/s00221-009-1857-7>
- Brenner E, Smeets JBJ (2015a) How people achieve their amazing temporal precision in interception. *J Vis* 15:1–21. <https://doi.org/10.1167/15.3.8>
- Brenner E, Smeets JBJ (2015b) How moving backgrounds influence interception. *PLoS ONE* 10:14–16. <https://doi.org/10.1371/journal.pone.0119903>
- Brenner E, Smeets JBJ (2018) Continuously updating one's predictions underlies successful interception. *J Neurophysiol* 120:3257–3274. <https://doi.org/10.1152/jn.00517.2018>

- Brenner E, Smeets JBJ, De Lussanet MHE (1998) Hitting moving targets. Continuous control of the acceleration of the hand on the basis of the target's velocity. *Exp Brain Res* 122:467–474. <https://doi.org/10.1007/s002210050535>
- Brenner E, Cañal-Bruland R, Van Beers RJ (2013) How the required precision influences the way we intercept a moving object. *Exp Brain Res* 230:207–218. <https://doi.org/10.1007/s00221-013-3645-7>
- Brouwer AM, Brenner E, Smeets JBJ (2000) Hitting moving objects. The dependency of hand velocity on the speed of the target. *Exp Brain Res* 133:242–248. <https://doi.org/10.1007/s002210000371>
- Brouwer AM, Brenner E, Smeets JBJ (2002) Hitting moving objects: is target speed used in guiding the hand? *Exp Brain Res* 143:198–211. <https://doi.org/10.1007/s00221-001-0980-x>
- De la Malla C, Rushton SK, Clark K et al (2019) The predictability of a target's motion influences gaze, head, and hand movements when trying to intercept it. *J Neurophysiol* 121:2416–2427. <https://doi.org/10.1152/jn.00917.2017>
- De Lussanet MHE, Smeets JBJ, Brenner E (2001) The effect of expectations on hitting moving targets: Influence of the preceding target's speed. *Exp Brain Res* 137:246–248. <https://doi.org/10.1007/s002210000607>
- Domenici P (2002) The visually mediated escape response in fish: predicting prey responsiveness and the locomotor behaviour of predators and prey. *Mar Freshw Behav Physiol* 35:87–110. <https://doi.org/10.1080/10236240290025635>
- Engel KC, Soechting JF (2000) Manual tracking in two dimensions. *J Neurophysiol* 83:3483–3496. <https://doi.org/10.1152/jn.2000.83.6.3483>
- Engel KC, Anderson JH, Soechting JF (1999) Oculomotor tracking in two dimensions. *J Neurophysiol* 81:1597–1602. <https://doi.org/10.1152/jn.1999.81.4.1597>
- Fajen BR, Warren WH (2004) Visual guidance of intercepting a moving target on foot. *Perception* 33:689–715. <https://doi.org/10.1068/p5236>
- Fiehler K, Brenner E, Spering M (2019) Prediction in goal-directed action. *J Vis* 19:1–21. <https://doi.org/10.1167/19.9.10>
- Franklin DW, Wolpert DM (2011) Computational mechanisms of sensorimotor control. *Neuron* 72:425–442. <https://doi.org/10.1016/j.neuron.2011.10.006>
- Fujii K, Shinya M, Yamashita D et al (2014) Anticipation by basketball defenders: an explanation based on the three-dimensional inverted pendulum model. *Eur J Sport Sci* 14:538–546. <https://doi.org/10.1080/17461391.2013.876104>
- Ghose K, Horiuchi TK, Krishnaprasad PS, Moss CF (2006) Echolocating bats use a nearly time-optimal strategy to intercept prey. *PLoS Biol* 4:865–873. <https://doi.org/10.1371/journal.pbio.0040108>
- Glorot X, Bordes A, Bengio Y (2011) Deep sparse rectifier neural networks. In: Proceedings of the international conference on artificial intelligence and statistics, pp 315–323
- Graves A (2013) Generating sequences with recurrent neural networks. *arXiv Prepr arXiv13080850* 1–43
- Hayhoe MM (2017) Vision and action. *Annu Rev Vis Sci* 3:389–413. <https://doi.org/10.1146/annurev-vision-102016-061437>
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* 6:65–70. <https://doi.org/10.2307/4615733>
- Imamizu H, Miyauchi S, Tamada T et al (2000) Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* 403:192–195. <https://doi.org/10.1038/35003194>
- Kane SA, Fulton AH, Rosenthal LJ (2015) When hawks attack: animal-borne video studies of goshawk pursuit and prey-evasion strategies. *J Exp Biol* 218:212–222. <https://doi.org/10.1242/jeb.108597>
- Kawato M (1999) Internal models for motor control and trajectory planning. *Curr Opin Neurobiol* 9:718–727. [https://doi.org/10.1016/S0959-4388\(99\)00028-8](https://doi.org/10.1016/S0959-4388(99)00028-8)
- Kingma DP, Ba JL (2015) Adam: a method for stochastic optimization. In: 3rd international conference on learning representations ICLR 2015—conference track proceedings, pp 1–15
- Kowler E (1989) Cognitive expectations, not habits, control anticipatory smooth oculomotor pursuit. *Vis Res* 29:1049–1057. [https://doi.org/10.1016/0042-6989\(89\)90052-7](https://doi.org/10.1016/0042-6989(89)90052-7)
- Kowler E, Aitkin CD, Ross NM et al (2014) Davida Teller award lecture 2013: the importance of prediction and anticipation in the control of smooth pursuit eye movements. *J Vis* 14:1–16. <https://doi.org/10.1167/14.5.10>
- Kowler E, Rubinstein JF, Santos EM, Wang J (2019) Predictive smooth pursuit eye movements. *Annu Rev Vis Sci* 5:223–246. <https://doi.org/10.1146/annurev-vision-091718-014901>
- Lacquaniti F, Maioli C (1989) The role of preparation in tuning anticipatory and reflex responses during catching. *J Neurosci* 9:134–148. <https://doi.org/10.1523/jneurosci.09-01-00134.1989>
- Land MF, Collett TS (1974) Chasing behaviour of houseflies (*Fannia canicularis*). *J Comp Physiol* 89:331–357. <https://doi.org/10.1007/BF00695351>
- Land MF, McLeod P (2000) From eye movements to actions: how batsmen hit the ball. *Nat Neurosci* 3:1340–1345. <https://doi.org/10.1038/81887>
- Lecun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521:436–444. <https://doi.org/10.1038/nature14539>
- López-Moliner J, Brenner E (2016) Flexible timing of eye movements when catching a ball. *J Vis* 16:1–11. <https://doi.org/10.1167/16.5.13>
- López-Moliner J, Brenner E, Louw S, Smeets JBJ (2010) Catching a gently thrown ball. *Exp Brain Res* 206:409–417. <https://doi.org/10.1007/s00221-010-2421-1>
- Mischiati M, Lin H-T, Herold P et al (2014) Internal models direct dragonfly interception steering. *Nature* 517:1–13. <https://doi.org/10.1038/nature14045>
- Mrotek LA, Soechting JF (2007a) Target interception: hand-eye coordination and strategies. *J Neurosci* 27:7297–7309. <https://doi.org/10.1523/JNEUROSCI.2046-07.2007>
- Mrotek LA, Soechting JF (2007b) Predicting curvilinear target motion through an occlusion. *Exp Brain Res* 178:99–114. <https://doi.org/10.1007/s00221-006-0717-y>
- Olberg RM (2012) Visual control of prey-capture flight in dragonflies. *Curr Opin Neurobiol* 22:267–271. <https://doi.org/10.1016/j.conb.2011.11.015>
- Olberg RM, Worthington AH, Venator KR (2000) Prey pursuit and interception in dragonflies. *J Comp Physiol A* 186:155–162. <https://doi.org/10.1007/s003590050015>
- Pöppel E (1997) A hierarchical model of temporal perception. *Trends Cogn Sci* 1:56–61. [https://doi.org/10.1016/S1364-6613\(97\)01008-5](https://doi.org/10.1016/S1364-6613(97)01008-5)
- Senot P, Zago M, Lacquaniti F, McIntyre J (2005) Anticipating the effects of gravity when intercepting moving objects: differentiating up and down based on nonvisual cues. *J Neurophysiol* 94:4471–4480. <https://doi.org/10.1152/jn.00527.2005>
- Senot P, Zago M, Le Se A et al (2012) When up is down in 0g: how gravity sensing affects the timing of interceptive actions. *J Neurosci* 32:1969–1973. <https://doi.org/10.1523/JNEUROSCI.3886-11.2012>
- Smeets JBJ, Brenner E (1994) The difference between the perception of absolute and relative motion: a reaction time study. *Vis Res* 34:191–195. [https://doi.org/10.1016/0042-6989\(94\)90331-X](https://doi.org/10.1016/0042-6989(94)90331-X)
- Tsutsui K, Shinya M, Kudo K (2019a) Spatiotemporal characteristics of an attacker's strategy to pass a defender effectively in a computer-based one-on-one task. *Sci Rep* 9:1–9. <https://doi.org/10.1038/s41598-019-54012-5>

- Tsutsui K, Shinya M, Kudo K (2019b) Human navigational strategy for intercepting an erratically moving target in chase and escape interactions. *J Mot Behav*. <https://doi.org/10.1080/00222895.2019.1692331>
- Wolpert DM, Flanagan JR (2001) Motor prediction. *Curr Biol* 11:R729–R732. [https://doi.org/10.1016/S0960-9822\(01\)00432-8](https://doi.org/10.1016/S0960-9822(01)00432-8)
- Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. *Trends Cogn Sci* 2:338–347. [https://doi.org/10.1016/S1364-6613\(98\)01221-2](https://doi.org/10.1016/S1364-6613(98)01221-2)
- Wolpert DM, Doya K, Kawato M (2003) A unifying computational framework for motor control and social interaction. *Philos Trans R Soc B Biol Sci* 358:593–602. <https://doi.org/10.1098/rstb.2002.1238>
- Wolpert DM, Diedrichsen J, Flanagan JR (2011) Principles of sensorimotor learning. *Nat Rev Neurosci* 12:739–751. <https://doi.org/10.1038/nrn3112>
- Yarrow K, Brown P, Krakauer JW (2009) Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nat Rev Neurosci* 10:585–596. <https://doi.org/10.1038/nrn2672>
- Zago M, Lacquaniti F (2005) Internal model of gravity for hand interception: parametric adaptation to zero-gravity visual targets on Earth. *J Neurophysiol* 94:1346–1357. <https://doi.org/10.1152/jn.00215.2005>
- Zago M, Bosco G, Maffei V et al (2004) Internal models of target motion: expected dynamics overrides measured kinematics in timing manual interceptions. *J Neurophysiol* 91:1620–1634. <https://doi.org/10.1152/jn.00862.2003>
- Zago M, McIntyre J, Senot P, Lacquaniti F (2008) Internal models and prediction of visual gravitational motion. *Vis Res* 48:1532–1538. <https://doi.org/10.1016/j.visres.2008.04.005>
- Zago M, McIntyre J, Senot P, Lacquaniti F (2009) Visuo-motor coordination and internal models for object interception. *Exp Brain Res* 192:571–604. <https://doi.org/10.1007/s00221-008-1691-3>
- Zago M, Iosa M, Maffei V, Lacquaniti F (2010) Extrapolation of vertical target motion through a brief visual occlusion. *Exp Brain Res* 201:365–384. <https://doi.org/10.1007/s00221-009-2041-9>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.