Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Biological kinematics drive action prediction regardless of effector appearance

Ghislain Saunier

ghislain@ikmail.com

Universidade Federal do Pará

Isabella Rolo Sarrazin

Universidade Federal do Pará

Priscila da Silva Azevedo

Universidade Federal do Rio de Janeiro

Catarine de Nazaré Moreira Saunier

Universidade Federal do Sul e Sudeste do Pará

Nicolas Gueugneau

Université Marie et Louis Pasteur

Research Article

Keywords: action prediction, biological kinematic, effector appearance, time-to-contact estimates

Posted Date: October 22nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11271681/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The ability to predict the outcome of others' actions is a fundamental skill that relies on action-perception coupling. However, our understanding of the role of factors such as kinematics and the appearance of effectors in this process is limited. The present study sought to investigate the respective contributions of kinematics and visual appearance of the effectors to action prediction. To this end, healthy participants were asked to estimate the time-to-contact between a human hand and objects exhibiting either biological or non-biological kinematics, under two visual conditions: occlusion (i.e. the final 50% of the action trajectory was occluded) and full vision. We found that prediction accuracy was higher for biological kinematics (independent of effector appearance) and was insensitive to visual condition, unlike for non-biological kinematics. Conversely, participants' responses were less variable for human biological kinematics, particularly in the occlusion condition. In contrast, responses were less variable for non-biological kinematics in the full vision condition. Furthermore, biological kinematics applied to a moving object exhibited higher variability in both visual conditions. These findings therefore reveal that biological kinematics are essential for accurate outcome prediction, whereas response variability depends on the interaction between motion kinematics and effector appearance.

Introduction

The capacity to predict the outcomes of others' actions is a prerequisite for effective social interaction and adaptive behaviour. Accurate prediction enables individuals to foresee the consequences of observed movements despite sensory delays and transient sensory impairments (Urgen and Saygin, 2020). A seminal account posits that action prediction relies on action-perception coupling, whereby observed movements are matched with their corresponding motor representations in the observer’s brain through the involvement of internal models of action (Wolpert and Flanagan, 2001).

Indeed, an inverse internal model would enable the mapping of action observation onto its motor counterpart (Erlhagen, 2003). This would permit a comparison of the kinematic features extracted in visual input with an endogenous knowledge of the observed motion kinematic (Hilt et al., 2026). Such a processing likely relies on the elaboration of an efference copy within the observer’s motor system, which would then be used to predict the outcomes of the observed action through a forward internal model (Miall, 2003). The interaction between inverse and forward internal models would permit an implicit process of motor simulation when an observer tries to predict the future states of an observed action (Grush, 2004; Jeannerod, 2006).

The proposal of an action-perception coupling involvement to action prediction is substantiated by a series of behavioural evidences (Craighero et al., 2008; de Assis Chagas et al., 2021; Graf et al., 2007; Hilt et al., 2026; Saunier et al., 2021, 2008; Stapel et al., 2012). For instance, Saunier et al. (2008; 2021) demonstrated that the estimate of the final position of a unique point-light dot depicting a complex body trajectory and occluded in its final part was dependent on the motion kinematic features. In a recent study, Saunier et al. (2023) observed that participants could accurately predict the time-to-contact

between the observed hand and an object during a reaching-to-grasping action even if the final part of the action was occluded (see also Craighero et al., 2008). Recently, Hilt et al. (2026) demonstrated, using an inference paradigm, that prediction in action observation involves a dynamic process. This process allows internal forward predictions to be updated throughout a continuous comparison between observed kinematics and the implicit knowledge of their internal counterparts. In addition to these behavioural evidences, neurophysiological studies have demonstrated that the mere observation of human actions is sufficient to engage the sensorimotor system, as well as visual areas, in order to infer its outcomes (Abreu et al., 2012; Rangel et al., 2021). The premotor cortex, the parietal lobule, the superior temporal sulcus plus the cerebellum comprise such a broad network (Calvo-Merino et al., 2006; Gazzola and Keysers, 2009; Urgen and Saygin, 2020). This network promotes a functional mapping of the visual input to the corresponding motor representations (Giorello and Sinigaglia, 2007).

Nonetheless, studies investigating the interaction between kinematic features and the appearance of the effector in order to predict the consequences of observed actions remain sparse. For instance, Stadler et al. (2012) utilised a classical point-light display of biological motion (Johansson, 1973) in a task involving predicting the coherent continuity of actions after a brief period of occlusion. The authors demonstrated that participants were more sensitive to human velocity profiles than artificial ones when solving the task (Stadler et al., 2012). However, when an ecological visual scene was employed, the appearance of the observed agents interfered with the judgements about action continuity after a transient occlusion (Saygin and Stadler, 2012). Therefore, the specific influence of appearance and kinematics on the accurate prediction of upcoming actions remains unclear. Acknowledging the interaction between kinematics and effector appearance is particularly relevant given the increasing prevalence of human interaction with animated objects in their daily life. Indeed, many domains of sensorimotor optimisation currently rely on technologies using artificial agents and/or virtual objects – e.g. sport training and sensorimotor rehabilitation (Payedimarri et al., 2022; Witte et al., 2025); as such, a better interactions comprehension of the underlying sensorimotor mechanisms will enable to improve these methodologies and their translational applications.

In the present study, we sought to disentangle the respective contributions of kinematic features and stimulus appearance using an action prediction paradigm. This paradigm involved estimating the time-to-contact (TTC) between the hand and an object when the final part of the effector's trajectory was occluded. The hand followed a biological velocity profile, whereas the moving object could be displaced according to a biological or non-biological (i.e., constant velocity) velocity profile. Thus, if action prediction mainly relies on kinematics, biological motion should produce higher accuracy of TTC prediction regardless of the effector's appearance. Otherwise, if action prediction mainly relies on the effector's appearance, the moving object should produce similar TTC prediction regardless of kinematic features. On other hand, if both factors interact (kinematic and effector's appearance), the human condition should produce the most accurate TTC predictions.

Materials and Methods

Participants

Thirty-one undergraduate students (22 females; mean age 24.1 ± 3.7 years) from the Federal University of the Pará (Brazil) participated in this study. All subjects exhibited normal or corrected visual acuity, had no history of any psychiatric or neurological impairment and were right-handed according to short online version of the Edinburg Handedness Inventory (https://www.brainmapping.org/shared/Edinburgh.php; Oldfield, 1971). The participants were naïve with regard to the objectives and hypotheses of the study. They gave their informed consent before the beginning of the experiment, which was carried out in accordance with the ethical standards of the Declaration of Helsinki (World Medical Association, 2013). The local ethical committee approved the procedures (CAAE: 64960417.8.0000.5147).

Apparatus and stimuli

The stimuli comprised four videos (resolution: 640 x 480 pixels; 30 frames/s) depicting either a left hand reaching-to-grasp an object or the object moving towards a static hand until contact was made. All stimuli were displayed on a 21” colour flat screen (black background, resolution 1024 x 768 pixels with a refresh rate of 60 Hz) connected to a PC (Presentation, Neurobehavioral Systems, Inc.). The hand motion followed a biological kinematic while the objects (a ball, a cube, a cup, or a car toy) could move in accordance with either biological or non-biological kinematic features. Accordingly, the stimuli were categorised as follows: Hand Motion (HM) - hand reaching-to-grasp an object; Biological Object (BO) - object moving towards the hand according to the biological kinematic of horizontal movement, where the acceleration time is equal to the deceleration time (i.e., the time to peak velocity ≈ 50% of the movement time; Gentili et al., 2007); and Non-Biological Object (NBO) - object moving towards the hand at constant velocity (i.e., ≈ the mean of BO velocities). For all motions, two visual conditions were displayed: an occlusion condition where the 50% of the final motion trajectory was hidden by a black rectangle, or a full vision condition of the visual scene. The stimuli were created using Adobe Premiere Pro software (Adobe, Inc). All videos were presented in sagittal view with the movement of the hand or object from left to right.

Procedure and design

Each participant sat at a comfortable viewing distance (≈ 60 cm) from the screen in a dimly lit room. Their task was to predict the time-to-contact (TTC) between the hand and the object in both visual contexts (occlusion and full vision) by pressing the space bar on the keyboard. The experiment comprised four blocks of twenty-four repetitions each (HM vs. BO and HM vs. NBO), always initiated by the two occluded blocks. The trials were pseudo-randomised within each block. One group of participants initiated the occlusion and full vision blocks with HM and BO, while the second group initiated the occlusion and full vision blocks with HM and NBO. To familiarise themselves with the experimental design, the subjects participated in a practice session. This session comprised four occluded trials (prior to the occlusion block) and four full-vision trials (before the full vision block) for

each condition (HM, BO and NBO). The objects used in the practice session were different to those used in the experimental session. Each trial began with a fixation cross appearing in the centre of the screen for a duration of 500 ms. This was followed by the presentation of the video. To indicate the effector (i.e., hand or object), a red arrow appeared above it for 500 ms prior to the initiation of motion. Each video was repeated three times within each block, for a total of twelve repetitions for each condition (4 videos x 3 trials). The participant had until 1500 ms to respond following the presentation of the video, with the subsequent trial commencing immediately thereafter (Fig. 1).

Figure 1 Experimental design. A. From top to bottom, an example of video sequence for each kinematic feature (i.e., hand motion (HM), biological object (BO) and non-biological object (NBO)) for both visual conditions (occlusion and full vision). The transparent rectangles represent the occluded part of the video. In the experiment, a black rectangle was used to provide a complete occlusion. The abscissa axis expresses the action timing as a the percentage of their total duration (i.e., normalised time). B. Timeline of the experiment with four successive blocks. The first two blocks were occlusion blocks, and the last two were full vision blocks. The sequence of blocks (i.e., black or red) was randomised between the participants.

A question was presented ("What object appeared in the previous video?") after six successive trials to control the subjects' attention. They were instructed to respond as fast as possible, and the number of wrong responses were recorded by the experimenter.

Data analysis

Kinematic analysis

A kinematic analysis of the different effectors was performed using a semi-automatic visual tracking (Kinovea software vers. 2025.2; https://www.kinovea.org). The x and y coordinates of the effectors were extracted. These coordinates were filtered with a second-order Butterworth low pass filter at 5 Hz. The wrist trajectory and tangential velocity profile were analysed for the human motion (HM; Fig. 2A). Displacement of the centre of the object and its tangential velocity profile were analysed for both the biological (Fig. 2B) and the non-biological (Fig. 2C) objects.

Figure 2 Kinematic features of effectors. The solid lines of the left panel and right panel represent, respectively, the mean of effector trajectory and tangential velocity profile (A. wrist for HM; B. and C. centre of the object for BO and NBO). The shadow areas represent the standard deviation. The time was normalised as the percentage of action duration for the tangential velocity profile.

The following parameters were computed: mean velocity (px/s), motion duration (ms) and displacement (pixels); peak velocity value (px/s) and time to peak velocity (TPV, expressed as a percentage of the acceleration time divided by the action duration), for HM and BO only (Table 1).

Table 1
Kinematic features (mean ± standard deviation) human motion (HM), biological object (BO) and non-biological object (NBO)

Kinematic parameters	HM	BO	NBO	Statistical results
Time to peak velocity (%)	47.73 (± 5.04)	50.25 (± 2.39)		p > 0.05
Peak velocity (px/s)	586.11 (± 58.45)	503.56 (± 73.27)		p > 0.05
Mean velocity (px/s)	313,83 (± 16.78)	218.93 (± 12.98)	167.77 (± 12.73)	p < 0.01
Motion duration (ms)	1250 (± 88.2)	1058.3 (± 50)	1216.7 (± 110.6)	p < 0.05
Distance (px)	375.83 (± 18.33)	231.72 (± 4.71)	233.87 (± 5.6)	p < 0.05

A Kruskal–Wallis test was conducted to compare the kinematic parameters of mean velocity, motion duration and distance, followed by a Dunn's post hoc comparison with a Holm correction when $p < 0.05$. A Mann-Whitney test was conducted for the peak velocity and TPV values, with a significant threshold of $p < 0.05$. Notably, no statistical significance was observed between the TPV of HM (47.73 ± 5.04%) and BO (50.25 ± 2.39%), both of which reached ≈ 50% ($U = 10$, $p > 0.05$, $r_{rb} = -0.25$). Peak velocity values also did not differ significantly ($U = 1$, $p > 0.05$, $r_{rb} = 0.875$) between HM (586.11 ± 58.45px/s) and BO (503.56 ± 73.27px/s). However, the mean velocity was found to have a significant effect ($H(2) = 9.85$, $p < 0.01$, $\eta^2 = 0.87$; Kruskal-Wallis test), with a lower mean velocity observed for NBO than for HM ($p < 0.01$, $r_{rb} = -1$). A distance effect was also observed ($H(2) = 8.346$, $p < 0.05$, $\eta^2 = 0.705$; Kruskal-Wallis test), as the HM path was longer than the BO path ($p < 0.05$). Finally, the motion duration showed a significant difference ($H(2) = 6.851$, $p < 0.05$, $\eta^2 = 0.539$; Kruskal-Wallis test), with a shorter duration observed for BO than HM ($p < 0.05$).

Time-to-Contact predictions

The accuracy of the TTC prediction was calculated for each trial as the difference in milliseconds between the participant's response time and the real TTC (constant error, CE). Regarding the hand motion, the real TTC was defined as the initial contact between the object and the manual thumb-index finger pinch. With regard to the object motion, the real TTC was defined as the first contact between the object and the third finger. Furthermore, the variability of participants' responses is expressed as the mean standard deviation (VE: variable error). Erroneous recordings, such as those resulting from multiple keyboard taps, and participants' responses that exceeded two standard deviations above or below the group mean for each experimental condition, were discarded in the final statistical analysis. Moreover, participants who had undergone more than four trials discarded were excluded from the final analysis. Consequently, the results of two participants (S13 and S23) were excluded. For the remaining

participants, 1.7% of repetitions were excluded for HM block 1 (occlusion: 1.7%; full vision: 1.7%), 1.4% of repetitions for HM block 2 (occlusion: 0.9%; full vision: 2%), 4.2% of repetitions for BO (occlusion: 2.6%; full vision: 5.7%) and 5% of repetitions for NBO (occlusion: 6.6%; full vision: 3.4%). The CE and VE dataset was found to be normally distributed ($p > 0.05$, Shapiro-Wilk test). No statistically significant differences were observed between the first and second blocks for the hand motion condition in the two visual conditions ($t(28) = 0.506$, $p > 0.05$ and $t(28) = -1.957$, $p > 0.05$ for the occlusion and full vision conditions, respectively). Furthermore, no learning intra-block effect was observed for any of the experimental conditions when comparing the mean of the first three trials with the mean of the last three (HM1 occlusion: $t(28) = -0.608$, $p > 0.05$; HM2 occlusion: $t(28) = -0.567$, $p > 0.05$; BO occlusion: $t(28) = -1.118$, $p > 0.05$; NBO occlusion: $t(28) = -0.882$, $p > 0.05$; HM1 full-vision: $t(28) = -1.838$, $p > 0.05$; HM2 full-vision: $t(28) = -1.705$; $p > 0.05$; BO full-vision: $t(28) = -1.893$, $p > 0.05$ and NBO full-vision: $t(28) = 1.811$, $p > 0.05$). Accordingly, the mean of the two HM blocks was utilised for subsequent analysis. The participants' CE and VE were compared using a repeated-measures ANOVA with the following factors: visual conditions (occlusion vs. full vision) and kinematics (HM, BO and NBO). The significance level was set to $p < 0.05$, and the Holm method was applied where significant effects were observed. A Greenhouse-Geisser correction was applied when the sphericity of the data was not respected (Mauchly's test, $p < 0.05$). Moreover, a multinomial logistic regression model (leave-one-subject-out, LOSO) was utilised to verify the predictability of kinematics from behavioural data. To avoid inflating classification accuracy due to within-subject dependencies, a LOSO cross-validation procedure was carried out. In each fold, all observations from one participant were excluded from model fitting and used exclusively for testing. This approach evaluates the model's ability to generalise to unseen participants. Finally, the attentional task was analysed using a t-test. No significant differences in attention were observed between the two kinematic blocks (HM vs. BO and HM vs. NBO; $t = 0.178$; $df = 77$; $p > 0.05$). Similar attentional results were observed when the visual conditions were compared (occlusion vs. full vision: $t = 0.297$; $df = 77$; $p > 0.05$). All statistical analyses were carried out using JASP (JASP, 2022) and R (R Core Team, 2021) software environments.

Results

Constant Error

The study revealed that neither visual condition had a statistically significant effect ($F(1, 28) = 0.83$, $p > 0.05$, $\hat{\omega}^2_G = 0.006$). Therefore, participants' accuracy was not dependent on the available visual information, since the accuracy of the occlusion condition was 110.98 ms (SE: ± 41.09) compared to 89.38 ms (SE: ± 30.76) for full vision condition ($p > 0.05$, *Cohen's dz* = 0.147). However, a significant kinematic effect was found ($F(1.597, 44.70) = 144.66$, $p < 0.001$, $\hat{\omega}^2_G = 0.459$) due to the discrepancy between the NBO kinematic (287.71 ms (SE: ± 25.00)) and both the HM (20.23 ms (SE: ± 25.01); $p < 0.001$, *Cohen's dz* = -1.82) and BO kinematic (-7.40 ms (SE: ± 31.38); $p < 0.001$, *Cohen's dz* = -2.008; Fig. 3A). It is noteworthy that no statistically significant differences were observed between the HM and BO kinematics ($p > 0.05$, *Cohen's dz* = 0.188). Finally, an interaction between visual conditions and

kinematics was observed (F(1.987, 55.628) = 7.26, p < 0.01, 2G = 0.012). In both visual conditions (occlusion and full vision), prediction accuracy for biological kinematics differed significantly from that for non-biological kinematics (HM_occ vs NBO_occ, p < 0.001, *Cohen's dz* = -2.067; BO_occ vs NBO_occ, p < 0.001, *Cohen's dz* = -2.222; HM_fv vs NBO_fv, p < 0.001, *Cohen's dz* = -1.572; BO_fv vs NBO_fv, p < 0.001, *Cohen's dz* = -1.794). However, no significant difference was found between biological kinematics (HM_occ vs BO_occ, p > 0.05, *Cohen's dz* = 0.155; HM_fv vs BO_fv, p > 0.05, *Cohen's dz* = 0.12). Interestingly, the visual condition significantly impacted the accuracy of CE, but only for non-biological kinematics (p < 0.05, *Cohen's dz* = 0.455; Fig. 3B). Indeed, the accuracy for non-biological kinematics was higher in the full vision condition (254.29 ms (SE: ± 16.28)) than in the occlusion condition (321.12 ms (SE: ± 30.45)).

Furthermore, in the context of non-biological kinematics, participants demonstrated a reactive rather than a predictive response. This assertion is substantiated by the mean CE, which exceeded 200 milliseconds, aligning with a simple reaction times of around 200ms (for a review see Silverman, 2010).

Variable Error

The VE of the participants depended on the visual conditions (F(1, 28) = 76.282, p < 0.001, 2G = 0.284). Indeed, participants demonstrated reduced variability when reaching to grasp an object with complete visual information (88.61 ms (SE: ± 6.62)) compared to the occluded condition (128.20 ms (SE: ± 8.06); p < 0.001, *Cohen's dz* = 1.239).

Furthermore, a significant kinematic effect on VE was identified (F(1.971, 55.194) = 63.652, p < 0.001, 2G = 0.275). The VE results of BO kinematics (135.55 ms (SE: ± 7.52)) differed significantly from those of HM (91.70 ms (SE: ± 4.99); p < 0.001, *Cohen's dz* = -1.372) and NBO kinematics (97.96 ms (SE: ± 8.44); p < 0.001, *Cohen's dz* = 1.176). No significant discrepancy was observed between the HM and NBO outcomes (p = 0.125, *Cohen's dz* = -0.196). Finally, the VE results revealed an interaction between visual conditions and kinematics (F(1.707, 47.798) = 6.92, p < 0.01, 2G = 0.067). Therefore, kinematics differed within each visual condition, such as occlusion (HM: 102.19 ms (SE: ± 5.74); BO: 153.58 ms (SE: ± 7.66); NBO: 128.83 ms (SE: ± 7.79)) and full vision (HM: 81.21 ms (SE: ± 3.15); BO: 117.53 ms (SE: ± 5.75); NBO: 67.09 ms (SE: ± 4.02); Fig. 4A). Furthermore, each kinematic interfered with the VE depending on the visual condition, since the occlusion condition of each kinematic demonstrated higher variability in participants' responses when compared to the full vision condition of the same kinematic (HM occ vs. HM fv, p < 0.01; BO occ vs. BO fv, p < 0.001; NBO occ vs. NBO fv, p < 0.001; Fig. 4B).

Probability results – LOSO model

To complement the previous results, a LOSO model was performed on constant error data to evaluate whether the kinematics could be discriminated from behavioural data (Fig. 5). Non-biological kinematics were classified above chance level (0.33) and were clearly discriminated in both occlusion (0.76; p < 0.001) and full vision (0.93; p < 0.001) conditions. Similar results were observed for biological object kinematics (0.69; p < 0.001, for both occlusion and full vision conditions).

For the hand motion, the model was unable to correctly discriminate it to biological object kinematics from the behavioural data. These results confirm that the key factor in accurately predicting the outcomes of an action is the motion kinematics (biological and non-biological), rather than the appearance of the effector.

Discussion

The present study sought to investigate the respective contributions of kinematics and visual appearance of effector to action prediction. Our results suggest that prediction accuracy is primarily driven by the kinematic features of the motion, regardless of the effector appearance. However, the response variability depended on an interaction between movement kinematics and effector appearance. Indeed, the consistency of these predictions appears to be determined by their congruence.

Accuracy of action prediction

The present findings support to the hypothesis that the temporal accuracy of action prediction relies on implicit motor knowledge. This knowledge is then compared to the spatio-temporal features of the visual input (Craighero et al., 2008; Saunier et al., 2023; Stapel et al., 2012), through the recruitment of internal models of action (Miall, 2003). Participants demonstrated a higher accuracy in their TTC estimates when the effector moved according to biological rather than non-biological kinematics. Furthermore, occluding the final part of the effector's trajectory did not affect prediction accuracy when the motion followed biological kinematics. It is therefore proposed that the matching of implicit motor knowledge (i.e. internal kinematic representations) with the kinematics extracted from visual input compensates for the loss of visual information during action observation. Sensorimotor predictions processed by forward internal models of action (Hilt et al., 2026), would compensate for transient loss of sensory input. This result aligns with the extensive literature on the predictive capabilities of the sensorimotor system in motion inference tasks (Pozzo et al., 2006; Saunier et al., 2021, 2015, 2008).

Nonetheless, in scenarios where there was a discrepancy between the visual input and internal representations of biological kinematics, participants tend to rely on reactive rather than predictive processing. Indeed, the participants exhibited a constant error exceeding 200 milliseconds under non-biological conditions. This delay is consistent with reaction times reported in visuomotor tasks (Silverman, 2010). Unlike to biological motion, the availability of visual information significantly impacted the accuracy of predictions for non-biological motion. Participants demonstrated a greater accuracy when they had a full vision of the non-biological action. However, despite the entire availability of the visual information, participants continued to exhibit a behavioural pattern consistent with reactive processing. We propose that the non-biological kinematics from the visual input, combined with the implicit knowledge of biological kinematics rules, would introduce a conflict that affects the ability to elaborate an accurate prediction of the action outcome.

Additionally, the impact of the movement's trajectory on the action prediction accuracy was examined. If TTC estimates depended on the simplicity of the motion's trajectory – defined as a straight horizontal path without a change of direction – it would be expected that more complex movements involving a change of direction would be less accurate. However, our results did not support to this hypothesis. Despite the hand's complex trajectory involving horizontal and vertical displacements (i.e., a parabolic trajectory) and a biological velocity profile (i.e., acceleration and deceleration), participants demonstrated greater accuracy in estimating the TTC than for an object moving horizontally at a constant velocity. It was hypothesised that a simple trajectory with a constant velocity would facilitate the visual pursuit of the target and consequently improve TTC accuracy estimation throughout visual extrapolation processing (Kerzel et al., 2001; Nijhawan, 2008). In contrast, the enhanced prediction accuracy observed for the biological effector (the hand or the object in the present study) suggests that action prediction cannot be explained solely by visual processing. The present findings support to the proposal that implicit motor simulation based on biological features, such as velocity and knowledge of biomechanical constraints (Saunier et al., 2015), contributes to the prediction of temporal outcomes, even when the observed actions follow a more complex kinematic pattern.

Furthermore, the LOSO model supports the predominance of kinematic features over effector appearance in order to facilitate accurate action prediction. As behavioural data demonstrated, the model could clearly distinguish between biological and non-biological kinematics. However, analysis of hand motion data showed that the model could not reliably differentiate between this and data obtained from biological objects. This finding further supports the proposal that the appearance of the effector is irrelevant in predicting the outcomes of a biological action.

Variability of action prediction

We found that the variability was lower for HM and NBO than for BO. This specific discrepancy in BO can be explained by the difference between the perceived kinematic and the effector appearance. Indeed, according to biological kinematics, the displacement of an object appears incongruous in the absence of human-object interaction. This incongruence may introduce noise when processing of internal models of actions, resulting in inconsistent responses from participants in both visual conditions (occlusion and full vision) compared to HM and NBO. Studies have demonstrated that discrepancies between the appearance of actors and their kinematics interfere with the activity of the parietal lobule (Saygin et al., 2012). The parietal lobule is notably associated with internal models of action, particularly the forward model (Blakemore and Sirigu, 2003; Fontana et al., 2012).

With concerning to the HM, we observed a greater consistency in the estimates in the occlusion condition than in the BO and NBO conditions. Given the HM's higher complexity, we expected this to be reflected in increased variability. However, this was not the case and supports the idea of consistent prediction for human motion, particularly in the absence of the complete sensory information (Saunier et al., 2008). In the full vision condition, the variability was lower for the NBO. For this specific experimental condition, we propose that the constant velocity of the object, when combined with full visual information, improves the consistency of participants' estimates.

Conclusion

The present study aimed to elucidate the relative contributions of movement kinematics and the effector's appearance to action prediction. It has been demonstrated that the temporal accuracy of action predictions is driven by the kinematic features. However, consistent estimates depend on congruence between kinematics and effector appearance. The results of the present study have direct implications for how we perceive the animate objects with which we interact daily, and appear useful in developing new protocols for sensorimotor rehabilitation or sports training involving human-object interaction.

Declarations

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Author Contribution

G.S. and I.R.S. conceptualised the study; G.S. and I.R.S. developed the methodology; I.R.S. and P.S.A. conducted the experiments; G.S. and I.R.S. analysed the data; G.S. wrote the original draft; G.S., C.N.M.S. and N.G. reviewed and edited the manuscript.

Acknowledgement

The authors would like to thank Ivete Furtado Ribeiro Caldas for her help with data acquisition, as well as the undergraduate students from the Universidade Federal do Pará (UFPA) who took part in the study.

Data Availability

The data that support the findings of this study are available upon request.

References

1. Abreu AM, Macaluso E, Azevedo RT, Cesari P, Urgesi C, Aglioti SM (2012) Action anticipation beyond the action observation network: a functional magnetic resonance imaging study in expert basketball players. Eur J Neurosci 35:1646–1654. https://doi.org/10.1111/j.1460-9568.2012.08104.x
2. Blakemore S-J, Sirigu A (2003) Action prediction in the cerebellum and in the parietal lobe. Exp Brain Res 153:239–245. https://doi.org/10.1007/s00221-003-1597-z

3. Calvo-Merino B, Grèzes J, Glaser DE, Passingham RE, Haggard P (2006) Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation. Curr Biol 16:1905–1910. https://doi.org/10.1016/j.cub.2006.07.065
4. Craighero L, Bonetti F, Massarenti L, Canto R, Fabbri Destro M, Fadiga L (2008) Temporal prediction of touch instant during observation of human and robot grasping. Brain Res Bull 75:770–774. https://doi.org/10.1016/j.brainresbull.2008.01.014
5. de Assis Chagas AM, Alves de Faria I, Saunier G, Bittencourt-Navarrete RE, Nogueira-Campos AA (2021) Implicit action prediction constrains observed biological action reconstruction. Heliyon 7:e06189. https://doi.org/10.1016/j.heliyon.2021.e06189
6. Erlhagen W (2003) Internal models for visual perception. Biol Cybern 88:409–417. https://doi.org/10.1007/s00422-002-0387-1
7. Fontana AP, Kilner JM, Rodrigues EC, Joffily M, Nighoghossian N, Vargas CD, Sirigu A (2012) Role of the parietal cortex in predicting incoming actions. NeuroImage 59:556–564. https://doi.org/10.1016/j.neuroimage.2011.07.046
8. Gazzola V, Keysers C (2009) The Observation and Execution of Actions Share Motor and Somatosensory Voxels in all Tested Subjects: Single-Subject Analyses of Unsmoothed fMRI Data. Cereb Cortex 19:1239–1255. https://doi.org/10.1093/cercor/bhn181
9. Giorello G, Sinigaglia C (2007) Perception in action. Acta Bio-Medica Atenei Parm 78(Suppl 1):49–57
10. Graf M, Reitzner B, Corves C, Casile A, Giese M, Prinz W (2007) Predicting point-light actions in real-time. NeuroImage 36 Suppl 2. T22–32. https://doi.org/10.1016/j.neuroimage.2007.03.017
11. Grush R (2004) The emulation theory of representation: motor control, imagery, and perception. Behav Brain Sci 27:377–396 discussion 396–442. https://doi.org/10.1017/s0140525x04000093
12. Hilt PM, Gueugneau N, Demougeot L, Fadiga L, Pozzo T, Papaxanthis C (2026) Evidence for a sensorimotor prediction in action observation. iScience 0. https://doi.org/10.1016/j.isci.2026.115876
13. JASP T (2022) JASP
14. Jeannerod M (2006) The simulation hypothesis of motor cognition. In: Jeannerod M (ed) Motor Cognition: What Actions Tell the Self. Oxford University Press, p 0. https://doi.org/10.1093/acprof:oso/9780198569657.003.0006
15. Johansson G (1973) Visual perception of biological motion and a model for its analysis. Percept Psychophys 14:201–211. https://doi.org/10.3758/BF03212378
16. Kerzel D, Jordan JS, Müsseler J (2001) The role of perception in the mislocalization of the final position of a moving target. J Exp Psychol Hum Percept Perform 27:829–840. https://doi.org/10.1037/0096-1523.27.4.829
17. Miall RC (2003) Connecting mirror neurons and forward models. NeuroReport 14:2135–2137. https://doi.org/10.1097/00001756-200312020-00001

18. Nijhawan R (2008) Visual prediction: psychophysics and neurophysiology of compensation for time delays. Behav Brain Sci 31:179–198 discussion 198–239. https://doi.org/10.1017/S0140525X08003804
19. Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113. https://doi.org/10.1016/0028-3932(71)90067-4
20. Payedimarri AB, Ratti M, Rescinito R, Vanhaecht K, Panella M (2022) Effectiveness of Platform-Based Robot-Assisted Rehabilitation for Musculoskeletal or Neurologic Injuries. Syst Rev Bioeng 9:129. https://doi.org/10.3390/bioengineering9040129
21. Pozzo T, Papaxanthis C, Petit JL, Schweighofer N, Stucchi N (2006) Kinematic features of movement tunes perception and action coupling. Behav Brain Res 169:75–82. https://doi.org/10.1016/j.bbr.2005.12.005
22. Core Team R (2021) R: A Language and Environment for Statistical Computing
23. Rangel ML, Souza L, Rodrigues EC, Oliveira JM, Miranda MF, Galves A, Vargas CD (2021) Predicting Upcoming Events Occurring in the Space Surrounding the Hand. Neural Plast. 2021, 6649135. https://doi.org/10.1155/2021/6649135
24. Saunier G, da Silva Azevedo P, Labiapari Pinto V, Nogueira-Campos AA (2023) The pleasantness and unpleasantness of an object distinctively drives its grasping prediction: behavioral evidence. Psychol Res 87:1491–1500. https://doi.org/10.1007/s00426-022-01765-w
25. Saunier G, Fontana AP, De Oliveira JM, Py MO, Pozzo T, Vargas CD (2021) Cerebellar damage affects the inference of human motion. Neurocase 27:169–177. https://doi.org/10.1080/13554794.2021.1905853
26. Saunier G, Paillard AC, Vargas CD, Pozzo T (2015) Visual inference of arm movement is constrained by motor representations. Behav Brain Res 290:197–200. https://doi.org/10.1016/j.bbr.2015.04.056
27. Saunier G, Papaxanthis C, Vargas CD, Pozzo T (2008) Inference of complex human motion requires internal models of action: behavioral evidence. Exp Brain Res 185:399–409. https://doi.org/10.1007/s00221-007-1162-2
28. Saygin AP, Chaminade T, Ishiguro H, Driver J, Frith C (2012) The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. Soc Cogn Affect Neurosci 7:413–422. https://doi.org/10.1093/scan/nsr025
29. Saygin AP, Stadler W (2012) The role of appearance and motion in action prediction. Psychol Res 76:388–394. https://doi.org/10.1007/s00426-012-0426-z
30. Silverman IW (2010) Simple reaction time: it is not what it used to be. Am J Psychol 123:39–50. https://doi.org/10.5406/amerjpsyc.123.1.0039
31. Stadler W, Springer A, Parkinson J, Prinz W (2012) Movement kinematics affect action prediction: comparing human to non-human point-light actions. Psychol Res 76:395–406. https://doi.org/10.1007/s00426-012-0431-2
32. Stapel JC, Hunnius S, Bekkering H (2012) Online prediction of others' actions: the contribution of the target object, action context and movement kinematics. Psychol Res 76:434–445.

https://doi.org/10.1007/s00426-012-0423-2

33. Urgen BA, Saygin AP (2020) Predictive processing account of action perception: Evidence from effective connectivity in the action observation network. Cortex J Devoted Study Nerv Syst Behav 128:132–142. https://doi.org/10.1016/j.cortex.2020.03.014
34. Witte K, Bürger D, Pastel S (2025) Sports training in virtual reality with a focus on visual perception: a systematic review. Front Sports Act Living 7:1530948. https://doi.org/10.3389/fspor.2025.1530948
35. Wolpert DM, Flanagan JR (2001) Motor prediction. Curr Biol CB 11. https://doi.org/10.1016/s0960-9822(01)00432-8. R729-732
36. World Medical Association (2013) World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA 310:2191–2194. https://doi.org/10.1001/jama.2013.281053

Figures

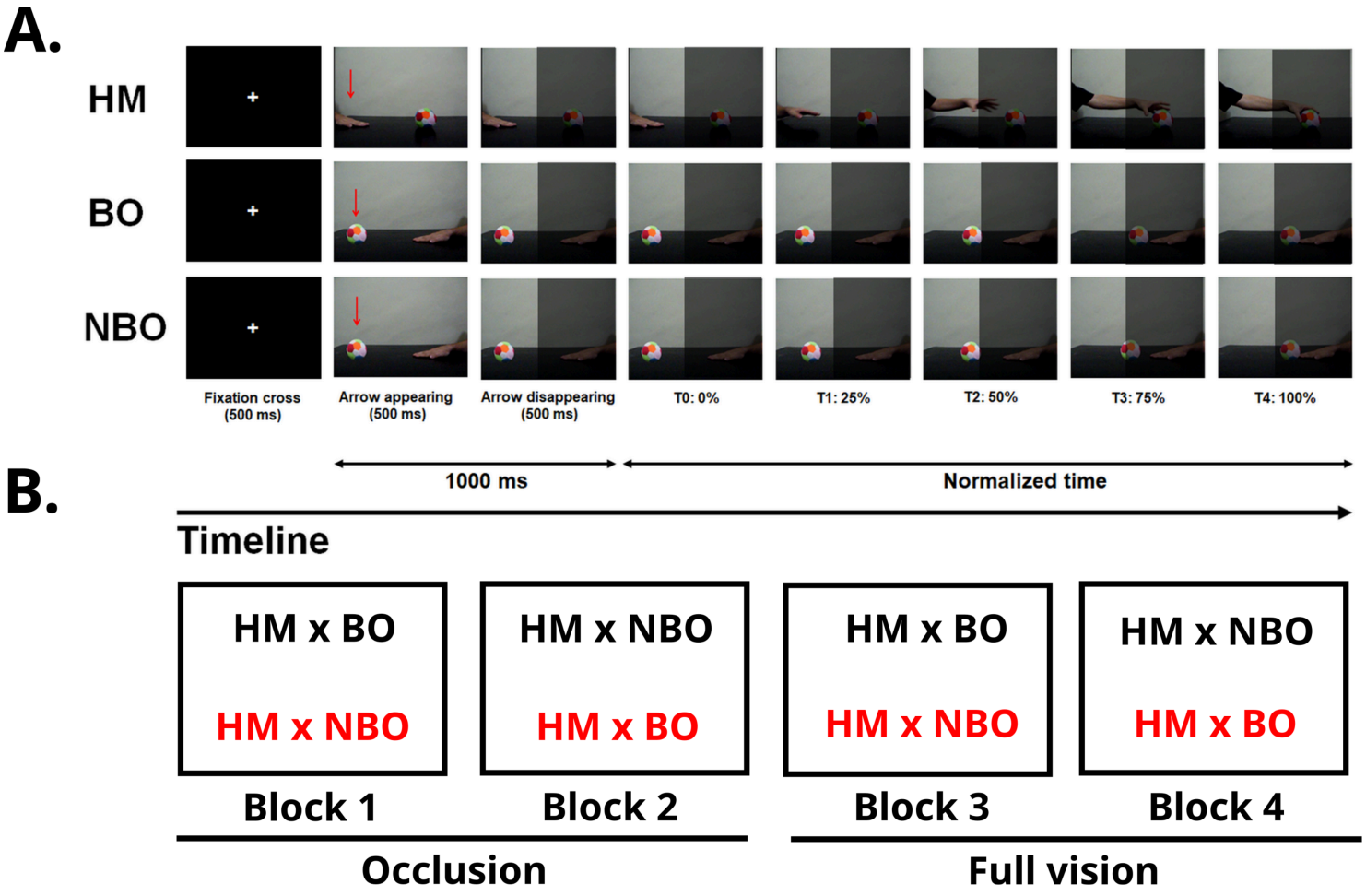


Figure 1

Experimental design. A. From top to bottom, an example of video sequence for each kinematic feature (i.e., hand motion (HM), biological object (BO) and non-biological object (NBO)) for both visual conditions (occlusion and full vision). The transparent rectangles represent the occluded part of the

video. In the experiment, a black rectangle was used to provide a complete occlusion. The abscissa axis expresses the action timing as a the percentage of their total duration (i.e., normalised time). B. Timeline of the experiment with four successive blocks. The first two blocks were occlusion blocks, and the last two were full vision blocks. The sequence of blocks (i.e., black or red) was randomised between the participants.

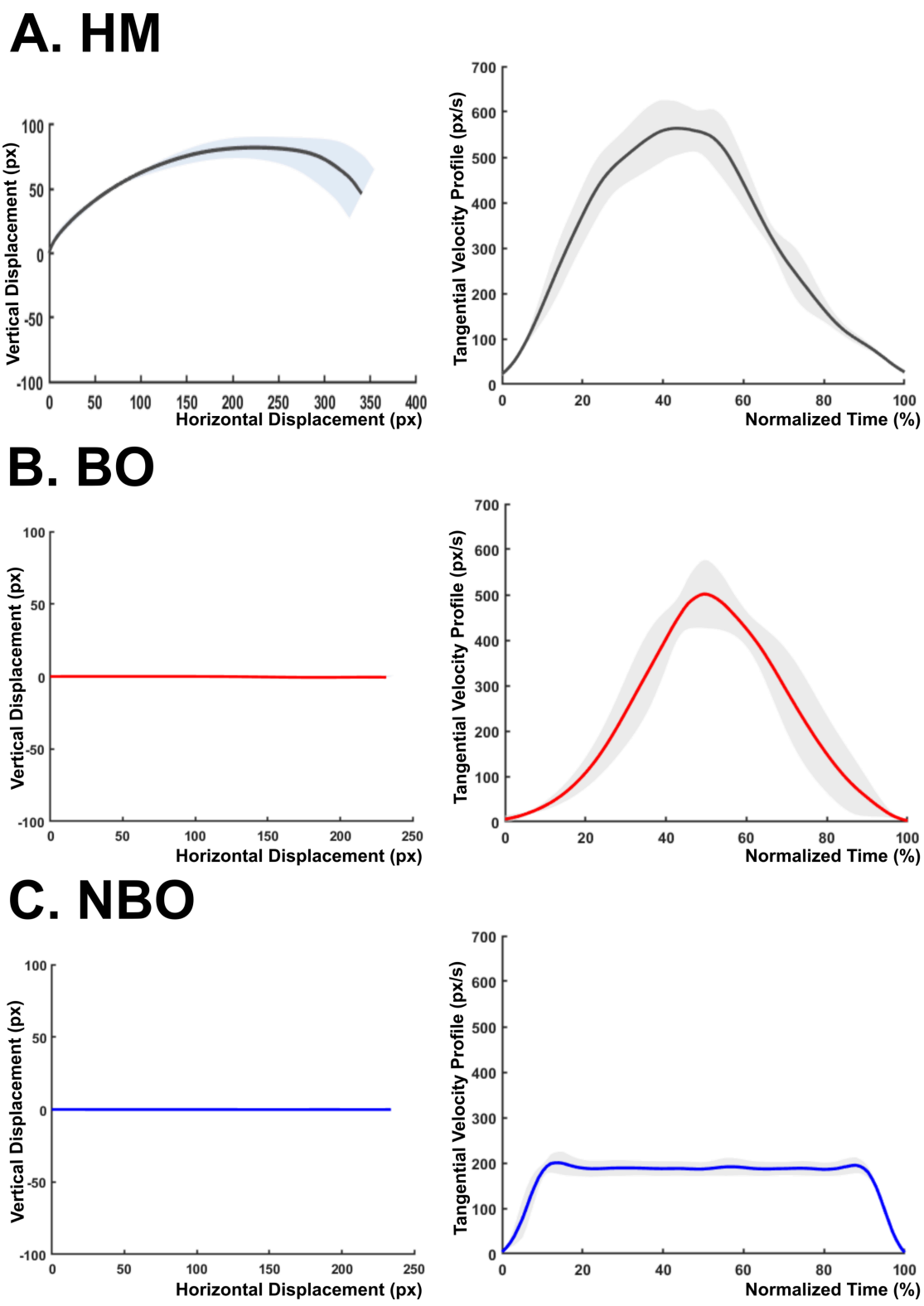


Figure 2

A kinematic analysis of the different effectors was performed using a semi-automatic visual tracking (Kinovea software vers. 2025.2; https://www.kinovea.org). The x and y coordinates of the effectors were extracted. These coordinates were filtered with a second-order Butterworth low pass filter at 5 Hz. The wrist trajectory and tangential velocity profile were analysed for the human motion (HM; Fig. 2A). Displacement of the centre of the object and its tangential velocity profile were analysed for both the biological (Fig. 2B) and the non-biological (Fig. 2C) objects.

Kinematic features of effectors. The solid lines of the left panel and right panel represent, respectively, the mean of effector trajectory and tangential velocity profile (A. wrist for HM; B. and C. centre of the object for BO and NBO). The shadow areas represent the standard deviation. The time was normalised as the percentage of action duration for the tangential velocity profile.

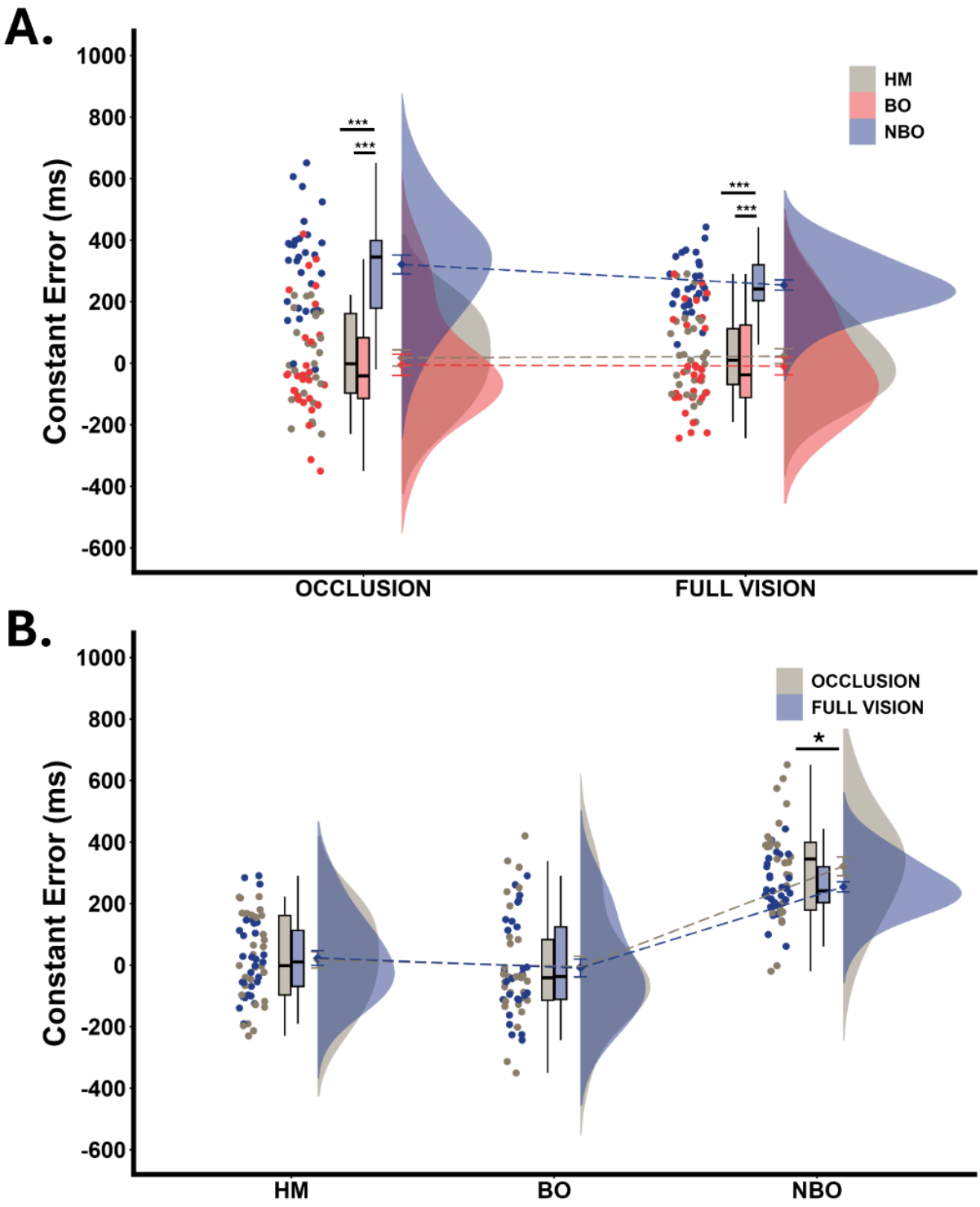


Figure 3

Raincloud plots for the interactions between visual conditions and kinematics (panel A) and between kinematics and visual conditions (panel B). Each dot represents the mean constant error for each participant and each kinematic: Hand Motion (HM), Biological Object (BO) and Non-Biological Object (NBO). Dash lines link the mean constant error values (± SEM). The box plots depict the 95% confidence

intervals and, the data distributions were observed for each experimental condition. *** p < 0.001; * p < 0.05.

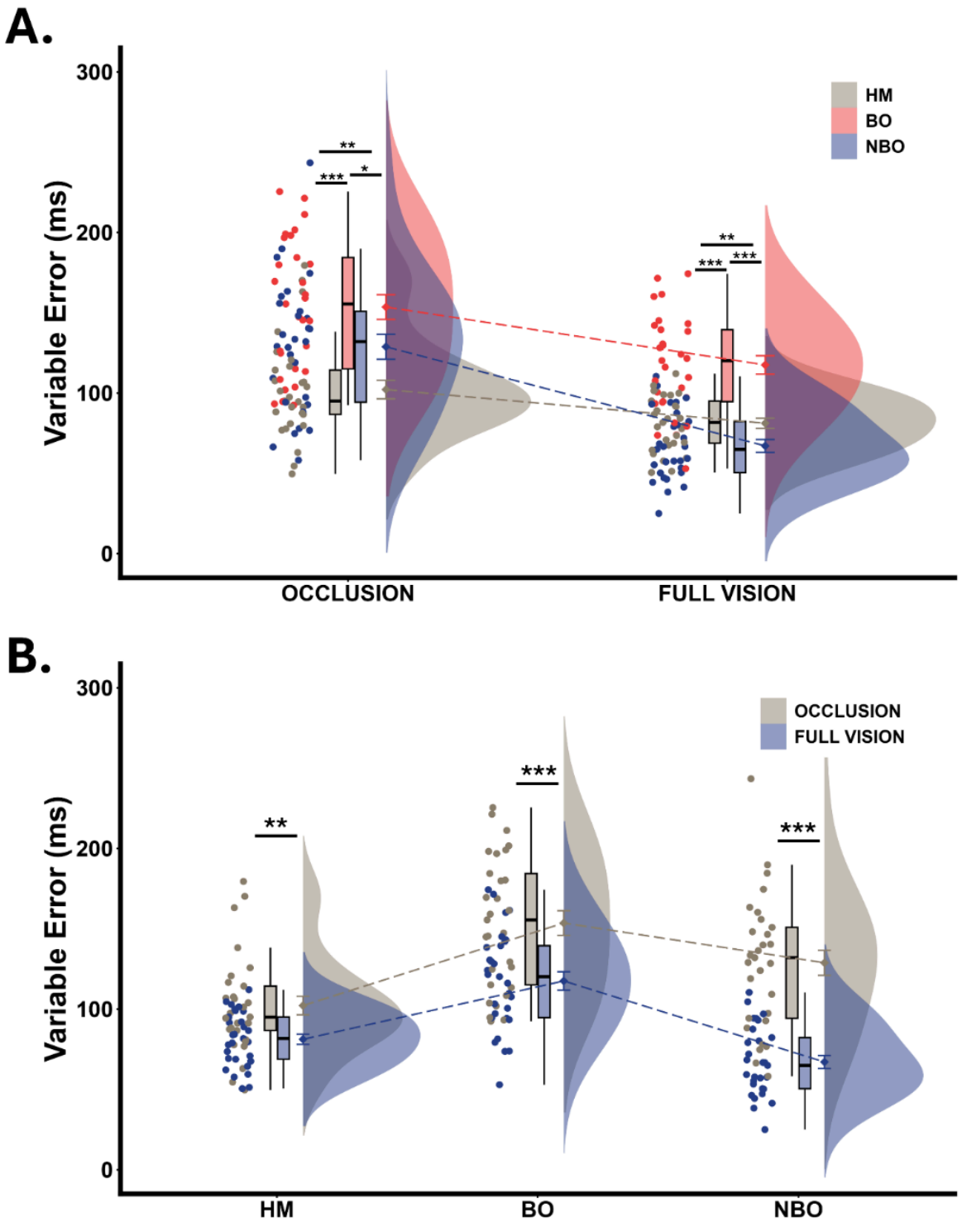


Figure 4

Raincloud plots for interactions visual conditions versus kinematics (A) and kinematics versus visual conditions (B). Each dot represents the mean constant error of each participant for each kinematic:

Hand Motion (HM), Biological Object (BO) and Non-Biological Object (NBO). The dash lines link the mean values of prediction error (± SEM). The box plots depict 95% confidence intervals and the data distributions for each experimental condition were observed. ***, p < 0.001; **, p < 0.01; *, p < 0.05.

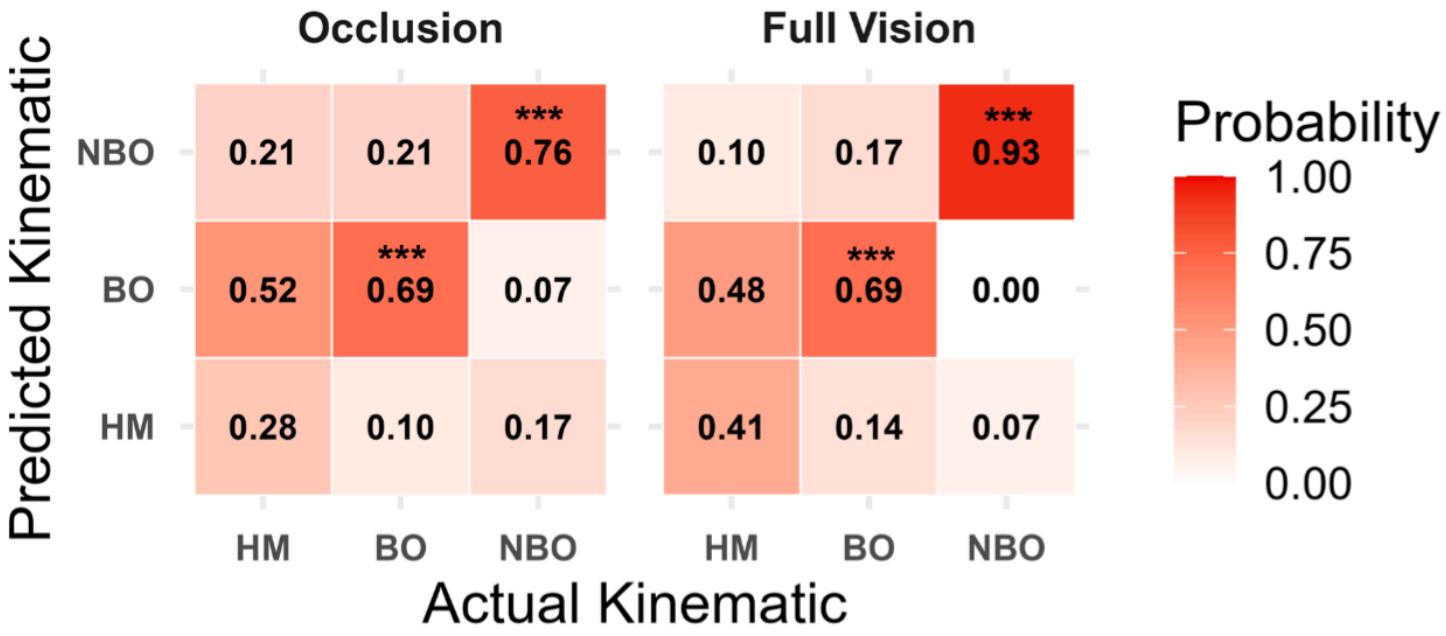


Figure 5

Matrices of the LOSO model derived from behavioural data in the occlusion (left panel) and full vision (right panel) conditions. The columns represent the actual kinematic, whilst the rows represent the predicted kinematic. The numerical values represent the ratio of the predicted kinematic to the actual kinematic (column normalised). The colour intensity of red is indicative of the classification probability, ranging from 0 to 1. The chance level was set to 0.33. A binomial test was conducted to verify the classification probability of kinematic that differed to chance level. ***, p < 0.001.