

Current Biology

Fast prediction in marmoset reach-to-grasp movements for dynamic prey

Highlights

- Video tracking marmoset reach-to-grasp behaviors in a cricket-hunting task
- Marmosets open grasp aperture earlier for faster velocity reaches
- Short visuo-motor delays in marmoset reaching for moving crickets
- Marmosets predict moving target positions to compensate for visuo-motor delays

Authors

Luke Shaw, Kuan Hong Wang,
Jude Mitchell

Correspondence

kuanhong_wang@urmc.rochester.edu
(K.H.W.),
jmitch27@ur.rochester.edu (J.M.)

In brief

Shaw et al. capture marmoset cricket-hunting behaviors by applying machine vision to high-speed video and find that reach-to-grasp movements for dynamic prey operate at short visuo-motor delays that rival oculomotor systems. They further show that visual prediction helps to compensate for the visuo-motor delays in reaching.



Report

Fast prediction in marmoset reach-to-grasp movements for dynamic prey

Luke Shaw,¹ Kuan Hong Wang,^{1,*} and Jude Mitchell^{2,3,*}¹Department of Neuroscience, University of Rochester Medical Center, Rochester, NY 14642, USA²Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14611, USA³Lead contact*Correspondence: kuanhong_wang@urmc.rochester.edu (K.H.W.), jmitch27@ur.rochester.edu (J.M.)<https://doi.org/10.1016/j.cub.2023.05.032>

SUMMARY

Primates have evolved sophisticated, visually guided reaching behaviors for interacting with dynamic objects, such as insects, during foraging.^{1–5} Reaching control in dynamic natural conditions requires active prediction of the target's future position to compensate for visuo-motor processing delays and to enhance online movement adjustments.^{6–12} Past reaching research in non-human primates mainly focused on seated subjects engaged in repeated ballistic arm movements to either stationary targets or targets that instantaneously change position during the movement.^{13–17} However, those approaches impose task constraints that limit the natural dynamics of reaching. A recent field study in marmoset monkeys highlights predictive aspects of visually guided reaching during insect prey capture among wild marmoset monkeys.⁵ To examine the complementary dynamics of similar natural behavior within a laboratory context, we developed an ecologically motivated, unrestrained reach-to-grasp task involving live crickets. We used multiple high-speed video cameras to capture the movements of common marmosets (*Callithrix jacchus*) and crickets stereoscopically and applied machine vision algorithms for marker-free object and hand tracking. Contrary to estimates under traditional constrained reaching paradigms, we find that reaching for dynamic targets can operate at incredibly short visuo-motor delays around 80 ms, rivaling the speeds that are typical of the oculomotor systems during closed-loop visual pursuit.¹⁸ Multivariate linear regression modeling of the kinematic relationships between the hand and cricket velocity revealed that predictions of the expected future location can compensate for visuo-motor delays during fast reaching. These results suggest a critical role of visual prediction facilitating online movement adjustments for dynamic prey.

RESULTS AND DISCUSSION

Tracking of marmoset reaches

Primates exhibit a sophisticated repertoire of visually guided reaching behaviors that enable them to interact with objects and navigate arboreal environments. Although the motor aspects of these reaching movements are largely conserved across mammalian species, primates are unique in their integration of visual information to guide reaching movements.^{19–23} A recent study used video in the field to examine reach-to-grasp movements for flying insects among wild marmoset monkeys.⁵ They found evidence to support an online active vision strategy during prey capture in which the head angle of marmosets actively tracked insects during reaching. However, the video resolution in the wild does not afford enough precision either to measure visuo-motor delay or to test if the motor control strategy exploits predictions of future target location to compensate for processing delays. Here, we establish an ecologically motivated reaching task within a laboratory setting using high-speed video to test the kinematics of reach-to-grasp movements for dynamic prey.

We used high-speed, multi-camera video and machine learning methods to track marmoset reaches for live, moving

crickets. Reaching experiments were conducted by docking a mobile video-recording platform to marmoset family cages within the marmoset colony and recording reaches using 3 GoPro cameras (Figures 1A and 1B). Hands, fingers, crickets, and the experimental apparatus were tracked from three cameras using DeepLabCut²⁴ and 3D position triangulated using Anipose²⁵ (Figures 1C–1G). DeepLabCut performance using the labels illustrated in Figures 1C and 1D is quantified in Figure S1. Although we were able to reconstruct 3D trajectories of reaching, most subsequent analyses focused on 2D tracking from the top-view camera angle depicted in Figure 1F because crickets mainly limited their motion to the horizontal plane during the reaching tasks.

Marmoset reach-to-grasp dynamics

Earlier work in laboratory settings have shown that marmosets are equally likely to prefer using their left or right hands in reaching,²⁶ and they typically employ a power grasp strategy to reach items in which all fingers open in unison and close around the target to press it within the palm of the hand, which has been shown in marmoset²⁷ and macaque.^{28,29} A key difference from macaques is that marmosets do not use a precision grip. Old World primates have opposable thumbs that allow them to grasp



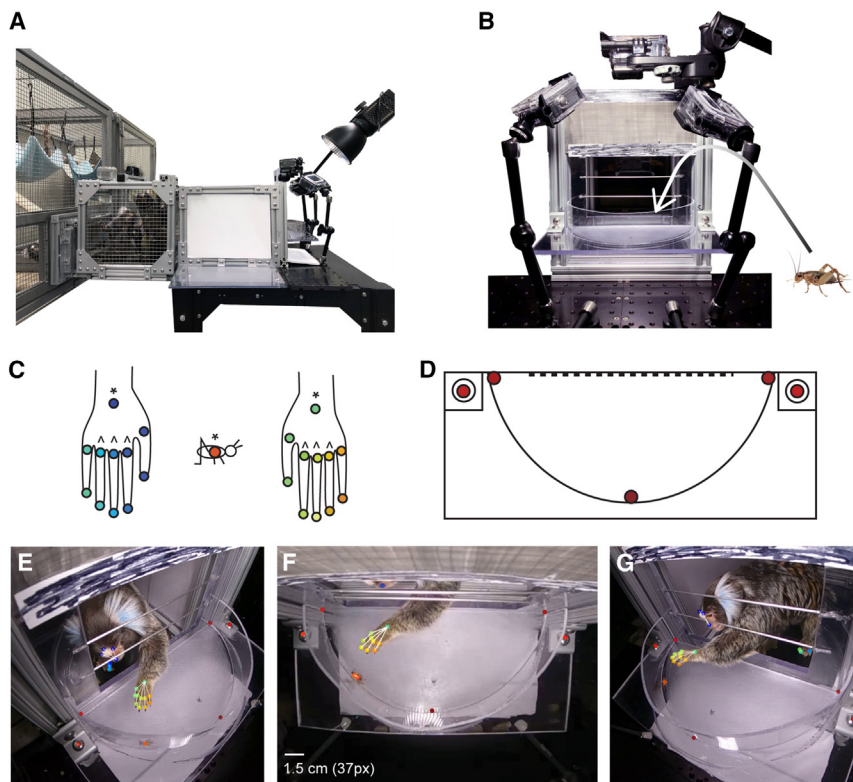


Figure 1. High-speed, multi-camera video reconstructs reaching for dynamic prey

(A) A wheeled video platform allowed tests of individual marmosets from inside their housing in a colony.

(B) Three GoPro cameras were calibrated around a semi-circular reaching arena where a live cricket was placed. Marmosets reached through a rectangular aperture to acquire crickets.

(C) Feature points were manually labeled in a subset of video frames to train DeepLabCut. Hand and cricket markers (*) were used for reaching analyses and knuckle markers (^) for grasping analyses.

(D) Arena markers were used to validate spatial coordinates and correct for any post-calibration camera movement.

(E–G) DeepLabCut labeling of synchronized frames of video from three camera angles. DeepLabCut tracked labeled points with a RMSE of ~ 3 and ~ 7 pixels for the cricket and hand, respectively, compared with ~ 37 pixels for the size of the marmoset wrist.

See also [Figure S1](#).

small objects between the thumb and fingers,³⁰ whereas most New World primates, including the marmoset, use a power grasp.³¹ We found that marmosets in our experiments predominantly employed a left or right single-handed (LH or RH) reaching strategy akin to a pause and lunge approach utilized by marmosets in the wild.^{5,32} This strategy may have been favored due to the reaching aperture utilized in our setup, which prevented marmosets from using their mouths or a two-handed pounce (P) strategy. We recorded a total of 271 reaches for live crickets (137 LH, 134 RH, and 10 P) from 10 marmosets. Gross quantification of these reaches through a left-handedness index (LHI) (LH reaches/total reaches) indicates an evenly split mixed hand preference over the population of reaches. Assessing this quantification at the marmoset participant level shows two marmosets with a strong right-handed preference (LHI = 0.31 and 0.04) and three marmosets with a strong left-handed preference (0.83, 0.87, and 0.92), whereas the other 5 marmosets had little preference (0.60, 0.68, 0.64, 0.66, 0.44, and 0.40). Of the total recorded set of 271 reaches for crickets, 78 reaches contained crickets that moved during or before the reach, which were selected for subsequent analyses (see [STAR Methods](#) for inclusion criteria).

Marmosets performed reach-to-grasp movements for moving crickets using a power grasp to enclose crickets within their hand at the end of the reach ([Figure 2](#)). In an individual reaching trial, the opening of grasp aperture begins as hand velocity drops with increasing proximity to target ([Figures 2A–2C](#)), which is consistent with findings in previous NHP studies from both marmosets and macaques.^{27–29} Finger separation distance was calculated as the average of the distances between the index finger and middle finger, and the middle finger and ring finger,

using knuckle markers ([Figure 2B](#), dark blue points). For this reach, there is a clear temporal order with the peak in hand speed followed by an increase in finger spread as the hand speed decreases ([Figure 2C](#)). This pattern was typical of the population of reaches. A frontal view of a typical reach ([Figure S2](#)) shows the fingers wrapping around the cricket.

We used dimensionality reduction methods to further characterize the grasping strategies employed in the population of reaches. For each reach, a time series of finger spread distance was calculated based on the knuckle markers. The distance values were interpolated over a fixed number of time steps over the total reach duration, with the reach start defined from the moment the hand first entered the arena to the end being when the fingers first began to close around the cricket. Principal component analysis of these reach trials captured 69.1% of the variance within the first two components. We found a continuum of points, where each point corresponds to one time-normalized reach, which extended along the first principal component as opposed to discrete clusters ([Figure 2D](#)). To visualize the variations in finger separation across the population of reaches, we performed a k-means clustering and split the space into the best fitting three clusters.

The power grasp strategy shows a temporal pattern of earlier grasp opening for faster movements. Cluster 1 shows an early finger spread after reach initiation, which continues throughout the reach ([Figure 2E1](#)). By comparison, in cluster 2, the finger spread initiates later during hand transport ([Figure 2E2](#)), and in cluster 3, finger spread initiates only at the very end of the reach ([Figure 2E3](#)). Plotting hand velocity time series for the same clusters, we find higher and earlier peak velocities for cluster 1 compared with cluster 2, and likewise for clusters 2 and 3 ([Figures 2F1–2F3](#)). Thus, the timing of grasp opening varies with reach speed. Video frames taken from an example reach

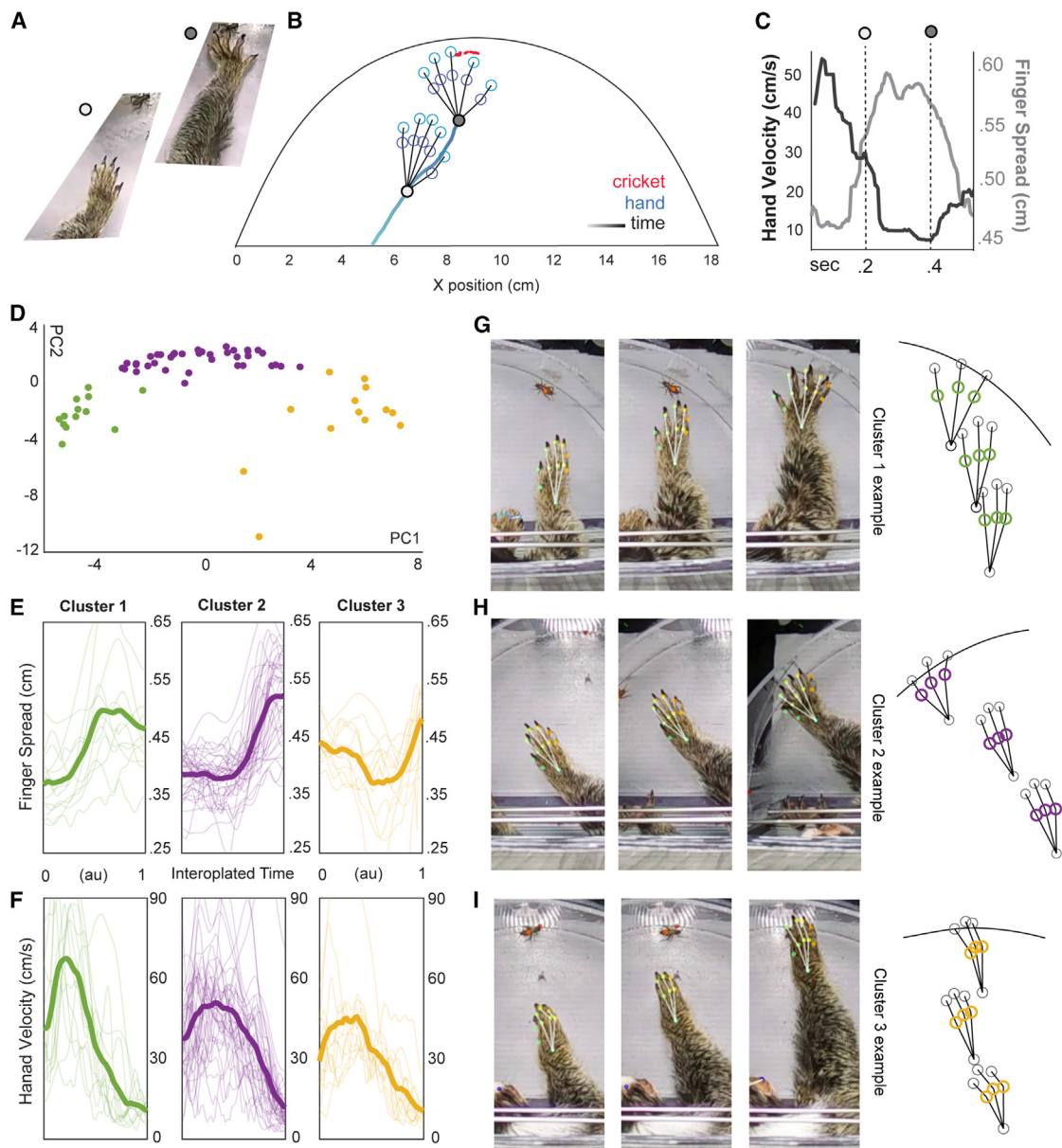


Figure 2. Hand aperture opens earlier with increasing hand speed in a power grasp

(A) Cropped video frames show a single reaching trial at two time points indicated by light and dark gray points.
 (B) The hand position for the corresponding time points in (A) is shown from the top camera view. The underlying light blue trace shows the trajectory of the reach from the central hand marker and the red traces indicates cricket motion.
 (C) Within the example reach the speed of the hand (shown in black, left vertical scale) peaked and began to fall before the finger spread increased, depicted in gray (right vertical scale).
 (D) The finger spread over time was submitted to a cluster analysis (STAR Methods) which revealed a continuum along the principal component axes that we split into three clusters (green, purple, and orange) for visualization. The two outliers present in cluster 3 (yellow, at the bottom) were due to finger contact against the retaining wall at the end of the reach.
 (E) Finger separation peaked at longer delays in the reach for the three identified clusters when plotted over normalized time from reach start to cricket occlusion.
 (F) The corresponding hand velocity traces for the same clusters reveal slower peak velocities correlating with later grasp opening.
 (G–I) Three frames from an example reach to grasp for each of the clusters reveals a power grasp enclosing the cricket; on the right, plots of the three knuckle markers used to measure finger spread at corresponding times.
 See also Figure S2.

in each of the three clusters illustrate the power grasp, with variation in the timing of grasp opening across the examples (Figures 2G–2I). In the third case with the slowest speed, the

fingers have still not fully separated by movement end (Figure 2I). In contrast, rodents often use an arpeggio grasp to first contact the object with fingers extended flatly, and then based on

somatosensory feedback, close around it.³³ Here, even the slowest reaches appear to involve opening of the grasp near contact (Figure 2E3) and thus are more consistent with a vision-guided power grasp strategy typical of macaque reaching.

Hand velocity converges to match lateral cricket velocity during reaching

Movement dynamics during prey capture have been found to follow control strategies for target interception, such as pure pursuit (PP) and proportional navigation (PN), across a range of species, including dragonflies,¹⁰ hawks,³⁴ and humans.³⁵ In PP, the rotational acceleration, or steering, of the body (or hand) is dictated by the instantaneous offset from the target location, called the range vector (RV). PP seeks to close the distance of the RV but ignores motion of the target and is not anticipatory of its future location. By contrast, PN aims to maintain a constant bearing angle of the RV as the pursuer closes the distance. One way to visualize the distinction between these strategies is to decompose velocity into a component that lies along the RV at each movement, the direct axis, and another component that lies on its orthogonal projection, the lateral axis. Both strategies predict changes in velocity related to closing distance on the direct axis, but in addition, PN predicts a matching of the velocity of the body (or hand) with the lateral velocity of the target to maintain a constant bearing angle. Another feature characteristic of PN is that the derivative of the RV over time will be negatively correlated at -1 with the RV itself.¹⁰

We examined the direct and lateral components of hand velocity as marmosets performed reaches to moving crickets in the arena (Figure 3). Although we are able to reconstruct the full 3D kinematics of reaching movements (Figure 3A), we focus our subsequent analyses on the horizontal plane seen from the top camera view, which captured most of the cricket motion. To examine the influence of cricket velocity on hand velocity, we decomposed velocity of the hand (H) and cricket (C) into direct and lateral components relative to the RV that connects the instantaneous position of the two (Figure 3B). In the set of reaches, most of the hand motion falls along the direct axis, Dh , which is necessary to traverse the arena to close on the cricket and exhibits a peaked shape typical of ballistic reaching movements (Figure 3C, dashed lines). Conversely, most of cricket motion falls along the lateral axis, Lc , due to the tendency of crickets to crawl along the perimeter of the reaching arena. The lateral component of absolute hand velocity, Lh , was closely matched in magnitude to cricket lateral velocity, Lc (Figure 3C, solid lines). We thus sought to determine to what extent the lateral hand velocity might match cricket velocity, both in sign and magnitude over the reach, as predicted during PN.

A typical reaching trial demonstrates that hand velocity tracks the lateral motion of the cricket target in a manner more consistent with a PN strategy for interception (Figures 3D–3F). In this trial, the cricket moves in the horizontal plane along the boundary of the arena and the hand adopts a curved trajectory in pursuit (Figure 3D). We applied computational methods from a previous study of hawk prey pursuit³⁴ to simulate the hand's trajectory using either a PP or PN steering model. For this trial, the PN model provides a closer match to the actual hand trajectory ($R^2 = 0.99$)

compared with the PP model ($R^2 = 0.94$) (Figure 3E). In line with the better fit of the PN model, the correlation of the RVs to their derivatives over time for this example trial takes on values near -1 , especially near the end of the reach as the hand converges upon the target. In PP, the lateral velocity is not necessarily matched, leading to overall weaker negative correlations. We find that the correlation increasingly converges toward -1 up to interception in this reach (Figure 3F).

When steering strategies are examined over the set of reaches for moving crickets, the PN model better explains the hand kinematics than the PP model. We applied PN and PP model simulations to the population of reaches and compared the RV correlations of the models against the actual movements. We find that the mean correlation coefficient converges toward -1 over the duration of reaches, as in the previous example trial, and that the correlation trace of the PN steering model better matches that of the hand than the PP model, especially as the hand nears interception (Figure 3G). To examine if the lateral velocity of the cricket was matched by hand velocity, we examined the mean of the lateral velocity of the hand multiplied by the sign of instantaneous cricket velocity, such that correspondence would give similar mean velocity and no correlation in direction would converge toward zero. This process was repeated for the lateral velocity components of PP and PN simulations. The mean lateral velocity of the hand, Lh , increasingly matches that of the cricket, Lc , over the time-course of reaches, which is very closely approximated by the PN model, but not the PP model (Figure 3H). Together these results provide evidence to support a PN strategy of hand guidance. However, a weakness of any of these steering models stems from their ability to compensate for visual processing delays, which represent a critical feature of any real biological or engineered system. Indeed, when we incorporate a visual delay of cricket position into the PP and PN models we find the goodness of fit for simulated kinematics break down progressively with visuo-motor delay (Figure 3I). A previous study found that a mixed pursuit strategy, including an additive combination of PP and PN guidance commands, was more robust than single strategy models.³⁴ However, even when implementing a mixed PP + PN strategy here, we find that performance matching the hand kinematics breaks down progressively with delay. Therefore, we sought to determine visuo-motor delay in marmoset reaching for this task and, furthermore, to test if an explicitly predictive control strategy would be able to better explain observed reach kinematics while using realistic delay values.

Hand velocity incorporates target predictions to compensate for visuo-motor delays

A key problem in guiding the hand during pursuit is that positional and velocity information about the target arrives at a processing delay in the visual system before it can influence actions. Motor control theories often assume predictive models that estimate future body and target positions to compensate for such delays in planning movements.^{12,36} In human experiments using double-step reaching tasks, in which an initial target turns off and a new illuminated target turns on mid-reach, there are visuo-motor delays ranging from 125 to 225 ms.^{37–40} By contrast, studies with head-restrained macaques performing double-step tasks have found substantially longer motor delays around 200–

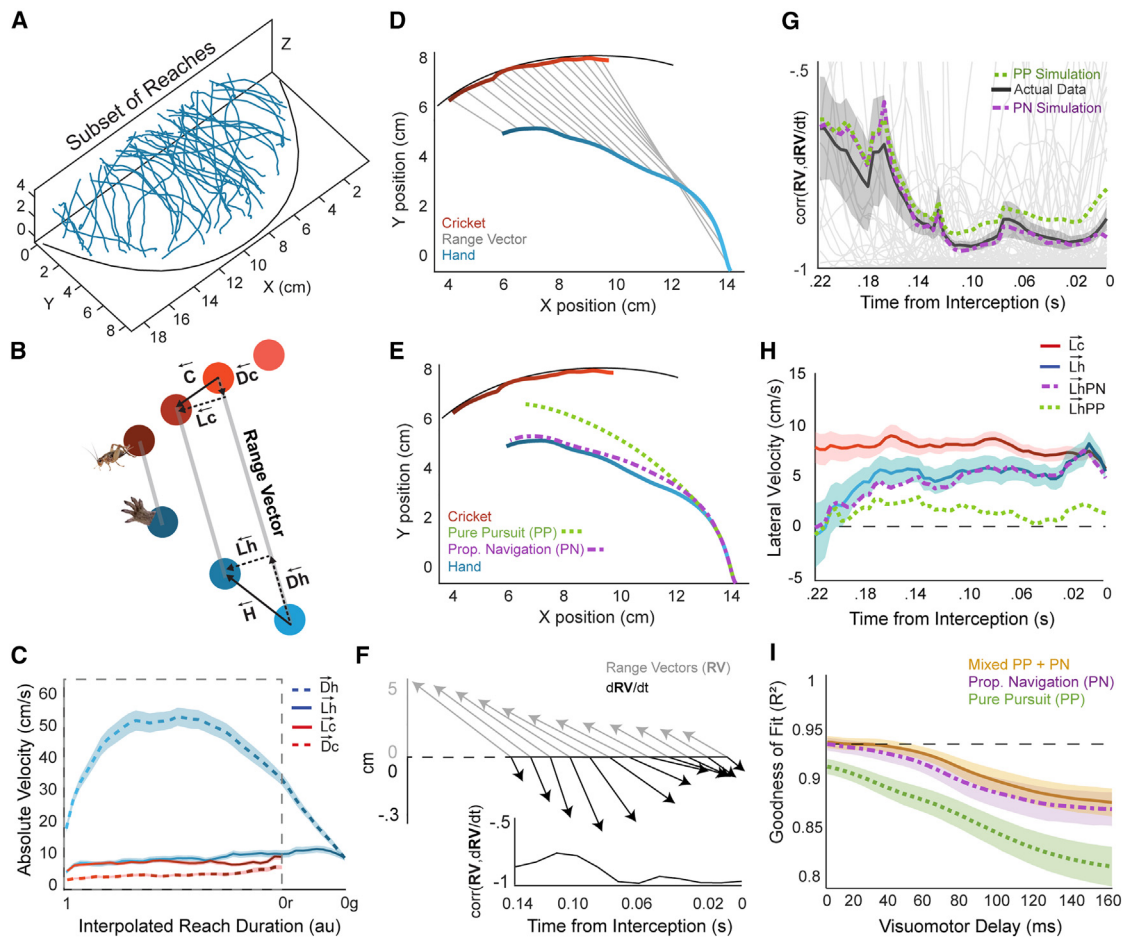


Figure 3. Hand velocity during reaching is consistent with proportional navigation

(A) Representative 3D reconstructions of the population of reaches for moving crickets.
 (B) At any moment, the hand and cricket velocity can be broken down into two components, the direct and lateral velocity. Direct velocity is the component projected along the range vector, whereas lateral is the orthogonal projection.
 (C) The average absolute cricket and hand velocity for the direct and lateral components as illustrated in (B) over interpolated time for the population of reaches. The dashed box indicates the period in which the cricket was visible. All error fields are ± 1 SEM.
 (D) Path of the hand (blue) and cricket (red) in an individual example reaching trial. Lighter points indicate the beginning of the reach. Gray lines indicate line of sight, or range vectors, for corresponding hand and cricket points, 1:4 downsampled.
 (E) A pure pursuit (PP) steering strategy aligns hand velocity with the range vector. Proportional navigation (PN) rotates hand velocity to mirror range vector rotations. Overlaid simulations of PP and PN show that PN best approximates this reach.
 (F) PN strategies predict matching the lateral component of the hand and cricket velocity while the range to the target is close, and if followed, the derivative of the range vector during pursuit should be in opposite direction of the range vector. For the example reach, the range vectors are shown in gray (matching gray lines in D) and their derivatives in black, with the correlation between vectors over time quantified below (black). The trace is aligned on the right to the moment of interception.
 (G) The correlations between the range vectors and their derivatives converge toward -1 on average across the population of 78 reaches (individual reaches shown as gray lines, average in black, ± 1 SEM in gray shade). This analysis was also performed for PP and PN simulations, showing that PN is a better match to actual data.
 (H) Lateral velocities from the hand and pursuit simulations were made positive when traveling in the same direction as cricket velocity. Hand lateral velocity increasingly approaches cricket lateral velocity throughout the reach. This is better matched by a PN simulation than a PP simulation. All error fields are ± 1 SEM.
 (I) Pursuit strategy simulations perform worse when a delay is introduced. We include a mixed model simulation, which uses a combination of PN and PP to steer hand velocity and is moderately more robust than proportional navigation alone. Goodness of fit quantified by R^2 values that compare model simulations with actual hand data. All error fields are ± 1 SEM.

300 ms.^{4,14–17} This slower timing may reflect constraints due to head-restraint and the artificial nature of the tasks. Thus, it was critical to first assess the visuo-motor delays in the current task.

Visuo-motor delays in dynamic reaching for cricket prey were much shorter than those found in previous non-human

primate studies, ranging between 80 and 100 ms. To estimate visuo-motor delay, we identified discrete cricket acceleration events across task trials in which the cricket increased its motion while the reach was ongoing. The lateral speed of the cricket was filtered to flag the top 5% of moments where

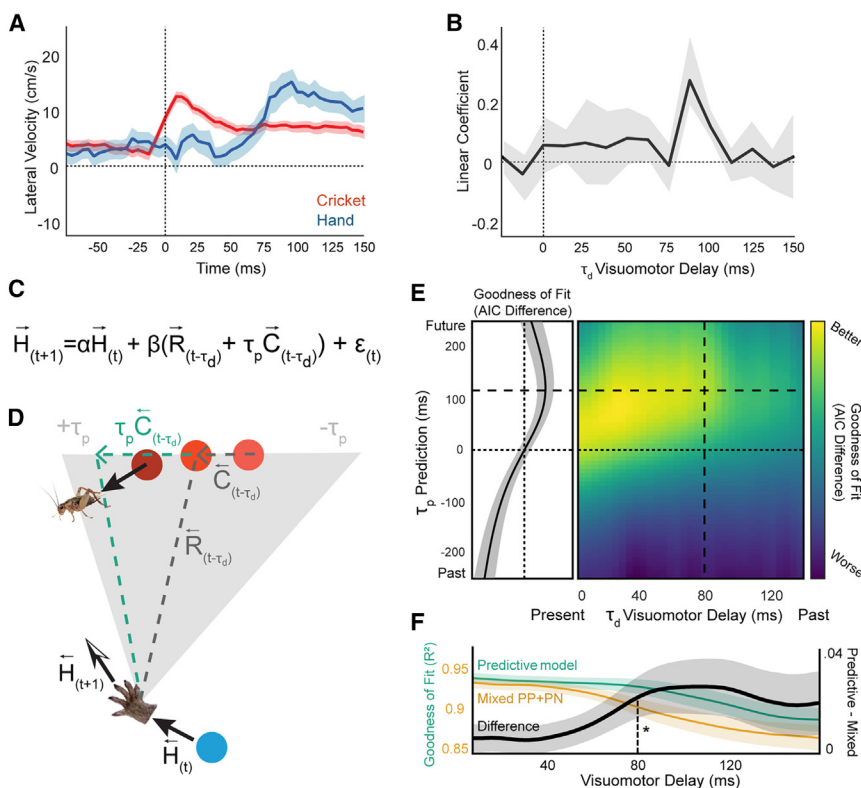


Figure 4. Cricket velocity predictively guides hand velocity to compensate for visuo-motor delay

(A) Increases in the lateral cricket velocity are followed by increases in lateral hand velocity at a latency of roughly 80 ms. A Difference of Gaussian filter was applied to the absolute velocity of cricket motion to identify discrete acceleration events ($n = 68$). The average cricket velocity (in red) time-locked to those events reflects increases in speed at time zero, as expected, whereas hand velocity follows with a comparable acceleration at a delay (blue), shown as ± 1 SEM.

(B) A linear regression of lateral hand velocity with the offset lateral cricket velocity from delays ranging from -25 to 150 ms reveals a significant regression coefficient between 80 and 100 ms. Error shown in 1 SEM estimated from a jackknife procedure.

(C) The multivariate linear model equation used to test the influence of visuo-motor delay τ_d and prediction time constant τ_p on the fit of hand velocity. (D) The model estimates the hand velocity at the next time step ($t + 1$) based on the range vector (R) at time $t - \tau_d$ and momentum from the previous hand velocity. A predictive term based on cricket velocity (C) at time $t - \tau_d$ (teal dashed lines) is multiplied by the prediction term τ_p to project the range vector into its estimated future location (teal dashed lines). The gray field represents the range of tested prediction τ_p values.

(E) The goodness of fit as a function of τ_d and τ_p is shown in color scale. Upon setting τ_d to 80 ms, the

best fit of this model is achieved using a τ_p prediction = 115 ms (indicated by intersection of black dashed lines). Color scale indicates the goodness-of-fit surface measured by a difference of AIC values between the full predictive model (equation as in C) and a model that only uses hand velocity (momentum term) as a predictor. Left: the goodness of fit with visuo-motor delay τ_d fixed at 80 ms and varying the prediction time constant τ_p is shown with 1 SEM estimated from a jackknife procedure.

(F) When this predictive pursuit model is applied in an approach to simulate hand trajectories (as in Figure 3), it provides a better fit (teal curve) that degrades more gradually than the best mixed PP + PN steering models (orange curve, left vertical axis). Colored traces indicate mean R^2 of simulation fits with 1 SEM fields. The model improvement for the predictive vs. best steering model (black curve, right vertical axis) at 80 ms delay and beyond is significantly better for the predictive models (sign rank test, $p = 0.0105$).

lateral velocity increased. We observed that immediately following an increase in lateral cricket velocity, there is a corresponding increase in lateral hand velocity in the same direction that begins around 80 ms (Figure 4A). This analysis, however, relies on discrete events in the total dataset. To determine if it was consistent throughout the entire dataset, we applied a simple linear regression on all the velocity data to predict lateral hand velocity based on the recent history of cricket lateral velocity (STAR Methods; Equation 4). This enabled us to fit a linear kernel for hand velocity based on cricket velocity (i.e., the impulse response function corrected for autocorrelations in cricket velocity). Consistent with the visuo-motor delay observed due to sudden jumps, the linear prediction found a significant prediction coefficient near 80 ms ($p < 0.0001$ at 87.5 ms and $p = 0.0559$ at 100 ms), but not at other delays (Figure 4B). This suggests that changes in hand velocity track the cricket velocity at relatively short delays around 80 – 100 ms, a value much faster than the 200 – 300 ms reported in constrained tasks.^{14–17} However, this value is within range of the delay seen in ocular following responses in marmosets, which is close to 75 ms.¹⁸ Of interest, similar to the reaches for live crickets, studies of oculomotor pursuit employ fixation targets that are moving continuously,

whereas the double-step reaching studies in macaques used instantaneous jumps in target positions.

Although the visuo-motor delays are briefer in the current task than previous studies, there may still be significant advantages for taking them into account by planning predictive movements that incorporate the cricket's velocity. To test the role of prediction in guiding the hand, we fit an auto-regressive model of hand velocity that included terms for the visuo-motor delay, τ_d , as well as a prediction time constant, τ_p , that dictated how far into the future the cricket's position was projected based on its velocity. We followed a generative model formulation from a previous study in macaques,⁴¹ which described the hand velocity in terms of a momentum term plus a scaled version of a RV, $R(t - \tau_d)$, at some visuo-motor delay (see equation; Figure 4C). To compensate for delays in the RV (cyan dashed line, Figure 4D) the model includes a term to estimate the cricket's future position, $\tau_p C(t - \tau_d)$, where τ_p is a linear temporal weight that determines how far in the future to adjust for the cricket's future position (dark gray dashed line, Figure 4D). We pooled data from 78 reaches for moving crickets and tested the predictive model fit for visuo-motor delays τ_d over a range from 0 to 145 ms in the past and prediction time constant τ_p over a range from 250 ms in the past up to 250 ms in the future. We assessed model fit

by the difference of Akaike information criterion (AIC)⁴² values for the full model (including cricket velocity prediction) against a baseline model that only contains information about hand velocity (momentum term).

Auto-regressive modeling indicates that hand velocity is best explained by a model in which visuo-motor delay is compensated for by using predictions based on cricket velocity. The AIC difference, or goodness of fit, between the full predictive model and the hand velocity model based on momentum showed the best fit with a positively correlated diagonal structure, reflecting that positive temporal prediction constants were matched to compensate for visuo-motor delays (yellow band in Figure 4E). When we set delay to 80 ms as determined above (Figures 4A and 4B), the best fit is achieved using a prediction constant of 115 ms into the future, (vertical dashed line, Figure 4E). To visualize standard error in these surface estimates, the cross-section at that 80 ms visuo-motor delay was examined as a function of the temporal prediction parameter, with its confidence intervals estimated using a jackknife procedure (left panel, Figure 4E). The prediction model demonstrated a significant improvement in model performance for temporal constants weighted into the future. To obtain a more direct comparison of this model against the PP, PN, and PP + PN mixed steering models (Figure 3I), we tested this prediction model using a simulation approach directly comparable to earlier analyses fitting hand kinematics, and we tested a range of visuo-motor delays. When compared against the best steering model, the PP + PN mixed strategy, the predictive model outperformed it at each delay, showing a more gradual breakdown in performance for longer visuo-motor delays (Figure 4F, cyan versus orange curves). Taking the difference in performance between models, the full predictive pursuit model significantly outperformed the mixed PP + PN steering model at 80 ms (sign rank test, $p = 0.0105$) and longer delays (Figure 4F, black trace). These findings support that a predictive term compensating for visuo-motor delay provides a better description of reaching kinematics at realistic delays. Furthermore, although steering models such as PN successfully describe marmoset reaching for moving crickets when they are formulated with no processing delays (Figure 3), they lack the ability to do so given systematic delays characteristic of the real system.

DISCUSSION

The present findings demonstrate that, during an ecologically motivated reaching paradigm for moving crickets, marmoset monkeys utilize a power grasp that opens earlier in the reach for faster velocity reaches. Furthermore, marmosets guide their hand movements at very brief visuo-motor delays, comparable to oculomotor pursuit. In an examination of hand guidance steering models, we find that PN explains hand motion reasonably well and provides better fits than a PP model (Figure 3). However, these models perform worse in the presence of realistic visuo-motor delays. To explain the reach kinematics given a realistic visuo-motor delay, which we estimate to be around 80 ms, we find that it is necessary to use a predictive pursuit model that incorporates a term to extrapolate the future location of the target based on cricket velocity. This predictive pursuit model outperforms the PN or mixed strategy model³⁴ at the estimated delay.

We find that the prediction constant closely matches the scale of the visuo-motor delay, extending slightly beyond it about 35 ms into the future. Such a prediction constant, which remains close to or slightly beyond the visuo-motor delay, is often close to optimal for maintaining stability and achieving optimal solutions in Bayesian integration⁴³ or Smith prediction.⁴⁴

The current finding emphasizes that moving toward realistic dynamics in experimental designs can⁴⁵ better inform our understanding of reaching behavior and provide insights into the underlying neural correlates and the evolutionary contexts that have driven their development. A previous study also found that primates use prediction in a pursuit task involving a simulated evading virtual target that was captured by a screen cursor under joystick control.⁴¹ In our study we have adapted their approach to modeling prediction, but have done so within a task in which reaching movements unfold at natural timescales. The former study identified predictive terms beyond 500 ms into the future, which match the delays imposed by using a joystick for control in their task. Here, we find that similar predictive strategies are involved to compensate delays when reaching for live prey, but that they occur at much shorter latencies where visuo-motor delays are comparable to oculomotor pursuit and under 100 ms. Although the present findings were sufficient to characterize the visuo-motor strategies in the current reaching task, future studies could leverage recent advances in head mounted eye tracking^{45,46} to characterize coordination between head and eye gaze in guiding predictive reaching.

Visually guided insect predation and fruit/leaf eating in an arboreal environment are hypothesized to play a role in the evolution of primates.^{47,48} Although not ubiquitous, insect hunting is not only observed in New World monkeys, such as marmosets, but also in Old World monkeys and hominoids.^{49,50} Fronto-parietal neural circuits for motor control in primates are thought to have evolved in primates to optimize visuo-motor strategies for coordination and speed of reach-to-grasp movements.⁵¹ Much evidence suggest that motor commands are sent via feedback pathways from premotor cortex to higher order sensory areas in parietal cortex and could mediate the forward models necessary to implement predictive on-line control.^{36,52} The current experimental paradigm shows that predictive strategies in reaching emerge naturally, even within a laboratory context for dynamic targets, such as crickets, affording new opportunities to study the underlying neural mechanisms within an ethologically motivated paradigm.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Reaching task and video recording
 - Video preparation and DeepLabCut pose estimation

● QUANTIFICATION AND STATISTICAL ANALYSIS

- Reaching Kinematic Data Preparation
- Grasp cluster analyses
- Pursuit models and velocity decomposition
- Predictive modeling and reaching analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cub.2023.05.032>.

ACKNOWLEDGMENTS

We would like to thank Dina Graf and members of the Mitchell lab for help with marmoset care and handling. We also thank Cory Miller, Alex Huk, and Cris Niell for comments on this manuscript. This work was supported by NIH grant UF1NS116377 (J.M. and L.S.) and a grant from the Schmitt Program on Integrative Neuroscience (K.H.W. and J.M.) and the Del Monte Institute for Neuroscience at the University of Rochester (K.H.W. and L.S.).

AUTHOR CONTRIBUTIONS

Conceptualization, formal analysis, investigation, and methodology, all authors; data acquisition, L.S.; writing – review & editing, all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: October 14, 2022

Revised: March 31, 2023

Accepted: May 15, 2023

Published: June 5, 2023

REFERENCES

1. Sustaita, D., Pouydebat, E., Manzano, A., Abdala, V., Hertel, F., and Herrel, A. (2013). Getting a grip on tetrapod grasping: form, function, and evolution. *Biol. Rev. Camb. Philos. Soc.* 88, 380–405. <https://doi.org/10.1111/brv.12010>.
2. Smith, J.M., and Smith, A.C. (2013). An investigation of ecological correlates with hand and foot morphology in callitrichid primates. *Am. J. Phys. Anthropol.* 152, 447–458. <https://doi.org/10.1002/ajpa.22378>.
3. César, J.C., and Bicca-Marques, J.C. (1999). Hand specialization, sympatry, and mixed-species associations in callitrichines. *J. Hum. Evol.* 36, 349–378. <https://doi.org/10.1006/jhev.1998.0272>.
4. Archambault, P.S., Ferrari-Toniolo, S., and Battaglia-Mayer, A. (2011). Online control of hand trajectory and evolution of motor intention in the parietofrontal system. *J. Neurosci.* 31, 742–752. <https://doi.org/10.1523/JNEUROSCI.2623-10.2011>.
5. Ngo, V., Gorman, J.C., De la Fuente, M.F., Souto, A., Schiel, N., and Miller, C.T. (2022). Active vision during prey capture in wild marmoset monkeys. *Curr. Biol.* 32, 3423–3428.e3. <https://doi.org/10.1016/j.cub.2022.06.028>.
6. Catania, K.C. (2009). Tentacled snakes turn C-starts to their advantage and predict future prey behavior. *Proc. Natl. Acad. Sci. USA* 106, 11183–11187.
7. Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.* 4, 423–431. [https://doi.org/10.1016/s1364-6613\(00\)01537-0](https://doi.org/10.1016/s1364-6613(00)01537-0).
8. Fujioka, E., Aihara, I., Sumiya, M., Aihara, K., and Hiryu, S. (2016). Echolocating bats use future-target information for optimal foraging. *Proc. Natl. Acad. Sci. USA* 113, 4848–4852. <https://doi.org/10.1073/pnas.1515091113>.
9. Merchant, H., and Georgopoulos, A.P. (2006). Neurophysiology of perceptual and motor aspects of interception. *J. Neurophysiol.* 95, 1–13. <https://doi.org/10.1152/jn.00422.2005>.
10. Mischiati, M., Lin, H.T., Herold, P., Imler, E., Olberg, R., and Leonardo, A. (2015). Internal models direct dragonfly interception steering. *Nature* 517, 333–338. <https://doi.org/10.1038/nature14045>.
11. Shadmehr, R., Smith, M.A., and Krakauer, J.W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* 33, 89–108. <https://doi.org/10.1146/annurev-neuro-060909-153135>.
12. Wolpert, D.M., and Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Netw.* 11, 1317–1329. [https://doi.org/10.1016/s0893-6080\(98\)00066-5](https://doi.org/10.1016/s0893-6080(98)00066-5).
13. Georgopoulos, A.P., Kalaska, J.F., Caminiti, R., and Massey, J.T. (1983). Interruption of motor cortical discharge subserving aimed arm movements. *Exp. Brain Res.* 49, 327–340. <https://doi.org/10.1007/BF00238775>.
14. Georgopoulos, A.P., Kalaska, J.F., and Massey, J.T. (1981). Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *J. Neurophysiol.* 46, 725–743. <https://doi.org/10.1152/jn.1981.46.4.725>.
15. Battaglia-Mayer, A., Ferrari-Toniolo, S., Visco-Comandini, F., Archambault, P.S., Saberi-Moghadam, S., and Caminiti, R. (2013). Impairment of online control of hand and eye movements in a monkey model of optic ataxia. *Cereb. Cortex* 23, 2644–2656.
16. Archambault, P.S., Caminiti, R., and Battaglia-Mayer, A. (2009). Cortical mechanisms for online control of hand movement trajectory: the role of the posterior parietal cortex. *Cereb. Cortex* 19, 2848–2864. <https://doi.org/10.1093/cercor/bhp058>.
17. Dickey, A.S., Amit, Y., and Hatsopoulos, N.G. (2013). Heterogeneous neural coding of corrective movements in motor cortex. *Front. Neural Circuits* 7, 51. <https://doi.org/10.3389/fncir.2013.00051>.
18. Cloherty, S.L., Yates, J.L., Graf, D., DeAngelis, G.C., and Mitchell, J.F. (2020). Motion perception in the common marmoset. *Cereb. Cortex* 30, 2658–2672. <https://doi.org/10.1093/cercor/bhz267>.
19. Kaas, J.H., Gharbawie, O.A., and Stepniewska, I. (2011). The organization and evolution of dorsal stream multisensory motor pathways in primates. *Front. Neuroanat.* 5, 34. <https://doi.org/10.3389/fnana.2011.00034>.
20. Kaas, J.H., and Stepniewska, I. (2016). Evolution of posterior parietal cortex and parietal-frontal networks for specific actions in primates. *J. Comp. Neurol.* 524, 595–608. <https://doi.org/10.1002/cne.23838>.
21. Sacrey, L.A.R., and Whishaw, I.Q. (2012). Subsystems of sensory attention for skilled reaching: vision for transport and pre-shaping and somatosensation for grasping, withdrawal and release. *Behav. Brain Res.* 231, 356–365.
22. De Bruin, N., Sacrey, L.R., Brown, L.A., Doan, J., and Whishaw, I.Q. (2008). Visual guidance for hand advance but not hand withdrawal in a reach-to-eat task in adult humans: reaching is a composite movement. *J. Mot. Behav.* 40, 337–346.
23. Wang, X., Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Y., Williams, P.R., Alwahaib, N.S.A., Kapur, K., Yu, B., et al. (2017). Deconstruction of corticospinal circuits for goal-directed motor skills. *Cell* 171, 440–455.e14. <https://doi.org/10.1016/j.cell.2017.08.014>.
24. Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21, 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>.
25. Karashchuk, P., Rupp, K.L., Dickinson, E.S., Walling-Bell, S., Sanders, E., Azim, E., Brunton, B.W., and Tuthill, J.C. (2021). Anipose: a toolkit for robust markerless 3D pose estimation. *Cell Rep.* 36, 109730. <https://doi.org/10.1016/j.celrep.2021.109730>.

26. Hook, M.A., and Rogers, L.J. (2008). Visuospatial reaching preferences of common marmosets (*Callithrix jacchus*): an assessment of individual biases across a variety of tasks. *J. Comp. Psychol.* *122*, 41–51. <https://doi.org/10.1037/0735-7036.122.1.41>.
27. Fox, D.M., Mundinano, I.C., and Bourne, J.A. (2019). Prehensile kinematics of the marmoset monkey: implications for the evolution of visually-guided behaviors. *J. Comp. Neurol.* *527*, 1495–1507.
28. Castiello, U., and Dadda, M. (2019). A review and consideration on the kinematics of reach-to-grasp movements in macaque monkeys. *J. Neurophysiol.* *121*, 188–204. <https://doi.org/10.1152/jn.00598.2018>.
29. Sartori, L., Camperio Ciani, A., Bulgheroni, M., and Castiello, U. (2013). Reaching and grasping behavior in *Macaca fascicularis*: a kinematic study. *Exp. Brain Res.* *224*, 119–124. <https://doi.org/10.1007/s00221-012-3294-2>.
30. Marzke, M.W. (1997). Precision grips, hand morphology, and tools. *Am. J. Phys. Anthropol.* *102*, 91–110. [https://doi.org/10.1002/\(SICI\)1096-8644\(199701\)102:1<91::AID-AJPA8>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1096-8644(199701)102:1<91::AID-AJPA8>3.0.CO;2-G).
31. Bishop, A. (1962). Control of the hand in lower primates. *Ann. NY Acad. Sci.* *102*, 316–337. <https://doi.org/10.1111/j.1749-6632.1962.tb13649.x>.
32. Schiel, N., and Souto, A. (2017). The common marmoset: an overview of its natural history, ecology and behavior. *Dev. Neurobiol.* *77*, 244–262.
33. Whishaw, I.Q., and Gorny, B. (1994). Arpeggio and fractionated digit movements used in prehension by rats. *Behav. Brain Res.* *60*, 15–24.
34. Brighton, C.H., and Taylor, G.K. (2019). Hawks steer attacks using a guidance system tuned for close pursuit of erratically manoeuvring targets. *Nat. Commun.* *10*, 2462. <https://doi.org/10.1038/s41467-019-10454-z>.
35. Lenoir, M., Musch, E., Janssens, M., Thiery, E., and Uyttenhove, J. (1999). Intercepting moving objects during self-motion. *J. Mot. Behav.* *31*, 55–67. <https://doi.org/10.1080/00222899909601891>.
36. Shadmehr, R., and Krakauer, J.W. (2008). A computational neuroanatomy for motor control. *Exp. Brain Res.* *185*, 359–381. <https://doi.org/10.1007/s00221-008-1280-5>.
37. De Brouwer, A.J., and Spering, M. (2022). Eye-hand coordination during online reach corrections is task-dependent. *J. Neurophysiol.* *127*, 885–895.
38. Brenner, E., and Smeets, J.B. (1997). Fast responses of the human hand to changes in target position. *J. Mot. Behav.* *29*, 297–310. <https://doi.org/10.1080/00222899709600017>.
39. Day, B.L., and Lyon, I.N. (2000). Voluntary modification of automatic arm movements evoked by motion of a visual target. *Exp. Brain Res.* *130*, 159–168. <https://doi.org/10.1007/s002219900218>.
40. Prablanc, C., and Martin, O. (1992). Automatic control during hand reaching at undetected two-dimensional target displacements. *J. Neurophysiol.* *67*, 455–469.
41. Yoo, S.B.M., Tu, J.C., Piantadosi, S.T., and Hayden, B.Y. (2020). The neural basis of predictive pursuit. *Nat. Neurosci.* *23*, 252–259.
42. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceedings of the 2nd International Symposium on Information Theory*, B. Petrov, and F. Csaki, eds. (Akadémiai Kiadó), pp. 267–281.
43. Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* *427*, 244–247. <https://doi.org/10.1038/nature02169>.
44. Read, J.C.A., Kaspiris-Rousellis, C., Wood, T.S., Wu, B., Vlaskamp, B.N.S., and Schor, C.M. (2022). Seeing the future: predictive control in neural models of ocular accommodation. *J. Vis.* *22*, 4. <https://doi.org/10.1167/jov.22.9.4>.
45. Parker, P.R.L., Abe, E.T.T., Leonard, E.S.P., Martins, D.M., and Niell, C.M. (2022). Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron* *110*, 3897–3906.e5. <https://doi.org/10.1016/j.neuron.2022.08.029>.
46. Meyer, A.F., Poort, J., O’Keefe, J., Sahani, M., and Linden, J.F. (2018). A head-mounted camera system integrates detailed behavioral monitoring with multichannel electrophysiology in freely moving mice. *Neuron* *100*, 46–60.e7. <https://doi.org/10.1016/j.neuron.2018.09.020>.
47. Sussman, R.W., Tab Rasmussen, D., and Raven, P.H. (2013). Rethinking primate origins again. *Am. J. Primatol.* *75*, 95–106. <https://doi.org/10.1002/ajp.22096>.
48. Cartmill, M. (1974). Rethinking primate origins. *Science* *184*, 436–443. <https://doi.org/10.1126/science.184.4135.436>.
49. McGrew, W.C. (2014). The ‘other faunivory’ revisited: insectivory in human and non-human primates and the evolution of human diet. *J. Hum. Evol.* *71*, 4–11. <https://doi.org/10.1016/j.jhevol.2013.07.016>.
50. Raubenheimer, D., and Rothman, J.M. (2013). Nutritional ecology of entomophagy in humans and other primates. *Annu. Rev. Entomol.* *58*, 141–160. <https://doi.org/10.1146/annurev-ento-120710-100713>.
51. Karl, J.M., and Whishaw, I.Q. (2013). Different evolutionary origins for the reach and the grasp: an explanation for dual visuomotor channels in primate parietofrontal cortex. *Front. Neurol.* *4*, 208. <https://doi.org/10.3389/fneur.2013.00208>.
52. Wagner, M.J., and Smith, M.A. (2008). Shared internal models for feedforward and feedback control. *J. Neurosci.* *28*, 10663–10673.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCES	IDENTIFIERS
Software and algorithms		
MATLAB	The MathWorks	9.13.0 (R2022b)
Marmoset Fast Reach custom analysis scripts	This paper; Github	https://doi.org/10.5281/zenodo.7884202
Hawk Pursuit Simulation Scripts	Brighton and Taylor ³⁴	Data S1
DeepLabCut	Mathis et al. ²⁴	https://github.com/DeepLabCut
Anipose	Karashchuk et al. ²⁵	https://github.com/lambdaloop/anipose
FFmpeg	FFmpeg Developers	ffmpeg.org
Deposited data		
Marmoset reaching data	This paper; Zenodo	https://doi.org/10.5281/zenodo.7869286

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Jude Mitchell (jmitch27@ur.rochester.edu), kuanhong_wang@urmc.rochester.edu.

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All data reported in this paper is publicly available at Zenodo: <https://doi.org/10.5281/zenodo.7869286>
- All original code in this paper is publicly available at Zenodo: <https://doi.org/10.5281/zenodo.7884202> and at Github: <https://github.com/lukehshaw/marmoset-fast-reach/tree/v1.4>

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All experimental procedures described here are approved by the University Committee on Animal Resources at University of Rochester. 10 common marmoset monkeys (*Callithrix jacchus*: 5 females, 5 males) were recorded by multiple cameras while reaching for live crickets (*Acheta domesticus*) and cheerios. Their ages ranged from 1.2 years old to 7.1 years old. The crickets used in experiments served as additional enrichment to their daily diet.

METHOD DETAILS

Reaching task and video recording

Reaching experiments were conducted by docking a mobile video recording platform to marmoset family cages within the marmoset colony and recording reaches using 3 GoPro Hero 7 Black cameras (Figure 1B). Marmosets were acclimated to the experimental apparatus and allowed to become comfortable reaching for cheerios as food treats, and then also for crickets. Acclimation was typically less than 30 minutes and was necessary only on the first day of experiments. To calibrate cameras prior to reaching experiments, we recorded a brief video of an 8x6 checkerboard array that was moved and rotated continuously within the reach arena (Figures 1B and 1D). Frames of this video were extracted and fed into triangulation algorithms provided in the Anipose toolkit.²⁵ A manually triggered photography flash aimed at the cameras was used to deliver a synchronization signal to all three cameras. This created 1-2 frames of video that were significantly over-exposed and easily identified using custom Matlab (Mathworks) and ffmpeg (Ffmpeg Developers) scripts.

Up to 15 cricket reaches were recorded from an individual marmoset on a given recording day. A sliding plate at the entry of the reaching enclosure was closed to isolate individual marmosets briefly from their family groups on the test platform. Reaching trials began by sliding an opaque acrylic plate upwards that occluded the view of the reaching platform, thus allowing the marmoset to view the target and begin hunting. The marmoset was able to reach through a large rectangular aperture (9cm x 13.5cm) that contained two laterally traversing 2mm stainless-steel bars at 3cm and 6cm height and was built into an acrylic plate. This aperture allowed for a

wide range of reach start points while also preventing the animal from getting outside the test chamber (Figure 1B). Before lifting the clear acrylic sheet to allow reaching to begin, the crickets were dropped onto the platform and allowed to move freely. Crickets were kept from crawling off the platform with a 4cm tall acrylic wall molded into an arc and cemented to the platform with a center maximum y-dimension of 9cm and base maximum x-dimension of 20cm. The platform was sanded to eliminate reflections and increase hand and target contrast. Reflection reduction was useful due to the competition of the features of a tracked subject's reflection against the features of the subject itself.

Our video recordings were optimized to capture high velocity ballistic reaches. The GoPro Hero 7 Black was specifically chosen for its high-speed capture capability and manual control of recording settings. High frame rate (1/240s) and high shutter speeds (1/960s) were sufficient to render all hand features in full detail without observable motion blur. To attain proper exposure at this high shutter speed and to avoid shining a bright light directly into the eyes of our subjects, a LED light source (Nanlite Forza 60), was positioned to project light at a 65-degree angle 45cm from the reaching platform (Figure 1B). Light intensity at the platform was measured to be 4.6E5 Lux, comparable to direct sun on a bright day. The cameras were arranged around the reaching arena to provide a view from above and a downward angled view from the right and left (Figures 1E–1G). The cameras were fixed in position using articulating camera brackets mounted to an aluminum breadboard.

Video preparation and DeepLabCut pose estimation

Videos were preprocessed by the experimenter to indicate which video frames contained reaches for crickets and to provide meta-data for each reach. Reach success, handedness, and whether reaches were composed of a single volley or multiple volleys were documented for each reach in the data workflow. We defined volley multiplicity by whether the hand returned fully to the reach aperture opening on the platform before extending again. Selected frames spanned from ~200ms prior to the hand moving through the reaching aperture to full retraction of the arm, which was typically accompanied by delivery of the target to the mouth. These data were used to generate standalone video files for each reach that were synchronized for all camera views. Adobe Premiere Pro was used for video viewing and scoring. Videos were cropped to 760x720 pixels to remove pixel areas with irrelevant information. DeepLabCut 2.1.8.2 performed feature tracking from the reaching videos.

DeepLabCut labeling utilized a marker set that was chosen to track the position of fingers, hands, cricket, and features of the reaching platform. We trained the pre-trained ResNet50 network supplied by DeepLabCut using 574 manually labeled, manually selected frames of video. The set of frames was selected to fully characterize the diversity of marmosets, camera angles, hand positions, and targets of our reaching video data. The final model was trained for 800k iterations and used the imgaug augmentation pack. DeepLabCut was deployed through Anipose, which generated unfiltered 2D tracking files using DeepLabCut algorithms. In some later recordings, an unexpected minor shift in the top-view camera position (Figure 1F) offset the accuracy of the original triangulation calibration. To correct for this, we re-aligned the 2D DeepLabCut tracking data from each reach video segment using a linear transformation matrix generated by comparing the arena feature markers (Figure 1D) taken from the original calibration videos to the arena feature markers of each reaching video. We measured the distance each point was moved to confirm that transformation did not introduce unintended scaling artifacts. These steps were performed using custom Matlab scripts after converting tracking data files to .csv files. All tracked points were transformed accordingly, re-packaged into .hd5 files using custom Python scripts, and then triangulated into 3D space using Anipose.

QUANTIFICATION AND STATISTICAL ANALYSIS

Reaching Kinematic Data Preparation

Manual reach scoring provided metadata for each reach, which guided the identification of video segments with hand extension and target motion position data. Position estimates with less than a .7 estimation likelihood were excluded from analyses, which is a metric provided by DeepLabCut indicating label accuracy probability. Hand and cricket tracking points were filtered using a 5 sample Gaussian smoothing window. We applied a 180-degree rotation of right-handed reaches and the cricket position about the reach platform midline to reduce any potential effects of any handedness-related directionality biases.

We defined a reach as the first tracked appearance of the central hand point beyond the reach aperture plane to the last tracked point of the cricket. Cricket tracking at the very end of reaches, during deceleration and grasp, was typically lost due to occlusion by the fingers or hand. Windows of reaching data were defined by viewing videos with overlaid tracked points and marking the beginning and end frame number. These frame number values were used to refer directly to the .csv tracking data files in custom Matlab analysis scripts. Reaches for crickets that moved greater than a 0.5 cm distance during the reach and moved continuously were included in analyses. Crickets that moved this distance by a jump or that ceased motion during the reach were not included in the data set, nor were crickets that moved less than 0.5cm in total.

Grasp cluster analyses

Reaches were selected by visualizing all the tracked points of each of the 78 selected reaches with moving crickets and excluding 14 reaches with poor knuckle and finger tracking. Each reach was interpolated to a 51-sample vector where each point is equal to the mean distance between the pointer and middle knuckle and middle and ring knuckle markers. Knuckle markers were chosen because they were better tracked and less subject to perturbation by the cricket retaining wall. Each finger separation vector was z scored and the population was subjected to a kmeans cluster analysis. The maximum number of clusters was set manually based

on testing 2-6 clusters. We repeated clustering using different initial centroids 10 times. The results of this process were nearly identical to a hierarchical wavelet-based clustering process performed using the `mdwtcluster` function provided by Matlab.

Pursuit models and velocity decomposition

To investigate the degree to which marmoset reaching adheres to steering guidance laws such as pure pursuit and proportional navigation, we adopted Matlab analysis scripts used to investigate hawk pursuit of moving targets.³⁴ These scripts provided the framework to simulate marmoset reaches by applying the sample-by-sample steering commands while matching the magnitude of each step to the measured actual hand speed. Simulation fits to tracked hand data were measured using the R-squared between the simulated and actual hand velocity across each reach, then pooled across all reaches. To set initial conditions for simulated reaches, we matched simulated hand velocity in the first 75ms of the reach to the mean of the actual hand velocity averaged over the same period. Following Brighton and Taylor,³⁴ the proportional navigation guidance law utilized by these scripts is defined as

$$\dot{\gamma}(t)_{pn} = N\lambda\lambda\lambda\lambda \quad (\text{Equation 1})$$

Where $\dot{\gamma}(t)_{pn}$ is turning rate, N is an experimentally derived gain constant, $\dot{\lambda}(t)$ is the line of sight rate, t is time, and τ is a delay parameter. Proportional pursuit is defined as

$$\dot{\gamma}(t)_{pp} = -K\delta(t - \tau) \quad (\text{Equation 2})$$

Where $\dot{\gamma}(t)_{pp}$ is turning rate, K is an experimentally derived gain constant, δ is the range vector angle with respect to the hand's velocity, t is time, and τ is a delay parameter. A mixed strategy is defined as the sum of the two above terms as follows:

$$\dot{\gamma}(t)_{pnp} = \dot{\gamma}(t)_{pn} + \dot{\gamma}(t)_{pp} \quad (\text{Equation 3})$$

Model parameters for gain were set by maximizing the total R-squared for hand velocity against model hand velocity. All reach trials ($N=78$) were included to fit a gain parameter that was fixed across trials. Confidence intervals on the R-squared performance were computed from the standard deviation of trial-by-trial R-squared performance.

Proportional navigation strategies yield parallel pursuit, which can be described as matching velocity orthogonal to the line of sight, or range vector. To investigate the behavior of range-vector-relative decomposed velocities in the above steering model frameworks, target and hand velocity vectors were decomposed into orthogonal velocity components relative to a range vector, which connects hand position to cricket position. We refer to the velocity component parallel to the range vector as *direct* velocity. We refer to the velocity component orthogonal to the range vector as *lateral* velocity. This distinction is important due to the typical movement of crickets along the perimeter of the reaching platform. When cricket velocity is decomposed in this way, the majority of cricket velocity is contained within the *lateral* velocity component. A summary of this decomposition is provided in Figure 3D.

A proportional navigation analysis was conducted according to methods used in dragonfly pursuit.¹⁰ Range vectors were calculated as a subtraction of hand position from cricket position. The time derivative of the series of range vectors was calculated and correlated with the range vector itself using cosine vector correlation. Proportional navigation predicts a negative correlation of -1 between the range vector and its derivative.

Predictive modeling and reaching analyses

We examined if sudden increases in cricket velocity during reaching led to corresponding increases in hand velocity. To identify increases in cricket velocity we filtered the lateral speed of the cricket (absolute value of lateral velocity) using a difference filter that was defined by a Gaussian kernel ($\sigma = 20\text{ms}$) with negative amplitude at negative lags and positive amplitude as positive lags. This filter produces the largest values for epochs where low speed is directly followed by an increase in speed, regardless of whether speed increase is due to the cricket moving to the left or right. We used a threshold to flag the top 5% of the filtered values and identify peaks, ensuring that no two peaks occurred within 100 ms of each other. From 78 total reach trials we flagged 68 events. We averaged the lateral cricket velocity and hand velocity traces from -100 to 150 ms time-locked around these events (Figure 4A). Results were similar if we relaxed our acceleration threshold to include the top 10% or 20% of events (not shown).

Decomposed hand and cricket velocity components were used to investigate the predictive influence of cricket velocity on hand velocity. We first tested a linear prediction in which the lateral hand velocity was predicted by a weighted linear sum of the lateral cricket velocity history. The linear predictor equations was defined as

$$H(t) = \sum^{\tau} \kappa_{\tau} C(t - \tau) + \epsilon(t) \quad (\text{Equation 4})$$

where $H(t)$ represents lateral hand velocity at time t , C is the lateral cricket velocity at time t minus time delay τ and κ_{τ} is a linear prediction coefficient at this time delay. To reduce the number of parameters we down-sampled the velocity traces from 240 to 80 hertz, giving 12.5ms steps. We tested a range of τ values from -25 to 150ms. We defined the range vector used to separate direct and lateral velocities as the vector from the hand to cricket at the start of the reach (when the hand first entered the arena). We considered the case of updating the range vector at each delay during the reach as in the full auto-regressive model discussed below and settled upon this approach for the sake of simplicity.

A recent study examined a pursuit strategy embedded in a 2D coordinate system with a term incorporating prey velocity to test for the role of prediction.⁴¹ To apply the same model, we employed an auto-regressive linear regression to fit parameters in Equation 5

shown below which includes a momentum term for hand velocity, and a term based on the predicted range vector. The parameters for visuo-motor delay, τ_d , and a prediction time constant, τ_p , were fit to optimize the match to actual hand velocity. This analysis was evaluated using the Matlab *mvregress* function. Concatenated velocity and range vector magnitude values from each reach were supplied to the regression function. Time is defined by frames of video, each of which captures 1/240 second, or 4.2ms of behavior.

$$H_{(t+1)} = \alpha H_{(t)} + \beta (R_{(t-\tau_d)} + \tau_p C_{(t-\tau_d)}) + \varepsilon_{(t)} \quad (\text{Equation 5})$$

In the above Equation 5, $H_{(t+1)}$ represents hand velocity one frame of video into the future (t+1) relative to the current $H_{(t)}$ velocity at time t. The velocity of the cricket C at a delayed time, t- τ_d , is multiplied by the prediction time constant, τ_p , to project cricket position into the future or past. We also include a variable representing the range vector from hand to the cricket position as R. This vector connects hand position at time t and cricket position at t- τ_d . The sequence of operations to generate these variables was to first temporally offset tracked cricket position values according to t- τ_d . Velocities and range vectors were calculated. Finally cricket velocities were multiplied by τ_p . The predictor variables are weighted by a set of coefficients α and β , respectively. All terms in this equation are composed of cartesian x and y constituents and thus are bivariate. These data were submitted to a least-square regression where $\varepsilon_{(t)}$ represents the residual error.

We varied the values of τ_p and τ_d to determine the effects of visuomotor delay and prediction on model fit. The parameter τ_d was varied from 0 to 35 frames of video (equivalent to 0ms to 146ms) such that time shifted cricket position values were aligned to present hand values. Velocities and range vectors were calculated after this offset was applied. Following this we multiplied cricket velocity by a range of τ_p values ranging from -60 to 60 frames of video (-250ms to 250ms). To quantitatively assess the influence of cricket velocity on model fit and to subsequently find offsets with significant cricket velocity influence, we compared the Akaike Information Criterion (AIC)⁴² of the model described in Eq5 to analogous models that exclude the range vector and cricket velocity predictor terms R and C. This was performed for each τ_p and τ_d in the above range. When the AIC value of one model is more negative than another, it can be interpreted to have a better fit. More negative AIC values in the cricket inclusion models would suggest that prediction based on cricket velocity is a significant contributing factor to reach kinematics. We plotted the AIC difference values and used a jackknife approach (leaving ~8 of 78 reaches out for 10 jackknife iterations) to estimate jackknife AIC mean and calculate standard error.

To compare the predictive pursuit model against steering models we applied the simulation approach discussed above in Equations 1, 2, and 3 and compared its R squared performance for matching hand velocity against those models. All features of the simulated reaches were applied as before for the steering models with the guidance law described by Equation 6 as

$$A_{\text{pred}} = \beta (R(t - \tau_d) + \tau_p C(t - \tau_d)) \quad (\text{Equation 6})$$

where A_{pred} is acceleration applied to hand velocity, β is an experimentally derived gain constant, $R(t - \tau_d)$ is the range vector and $C(t - \tau_d)$ is the cricket velocity, both at delay τ_d , and τ_p is the prediction time constant. All 78 trials were included to maximize the total R-squared fit across trials in fitting β and τ_p parameters for each tested visuo-motor delay τ_d . Confidence intervals on the R-squared performance were computed from the standard deviation of trial-by-trial R-squared values with statistical tests comparing trial R-squared between models (Sign Rank tests).

Current Biology, Volume 33

Supplemental Information

**Fast prediction in marmoset reach-to-grasp
movements for dynamic prey**

Luke Shaw, Kuan Hong Wang, and Jude Mitchell

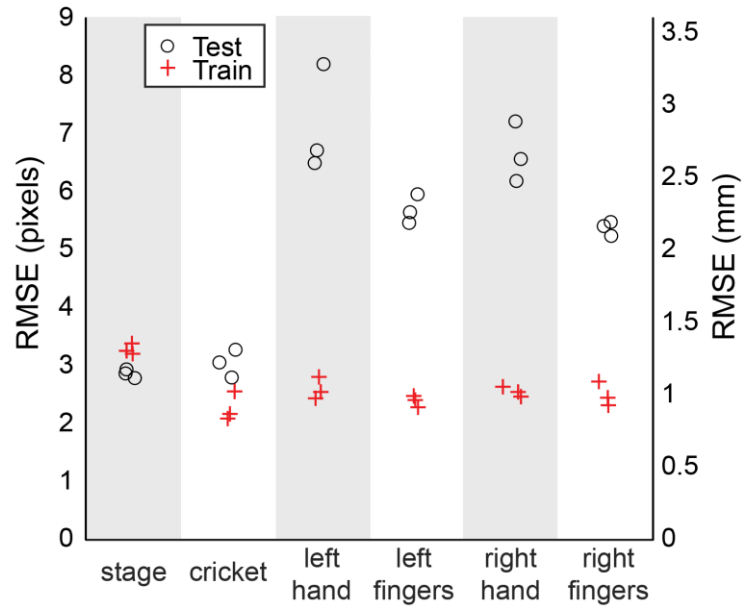


Figure S1. Tracking accuracy in DeepLabCut video analysis, related to Figure 1: Three DeepLabCut networks were trained to show tracking errors from the testing and training sets in RMSE quantified in pixels and converted to mm using the pixel to mm conversion respective to the above-view camera (Figure 1F). The label 'stage' on the x-axis refers to the reach platform stage and corresponding arena points (Figure 1D). Fingers refers to the average of both the fingertip and knuckle points, which equals 10 points for each hand. The average RMSE values depict the stability of the training approach over 3 different training iterations. Each point depicts the labeling error of a model trained with a different "shuffle" of the test and training data. The model used for all tracking analyses performed with a testing RMSE of 2.79 pixels for the stage and arena markers, 3.28px for the cricket marker, 8.21px for the left hand, 5.1px for the left fingertips, 5.9px for the left knuckles, 6.2px for the right hand, 5.8px for the right fingertips, and 5.0px for the right knuckles. For comparison, the width of the wrist is typically around ~1.5cm, which equates to 37px. The length of the cricket ranged from ~1.25cm to ~3cm. We utilize the hand marker for reach modeling and correlation analyses because of its tracking reliability across all three camera views and its reduced susceptibility to occlusion-related tracking drops, which were more common in the finger marker.

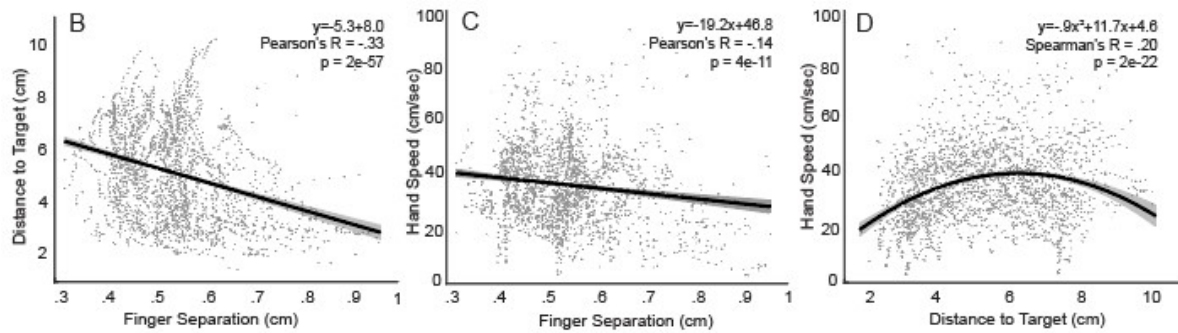


Figure S2. Further grasp characterization in relation to hand speed and distance, related to Figure 2. **A:** Grasp visualization from a low camera angle to show parallel hand orientation and simultaneous finger closure. **B-D:** Scatter plots illustrate the relationships between distance to target, hand speed, and finger separation across reaching movements. Points show a 1/4 subsample of all tracked points from the population of reaches. Each plot is fit with a line (B,C) or 2nd order polynomial (D), which is shown with a 95% confidence interval. **B:** Across many reaches the finger separation increases with proximity to target. **C:** Finger separation increases with decreased hand speed. **D:** Hand speeds are lower when at greater distances from target and when in close proximity to the target.