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

Luke Shaw[1], Kuan Hong Wang[1*], Jude Mitchell[2*],

[1] Department of Neuroscience, University of Rochester Medical Center, Rochester, NY
[2] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY
* Correspondence to: kuanhong_wang@urmc.rochester.edu, jmitch27@ur.rochester.edu.

Abstract

Primates have evolved sophisticated visually guided reaching behaviors for interacting with dynamic objects, such as insects during foraging. Reaching control in dynamic natural conditions requires active prediction of the target's future position in order to compensate for visuo-motor processing delays and enhance online movement adjustments. 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 step once during the movement. However, those highly constrained approaches also 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 (*Ngo et. al. 2022*). To examine the complementary dynamics for 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 marmosets 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 prediction for dynamic targets can operate at incredibly short visuo-motor delays, rivaling the speeds that are typical of the oculomotor systems during closed-loop visual pursuit. Multivariate linear regression modeling of the kinematic relationships between the hand and cricket velocity revealed a visuo-motor delay below 100ms. Predictive corrections for moving crickets were incorporated into hand motion especially during the latter stage of the reach-to-grasp trajectory. This approach reveals that the marmoset motor system for reaching can operate at short visuo-motor delays comparable to the oculomotor system under natural dynamic conditions.

New and Noteworthy

Deeplabcut, marmosets, naturalistic paradigm, crickets, measurement of visuomotor delay in marmosets, evidence for fast predictive reaching.

INTRODUCTION

Primates exhibit a sophisticated repertoire of visually-guided reaching behaviors that enable them to interact with objects and navigate arboreal environments. These behaviors are thought to provide a foundation for the evolution of more complex motor behaviors such as tool use (Peeters *et al.*, 2009; Kaas & Stepniewska, 2016; Kastner *et al.*, 2017). While the motor aspects of these reaching movements are well conserved across mammalian species, primates are unique in their integration of visual information to guide reaching movements (De Bruin *et al.*, 2008; Kaas *et al.*, 2011; Sacrey & Whishaw, 2012; Kaas & Stepniewska, 2016; Wang *et al.*, 2017). Previous research in humans and non-human primates has largely focused on experimental paradigms that involve highly repeated ballistic movements in constrained positions (Georgopoulos *et al.*, 1983; D'Mello *et al.*, 1985; Buneo *et al.*, 2002; Saunders & Knill, 2003; Elliott *et al.*, 2010; Battaglia-Mayer *et al.*, 2013; Dekleva *et al.*, 2018; Avraham *et al.*, 2019). While this classical experimental approach provides a foundation for understanding the neural mechanisms of reaching control and helps to delineate the roles of frontal-parietal networks, much remains unknown about reaching control in the unrestrained and dynamic conditions typical of natural movements (Datta *et al.*, 2019; Nastase *et al.*, 2020).

Motor control in dynamic conditions requires active prediction of the target's future position to compensate for visuo-motor processing delays. For example, for an organism to successfully intercept a moving target, it must be able to account for the movement of the target, the movement of its own body, as well as neural conduction and processing delays. Furthermore, it must be capable of accommodating unexpected perturbations of planned behavior. Most theoretical models of motor control are based on a predictive model of the future body and target positions which are incorporated into an update mechanism based on real time visual and proprioceptive information (Wolpert & Kawato, 1998; Shadmehr & Krakauer, 2008). These two control features, prediction and motor update, have been studied extensively in human and animal models, but there is limited research involving primates in freely reaching settings. Researchers have estimated motor update delay to range from 125-225ms in human experiments where the location of a simple target is moved unexpectedly during free reaching or pointing tasks (Prablanc & Martin, 1992; Brenner & Smeets, 1997; Day & Lyon, 2000; De Brouwer & Spering, 2021). The limited number of studies with macaques that have used double-step reaching tasks to probe online correction have found substantially longer motor delays around 200-300ms (Georgopoulos *et al.*, 1981; Battaglia-Mayer *et al.*, 2013); Archambault et al, 2008; 2011; Dickey, Amit, Hatsopoulos et al., 2013) Of note, the redirected reaches are typically preceded by a saccadic eye movement at 100-200ms latency to the new jumped location followed by changes in the hand trajectory at 50-100ms (Archambault et al, 2015). Further, the jumps of targets used in non-human primate studies have often been much larger than those in human studies, which could impose additional delays in redirecting attention before movement adjustment. It is also the case that non-human primate studies have been highly constrained, involving head-restraint and contingencies for juice rewards that could favor conservative strategies biased for the original target (Archambault et al, 2015). It remains unclear what are the dynamics for making online adjustments in unrestrained reaching movements for natural moving targets, such as moving insects. Because natural prey move in a continuous way they would likely pose different constraints on the coordination between attention and the hand.

Two recent studies in non-human primates have examined movements for continuous moving dynamic targets, though without estimating the visuo-motor delay of the hands. The first of these studies used video in the field to examine reach-to-grasp movements for flying insects among wild marmoset monkeys (Ngo et al., 2022). While they find evidence to support that the head angle of marmosets actively tracks insects during reaching, thus supporting an online active vision strategy, the video resolution does not afford a precise measure of visuo-motor delay. Another study in macaques examined pursuit of continuous moving targets using joystick controlled cursor on a computer screen (Yoo et al., 2020). They found that macaques take into account a target's position and velocity to guide interception ahead of the target on the screen. However, the use of a joystick, which included momentum and other terms to control the cursor, imposes additional sensory-motor delays in the task that cannot be separated from the dynamics of the hand. The predictive nature of hand guidance during interceptive reaching tasks is, to date, an understudied topic in non-human primate research in freely-reaching or naturalistic contexts.

Advances in machine vision have provided tools to track complex behavior without the need for optical markers and controlled settings. DeepLabCut, for example, is a machine learning toolkit that has been optimized for flexible and convenient scientific use, which can track a broad scope of user-trained features captured in video. Integrative packages such as Anipose have also been developed to use DeepLabCut's back-end to generate 3D position estimates from multi-camera recordings with minimal user effort and low tracking error (Mathis *et al.*, 2018; Karashchuk *et al.*, 2021; Moore *et al.*, 2022). Recently these methods have further been applied and validated for marmoset monkeys in controlled laboratory conditions (*Moore et al.,* 2022). These tracking systems provide a wide range of experimental flexibility, which is critical for expanding reaching experiments beyond more constrained classical approaches.

We examined predictive reaching movements in the marmoset monkey. The marmoset monkey is a small New World primate that has drawn interest as an emerging animal model in neuroscience due to the similarity of its brain organization with larger primates. It offers a few experimental advantages over larger primates. First, they breed well in captivity giving birth to pairs 2-3 times each year to provide larger animal numbers. Their cortical surface is lissencephalic, which facilitates access to visual, parietal, and frontal brain areas for array recordings, imaging, and optogenetic stimulation. Marmosets perform visually-guided reaching in natural conditions during foraging, locomotion, social interaction, and hunting (Schiel *et al.*, 2010; Mitchell & Leopold, 2015; Schiel & Souto, 2017; Mundinano *et al.*, 2018; Hori *et al.*, 2021). While marmoset reaching behavior has been studied in the field (Schiel *et al.*, 2010) and using other tracking approaches (Takemi *et al.*, 2014; Moore *et al.*, 2022), only recently has there been any progress examining reaching for dynamic targets (insects) in the field (Ngo *et al.*, 2022). Furthermore, general studies of nonhuman primate prehensile kinematics are surprisingly few in number (Schiel & Souto, 2017).

Here we describe an ecologically motivated reaching task for dynamic targets (insects) that provides accurate 3D kinematics within a laboratory setting in marmoset monkeys. The analyses used in this study aim to determine whether target motion becomes incorporated into hand motion, resulting in predictive reaching of the hand. Our results provide kinematic descriptions of marmoset reaching for moving crickets, suggest an estimation of marmoset visuo-motor delay, and offer evidence for the predictive nature of reaching. These unrestrained reaching paradigms provide first steps towards future electrophysiological and optogenetic experiments to probe the role of frontal-parietal circuits in predictive control.

METHODS

Animals

All experimental procedures described here are approved by the University Committee on Animal Resources at University of Rochester. 10 marmoset monkeys (5 females, 5 males) were recorded by multiple cameras while reaching for live crickets 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. A mobile experimental apparatus enabled us to test if marmosets would reach directly or predictively to moving crickets **(Figure 1A)**. Marmosets were acclimated to the experimental apparatus **(Figure 1B)** 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. No marmosets were reluctant to reach and acquire cheerios. Any initial apprehension for eating live crickets was typically overcome by providing enrichment food items such as marshmallows alongside the crickets. In some cases, viewing other marmosets from their family group reaching for crickets was sufficient to provide interest and motivation for participation in the experiments. Crickets demonstrated a broad range of activity from no locomotion, to crawling rapidly around the perimeter of the reaching platform as described below, to jumping to escape capture.

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)**. 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 **(Figure 1C)**. Frames of this video were extracted and fed into triangulation algorithms provided in the Anipose toolkit (Karashchuk *et al.*, 2021). 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 1D)**. 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 **(Figure 1E)**. Reflection reduction was

useful due to the competition of the features of a tracked subject's reflection against the features of the subject itself.

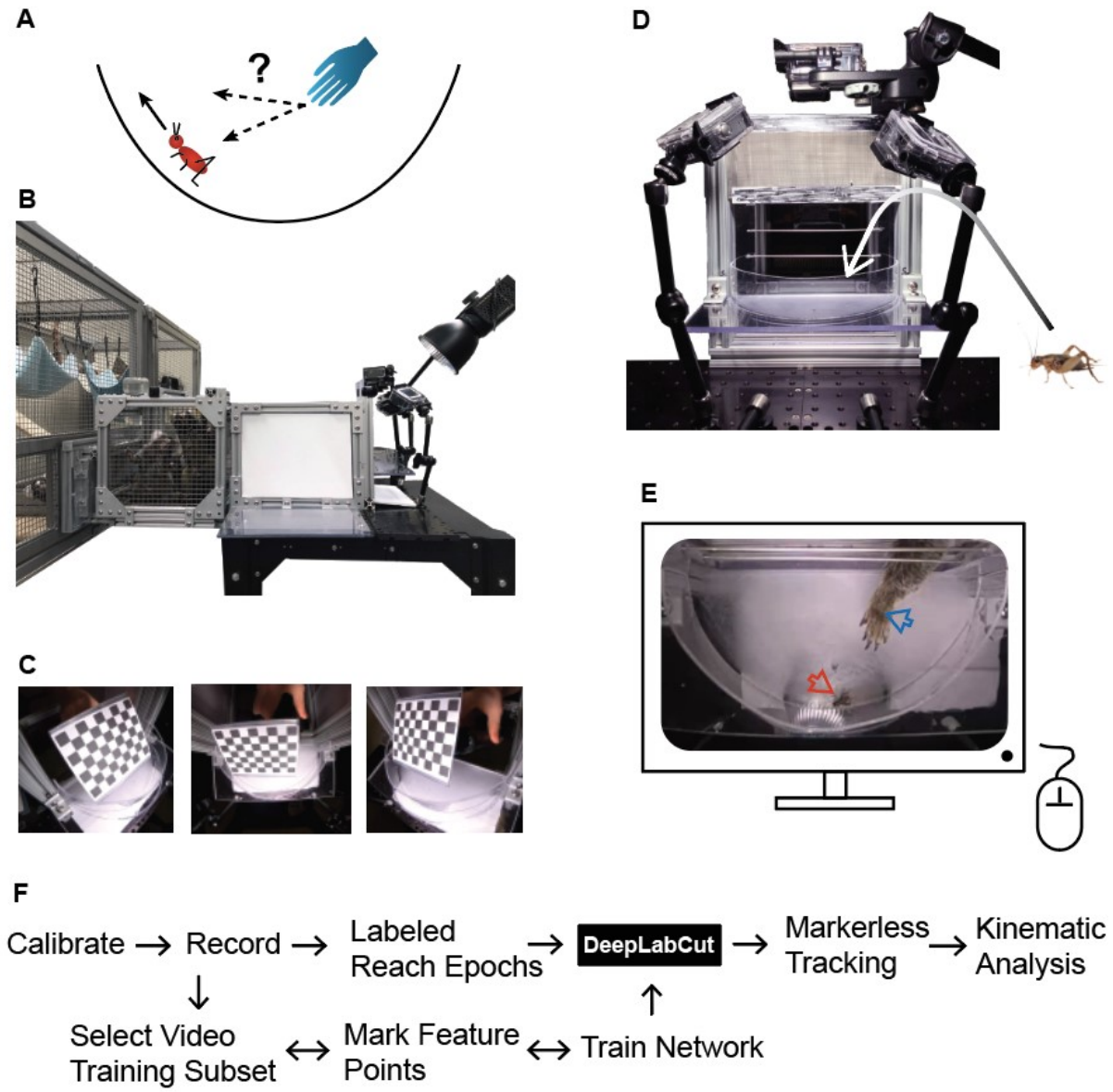

Figure 1. Reaching experimental design and DeepLabCut video analysis workflow. A: Experiments and analyses were designed to ask whether marmosets reach directly or predictively for moving crickets. **B:** Video recording was performed in the marmoset colony by docking a wheeled platform containing a box to isolate individual marmosets from their family groups for reaching experiments on the platform. **C:** The three cameras were spatially calibrated using a checkerboard pattern prior to experiments for subsequent 3D triangulation. **D:** The platform opened to a reaching aperture and half-circle arena where crickets were placed. **E:** Feature points were manually labeled on a small subset of frames of video for DeepLabCut training. **F:** The data collection and DeepLabCut analysis workflow is indicated by arrows. Dual-direction arrows indicate repeated instances of DeepLabCut network refinement to converge on a solution that generalizes well to novel data.

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 **(Figure 1D, 2A-C)**. 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 metadata 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 **(Figure 1F)**. 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. The full workflow of video data preparation for DeepLabCut network training and video analysis is outlined in **Figure 1F**. 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 **(Figure 2)**. 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 2B)** 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 2D)** 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.

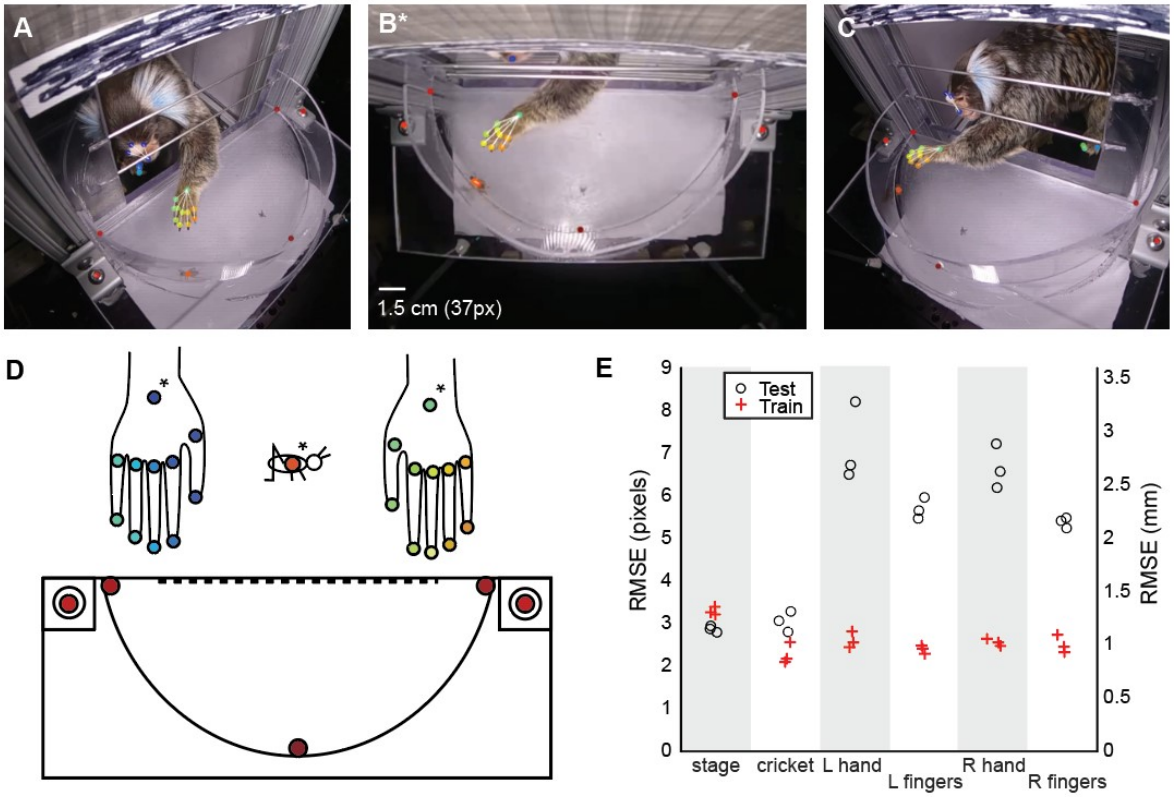

Figure 2. Feature points and tracking accuracy in DeepLabCut video analysis. **A-C:** Views from the left (A), above (B) and right (C) cameras with DeepLabCut tracked points. **D:** DeepLabCut was trained to identify points of the reach stage and arena (shown in red), a central hand point, fingertips, knuckles, and cricket targets. Central hand and cricket markers tracked by the above-view camera were the focus of subsequent predictive analyses (*). **E:** 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 (B). The label 'stage' on the x-axis refers to the reach platform stage and corresponding arena points. Fingers refers to the average of both the fingertip and knuckle points (10 points for each hand).

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 **(Figure 2D)** beyond the reach aperture plane to the last tracked point of the cricket. Cricket tracking at the very end of reaches, during deacceleration and grasp, was typically lost due to occlusion by the fingers or hand. Windows of reaching data were defined by viewing videos with overlayed 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.

Predictive Modeling and Reaching Analyses

Target and hand velocity vectors were decomposed into orthogonal velocity components relative to a displacement line segment, which connects hand position to cricket position. We refer to the velocity component parallel to the displacement segment as *direct* velocity. We refer to the velocity component orthogonal to the displacement segment 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 4D**.

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 (σ = 20ms) 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 5A**). 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

Eq 1
$$X_t^h = \sum_{\tau=-25}^{150} P^c X_{t-\tau}^c + E_t$$

where X_t^h represents lateral hand velocity at time t, X^c is the lateral cricket velocity at time t minus time delay τ, and P^c 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 in lag from -25 to 150ms. We defined the displacement 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 displacement 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.

To determine whether cricket velocity terms were significant determinants of future hand velocity, we employed an auto-regressive linear regression model formatted as shown in Eq 2, which was evaluated using the Matlab *mvregress* function. Concatenated velocity and displacement 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.

Eq 2 $$X^h_{t+1} = B^h X^h_t + B^c X^c_{t-\boldsymbol{\tau}} + B^p X^p_{t-\boldsymbol{\tau}} + E_t$$

In the above Equation 2, X^h_{t+1} represents hand velocity one frame of video into the future (t+1) relative to the current X^h hand velocity at time t. This velocity is a bivariate variable containing both lateral and direct components (L^h, D^h). The velocity of the cricket X^c at time t – τ, where τ represents delay, can guide predictions of the future hand position. This velocity is also a bivariate variable containing both lateral and direct components for the cricket (L^c, D^c). We also include a variable representing the magnitude of the displacement vector from hand to the cricket position as X^p. This vector connects hand position at time t and cricket position at t- τ. The predictor variables are weighted by a set of coefficients, B^h, B^c, and B^p, respectively. These data were submitted to a least-square regression where E_t represents the residual error.

To determine the visuomotor delay involved in prediction, we explored a delay parameter in the linear prediction equation. The parameter τ in Eq2 was varied from 0 to 36 frames of video (equivalent to 0 to 150ms) such that "past" cricket velocities were aligned to present hand velocities. To quantitatively assess the influence of cricket velocity on model fit and to subsequently find offsets with significant cricket velocity influence, we compared the Aikake Information Criterion (AIC) (Aikake, 1973) of the model described in Eq2 to analogous models that exclude the cricket velocity predictor terms (B^cX^c). This was performed for each τ offset 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 one-reach-out for each jackknife iteration) to estimate AIC and predictor coefficients *B*.

To compare the velocities of hand and cricket across multiple reaches of varying distance and duration, interpolation was conducted using two methods. Each method compressed each reach into 30 points. Distance to target interpolation was conducted using 30 points evenly distributed across the total distance of each reach. Reach duration interpolation was conducted using 30 points evenly distributed across the total duration of each reach.

To investigate the evolution of prediction within reaches using a direct, model-free approach, we tested the correlation of lateral hand velocity with lateral cricket velocity at different points within the reach. Using both interpolation strategies, we calculated the T_{max} corrected (Blair *et al.*, 1994) Pearson's correlation of cricket velocity to hand velocity at each interpolation point over the population of reaches using a permutation approach. These corrected correlation tests were conducted using the PERMUTOOLS (Crosse *et al.*, 2020) open-source package available at https://github.com/mickcrosse/PERMUTOOLS. We performed this test for each τ offset. Confidence intervals were calculated from corrected correlation coefficients using a Fisher transformation. We show data from the distance to target interpolation method alone, as both interpolation methods provided similar results.

RESULTS

To study visually guided reaching behavior in a naturalistic task, we used high-speed, multi-camera video and DeepLabCut to track marmoset reaches for live, moving crickets. We first examined the kinematics of 3D reach to grasp movements to quantify the relative timing of the reach and grasp closure. Then we focused on an analysis of the 2D position as imaged from above the reaching arena to test the hypothesis that marmosets reach predictively for moving crickets and to test visuomotor delay. We first used a simple linear predictor to estimate the hand's response to increases in cricket velocity. Following this, we applied a multivariate linear regression model composed of hand and cricket velocities to determine whether cricket velocity influences hand velocity, or alternatively, if simple target displacement determines movement. We also examined the optimal visuomotor delay in prediction and determined at what moment in the reach prediction plays a more significant role.

DeepLabCut Performance

DeepLabCut provided markers that generalized well with comparable measurement accuracy in cross-validated datasets. To perform model testing using DeepLabCut's estimates of marker positions, 95% of the manually labeled data set were used in training, and 5% were held out to test generalization performance. The model was trained 3 times leaving out a different 5% each time to estimate the precision of training. Training error asses the model's accuracy on the labeled training data set. Test error assess the model's accuracy on novel labeled frames of video. For reference in the following error estimates, 1 pixel from the above view camera **(Figure 2B)** is equivalent to ~0.42mm. Our trained model achieved a root mean squared error (RMSE) error of 2.63 pixels (px) and a test error of 4.63 px for all trained points. Our training approach generated effective model generalization of hand, finger, target, and arena markers to the testing set of video frames, which is summarized in **Figure 2E**. 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 markers.

Marmoset Reaching Characteristics

When hunting for prey in the wild, marmosets use a left-handed (LH), right-handed (RH), or a two-handed "pouncing" (P) strategy. We found that marmosets in our experiments predominantly employed a left or right-handed reaching strategy. We recorded a total of 271 reaches for live crickets (137 LH, 134 RH, 10 P) from 10 marmosets. Gross quantification of these reaches through a left-handedness index (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=.31 and .04) and three marmosets with a strong left-handed preference (.83, .87, and .92). Of the total recorded

set of 271 reaches for crickets, 78 reaches contained crickets that were sufficiently moving during and prior to the reach. These 78 reaches were utilized for subsequent analyses.

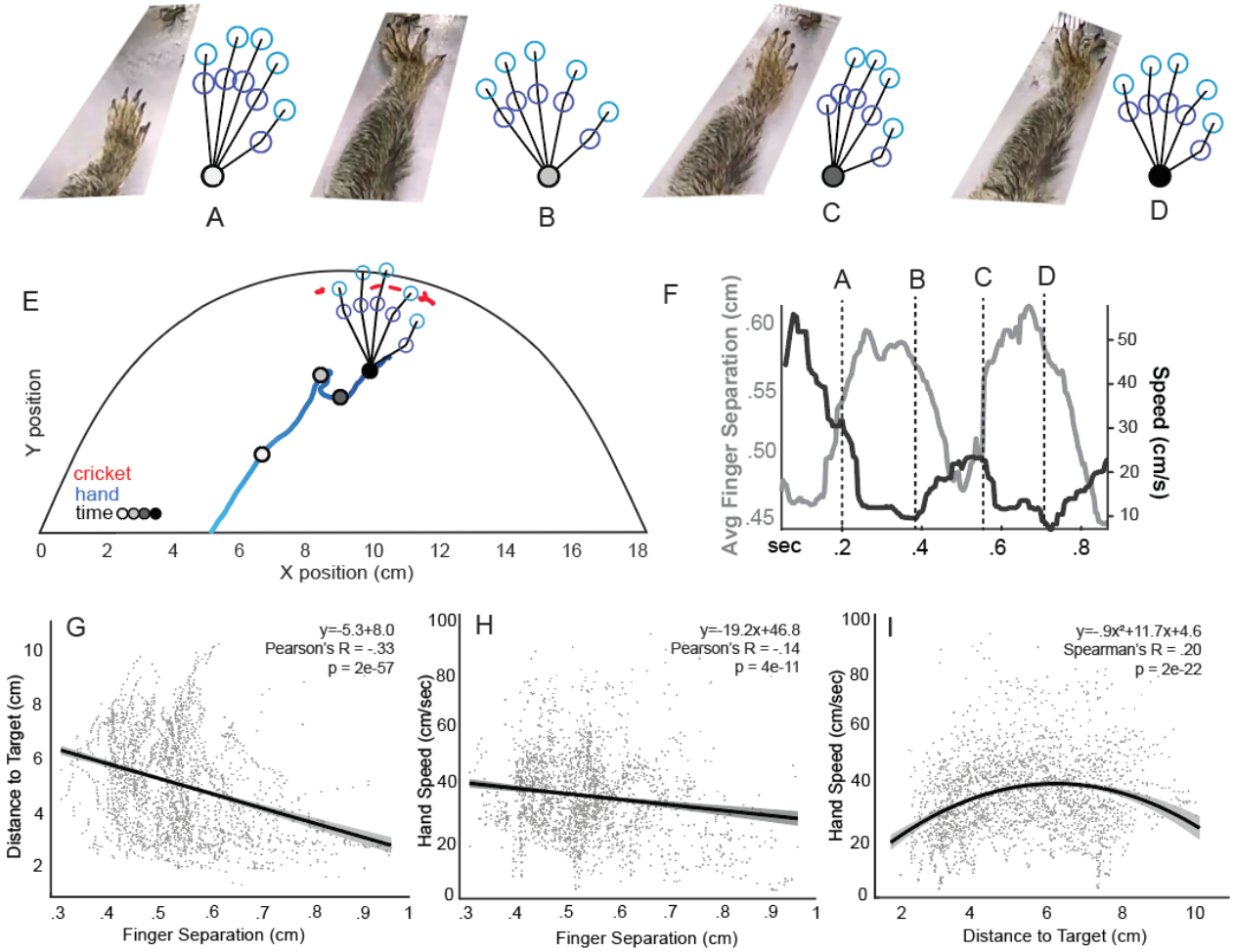

Figure 3. Hand shape evolves in relation to hand speed and target distance during reaching. A-D: Cropped video frames with corresponding 3D tracked hand and finger marker estimates at different points in an example reach. **E:** The example reach is shown from the top camera view. It contained two approaches to the target, the first in which the cricket was not acquired. Points labeled on the reach by white to black circles indicate corresponding points shown in A-D. **F:** Within the example reach there were two speeded approach portions in which the hand speed, depicted in black (right vertical scale), was anticorrelated with the finger separation, depicted in gray (left vertical scale). **G-I**: Scatter plots illustrate the relationships between distance to target, hand speed, and finger separation across reaching movements. Each plot is fit with a line (G,H) or 2nd order polynomial (I), which is shown with a 95% confidence interval. **G:** Across many reaches the finger separation increases with proximity to target. **H:** Finger separation increases with decreased hand speed. **I:** Hand speeds are lower when at greater distances from target and when in close proximity to the target.

Grasp Aperture Relationships to Hand Speed and Target Distance

Marmosets perform reach-to-grasp movements for moving crickets with generalizable relationships between finger spread, hand speed, and proximity to target **(Figure 3 A-E)**. Opening of grasp aperture with proximity to target, for example, is expected (Sartori *et al.*, 2013; Castiello & Dadda, 2019; Fox *et al.*, 2019). These data are shown here to confirm these expectations and

to demonstrate the tracking capability of DeepLabCut. Even for the smallest sized crickets, marmosets utilized a power grip approach (Macfarlane & Graziano, 2009) to grasping in that all fingers opened and closed upon target interception. **Figure 3E** demonstrates an individual reaching trial whereby the speeded initial hand volley did not result in target capture and was followed by a slower, small-scale extension volley. Illustrations of the hand **(Figure 3A-D)** demonstrate DeepLabCut tracking of hand features with Anipose triangulation at different points in this reaching trial. Cropped views of the corresponding frames of video are shown alongside these illustrations, which also show the cricket target. At point A, the hand is traveling at relatively high velocity, and so the fingers are positioned in a tight, parallel conformation. Point B shows the opening of the grasp aperture at the slowed-velocity endpoint of this reach. A similar pattern recurs in the second volley for the cricket with extended, aligned fingers at the speeded portion of this truncated volley (C), which is followed by an opening of grasp aperture prior to target capture (D). Finger separation distance was calculated for each pair of adjacent fingers at the knuckle markers. These values were averaged for each frame of video and plotted alongside the speed of the hand **(Figure 3E)**. For this example reach, there is a clear inverse relationship between finger spread and hand speed.

Across the population of reaches for moving crickets the reach-to-grasp dynamics are evidenced by increases in grip aperture occurring near the end of the reach as hand velocity drops **(Figure 3G-I)**. The hand data shown here extends from the appearance of the hand marker to the completion of the reach for each reach beyond the disappearance of the cricket marker, which is the cutoff for later modeling and correlational analyses. Reach completion is indicated by a combination of maximum arm extension, velocity minimum, and the disappearance of fingertip markers. Scatter plots show points sampled across the duration of reach movements with linear regression lines superimposed for the distance to target and hand speed against finger separation **(Figure 3G, H)**. When distance to target decreases, finger separation increases, yielding a significantly negative correlation (m=-5.3, Pearson's R=-.33, p<.0001), consistent with grip aperture opening at the end of the reach **(Figure 3G)**. This trend is also evident when correlating hand speed to average finger spread (m=-19.2, Pearson's R=-.14, p<.0001) with the grip aperture opening as hand speed drops at the end (**Figure 3H**). Consistent with a bell-shaped velocity profile the relationship between hand speed and distance to target is best modeled by a second order polynomial whereby lowest hand speeds are typically correlated with the smallest and greatest distances to target (Spearman's R=.20, p<.0001, **Figure 3I**). This non-linear relationship for hand speed over the reach explains why the correlation between grip aperture and hand speed averaged irrespective of distance **(Figure 3H)** is less strongly correlated than grip aperture to distance from target **(Figure 3G)**.

Lateral Components of Reach Velocity Suggest that Cricket Velocity Influences Reaching Online

A single reaching trial illustrates cricket velocity guiding hand velocity **(Figure 4A-C)**. Over the individual reaching trial, the earliest points have lighter shades to illustrate earlier time. To determine whether the hand motion is influenced by target velocity, we first decomposed the velocity of hand (h) and cricket (c) into components relative to the displacement vector (P) that connects the instantaneous position of the two. The hand and cricket direct velocity components, $\vec{D}h$ and $\vec{D}c$, are in line with P, the displacement vector whereas the lateral velocity components, $\vec{L}h$ and $\vec{L}c$, lie orthogonal to P **(Figure 4D)**. Traces of the lateral and direct velocity components for the example reach are shown in **Figure 4B and 4C**. The majority of hand velocity is captured

by the direct component, $\vec{D}h$, whereas the cricket velocity is largely characterized by lateral movement, $\vec{L}c$. If $\vec{D}h$ and $\vec{D}c$ or $\vec{L}h$ and $\vec{L}c$ are the same sign, it means they are traveling in the same direction. If the marmoset is reaching predictively, $\vec{L}h$ should align with $\vec{L}c$. Grey areas in

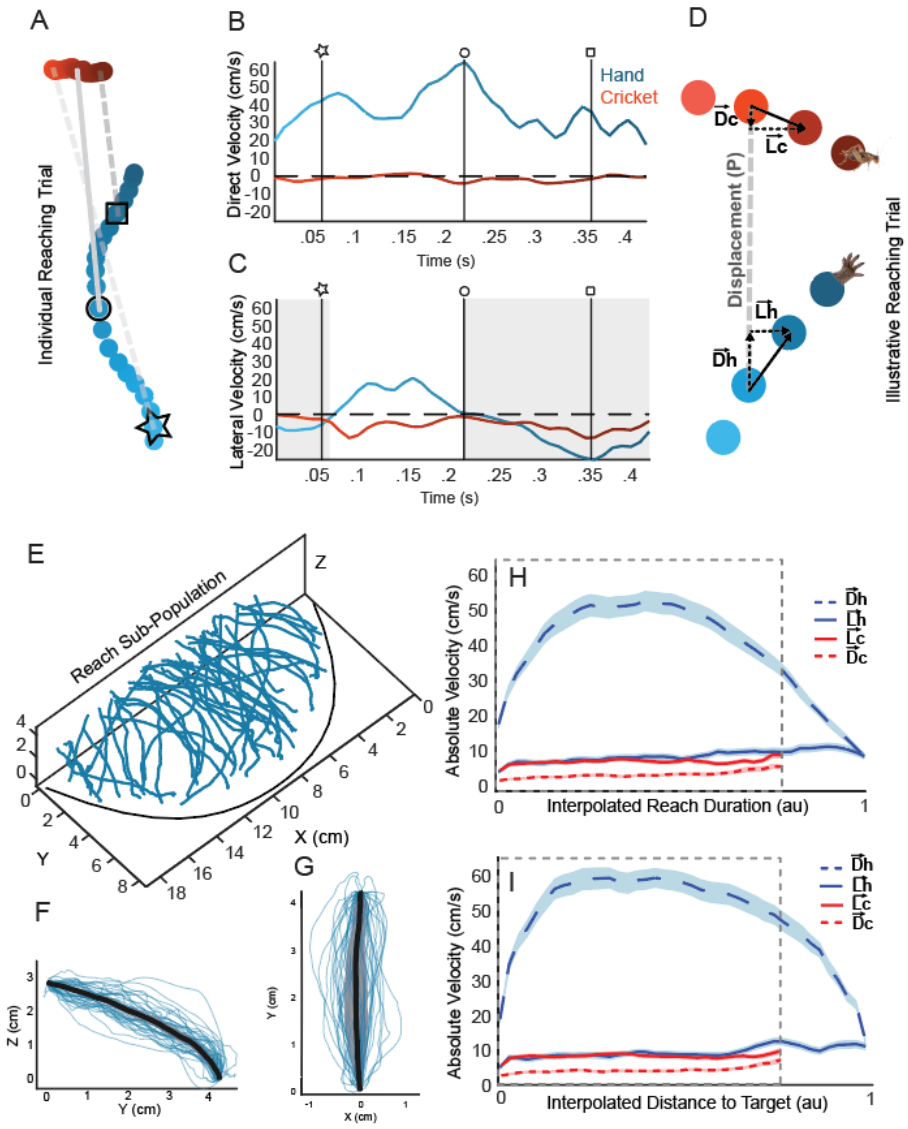

Figure 4. The lateral velocity of the hand and cricket suggests a predictive strategy in reaching. A: Path of the hand and cricket in an individual example reaching trial. Lighter colors indicate the beginning of the reach. The star, circle and square depicted here indicate three time points from early to late in the reach. The displacement vector at each moment in the reach is indicated by the gray lines. At any moment, the absolute velocity can be broken down into two components, the direct velocity (D) defined by velocity projected along the displacement vector and the lateral velocity (L) defined by the orthogonal projection. **B,C:** The direct (top graph) and lateral (bottom graph) velocity components corresponding to the reach movement shown in A. **D:** An illustration of the breakdown of velocity into direct velocity (D) along the displacement vector and lateral velocity (L) orthogonal to displacement. **E:** Representative subset of the population of reaches for moving crickets shown in 3D. **F,G:** The traces of each reach are superimposed in 2D plots after transforming and scaling each reach to have the same beginning and end point. The average trace is indicated by the black line. **H:** Component velocities taken from the full reach population shown on a normalized time scale from 0 to 1 where 0 is the moment the hand crosses the aperture of the isolation box and 1 is the moment of the grasp. The dotted box indicates the portion of the reach with cricket tracking prior to the hand occluding the cricket. Time was interpolated to sample each reach uniformly on the normalized scale. Each trace is shown with standard error of the mean. **I:** Same as in H, but with interpolation done as a function of the distance to target rather than time.

the $\vec{L}$ in **Figure 4C** denote areas where the $\vec{L}h$ and $\vec{L}c$ are pointed in the same direction. For this example, the greatest alignment in velocity occurred towards the end of the reach as the hand converged on the cricket. Specific time points in the reach are illustrated by the star, circle, and square markers. The star denotes a point where the hand transitions from traveling in the same lateral direction as the cricket to the opposite lateral direction, or behind, the cricket. The circle indicates the peak velocity of the hand, where the trajectory of the hand switches back to align with the lateral velocity of the cricket. The square indicates the near-end of the reach, where $\vec{L}h$ and $\vec{L}c$ are clearly congruent and the hand is reaching ahead of the cricket. In this illustration, note that the hand and cricket points do not converge completely at the reach end-point. This is due to the placement of the hand marker, which is closer to the wrist than the finger tips, which occlude the cricket from the top-view camera and prevent tracking.

The spatial position and velocity features of reaches to crickets reveal a consistent pattern of outward extension towards the cricket, with peaked velocity near the distance mid-point of the reach **(Figure 4E-I)**. Figure 4E shows the 3D reconstructions of a subset of reaches included in the moving cricket data set. To demonstrate the inherent variability of the hand's trajectory both as a function of behavior and tracking accuracy, panels 4F and 4G show this same set of reaches that have been transformed to have the same beginning and ending point. Variations around the mean reach may reflect a variety of sources of error, including measurement error, but one key source we consider later is adjustments of the reach that take into account the motion of the cricket and could be predictive of hand velocity. Averaged velocity representations of the population of reaches for moving crickets in panels **4H** and **4I** shows a dominance of direct velocity for hand motion. Conversely, crickets show lateral velocity dominance, which is expected, as the majority of their locomotion followed a path along the perimeter of the reaching platform retaining wall. The overall magnitude of cricket lateral velocity appears to be relatively matched with that of the hand lateral velocity. The dotted gray boxes in panels **4H** and **4I** indicate the portion of the reach prior to the occlusion of the cricket by the fingertips and hand.

Linear Prediction in Reaching Movements

According to a predictive reaching hypothesis, the lateral velocity of the cricket moving along the perimeter should correlate at some visuomotor delay with hand velocity. Alternatively, if the reach is driven purely by cricket position at any moment without taking its velocity into account, then cricket velocity will not correlate with hand velocity. To examine this possibility, we first identified discrete events across task trials in which the cricket suddenly accelerated while the reach was ongoing. The lateral speed of the cricket was filtered to flag the top 5% of moments where velocity increased, either moving laterally to the left or right in the arena (see Methods). Rightwards increases in speed were reflected in sign in order to pool together all velocity traces in averaging. We observed that immediately following an increase in lateral cricket velocity there is a corresponding increase in lateral hand velocity that begins around 80ms (**Figure 5A**). To determine if this delay was consistent throughout the entire data set, we applied a simple linear regression on all the velocity data to predict lateral hand-velocity based on the recent history of cricket velocity (see methods, Equation 1). This enabled us to fit a linear kernel for hand velocity based on input cricket velocity (i.e., the impulse response function corrected for auto-correlation in cricket velocity). Consistent with the visuo-motor delay observed due to sudden jumps in cricket velocity, the linear prediction based on the full data set finds a significant prediction coefficient near 80 ms ($p<.0001$ at 87.5ms and $p=.0559$ at 100ms), but not other delays (**Figure 5B**). This suggests that changes in hand velocity track the cricket velocity within a relatively brief window of

influence from 80-100 ms, and with a visuo-motor delay much faster than found previously in more constrained reaching tasks.

To consider the role of prediction and visuo-motor delay more thoroughly we developed a full auto-regressive model. We pooled data from 78 reaches for moving crickets to determine whether cricket velocity is a significant predictor of hand velocity using multivariate linear regression modeling. We furthermore used this method to test if there is an optimal time delay between hand and cricket velocities that yields the best model fit to estimate visuo-motor delay. The current hand velocity ($\vec{D}h$, $\vec{L}h$), the cricket velocity at a delay in the past ($\vec{D}c$, $\vec{L}c$), and the displacement from the target at the same delay in the past (*P*), were used as regressor variables to predict the hand velocity one video frame in the future ($\vec{D}h$ and $\vec{L}h$ at t+1). A simplified linear equation for this model is in Equation 2 (see methods).

The inclusion of cricket velocity improves model fitting for hand velocity indicating that reaching is predictively guided. We first tested whether the inclusion or exclusion of the cricket velocity terms in the linear model provides a better fit when all values are temporally aligned as described in Eq 2. This was assessed by testing for significance of the *B* values of $\vec{D}c$ and $\vec{L}c$ in the cricket inclusion model and by comparing the Aikake Information Criterion (AIC) values of the cricket inclusion and exclusion models (**Figure 5C-D**). Because adjacent samples in time are not independent, there is an inherent temporal correlational structure of data belonging to any particular reach that is not accounted for by the mvregress algorithm. Therefore, we performed a leave-one-reach-out jackknife procedure to estimate the variations in predictor coefficients, B^c, and AIC values. Without any introduced delay (tau=0), the $\vec{L}c$ term yielded predictor B^c values that reached significance (*B*=.046, p<.0001) indicating a positive predictive relationship between $\vec{L}h$ and $\vec{L}c$ even with no delay (left-most point in the plot, tau=0). Another significant predictor was the hand velocity terms ($\vec{D}h$: *B*=.97 p<.0001, $\vec{L}h$: B=.91, p<.0001) which simply reflects the hand's momentum. Consistent with the predictor coefficients, the AIC value of the inclusion model was significantly more negative than the AIC value of the cricket exclusion model even with no delay (AIC difference = -61.50, p<.0001). Having determined that cricket velocity is a significant predictor of marmoset hand velocity, we set out to ask whether we could infer the visuo-motor delay between hand and cricket velocity based on these model predictions.

Estimation of Visuomotor Delay

Varying the temporal delay (tau) between cricket and hand velocity in the autoregressive model also resulted in an estimated marmoset visuomotor delay in reaching around 80ms. We repeated regression model analyses iteratively by increasing the temporal offset between hand and cricket velocities and fitting our model accordingly. We performed this for 36 tau values, which encompassed a range of 0 to 150ms. From this we calculated *B* values and AIC differences through the leave-one-reach-out jackknife procedure described above. We then plotted the jackknife means as a function of temporal offset with 2x standard-error-of-the-mean error regions (**Figure 5C-D**). Repeated AIC model comparisons revealed two negative peaks in AIC difference first at 62.5ms (AIC difference = -102.12, p<.0001; AIC without $\vec{L}c$, $\vec{D}c$ terms = -28188.8, AIC with $\vec{L}c$, $\vec{D}c$ terms = - 28264.8) and then a slightly stronger peak at 79.2ms (AIC difference = -103.22, p<.0001; AIC without $\vec{L}c$, $\vec{D}c$ terms = -28141.4, AIC with $\vec{L}c$, $\vec{D}c$ terms = -28218.3). This is matched by a positive peak in the correlation of predictor coefficients *B*($\vec{L}h$ x $\vec{L}c$) at 79.2ms (*B*=.097, p<.0001). Given that this model is configured to predict hand velocity one frame of video

into the future, these values should be interpreted to include an additional 4.2ms, which sets the ideal model fit to a delay of 83.4ms. However overall, these results suggest an optimal visuo-motor delay around 80ms, the delay at which cricket velocity best predicts future hand velocity.

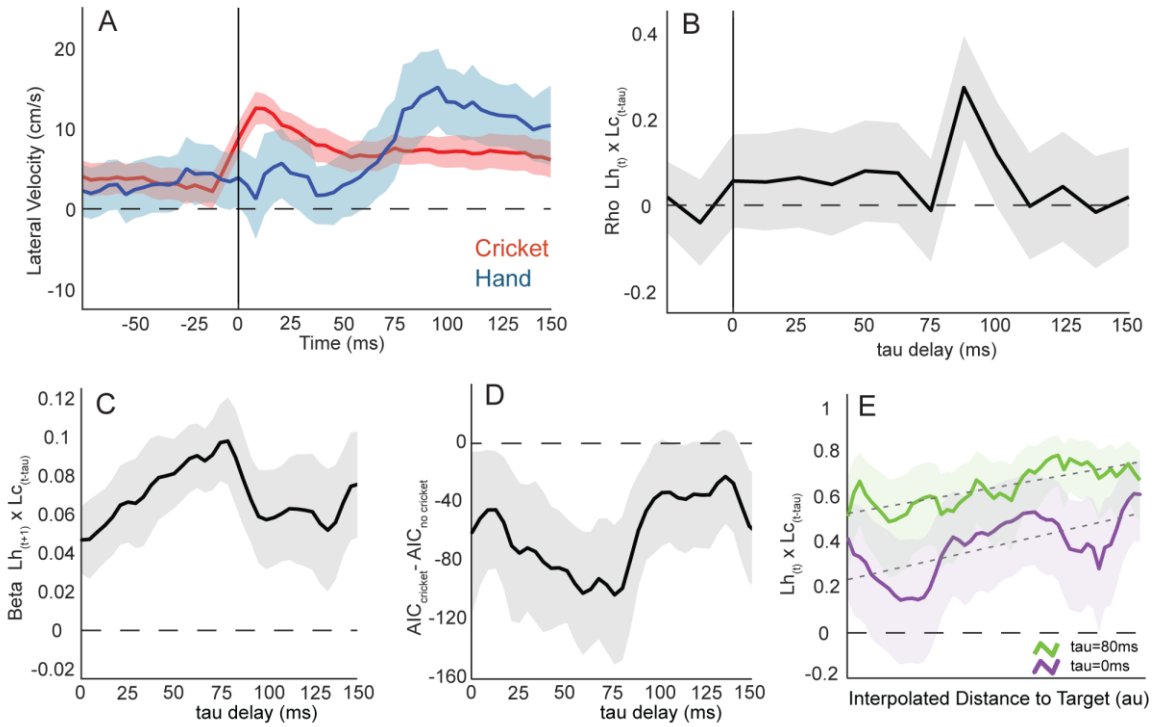

Figure 5. Cricket velocity predictively guides hand velocity at a visuomotor delay. : Increases in the lateral cricket velocity are followed by increases in lateral hand velocity at a latency of roughly 80ms. A Difference of Gaussian filter applied to the absolute velocity traces of cricket motion identified discrete acceleration events (N=68). The average cricket velocity time-locked to those events reflects increases in speed at time zero, as expected, while hand velocity follows at a delay (purple). Error bars show 2 SEM. **B.** A linear regression of lateral hand velocity with the recent history of lateral cricket velocity from time lags going from -25 to 150 ms reveals a significant prediction coefficient between 80-100 ms. **C:** The Beta value of the auto-regressive model of lateral hand velocity is shown with lateral cricket velocity being sampled at different delay values. The peak Beta values align near 80ms where linear prediction from cricket velocity is strongest. **D.** The difference in performance between the predictive model including cricket velocity and the model with pure displacement is shown as a function of visuo-motor delay. The Aikake Information Criterion (AIC) takes negative values to indicate superior performance for the model with cricket velocity as a predictor. The delay of cricket velocity was varied from 0 to 150 ms, with peak performance between 60-80ms. The black line is the average of a leave-one-reach-out jackknife process with the shaded region reflecting 2 SEM. **E:** The Pearson correlation between lateral hand velocity with lateral cricket velocity is shown normalized as a function of reach distance to target with 95% CI plotted first with no visuo-motor delay (purple; dashed line indicates linear fit) and then with an 80ms visuo-motor delay (green; dashed line indicates linear fit). Correlations for the predictor are greater with the 80ms visuo-motor delay, and can be seen to increase towards the end of the reach movement.

Within-Reach Progression of Predictive Reaching

To further investigate when predictive information contributes the most during the reach, we tested the structure of correlations between lateral cricket and hand velocity ($\vec{L}c$ and $\vec{L}h$) over the duration of the reach. We binned reaches into 30 points using reach distance interpolation methods and ran multiple comparisons corrected correlation analyses at each point (see

methods). When the reach was interpolated over reach distance there was an overall increase in the correlation between $\vec{L}c$ and $\vec{L}h$ throughout the reach even with no temporal delay (tau = 0 ms) between cricket and hand velocity (**Figure 5E**, purple). This relationship is highlighted by the dashed linear fit plotted over the correlation curve, which shows a significant linear correlation over the reach interpolation (slope=0.008,r=0.78,p<0.0001). A similar linear trend of increasing correlation over the reach was evident also for the optimal delay (tau = 80ms), though with overall correlations being shifted higher (**Figure 5E**, green; linear fit dashed line: slope=0.010, r=0.65, p<0.0001). This suggests that predictive information guiding the reach based on cricket velocity increases over the duration of the reach peaking at cricket acquisition.

DISCUSSION

The current findings demonstrate the effectiveness of DeepLabCut (DLC) to track marmoset reach-to-grasp movements and reveal the predictive nature of reaching for dynamic targets such as live crickets. We find that grasp aperture in marmoset reaching shows stereotypical relationships with distance from target and hand speed where it opens towards the end of the reach. To test if reaches to live crickets involve prediction based on cricket velocity we fit an auto-regressive multivariate regression model and showed that the inclusion of cricket velocity improved the model's ability to predict future hand velocity. This further enabled us to estimate the visuomotor delay at roughly 80ms. Additional analyses suggest that the influence of cricket velocity on prediction is greatest in the later portion of the reach as the hand closes towards the target, although a positive correlation is evident across the entirety of the reach.

The benefits of utilizing naturalistic behavior for neuroscience research are tempered by experimental variability. However, modeling approaches provide a tractable alternative to traditional trial averaging for addressing this variability issue. Naturalistic behavior can tap into more dynamic aspects of visually-guided reaching including prediction. Furthermore, this approach presents animals with tasks that they are already accustomed to performing. As a clear example, our research was conducted with none of the extensive animal training effort that is required for traditional vision task-related experiments with non-human primates. However, both the subject and the experimental environment (e.g., crickets) have larger intrinsic variability, which increases the difficulty of collecting trials that meet inclusion criteria and performing traditional averaging over repeated trials. We circumvented the lack of identical repeats by performing model fitting, which made it possible to determine the influence of prediction on the behavior.

Previous model-based approaches have shown that primates utilize simple physical parameters to predictively pursue an evading virtual target with a joystick (Yoo *et al.*, 2020). A linear prediction term in the model employed by Yoo and colleagues allowed them to estimate a visuo-motor delay in their task. Similar to our result, they find evidence for a predictive strategy, although their estimate of visuo-motor delay differs substantially from the value reported here. This can likely be attributed to differences in their task design because movements of the cursor on the screen were limited by the speed allowed under joystick control and momentum terms therein, with overall much slower timescales for target and cursor speeds on the order of seconds. By contrast the natural dynamics of marmoset reaches unfold typically in under 200ms. The approach taken in our study emphasizes the dynamics of natural behavior and motor control, which enables interrogation of features of the visuomotor system within the natural timescales of visually-guided reaching and target pursuit.

The derived estimate of marmoset visuomotor delay is consistent with experimentally derived estimates in humans given the quicker processing and neural transduction times of smaller primates. As referenced in the introduction, a widely studied paradigm that interrogates the on-line visuomotor response delay is the double step task. In this task, subjects reach for an initial target, which is extinguished and replaced by a second target during the flight of the hand. On-line corrective reaction times, which are analogous to an on-line visuomotor response delay, have been reported in humans from 125ms to > 200ms (Prablanc & Martin, 1992; Brenner & Smeets, 1997; Day & Lyon, 2000; Brouwer *et al.*, 2002; Reichenbach *et al.*, 2009; De Brouwer & Spering, 2021). This range is likely due to differences in task parameters (eg. pointing with a finger, holding a rod while reaching, using a pantograph, etc.), measurement methods (eg. touch plates, receiver operating characteristic analyses, scored deviation from control, etc.), and the influence of separate motor control pathways with different latencies (eg. tectospinal versus corticospinal). If we take the bottom line from one of these human experiments that derives delay from a deviation in hand kinematics (Day & Lyon, 2000), which is 125ms, and compare that to our estimate of marmoset reaching delay, which is 80ms, we are within range of a human/marmoset difference seen in a separate, but related comparison of human and marmoset kinematics. The ocular following response measures the automatic movement of the eye when presented with coherently moving dot patches. In a study of both humans and marmosets, humans demonstrate a 125ms delay in ocular following response while marmosets demonstrate a 75ms delay, which aligns well with our findings (Cloherty *et al.*, 2020). While the motion of the crickets in this reaching paradigm is not controlled in the same way as in double step psychophysics experiments, it does demand adjustments in reaching to account for its motion and thus provides sufficient variation to estimate visuomotor delay through model fitting.

The progression of predictive reaching analysis without the inclusion of a visuomotor delay (**Figure 5C,D**) suggests the correlation of lateral hand motion to lateral cricket motion increases throughout the reach, but with weaker correlation in the first quarter mark of the reach. This relationship is consistent with a reaching strategy whereby the hand is rapidly transported to target proximity using instantaneous target position and that a predictive, velocity-based targeting strategy comes on-line as the hand slows down. A second possibility is that the hand velocity in the initial extension of the arm is masked by other physical constraining variables on the reach such as posture, balance, and setting the shoulder and elbow joints into motion. These are in the most flux in the initial stages of the reach (Rouse & Schieber, 2015). Given the neural and biophysical activation energy of a reach, it seems reasonable that it is only after the system is fully set into motion that more subtle adjustments influenced by target velocity can be made. These questions could be addressed in future experiments where controlled perturbations in the trajectory of a real-world moving target are possible at different moments during the reach.

While insect hunting is effective for tests of marmoset reaching and has been quantified in natural behavior recently in wild marmosets (Ngo et al, 2022), we do find some limitations using this approach that could be easily addressed in future laboratory studies. In particular, for our study the number of reaching movements in which the cricket moved during the reach were relatively small in proportion (~30% of reaches). Further, the number of reaching movements that we could obtain in any single behavioral session was limited to roughly 15 crickets in a day, more than enough to satiate a marmoset. Given that only a small fraction of those movements involves cricket motion during the reach, we obtained relatively few dynamic reaches per day in any given animal. This limitation of data quantity motivated our pooled approach across 10 marmosets to analyses. By pooling movements across animals and days we were able to estimate the

visuomotor delay, but it would be difficult to employ this same paradigm to test causal manipulations of reaching behavior measurable within a single behavioral session. Future work may seek an intermediate solution without live crickets, such as reaching to small food treats that are moved on a robotic arm, in which the size of reward can be made smaller per reach and the movement of the target can be controlled under online feedback during the reach.

We found that DeepLabCut-powered tracking provided reliable descriptions of the trajectory of the hand supporting the reliability reported in a previous study (Moore et al, 2022). DeepLabCut affords a key advantage for studying natural movements over other methods that affix optical tracking markers to hands and arms, which may impose constraints on animal reaching and require additional time and care to install. We found that GoPro cameras were well suited for data collection in these reaching tasks due to their ability to capture high frame rate video with user-set shutter speeds that reduce motion blur. While the arrangement of cameras was optimized for a 3-D rendering of the top-down view of the hand, occlusion of the fingers was common due to the wide variability of reach trajectories, the pronation of the hand, and the tendency of the fingers to make contact with the cricket retaining wall. Regardless, it is impressive that DLC could provide tracking of features with such small degrees of error and could discriminate between the right and left hands without confusing their features.

The experimental approaches described here will have applications towards investigating reaching control neural circuitry in natural unconstrained tasks. Much evidence suggests that motor commands are sent via feedback pathways from premotor cortex to higher order sensory areas in parietal cortex. These signals are thought to facilitate a comparison of expected and actual sensory perception that is necessary for on-line movement guidance (Shadmehr & Krakauer, 2008; Wagner & Smith, 2008). Such on-line comparisons have been proposed to provide the most rapid and precise control of reaching (Desmurget & Grafton, 2000). Areas in the intraparietal sulcus are a likely hub of reaching planning and on-line control given the convergence of premotor feedback and sensory feedforward connectivity (Gharbawie *et al.*, 2011; Reser *et al.*, 2013; Majka *et al.*, 2016). However, parietal cortex is likely more broadly involved in more complex tasks than typically investigated classical constrained reaching tasks. Induction of optic ataxia through muscimol delivery to posterior parietal cortex or transcranial magnetic stimulation is one straightforward example where deficits are only fully revealed during complex, indirect targeting tasks. Participants in these experiments show deficits in reaching for dynamic targets in 3D double step tasks in addition to the canonical optic ataxia deficit in manually targeting objects in peripheral space (Desmurget *et al.*, 1999; Battaglia-Mayer *et al.*, 2013). Furthermore, in parietal cortex there is an incredible complexity and diversity of spatially tuned neurons that are centered in a variety of coordinate reference frames that are dynamically recruited during reaching (Buneo *et al.*, 2002; Bosco *et al.*, 2017). Thus moving towards real-world complexity and dynamics in experimental designs is likely essential to fully understand reaching behavior and the neural correlates of reaching control.

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 (JFM and LS) and a grant from the Schmitt Program on Integrative Neuroscience (KHW and JFM) and the Del Monte Institute for Neuroscience at the University of Rochester (KHW and LS).

REFERENCES

Aikake, H. (1973) Information theory and an extension of the maximum likelihood principle In Petrov, B., Csaki, F. (eds) *2nd International Symposium on Information Theory* Akadémiai Kiadó Budapest, Hungary pp. 267-281

Avraham, G., Sulimani, E., Mussa-Ivaldi, F.A. & Nisky, I. (2019) Effects of visuomotor delays on the control of movement and on perceptual localization in the presence and absence of visual targets *Journal of Neurophysiology*. American Physiological Society, pp. 2259-2271.

Battaglia-Mayer, A., Ferrari-Toniolo, S., Visco-Comandini, F., Archambault, P.S., Saberi-Moghadam, S. & Caminiti, R. (2013) Impairment of Online Control of Hand and Eye Movements in a Monkey Model of Optic Ataxia. *Cerebral Cortex*, **23**, 2644-2656.

Blair, R.C., Higgins, J.J., Karniski, W. & Kromrey, J.D. (1994) A Study of Multivariate Permutation Tests Which May Replace Hotelling's T2 Test in Prescribed Circumstances. *Multivariate Behav Res*, **29**, 141-163.

Bosco, A., Piserchia, V. & Fattori, P. (2017) Multiple Coordinate Systems and Motor Strategies for Reaching Movements When Eye and Hand Are Dissociated in Depth and Direction. *Front Hum Neurosci*, **11**, 323.

Brenner, E. & Smeets, J.B. (1997) Fast Responses of the Human Hand to Changes in Target Position. *J Mot Behav*, **29**, 297-310.

Brouwer, A.-M., Brenner, E. & Smeets, J.B.J. (2002) Hitting moving objects: is target speed used in guiding the hand? *Experimental Brain Research*, **143**, 198-211.

Buneo, C.A., Jarvis, M.R., Batista, A.P. & Andersen, R.A. (2002) Direct visuomotor transformations for reaching. *Nature*, **416**, 632-636.

Castiello, U. & Dadda, M. (2019) A review and consideration on the kinematics of reach-to-grasp movements in macaque monkeys. *J Neurophysiol*, **121**, 188-204.

Cloherty, S.L., Yates, J.L., Graf, D., DeAngelis, G.C. & Mitchell, J.F. (2020) Motion Perception in the Common Marmoset. *Cereb Cortex*, **30**, 2658-2672.

Crosse, M., Foxe, J. & Molholm, S. (2020) PERMUTOOLS: A MATLAB Package for Multivariate Permutation Testing.

D'Mello, G.D., Duffy, E.A.M. & Miles, S.S. (1985) A conveyor belt task for assessing visuo-motor coordination in the marmoset (Callithrix jacchus): Effects of diazepam, chlorpromazine, pentobarbital and d-amphetamine *Psychopharmacology*, pp. 125-131.

Datta, S.R., Anderson, D.J., Branson, K., Perona, P. & Leifer, A. (2019) Computational Neuroethology: A Call to Action *Neuron*. Cell Press, pp. 11-24.

Day, B.L. & Lyon, I.N. (2000) Voluntary modification of automatic arm movements evoked by motion of a visual target. *Experimental Brain Research*, **130**, 159-168.

De Brouwer, A.J. & Spering, M. (2021) Eye-hand coordination during online reach corrections is task-dependent. Cold Spring Harbor Laboratory.

De Bruin, N., Sacrey, L.A.R., Brown, L.A., Doan, J. & 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 *Journal of Motor Behavior*, pp. 337-346.

Dekleva, B.M., Kording, K.P. & Miller, L.E. (2018) Single reach plans in dorsal premotor cortex during a two-target task. *Nat Commun*, **9**, 3556.

Desmurget, M., Epstein, C.M., Turner, R.S., Prablanc, C., Alexander, G.E. & Grafton, S.T. (1999) Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nature Neuroscience*, **2**, 563-567.

Desmurget, M. & Grafton, S. (2000) Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, **4**, 423-431.

Elliott, D., Hansen, S., Grierson, L.E.M., Lyons, J., Bennett, S.J. & Hayes, S.J. (2010) Goal-Directed Aiming: Two Components but Multiple Processes *Psychological Bulletin*, pp. 1023-1044.

Fox, D.M., Mundinano, I.C. & Bourne, J.A. (2019) Prehensile kinematics of the marmoset monkey: Implications for the evolution of visually-guided behaviors *Journal of Comparative Neurology*, pp. 1495-1507.

Georgopoulos, A.P., Kalaska, J.F., Caminiti, R. & Massey, J.T. (1983) Interruption of motor cortical discharge subserving aimed arm movements. *Experimental brain research*, **49**, 327-340.

Georgopoulos, A.P., Kalaska, J.F. & Massey, J.T. (1981) Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology*, **46**, 725-743.

Gharbawie, O.A., Stepniewska, I. & Kaas, J.H. (2011) Cortical connections of functional zones in posterior parietal cortex and frontal cortex motor regions in new world monkeys. *Cereb Cortex*, **21**, 1981-2002.

Hori, Y., Cléry, J.C., Schaeffer, D.J., Menon, R.S. & Everling, S. (2021) Functional Organization of Frontoparietal Cortex in the Marmoset Investigated with Awake Resting-State fMRI. *Cerebral Cortex*.

Kaas, J.H., Gharbawie, O.A. & Stepniewska, I. (2011) The organization and evolution of dorsal stream multisensory motor pathways in primates. *Front Neuroanat*, **5**, 34.

Kaas, J.H. & Stepniewska, I. (2016) Evolution of posterior parietal cortex and parietal-frontal networks for specific actions in primates. *Journal of Comparative Neurology*, **524**, 595-608.

Karashchuk, P., Rupp, K.L., Dickinson, E.S., Walling-Bell, S., Sanders, E., Azim, E., Brunton, B.W. & Tuthill, J.C. (2021) Anipose: A toolkit for robust markerless 3D pose estimation. *Cell Reports*, **36**, 109730.

Kastner, S., Chen, Q., Jeong, S.K. & Mruczek, R.E.B. (2017) A brief comparative review of primate posterior parietal cortex: A novel hypothesis on the human toolmaker. *Neuropsychologia*, **105**, 123-134.

Macfarlane, N.B.W. & Graziano, M.S.A. (2009) Diversity of grip in Macaca mulatta. *Experimental Brain Research*, **197**, 255-268.

Majka, P., Chaplin, T.A., Yu, H.H., Tolpygo, A., Mitra, P.P., Wójcik, D.K. & Rosa, M.G.P. (2016) Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. *Journal of Comparative Neurology*, **524**, 2161-2181.

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W. & Bethge, M. (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, **21**, 1281-1289.

Mitchell, J.F. & Leopold, D.A. (2015) The marmoset monkey as a model for visual neuroscience. *Neuroscience Research*, **93**, 20-46.

Moore, D.D., Walker, J.D., MacLean, J.N. & Hatsopoulos, N.G. (2022) Validating markerless pose estimation with 3D X-ray radiography. *J Exp Biol*, **225**.

Mundinano, I.-C., Fox, D.M., Kwan, W.C., Vidaurre, D., Teo, L., Homman-Ludiye, J., Goodale, M.A., Leopold, D.A. & Bourne, J.A. (2018) Transient visual pathway critical for normal development of primate grasping behavior. *Proceedings of the National Academy of Sciences*, **115**, 1364-1369.

Nastase, S.A., Goldstein, A. & Hasson, U. (2020) Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, **222**, 117254.

Ngo, V., Gorman, J.C., De la Fuente, M.F., Souto, A., Schiel, N. & Miller, C.T. (2022) Active vision during prey capture in wild marmoset monkeys. *Curr Biol*, **32**, 3423-3428 e3423.

Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G. & Orban, G.A. (2009) The representation of tool use in humans and monkeys: common and uniquely human features. *J Neurosci*, **29**, 11523-11539.

Prablanc, C. & Martin, O. (1992) Automatic Control During Hand Reaching at Undetected Two-Dimensional Target Displacements. *Journal of Neurophsyiology*, **67**, 455-469.

Reichenbach, A., Thielscher, A., Peer, A., Bulthoff, H.H. & Bresciani, J.P. (2009) Seeing the hand while reaching speeds up on-line responses to a sudden change in target position. *J Physiol*, **587**, 4605-4616.

Reser, D.H., Burman, K.J., Yu, H.H., Chaplin, T.A., Richardson, K.E., Worthy, K.H. & Rosa, M.G. (2013) Contrasting patterns of cortical input to architectural subdivisions of the area 8 complex: a retrograde tracing study in marmoset monkeys. *Cereb Cortex*, **23**, 1901-1922.

Rouse, A.G. & Schieber, M.H. (2015) Spatiotemporal distribution of location and object effects in reach-to-grasp kinematics. *J Neurophysiol*, **114**, 3268-3282.

Sacrey, L.A.R. & Whishaw, I.Q. (2012) Subsystems of sensory attention for skilled reaching: Vision for transport and pre-shaping and somatosensation for grasping, withdrawal and release *Behavioural Brain Research*. Elsevier B.V., pp. 356-365.

Sartori, L., Camperio Ciani, A., Bulgheroni, M. & Castiello, U. (2013) Reaching and grasping behavior in Macaca fascicularis: a kinematic study. *Experimental Brain Research*, **224**, 119-124.

Saunders, J.A. & Knill, D.C. (2003) Humans use continuous visual feedback from the hand to control fast reaching movements. *Experimental Brain Research*, **152**, 341-352.

Schiel, N. & Souto, A. (2017) The common marmoset: An overview of its natural history, ecology and behavior *Developmental Neurobiology*, pp. 244-262.

Schiel, N., Souto, A., Huber, L. & Bezerra, B.M. (2010) Hunting strategies in wild common marmosets are prey and age dependent *American Journal of Primatology*, pp. 1039-1046.

Shadmehr, R. & Krakauer, J.W. (2008) A computational neuroanatomy for motor control. *Experimental Brain Research*, **185**, 359-381.

Takemi, M., Kondo, T., Yoshino-Saito, K., Sekiguchi, T., Kosugi, A., Kasuga, S., Okano, H.J., Okano, H. & Ushiba, J. (2014) Three-dimensional motion analysis of arm-reaching movements in healthy and hemispinalized common marmosets. *Behav Brain Res*, **275**, 259-268.

Wagner, M.J. & Smith, M.A. (2008) Shared Internal Models for Feedforward and Feedback Control *Journal of Neuroscience*. Society for Neuroscience, pp. 10663-10673.

Wang, X., Liu, Y., Li, X., Zhang, Z., Yang, H., Zhang, Y., Williams, P.R., Alwahab, N.S.A., Kapur, K., Yu, B., Zhang, Y., Chen, M., Ding, H., Gerfen, C.R., Wang, K.H. & He, Z. (2017) Deconstruction of Corticospinal Circuits for Goal-Directed Motor Skills. *Cell*, **171**, 440-455 e414.

Wolpert, D.M. & Kawato, M. (1998) Multiple paired forward and inverse models for motor control. *Neural Networks*, **11**, 1317-1329.

Yoo, S.B.M., Tu, J.C., Piantadosi, S.T. & Hayden, B.Y. (2020) The neural basis of predictive pursuit *Nature neuroscience*. Nat Neurosci, pp. 252-259.