

Current Biology

Active vision during prey capture in wild marmoset monkeys

Highlights

- Marmosets employ active visual strategies for effective prey capture in the wild
- Flexible biomechanical movements are used to maximize the visual field
- Visual behavior flexibly adapted to the challenges of the environment
- Visually guided feedback of the hands is needed to capture flying prey in 3D

Authors

Victoria Ngo, Julia C. Gorman, María Fernanda De la Fuente, Antonio Souto, Nicola Schiel, Cory T. Miller

Correspondence

corymiller@ucsd.edu

In brief

Ngo, Gorman et al. characterized the natural prey-capture behavior of wild marmoset monkeys to illustrate the integral role of active vision for effective hunting. Markerless-tracking analyses revealed how marmosets visually track insects and how visual feedback guides and adjusts rapid hand movements to capture flying insects.



Report

Active vision during prey capture in wild marmoset monkeys

Victoria Ngo,^{1,6} Julia C. Gorman,^{1,2,6} María Fernanda De la Fuente,^{3,4} Antonio Souto,⁵ Nicola Schiel,⁴ and Cory T. Miller^{1,2,7,8,*}

¹Cortical Systems and Behavior Laboratory, University of California, San Diego, La Jolla, CA 92039, USA

²Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA 92039, USA

³Programa de Pós-graduação em Etnobiologia e Conservação da Natureza, Universidade Estadual da Paraíba, Campina Grande, Paraíba 58429-500, Brazil

⁴Laboratório de Etologia Teórica e Aplicada, Departamento de Biologia, Universidade Federal Rural de Pernambuco, Recife, Pernambuco 52171-900, Brazil

⁵Laboratório de Etologia, Departamento de Zoologia, Universidade Federal de Pernambuco, Recife, Pernambuco 50670-901, Brazil

⁶These authors contributed equally

⁷Twitter: @corymillerarmmo

⁸Lead contact

*Correspondence: corymiller@ucsd.edu

<https://doi.org/10.1016/j.cub.2022.06.028>

SUMMARY

A foundational pressure in the evolution of all animals is the ability to travel through the world, inherently coupling the sensory and motor systems. While this relationship has been explored in several species,^{1–4} it has been largely overlooked in primates, which have typically relied on paradigms in which head-restrained subjects view stimuli on screens.⁵ Natural visual behaviors, by contrast, are typified by locomotion through the environment guided by active sensing as animals explore and interact with the world,^{4,6} a relationship well illustrated by prey capture.^{7–12} Here, we characterized prey capture in wild marmoset monkeys as they negotiated their dynamic, arboreal habitat to illustrate the inherent role of vision as an active process in natural nonhuman primate behavior. Not only do marmosets share the core properties of vision that typify the primate Order,^{13–18} but they are prolific hunters that prey on a diverse set of prey animals.^{19–22} Marmosets pursued prey using vision in several different contexts, but executed precise visually guided motor control that predominantly involved grasping with hands for successful capture of prey. Applying markerless tracking for the first time in wild primates yielded novel findings that precisely quantified how marmosets track insects prior to initiating an attack and the rapid visually guided corrections of the hands during capture. These findings offer the first detailed insight into the active nature of vision to guide multiple facets of a natural goal-directed behavior in wild primates and can inform future laboratory studies of natural primate visual behaviors and the supporting neural processes.

RESULTS AND DISCUSSION

Prey-capture strategies

Here we analyzed 288 4K UHD videos of marmosets engaged in active prey pursuit and capture of insects in northeastern Brazil to characterize the relationship between natural visual and positional behavior in a wild primate. Hunting behaviors were grouped into three of the more typical tactics distinguished by the sequence of visually guided motor actions employed to capture different prey types in the dynamic forest environment.¹⁹ These three strategies were largely determined by the prey itself, such as whether it was moving on a substrate or flying, or was stationary and relying on cryptic coloration, with variation within each strategy reflecting the nuanced interactions between prey behavior and the immediate substrate idiosyncrasies. Although our analyses highlight hunting of insects, marmosets also actively hunt other prey (e.g., lizards, N = 27 videos).

Figure 1A illustrates “mouth capture,” the strategy employed by marmosets to track small, mobile prey with limited capture avoidance behaviors (e.g., ants, termites, etc., N = 80 videos; Video S1) characterized by the overwhelming use of their mouth to capture insects rather than their hands (n = 73 observations; mouth, 90.4%; hands, 9.6%; p < 0.01, Fisher’s exact test). The prevalence of mouth use in this context contrasted with all other prey-capture strategies.

Figure 1B depicts “stalk/pause and lunge,” the strategy employed by marmosets to capture stationary prey that often relied on camouflage to avoid capture (e.g., stick bugs, moths, grasshoppers, N = 93 videos; Video S2). Notably, insects that evolved mimicry increase the difficulty of their detection and are akin to a natural visual pop-out task to marmosets.^{23,24} Some of these species, while stationary on a substrate, enact fast evasive behaviors when pursued by predators. Grasshoppers and dragonflies, for example, will rapidly jump or fly away if predators are



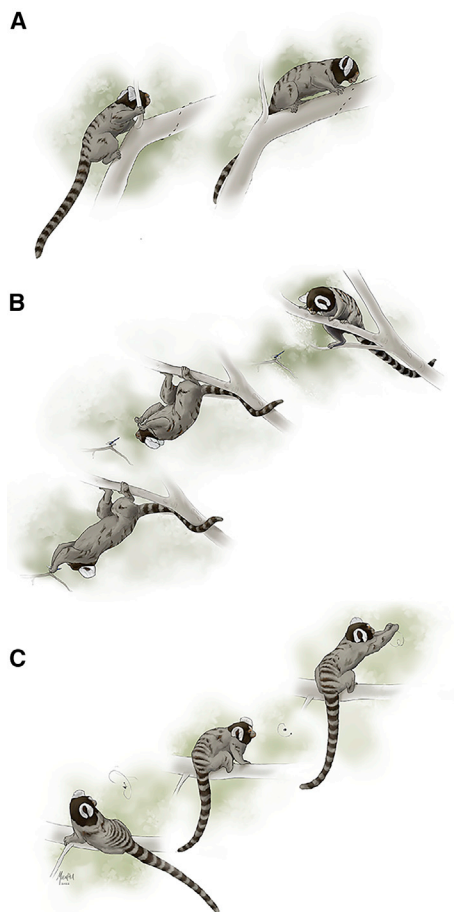


Figure 1. Marmoset prey-capture strategies

(A) “Mouth capture.” The marmoset stands over the prey and captures multiple insects with their mouth. See also [Video S1](#).

(B) “Stalk/pause and lunge.” The animal identifies the prey, gets into position, and pauses before making a ballistic grasp. See also [Video S2](#).

(C) “Capture in flight.” The monkey identifies and tracks the prey before capturing it with a visually guided grasp. See also [Video S3](#).

detected, while stick bugs evade predation by holding their legs against their body and drop to the ground to hide among the leaves.²⁵ As a result, when hunting these prey, marmosets position themselves for the attack, sometimes slowly stalking the prey ([Video S2](#)). Once in position, marmosets typically pause for 1–5 s before initiating a high-speed, ballistic grasp of the prey ([Video S2](#)).

Hunting flying insects is particularly challenging for marmosets with respect to both the demands on active vision and the role of sensory feedback to guide precise, high-speed hand movements necessary for successful capture ([Figure 1C](#)). In the recordings of marmosets pursuing flying beetles, their behavior typically abided by one of two strategies. It either involved a stationary monkey visually tracking the prey for a period of several seconds before initiating the ballistic grasp ($N = 47$ videos; [Video S3](#)) or the animal simultaneously visually tracking and physically following the prey through the arboreal substrate as the insect’s flight pattern changed ($N = 68$ videos; [Video S3](#)). Two-handed captures were more common for flying insects than one-handed

or mouth captures (64.7%, $n = 224$ observations), likely because this tactic yielded a notably higher success rate (82.4%) than one-handed captures ($\chi^2(1) = 9.5$, $p = 0.002$).

Positional behavior during arboreal prey capture

The stability and orientation of the substrate in which prey pursuit occurred were significant factors that affected the variability of prey capture. Marmosets countered these challenges with creative changes in positional behavior: the use of adaptive positioning with multiple limbs on horizontal branches, including reaching both above ([Figure 2A](#)) and below ([Figure 2B](#)) the branch, and vertical branches that, likewise, involved pursuing prey above ([Figure 2C](#)) and below ([Figure 2D](#)) the substrate. Notably, one shared similarity across all positional behaviors during prey capture was optimizing visual access of the target while balancing the need for stability, sometimes extending their body using only their hindlimbs as support at a range of angles from the substrate, often on small unstable branches. We quantified these positional behaviors by measuring the monkey’s angle of attack and the extent to which individuals extended their body in the attack ([Figure S1](#)). The average percent body length extension during prey capture for marmosets on top of the substrate was $54.3\% \pm 32.0\%$ (mean $\pm$ SD), while body extension when clinging under the substrate was $65.8\% \pm 36.3\%$ (mean $\pm$ SD). Marmosets hunting on a vertical substrate with the normal vector of the head pointing upward extended their body $57\% \pm 29.1\%$ (mean $\pm$ SD), while marmosets hunting on a vertical substrate with the normal vector of the head pointing downward extended their body $51.1\% \pm 39.7\%$ (mean $\pm$ SD) ([Figure 2E](#)). When marmosets are under a horizontal substrate, they will typically extend downward 95% of the time ($n = 20/21$, $N = 156$ observations) as they are already hanging from the branch with their lower limbs, whereas when they are on top of the horizontal substrate, they will extend downward only about 12% of the time ($n = 9/73$, $N = 156$ observations). These results illustrate how dynamic biomechanical movements are integral for active vision and successful prey capture in marmosets as these changes in positional behavior effectively function to optimize the view of the prey in a complex environment.

Gaze tracking of flying insects

To precisely quantify different facets of visual behaviors during prey capture in wild marmosets, we next applied the markerless tracking technology SLEAP^{26,27} to a subset of videos that met a set of criteria related to video recording quality and modeling accuracy ([STAR Methods](#); [Figure S2](#)). These analyses focused on hunts of flying insects because this context highlights the unique challenge of successfully capturing a moving insect in three-dimensional space and the role of vision as an active process for a goal-directed, high-precision motor action. We distinguished between two phases of the marmoset hunt based on the visual challenges. In the first phase, marmosets visually identify and track the prey prior to attacking. During the second phase, marmosets initiate the ballistic grasp based on the anticipated location of the prey.

We first analyzed head movements in the final seconds prior to the initiation of the ballistic grasp as a proxy for gaze tracking, including how they covaried with the flight path of the insect prey in a subset of 5 videos that met our criteria. Frame grabs from an

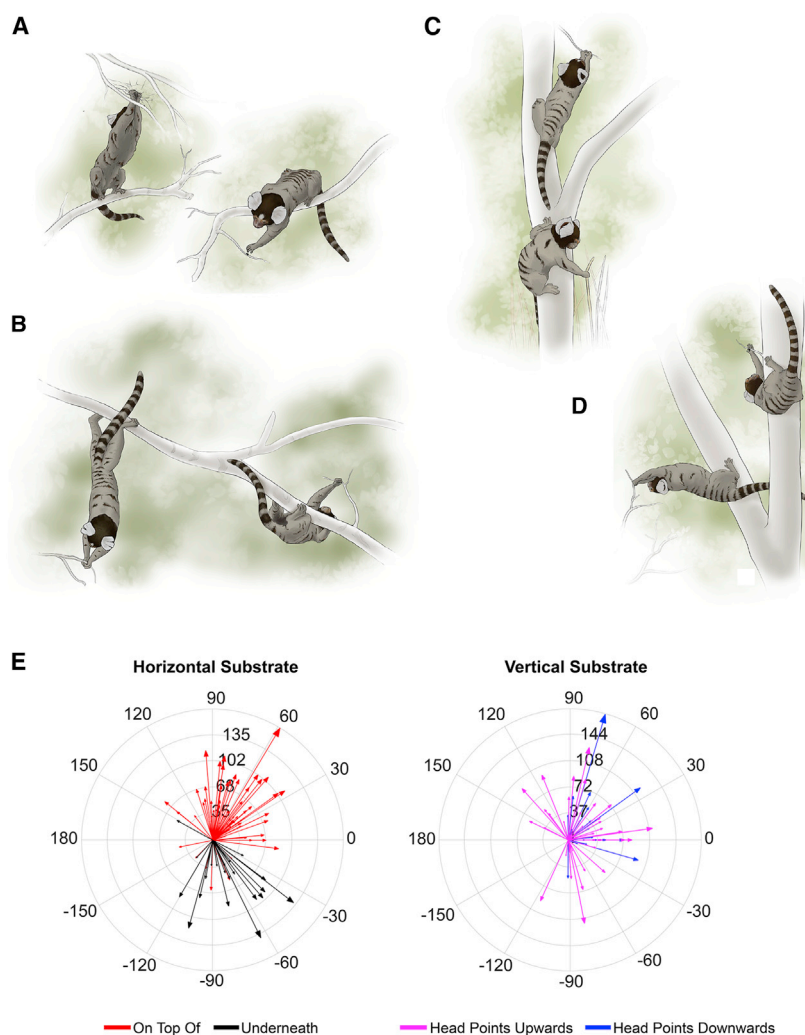


Figure 2. Positional behavior preferences based on substrate orientation

(A–D) Illustrations depicting the range of positional behaviors during hunting (reaching above or below) while the marmoset is situated (A) on horizontal substrate, (B) under horizontal substrate, (C) on vertical substrate with the normal vector of the marmoset's head pointing up, and (D) on vertical substrate with the normal vector of the marmoset's head pointing down.

(E) The length of each vector on the polar plots (shown in black, red, magenta, and blue) corresponds to the percent change in body length of extension calculated in pixel units. The polar plots also portray the marmoset's angle of extension when reaching for prey depending on the substrate being grasped. Vectors pointing below the horizontal axis refer to marmosets reaching downward while hunting. Horizontal substrates are shown to the left and vertical substrates to the right. See also Figure S1.

Visually guided prey capture

Because the position of flying insects is constantly changing, successful capture relies on a complement of two overlapping visually guided processes. First, marmosets must anticipate the likely position of the insect in three-dimensional space. Second, as insects frequently change their flight pattern, marmosets need to make quick adjustments in response to changes in insect trajectory after the ballistic grasp is initiated. Notably, these elements of the behavior may be unique to primates, as they involve visual control of coordinated hand movements and adjustment

exemplar video show three time points over 500 ms immediately before the ballistic grasp is initiated and highlight the close relationship between head and insect movements that occur during this time period (Figure 3A). Correlation between insect and head movements was largely limited to the final period—1.5 s—before initiating capture. Figure 3B shows the change in XY coordinates for the head and insect movements in the same exemplar video over a longer period of time. This observation was also true for all videos analyzed. The Pearson correlation coefficient between head and insect movements remained between 0.5 and 1 ($N = 5$ observations) for the entirety of this time period (Figure 3C). One other notable result from this analysis was the prodigious increase in head velocity over the 500 ms prior to initiating the capture action (Figure 3D). This likely suggests that marmosets were closely tracking the flight path, with vision providing important information about the speed and motion direction of the prey, and potentially reflecting a change in the attentional demands necessary to accurately capture the flying beetle. This analysis may also suggest that marmosets have a preference for a particular type of insect flight behavior when initiating an attack, when the insects are moving along a continuous path rather than at times when the path was less predictable.

of hand shape to effectively grasp prey (or other objects) in three-dimensional space.²⁸ To quantify the second phase of the hunt, visually guided prey capture, we again applied SLEAP to a subset of 15 videos that met our criteria (STAR Methods). Figure 4A depicts a parallel series of frame grabs from an exemplar video and the respective XY coordinates of the hand movements when a marmoset reaches for and grasps a flying insect that illustrate how the respective position of the hands and insect change over this time. Importantly, the hands did not follow the optimal trajectory in any of the videos (Figure 4B) and diverged in several quantifiable ways. First, the hand movements had several inflections reflecting the changes in trajectory during the motor action (average = 2.4, range = 0–7; Figure 4C). Second, the average tortuosity for the hand movement trajectory was 3.08, which significantly diverged from the optimal path ($t(24) = 5.31$, $p < 0.0001$; Figure 4D). These corrections in hand trajectory are occurring at high speeds, as the latency to peak velocity occurred at an average of 224.44 ms (Figure 4E), while the mean duration of the entire ballistic grasp during capture was 375.61 ms (Figure 4F; $N = 123$ observations). Our analyses indicate that marmosets made real-time visually guided

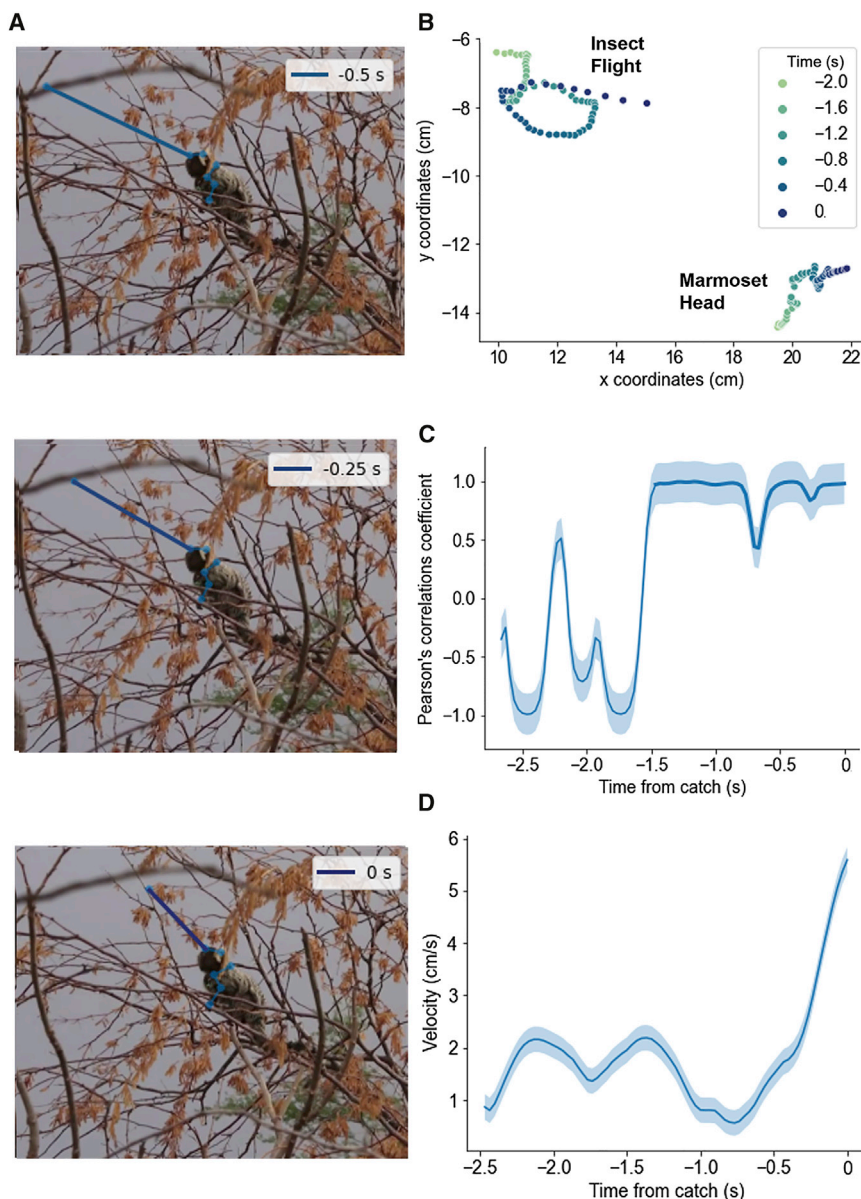


Figure 3. Gaze tracking flying insects during prey pursuit

Gaze tracking was characterized by the relationship between marmoset head and insect movements prior to initiating capture using SLEAP for accurate annotations. See also Figure S2.

(A) Three single frames from the final 500 ms before prey capture illustrate coordinated movements between the marmoset head and insect flight trajectory.

(B) Raw XY positional data of both the marmoset's head movement (bottom right) and insect's flight pattern (top left) from the example screen shown in (A) over the 2 s prior to initiating the ballistic grasp for prey capture.

(C) Pearson's correlation coefficient between the marmoset head and insect flight in a sliding window over the 2.5 s before prey capture was initiated. Zero (0) indicates the onset of the ballistic grasp.

(D) Average velocity (cm/s) of the marmoset head over the same time period as in (C).

95% confidence intervals are shown in shading for (C) and (D).

advantage of this natural behavior is that it comprises visual processes studied independently in primate vision for decades—discrimination, recognition, motor planning, decision-making, and visually guided selection, among others—within a single, cohesive visual behavior.³² These visual processes are in fact not separable, but operations within an integrated behavioral sequence that is representative of the distinct challenges that together have driven the evolution of the complementary mechanisms in the primate visual system, and in particular precise visually guided reaching and control of hand movements and articulation that may be unique to our Order. While marmoset hunting strategies were largely determined by the behavior of the prey type, in contrast to many classic neuroethological behaviors, prey capture

corrections to the path of the hand trajectory during the short interval of time from the initiation of the motor action until the insect was captured, likely due to changes in the prey's flight path or potentially other insects in the visual field. The trajectories of the left and right hands were not statistically different ($t = -0.56$, $p = 0.589$). This finding shows that, like other primates,^{29–31} wild marmoset hand movements are under continuous visual control when targeting the capture of flying beetles and can be modified through feedback in response to the exact types of natural challenges the visual system evolved mechanisms to overcome.

Here we demonstrate that active vision is integral to successful prey capture in wild marmosets. Far from being a passive process, vision is closely coupled to all elements of prey capture in marmosets, including creative biomechanical positioning on often precarious substrates, similarly to other species.^{7–11} A compelling

itself was far from stereotyped in these monkeys. Rather, the general hunting strategies were highly variable, likely reflecting the need to be adaptable to the immediate environment to find creative solutions for successful predation. The behavioral variability that emerges due to an interaction between a presumably ideal strategy for obtaining prey and the specific ecological challenges underscores the advantages of primate prey capture to elucidate the likely supporting perceptual and cognitive processing, including decision-making and cognitive control, under natural variable conditions, and can be used as principles that guide future experimental work on these key issues. As not all theoretical models of vision developed in more traditional laboratory settings are likely reflective of how this system functions under real-world scenarios,³³ natural primate behaviors such as prey capture will likely be necessary to elucidate remaining key questions about primate brain function.

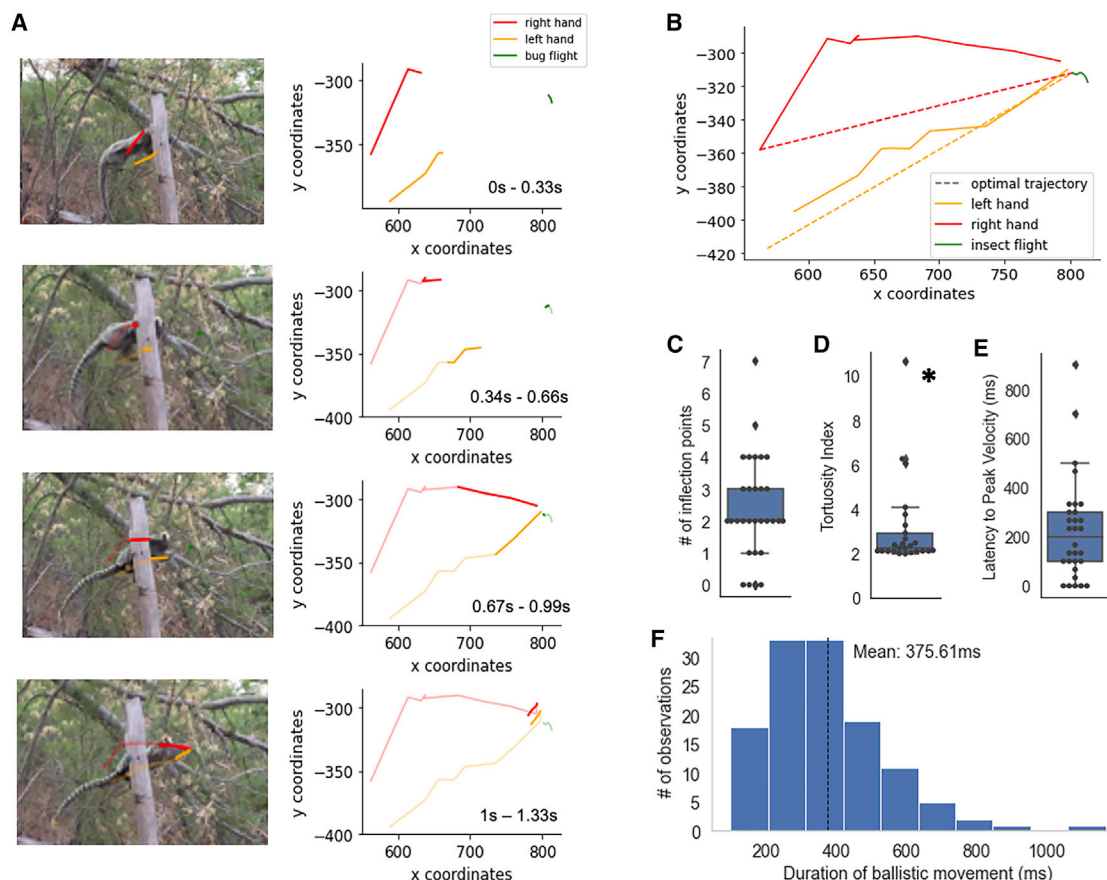


Figure 4. Visually guided prey capture

(A) Left plots four continuous time periods during prey capture from an exemplar video and the XY coordinates of the marmoset's right (red line) and left (yellow line) hands and insect (green line); right plots only the XY coordinates. The dark lines show the change in movement during the time interval plotted over that specific period of time, while lightly shaded lines indicate the preceding time periods.

(B) Summary of (A) shows the path of both hands and the insect, as well as the optimal trajectory in blue for each hand.

(C) Number of inflection points during reaches from all videos analyzed. Error bars represent SD.

(D) Plots the tortuosity index for all prey-capture reaches. Significance is $p < 0.0001$. Error bars represent SD.

(E) Latency to peak velocity (ms) during ballistic grasps for flying prey. Error bars represent SD.

(F) Duration of ballistic movement (ms) from initiation of the motor action to successful grasp of the prey.

See also Figure S2.

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
 - Study site and subjects
- METHOD DETAILS
 - Data analysis
 - Positional behavior during arboreal prey capture
 - Markerless tracking analyses
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Statistical analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at Database: <https://doi.org/10.1016/j.cub.2022.06.028>.

ACKNOWLEDGMENTS

We thank Drs. T. Perreira, A. Fishbein, D. Grijseels, and A. Lefevre for comments; Dr. G. Baracruhy for permission to conduct research at Baracruhy Biological Field Station; and Sarah Mientka for the marmoset illustrations. This work was supported by the NIH (UF1 NS116377) and AFOSR (19RT0316) to C.T.M.

AUTHOR CONTRIBUTIONS

V.N. managed the study, curated the data, analyzed the data, and wrote the paper. J.C.G. curated the data, developed analysis methods, analyzed the data, and wrote the paper. M.F.D.I.F. collected all the data. A.S. conceptualized the study and wrote the paper. N.S. conceptualized the study, supervised the field data collection, and wrote the paper. C.T.M. conceptualized the study, supervised the study, acquired the funding, and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 1, 2022

Revised: June 8, 2022

Accepted: June 9, 2022

Published: June 23, 2022

REFERENCES

- Kleinfeld, D., Ahissar, E., and Diamond, M.E. (2006). Active sensation: insights from the rodent vibrissa sensorimotor system. *Curr. Opin. Neurobiol.* 16, 435–444. <https://doi.org/10.1016/j.conb.2006.06.009>.
- Parker, P.R.L., Brown, M.A., Smear, M.C., and Niell, C.M. (2020). Movement-related signals in sensory areas: roles in natural behavior. *Trends Neurosci.* 43, 581–595. <https://doi.org/10.1016/j.tins.2020.05.005>.
- Wachowiak, M. (2011). All in a sniff: olfaction as a model for active sensing. *Neuron* 71, 962–973. <https://doi.org/10.1016/j.neuron.2011.08.030>.
- Gibson, E.J. (1979). *The Ecological Approach to Visual Perception* (Houghton Mifflin).
- Leopold, D.A., and Park, S.H. (2020). Studying the visual brain in its natural rhythm. *Neuroimage* 216, 116790. <https://doi.org/10.1016/j.neuroimage.2020.116790>.
- Schroeder, C.E., Wilson, D.A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* 20, 172–176. <https://doi.org/10.1016/j.conb.2010.02.010>.
- Palleroni, A., Miller, C.T., Hauser, M., and Marler, P. (2005). Predation: prey plumage adaptation against falcon attack. *Nature* 434, 973–974. <https://doi.org/10.1038/434973b>.
- Lin, H.-T., and Leonardo, A. (2017). Heuristic rules underlying dragonfly prey selection and interception. *Curr. Biol.* 27, 1124–1137. <https://doi.org/10.1016/j.cub.2017.03.010>.
- Mischianti, 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>.
- Michaels, A.M., Abe, E.T., and Niell, C.M. (2020). Dynamics of gaze control during prey capture in freely moving mice. *Elife* 9, e57458. <https://doi.org/10.7554/elife.57458>.
- Wagner, H., Kettler, L., Orlowski, J., and Tellers, P. (2013). Neuroethology of prey capture in the barn owl (*Tyto alba* L.). *J. Physiol. Paris* 107, 51–61. <https://doi.org/10.1016/j.jphysparis.2012.03.004>.
- Hoy, J.L., Yavorska, I., Wehr, M., and Niell, C.M. (2016). Vision drives accurate approach behavior during prey capture in laboratory mice. *Curr. Biol.* 26, 3046–3052. <https://doi.org/10.1016/j.cub.2016.09.009>.
- Mitchell, J.F., Reynolds, J.H., and Miller, C.T. (2014). Active vision in marmosets: a model system for visual neuroscience. *J. Neurosci.* 34, 1183–1194. <https://doi.org/10.1523/jneurosci.3899-13.2014>.
- Mitchell, J.F., and Leopold, D.A. (2015). The marmoset monkey as a model for visual neuroscience. *Neurosci. Res.* 93, 20–46. <https://doi.org/10.1016/j.neures.2015.01.008>.
- Mitchell, J.F., Priebe, N.J., and Miller, C.T. (2015). Motion dependence of smooth pursuit eye movements in the marmoset. *J. Neurophysiol.* 113, 3954–3960. <https://doi.org/10.1152/jn.00197.2015>.
- Coop, S.H., Yates, J.L., and Mitchell, J.F. (2021). Foveal remapping of motion in area MT of the marmoset monkey. *J. Vis.* 21, 2638. <https://doi.org/10.1167/jov.21.9.2638>.
- Hori, Y., Cléry, J.C., Selvanayagam, J., Schaeffer, D.J., Johnston, K.D., Menon, R.S., and Everling, S. (2021). Interspecies activation correlations reveal functional correspondences between marmoset and human brain areas. *Proc. Natl. Acad. Sci. USA* 118, e2110980118. <https://doi.org/10.1073/pnas.2110980118>.
- Hung, C.C., Yen, C.C., Ciuchta, J.L., Papoti, D., Bock, N.A., Leopold, D.A., and Silva, A.C. (2015). Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. *J. Neurosci.* 35, 1160–1172. <https://doi.org/10.1523/jneurosci.2659-14.2015>.
- Schiell, N., Souto, A., Huber, L., and Bezerra, B.M. (2010). Hunting strategies in wild common marmosets are prey and age dependent. *Am. J. Primatol.* 71, 1–8.
- Schiell, N., and Souto, A. (2017). The common marmoset: an overview of its natural history, ecology and behavior. *Dev. Neurobiol.* 77, 244–262. <https://doi.org/10.1002/dneu.22458>.
- Digby, L., and Barreto, C.E. (1998). Vertebrate predation in common marmosets. *Neotrop. Primates* 6, 124–126.
- Souto, A., Bezerra, B.M., Schiell, N., and Huber, L. (2007). Saltatory search in free-living *Callithrix jacchus*: environmental and age influences. *Int. J. Primatol.* 28, 881–893. <https://doi.org/10.1007/s10764-007-9165-1>.
- Bichot, N.P., and Schall, J.D. (1999). Saccade target selection in macaque during feature and conjunction visual search. *Vis. Neurosci.* 16, 81–89. <https://doi.org/10.1017/s0952523899161042>.
- Wang, Q., Cavanagh, P., and Green, M. (1994). Familiarity and pop-out in visual search. *Percept Psychophys* 56, 495–500. <https://doi.org/10.3758/bf03206946>.
- Markle, S. (2008). *Stick Insects: Masters of Defense* (Lerner Publications).
- Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.H., Murthy, M., and Shaeivitz, J.W. (2019). Fast animal pose estimation using deep neural networks. *Nat. Methods* 16, 117–125. <https://doi.org/10.1038/s41592-018-0234-5>.
- Pereira, T.D., Shaeivitz, J.W., and Murthy, M. (2020). Quantifying behavior to understand the brain. *Nat. Neurosci.* 23, 1537–1549. <https://doi.org/10.1038/s41593-020-00734-z>.
- Cartmill, M. (2005). New views on primate origins. *Evol. Anthropol.* 1, 105–111. <https://doi.org/10.1002/evan.1360010308>.
- Oostwoud Wijdenes, L., Brenner, E., and Smeets, J.B.J. (2011). Fast and fine-tuned corrections when the target of a hand movement is displaced. *Exp. Brain Res.* 214, 453–462. <https://doi.org/10.1007/s00221-011-2843-4>.
- Goodale, M.A., Pelisson, D., and Prablanc, C. (1986). Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* 320, 748–750. <https://doi.org/10.1038/320748a0>.
- Song, J.-H., Takahashi, N., and McPeck, R.M. (2008). Target selection for visually guided reaching in macaque. *J. Neurophysiol.* 99, 14–24. <https://doi.org/10.1152/jn.01106.2007>.
- Miller, C.T., Gire, D., Hoke, K., Huk, A.C., Kelley, D., Leopold, D.A., Smear, M.C., Theunissen, F., Yartsev, M., and Niell, C.M. (2022). Natural behavior is the language of the brain. *Curr. Biol.* 32, R482–R493. <https://doi.org/10.1016/j.cub.2022.03.031>.
- Matthis, J.S., Muller, K.S., Bonnen, K.L., and Hayhoe, M.M. (2022). Retinal optic flow during natural locomotion. *PLoS Comput. Biol.* 18, e1009575. <https://doi.org/10.1371/journal.pcbi.1009575>.
- De la Fuente, M.F., Sueur, C., Garber, P.A., Bicca-Marques, J.C., Souto, A., and Schiell, N. (2022). Foraging networks and social tolerance in a cooperatively breeding primate (*Callithrix jacchus*). *J. Anim. Ecol.* 91, 138–153. <https://doi.org/10.1111/1365-2656.13609>.
- Caselli, C.B., Ayres, P.H.B., Castro, S.C., Castro, S., Souto, A., Schiell, N., and Miller, C.T. (2018). The role of extra-group encounters in a neotropical cooperative breeding primate, common marmosets: field playback experiments. *Anim. Behav.* 136, 137–146. <https://doi.org/10.1016/j.anbehav.2017.12.009>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Data and Software	Dryad	https://doi.org/10.6076/D1P88

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Cory Miller (corymiller@ucsd.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All data and original code have been deposited in a Dryad repository Database: <https://doi.org/10.6076/D1P88X> and is publicly available as of the date of publication. DOI is listed in the [key resources table](#).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Study site and subjects

High-resolution videos were recorded of wild marmosets inhabiting the semiarid scrub-forests in the Baracuh Biological Field Station in Northeast Brazil (7°31'42"S, 36°17'50"W).^{34,35} Data collection was conducted between March 2020 and June 2021 following two social groups- House Group (mean group size of 9 animals) and Coqueiro Group (mean group size of 8 animals).

METHOD DETAILS

Data analysis

Analyses were performed on the 288 Ultra HD video recordings (4K; width: 3840 pixels; height: 2160 pixels) of marmoset prey-capture on the Sony FDR AX-53 camcorder at a frame rate of 29.970 fps. Videos may contain one or more observations of a hunting strategy by one or more subjects being filmed. 287 videos that were not reflected in the final analyses exhibit marmosets foraging for food by leaf manipulation or opportunistic prey capture.¹⁹ Videos recorded in the field were uploaded in Brazil onto a shared server and cataloged at UCSD.

Positional behavior during arboreal prey capture

Quantitative analysis shown in [Figure 2](#) was performed using Adobe Premiere Pro version 22.0 and Adobe Photoshop version 22.5.1. Premiere was used to quantify the duration of a subject's hunting action frame by frame up to hundredths of a second, including the pause before a ballistic movement, and the start to completion of a ballistic movement based on the type of prey the marmoset went after. Using Premiere, the start frame of the ballistic movement marks when the marmoset's body is at its original position. The end frame of the ballistic movement reflects when the marmoset is most fully extended, reaching out towards its prey. Frames depicting a side profile of the moment before and after a ballistic movement was achieved were exported into Adobe Photoshop for further examination.

Since body lengths of animals in the wild were not possible, the ruler tool in Adobe Photoshop was used to obtain measurements in pixels including the marmoset's body length pre- and post- ballistic movement in exported frames, allowing us to attain the percent length of extension. For the pre-ballistic measurement of the marmoset's body, it was measured from the base of the tail or foot to the top of their forehead or mouth, depending on the visibility of each body part in the frame. The post-ballistic measurement was measured from either the same tail base or foot—if visible and in instances where they are using their legs to extend—to the extended body part being used to grab (either the extended hand(s) or mouth). Additionally, the ruler tool was used to estimate the angle of extension from the substrate a subject opts for depending on how they are oriented on that substrate. The angle was measured from the parallel substrate, to the base of where the marmoset is clinging onto, and to the extended arms. Owing to the limitations of a 2D video, videos demonstrating body extension along the Z plane were not taken into consideration for measuring to minimize the chance of error.

The polar plot in Figure 2E was created using a custom MATLAB script to demonstrate the marmoset's angle of extension depending on the substrate they are hunting on (horizontal or vertical) and the length of the arrow depicts a visual representation of the percent body length of extension in pixels. The “-” sign refers to the downward direction the marmoset is reaching in (Figure 2E). As a way to classify differences in the marmoset's orientation relative to the vertical substrate, we used the direction of the normal vector extending from the top of their head, as this was the most visible point of reference in the frames, to define if the marmosets were extending upright (normal vector points up) or “supine” (normal vector points down). To obtain the normalized percent change in length of extension, the following formula was used:

$$= \frac{\text{Post Ballistic Body Length (pixels)} - \text{Pre Ballistic Body Length (pixels)}}{\text{Pre Ballistic Body Length (pixels)}} \times 100$$

Since only one camera angle was often available for analysis, the ability to measure velocity during a capture was not feasible. Instead, using the aforementioned formula, we examined the normalized percent change in the marmoset's body length over the duration of the ballistic movement during a hunting sequence to capture their change in motion.

Markerless tracking analyses

To more precisely quantify the marmoset's visuomotor integration that occurs when hunting, we employed computer vision technology for markerless tracking on one dynamic hunting strategy involving a one- or two-handed capture mid-air following rapid head-gaze tracking behavior of flying beetles (Coleoptera). Down-sampled videos (1280 x 720) were uploaded onto SLEAP (Social LEAP Estimates Animal Poses)^{26,27} where deep learning was employed for markerless tracking of the marmoset body, whereas the flying insect was hand annotated either on a separate project or following the inference process. The right hand, left hand, middle top of head, left ear tuft, and right ear tuft of the marmoset in the video were manually annotated in a small percentage (~15%) of frames. Then, active learning takes place as we train a neural network through the single animal pipeline to estimate positions of the marmoset's body parts by running inference until satisfied with the accuracy of the network's predictions. A custom MATLAB script was used to verify the accuracy of the computer's predictions by comparing the average distance between the hand labels to the predicted labels (Figure S2). While pose estimation was computed for the entire video clip (typically ~15s) to improve the robustness and accuracy of the model, analyses presented here focused on the few seconds before and during insect capture. These select frames were further refined for analysis by additional hand annotations. H5 files were then exported for use in additional analyses. The frame numbers for start and end behaviors, such as gaze tracking and ballistic movement, were hand marked and verified by two other people in the Miller Lab.

Gaze-tracking flying beetles

Analysis was performed on videos that met the following criteria: (1) Video maintained stable, continuous focus of both the insect and the animals head for at least 2 seconds prior to initiating capture, (2) the animal was not chasing the insect during this period or otherwise locomoting, but remained sitting on the substrate and (3) there were no obstructions of either the insect or marmoset that would affect annotation of the video. Five videos met these strict criteria and were used in the analysis typically due to the flying insect not remaining in focus for sufficient periods of time during the video. The frame number of both the start of the marmoset gaze and the end of the gaze/start of the ballistic movement was marked in each video. The start of the gaze was defined as when the marmoset's head movement stops scanning and the head moves in closer before the body starts to move in the same direction. The end of the gaze period is defined as when the marmoset's head movement shifts away from the targeted prey, usually to redirect its attention to the route they are trying to take as they move closer or to check if there are other competitors around. The gaze can then resume when the head movement turns back to the direction of where the target is/where the target is moving. Videos were processed in Python 3.6 and the aforementioned time periods mentioned during the gaze period were analyzed (Figure 3A). Gaze estimates in 3B were made between the SLEAP label that demarcated the middle of the head and the SLEAP bug label. A line was drawn from the two to show an approximate gaze line. In 3C, velocity of both the head SLEAP labels and the insect's SLEAP label were calculated. A rolling correlation was taken to quantify the relationship between the head and the bug in the final second. The Python pandas package outputs the Pearson's Correlation Coefficient in a sliding bin window with a bin size of 10. In order to characterize the period of high correlation before the bug catch, a rolling average of bin size 10 was taken and any correlation coefficient above 0.5 was considered significant. A 95% confidence interval was calculated using Python in order to show the variability of the data in both 3C and 3D.

Visually-guided prey capture

These analyses were performed on videos that maintained stable, continuous focus and/or accurate model prediction of both the insect and at least one of the animal's hands for the duration of time from the initiation of the ballistic grasp till prey capture. The frame number of both the start and end of the marmoset's ballistic movement was marked in each video. Start of ballistic grasp was defined as when the marmoset lifts its hand(s) from the branch or hand(s) move away from the body towards the target without stopping until the target is caught. End of ballistic grasp was defined as when the target prey is caught. Analyses were performed on 15 videos that met these criteria and had the least occlusion by vegetation. Lines were drawn from the start of the ballistic movement to the point of prey capture to demonstrate the most direct path - the optimal trajectory - of a reach if the insect's flight path was perfectly predicted. We next applied two different measurements to quantify how much marmoset hand trajectories deviated from the optimal trajectory: ‘Number of Inflection points’ and ‘Tortuosity’. The left and right hands were analyzed separately in each video because both hands were not always clearly visible. This yielded 15 left hand events and 9 right hand events from the videos used in analyses here.

- *Inflection points*. This analysis identified the number of instances that marmosets modified the direction of their hand movement during the ballistic grasp leading to prey capture. This measure was taken by first calculating the distances from the ideal line and then calculating the number of local maxima and minima. Values were validated by marking the point at which a local max and min were found.
- *Tortuosity*. - To better quantify the changes and curvature of the arm reaches during prey capture of flying insects, we measured the tortuosity of the reach trajectory. Tortuosity was calculated to show the ratio between the distance of the optimal trajectory or total length (L) and the actual movement of the hand or the path length (C). This ratio was then $\tau = C/L$. In order to test for significance, a paired t-test was performed against the null hypothesis of 1, since 1 would be the value of τ if $C = L$.
- *Latency to peak velocity*. We calculated the velocity of the hand over the course of the ballistic action to determine the time at which the velocity was at its maximum.

To calculate the duration of ballistic reaches to capture flying insects for [Figure 4E](#), we lowered our selection criteria to all videos in which the body was visible (N = 123 observations). The frame numbers were denoted for the duration of the ballistic movement and the average was calculated.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

Statistical analyses of SLEAP output were performed using Python. Excel (Microsoft Corporation) was used to conduct all other statistical analyses, including mean values and standard deviation. We used the Fisher Exact Test with probability mass function to compare methods of capture on small, mobile prey. To examine whether there was a preference for capturing small, flying beetles using one or two hands, a χ^2 test was applied. Statistical significance was set to $p < 0.05$ (two tailed) for all analyses. Statistical analysis of [Figure 4E](#) was a two sided t test performed in Python.