Independently learned behavioral strategies support coordination in common marmosets

Oviya Mohan (omohan@ur.rochester.edu)[1], Janejira Chapoomee[2], and Dora Biro[1]
[1]Department of Brain and Cognitive Sciences, University of Rochester
[2]Department of Neuroscience, University of Rochester Medical Center

Abstract

Coordination, where two or more individuals align their actions to achieve a common goal, has been documented in several species using the rope-pulling paradigm. Nonetheless, the extent to which subjects understand their partner's role in facilitating successful coordination remains unclear. We tested five pairs of common marmosets (*Callithrix jacchus*) to address this question using successive phases with increasing demand for coordination (decreasing rope length and delaying one monkey). Two of the tested pairs succeeded on all rope lengths as well as a 5s delay but failed a 10s delay. Behavioral analyses showed that these two pairs achieved success through different mechanisms employed by each subject, suggesting that individually learned behavior may be sufficient to support coordination in the absence of shared intentionality, at least in the context of a physically intuitive task. Our study yields novel insights into the cognitive underpinnings of coordination in a cooperatively breeding species.

Keywords: coordination, primates, marmosets, rope-pulling

Introduction

Coordination can be defined as an individual acting in accordance with the actions of one or more other individuals for the fulfilment of a common goal (Camerer, 2011). Often this involves individuals performing the same action, also known as pure coordination. For example, everyone driving on the left or right side of the road within a population (e.g., a country) solves the coordination problem of avoiding collisions. Explanations of how humans solve such problems often invoke shared intentionality, i.e., the ability to recognize other agents as sharing one's goals (Hawkins et al., 2019; Tomasello & Moll, 2010). However, coordination problems are not unique to humans and can be achieved through simpler heuristics-based mechanisms. For example, in the bidirectional traffic lanes of army ant (*Eciton burchelli)* foraging columns, the same problem of head-on collisions is avoided through asymmetrical preferences in in- and out-bound ants' turning angles as they approach an oncoming ant (Couzin & Franks, 2003). While this simple heuristic is likely genetically encoded in army ants, elsewhere the challenge of settling on one of several options at the population level requires flexibility and learning at the individual level. To what extent such learning relies on simple heuristics versus perceiving and understanding the problem as one involving others remains an open question.

To investigate the emergence of coordination in controlled settings, the string-pulling or rope-pulling paradigm (Crawford, 1937) has become a widely used approach. This physically intuitive task requires two individuals to pull on both ends of a rope simultaneously to obtain a reward. If only one individual pulls, the rope unravels from the apparatus resulting in neither individual getting a reward. Performance on this task and its variants varies significantly across species. Capuchins (*Cebus apella*), for example, in a similar handle-pulling task, often fail to account for their partner's readiness or presence, with success occurring primarily by chance rather than shared intentionality (Visalberghi et al., 2000). Chimpanzees (*Pan troglodytes*) demonstrate higher success, and even the ability to selectively recruit helpful partners (Melis et al., 2006), but analysis of gazing behavior suggests coordination may rely on one individual monitoring the other rather than a joint understanding of the task (Hirata & Fuwa, 2007). Elephants (*Elephas maximus*), on the other hand, reliably modify their behavior based on whether their partner is delayed or has access to rope, suggesting a more robust understanding of their partner's role (Plotnik et al., 2011). These divergent results raise a fundamental question about whether seemingly coordinated behavior emerges from an understanding of a partner's necessity, or if it can be explained by simpler, individually learned contingencies.

Common marmosets (*Callithrix jacchus),* a species of small-bodied New World monkey, represent intriguing subjects for exploring these questions. Like humans, they engage in pair bonding and cooperative breeding, where non-parent adults help take care of the young. This is theorized to have significantly shaped their socio-cognitive abilities (Burkart & Finkenwirth, 2015; Miller, 2025) including a propensity for prosocial behavior (Burkart & van Schaik, 2020), following gaze cues (e.g., Burkart & Heschl, 2006), and collective problem solving (e.g., Sehner et al., 2022) amongst others. A recent study tested coordination in marmosets using an automated lever-pulling task and demonstrated usage of flexible partner-dependent strategies and reliance on "social vision" (monitoring of the partner's behavior) to achieve successful coordination (Meisner et al., 2025). However, this task lacked the causal transparency of the rope-pulling paradigm which allows individuals to directly witness the consequences of their own as well as their partner's interactions with the apparatus. More generally, it is possible that the intuitiveness of the physical problem plays an important part in scaffolding subjects' understanding of the task and hence of the role of their partner.

In the present study, we developed a rope-pulling task to test pairs of common marmosets across phases with increasing demand for coordination implemented both by requiring more precisely timed actions between the individuals as well as delaying one of the individuals at random. We also analyzed gazing behavior to characterize

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

the mechanism by which successful coordination occurs. Our goal was to examine the cognitive underpinnings of coordination in a cooperatively breeding species.

Methods

Subjects

We tested five pairs of common marmosets ranging from 1 to 4 years of age in the rope-pulling paradigm (Table 1). All five pairs were cage-housed in a colony room with other marmoset families present; three were pair-housed with only their partner, while two lived in larger family groups (8-10 individuals). Marmosets had free access to laboratory chow (callitrichid diet) and water and were placed on a 12-hour light/dark cycle. None of the marmosets had participated in any prior experiments and were not food or water deprived before training or testing sessions. All experimental procedures were approved by the Institutional Animal Care and Use Committee at the University of Rochester.

Table 1. Subjects.

Pair	Subjects	Age	Sex	Relationship
AW	Apple Walnut	2y 3m 1y 4m	F M	Breeding pair (no offspring)
FZ	Flint Zebib	1y 7m 3y 1m	M F	Breeding pair (no offspring)
EK	Ester Kofi	3y 6m 2y 3m	F M	Non-breeding pair (not related)
PS	Poptart Strudel	1y 2m 1y 2m	F F	Twins (housed with family)
PC	Pecan Chestnut	1y 3m 1y 3m	M M	Twins (housed with family)

Individual Training

Each marmoset was familiarized with the apparatus and individually trained to pull rope to obtain a reward in their home cage. The individual being trained was separated from their partner/rest of the family into one section of the cage. The same apparatus used for the main task (Fig. 1) was used for individual training, except only a single slider was present with the rope threaded through it. Multiple stages were used to familiarize the marmosets with the task apparatus, starting with familiarization with the reward tube, slider, and rope, leading up to the final stage, which required the marmoset to pull the rope to move the slider with the reward (mini marshmallows) towards them to obtain it. Both marmosets in a pair had to pass this final stage of training to move onto pair testing. Once they passed individual testing, they were habituated to being transported to and being present in the testing rig (Fig. 1) as a pair before pair testing commenced.

Pair Testing

Pairs were tested across sessions of 10 trials each. At the start of each session, pairs were brought to the testing rig and allowed to habituate for 5 minutes. Habituation involved allowing the marmosets to freely walk back and forth across

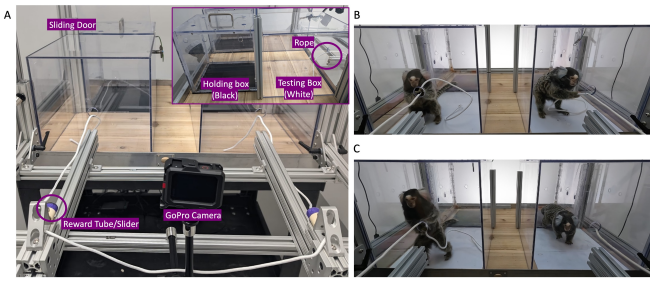

Figure 1. Experimental set-up. A. Subjects were released from the holding box into the testing box with access to the rope (inset – side view) via a sliding door; rope was threaded through reward tubes on both sides; pulling only one side caused the rope to unravel. **B.** Example of a successful trial with both marmosets pulling the rope simultaneously. **C.** Example of a failed trial where the marmoset on the left pulled the rope before the marmoset on the right (rope can be seen leaving the box) had reached it.

the holding and testing box to familiarize themselves with the set-up. Each trial began with both marmosets in the holding box with the sliding doors closed. Two experimenters, one on each side of the rig, opened the doors to begin the trial. At the end of a trial – whether successful (Fig. 1B) or failed due to the rope leaving the box (Fig. 1C) or due to inaction – the marmosets were led back into the holding box to reset for the next trial. Only one session was conducted per day, and pairs performed 3-4 sessions a week.

Pairs were tested across phases with increasing demand for coordination implemented both by requiring more precisely timed actions between the individuals by reducing the length of the rope (three non-delayed phases) and by delaying one of the individuals at random (two delayed phases), with each marmoset delayed on half the trials within each session (Table 2). To progress to the next phase, pairs had to achieve at least 70% success in two consecutive sessions. If a pair did not meet this criterion within 10 consecutive sessions of a phase, the pair did not progress to the next phase.

Table 2. Test phases. Rope length refers to the length of the section of rope available to each subject inside the testing box at the start of a trial; delay refers to the time interval between releasing the two subjects from their holding boxes.

Phase	1	2	3	4	5
Rope length	50 cm	25 cm	5 cm	25 cm	25 cm
Delay	0s	0s	0s	5s	10s

Behavioral Analysis

All trials were recorded using a GoPro Hero 10, and videos were coded using open-source software BORIS (Friard & Gamba, 2016) to create an ethogram of the behaviors of interest (Table 3). When a behavior was observed in the video, the corresponding timestamp was recorded in BORIS and used to calculate occurrence and duration metrics. Point events (e.g., Trial Start) were recorded as single points in

time, while state events were recorded as durations having a start and stop (e.g., $Gaze_{start}$, $Gaze_{stop}$). Modifiers were used to describe behaviors that had multiple types; for example, contact with the rope could either be the marmoset simply holding (Contact_hold), nibbling (Contact_nibble), or pulling (Contact_pull) the rope.

To assess which behavioral strategies may contribute to successful coordination, the *latency to first pull* variable (first occurrence of $Contact\_pull_{start}$ – TrialStart) was used to analyze performance across subjects, sessions, and phases as the most relevant, task-related, goal-directed behavior performed by subjects. Note that in some cases subjects stopped pulling after their first pull, resulting either in failure or in success if they later continued pulling, but such instances were rare. Hence, *latency to first pull* remains a reliable diagnostic of task learning by the marmosets. This variable also allows comparisons between successful and failed trials since subjects can pull the rope even if it does not lead to eventual success. All subsequent statistical analyses use *latency to first pull* as the behavioral outcome. In the delayed phases, *latency to first pull* for each subject was calculated from when that subject was released from the holding box (in the case of the non-delayed subject, this corresponded to TrialStart; for the delayed subject it was DelayedRelease).

Table 3. Ethogram used to code behavior.

Behavior	Type	Modifiers	Description
Trial Start	Point		**Non-delayed trials**: time when both holding boxes are opened simultaneously; **Delayed trials**: time when holding box for first marmoset is opened
Delayed Release	Point		**Delayed trials only**: time when holding box for second marmoset is opened
Contact	State	Pull, Hold, or Nibble	Marmoset makes contact with rope
Gaze	State		Marmoset looks at partner or partner's rope/reward tube/slider
Reward	Point		Marmoset obtains reward
Trial End	Point	Success or Failure	**Success**: both marmosets retrieve reward. Recorded as time when second marmoset obtains reward; **Failure**: rope leaves testing box of a marmoset or trial reaches maximum time of 2 minutes

Statistical Analysis

Non-delayed Phases. To examine the rates at which each individual subject modified their behavior (*latency to first pull*) to achieve success in our two successful pairs, we used a Generalized Linear Mixed Model (GLMM) from the glmmTMB package in R (R Core Team, 2023). Given the positive right-skewed nature of latency data, a Gamma distribution with a log-link function was used. The model evaluated the interactive effects of trial outcome (success vs. failure), session (to account for learning, session number was scaled within phase to facilitate model convergence) and phase (to account for changes in task by varying rope length). A likelihood ratio test confirmed that retaining this three-way interaction significantly improved fit over a reduced two-way interaction model ($\chi^2(2) = 35.62$, $p < 0.001$; $\Delta AIC = 31.6$). Pair was also included as a fixed effect to account for any baseline differences in latency between the pairs.

To account for repeated measures and individual differences across our subjects, the model also included random intercepts for subject and random slopes for session. Initial residual diagnostics (using DHARMa) revealed heterogeneity in the variance, likely driven by long latencies in some trials. However, removing these as "outliers" would remove trials that likely contributed to the learning for each individual subject. To retain these latencies as biologically meaningful while ensuring model accuracy, dispersion in the data was explicitly modeled, both as a function of session as well as trial outcome. This provided a better model fit compared to a model assuming constant dispersion ($\chi^2(2) = 62.29$, $p < 0.001$; $\Delta AIC = 58.3$).

Delayed Phases. In the delayed phases, to address whether individuals were modifying their behavior based on the type of delay (self vs partner) to achieve success, we employed a separate GLMM and modeled the *latency to first pull* in successful trials using a Gamma distribution with a log-link function. The model evaluated the interaction effect between type of delay and subject as well as main effects of session and phase. To prevent overfitting on this restricted dataset, we chose the most parsimonious model. To account for non-independence of data, random intercepts were included for each distinct session per subject. Dispersion parameters were omitted as model comparisons confirmed they provided no significant improvement to model fit.

Gazing Behavior. We analyzed the occurrence of social monitoring (gazing at the partner or the partner's apparatus) using a generalized linear model (GLM). Given the rarity of gaze occurrences (present in only 5.1% of the trials across the two successful pairs, AW and FZ), it was modeled as a binary outcome per subject per trial (0 = no gaze, 1 = one or more gazes, maximum observed per trial per subject was 3) using a Bernoulli distribution with logit-link function. This was preferred over a model using gaze count data since multiple gazes by the same subject occurred only in 1.1 % of the total trials. Fixed effects of the model included phase (both delayed and non-delayed), subject identity, trial outcome (success or failure) and session (scaled within phase). To account for varying trial durations (longer trials are more likely to have more gazing events), natural log of trial duration was included as an offset in the model. Random effects did not improve model fit and were excluded to obtain a more parsimonious model.

Results

Two of the tested pairs, AW and FZ, passed the non-delayed phases in 3 to 7 sessions (Fig. 2A and C). In the delayed phases, both succeeded with 25 cm rope and 5s delay within 7 sessions; however, neither passed the phase with 25 cm rope and 10s delay (Fig. 2B and D). The other three pairs, EK (Fig. 2E), PS (Fig. 2F), and PC (Fig. 2G), did not pass the first non-delayed phase with 50 cm rope within our 10-session criterion and were not tested further.

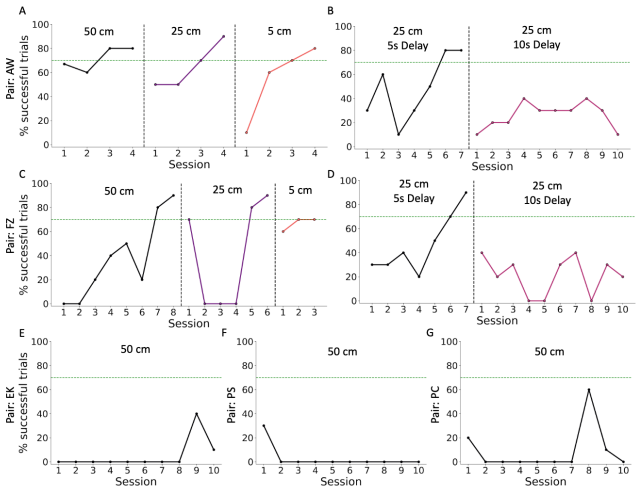

Figure 2. Marmoset pairs' performance across the different phases of the experiment. A and **B** show, for pair AW, progression through the non-delayed and delayed phases, respectively. **C** and **D** show the same progression for pair FZ. Pairs EK (**E**), PS (**F**) and PC (**G**) did not progress past the first non-delayed phase. Dashed green line represents 70% success.

Non-delayed Phases. Results of the GLMM showed the monkeys' learning across sessions, characterized as change in *latency to first pull*, was governed by both the outcome of the trial (success or failure) as well as the phase (length of rope). For the 50 cm phase, latencies remained relatively consistent, with only a marginal trend towards successful trials becoming faster than failed trials (b = -0.18, SE = 0.08, z = 2.27, p = 0.082). For the 25 cm phase, the model revealed a significant three-way interaction indicating that latencies reduced significantly across sessions for successful trials compared to failed trials (b = -0.85, SE = 0.21, z = -4.02, p < 0.001), potentially indicating learning. However, for the 5 cm phase, the three-way interaction revealed the opposite: successful trials were slower compared to failed trials (b = 1.23, SE = 0.29, z = 4.17, p < 0.001), suggesting that failures here resulted from monkeys being increasingly likely to pull too soon as the phase progressed.

Pair AW had faster baseline speed compared to pair FZ (b = 0.50, SE = 0.20, z = 2.47, p = 0.013). The random slopes showed stark differences between individual learning trajectories. While Apple was relatively consistent, Walnut exhibited a steep negative slope (i.e., he became faster). Flint and Zebib also remained relatively consistent with Flint having the shortest baseline latency (2.02s) and Zebib having the longest (3.16s) of the four monkeys. The raw data and individual trajectories from the model are visualized in Figure 3. The dispersion parameters revealed a significant effect of session on dispersion (b = 0.18, SE = 0.08, z = 2.27, p = 0.023), and showed successful trials exhibiting significantly lower dispersion compared to failed trials (b = 0.94, SE = 0.13, z = 7.22, p < 0.001).

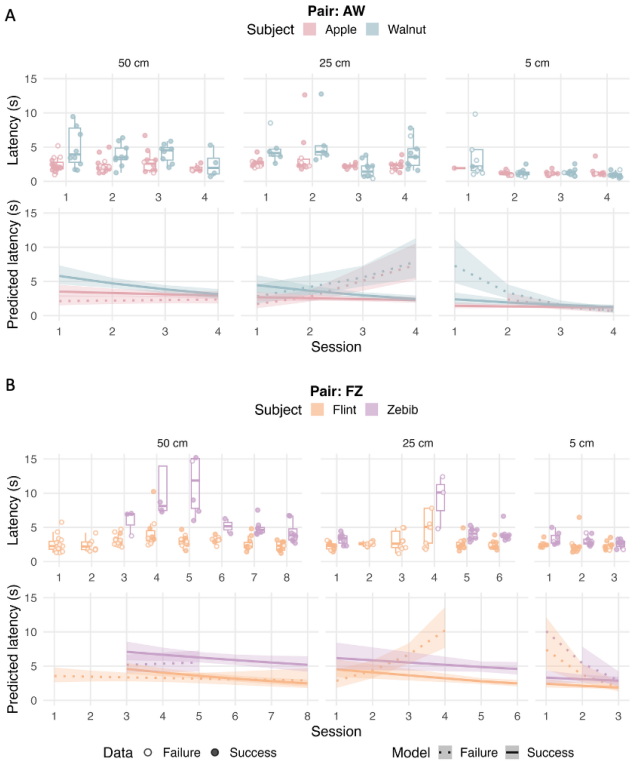

Figure 3. Change in latency to pull rope across consecutive sessions of the non-delayed phases. A and **B** show, for pairs AW and FZ respectively, raw data on top and model predictions below. Boxplots show medians and interquartile ranges (IQR) with whiskers extending to 1.5× IQR (for all trials); failed trials are visualized as hollow dots and successful trials as solid dots. Lines represent model-predicted latencies for success (solid) and failure (dotted) outcomes; shaded regions show 95% confidence intervals.

Delayed Phases. A separate GLMM was used to evaluate whether individual subjects modified their latencies based on who was being delayed (self vs. partner). After controlling for session-level noise (Variance = 0.167, SD = 0.408), analysis of the successful trials in the two delayed phases showed distinct mechanisms of success for the two pairs. For pair AW, Apple remained consistent in her latency regardless of who was delayed (b = 0.005, SE = 0.106, z = 0.049, p = 0.961) while Walnut exhibited a shift, significantly increasing his latency when Apple was delayed (b = 0.388, SE = 0.171, z = 2.262, p = 0.023). In pair FZ, neither Flint (b

= 0.171, *SE* = 0.149, *z* = 1.149, *p* = 0.250) nor Zebib (*b* = 0.235, *SE* = 0.164, *z* = 1.437, *p* = 0.150) significantly modified their latencies based on who was delayed. However, Zebib showed a significantly longer baseline latency (*b* = 0.672, *SE* = 0.197, *z* = 3.408, *p* < 0.001) compared to the others. Raw latencies for each subject on successful trials across the two delayed phases are shown in Figure 4.

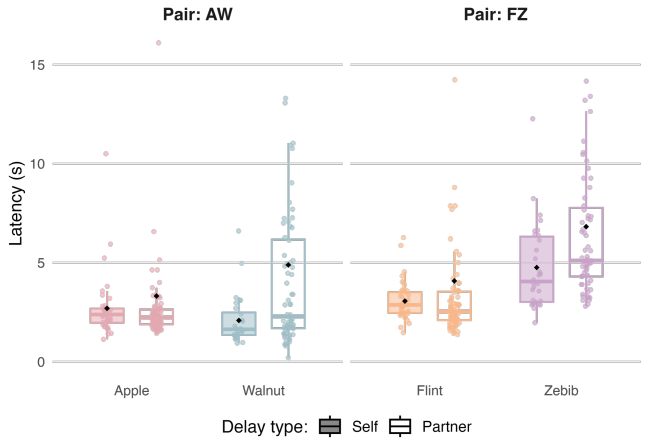

Figure 4. Latencies on successful trials in delayed phases as a function of who was delayed (self vs partner). Boxplots represent the median and interquartile range (IQR) with whiskers extending to 1.5× IQR, while black diamonds indicate the arithmetic mean. To prioritize visibility of data, the y-axis is truncated at 16s, leading to 12 trials being excluded from the plot (Apple Partner = 2; Walnut Partner = 4; Flint Partner = 3; Zebib Partner = 3); nonetheless, all successful trials were included in statistical calculations and summary markers (means/medians).

Gazing Behavior. The GLM revealed the effect of phase on the occurrence of gazing behavior. While there was no significant difference between the first two non-delayed phases (50 cm and 25c m rope; *b* = -0.24, *SE* = 0.40, *z* = -0.60, *p* = 0.548) subjects were significantly less likely to gaze at their partner in the third non-delayed phase (5 cm; *b* = -1.21, *SE* = 0.41, *z* = -2.94, *p* = 0.003) and the first delayed phase (5s delay; *b* = -1.13, *SE* = 0.39, *z* = -2.92, *p* = 0.003) with a marginal but non-significant decline continuing into the second delayed phase (10s delay; *b* = -1.10, *SE* = 0.65, *z* = -1.70, *p* = 0.089). There was a marginal positive trend for session (*b* = 0.31, *SE* = 0.16, *z* = 1.91, *p* = 0.056) and no significant effect of trial outcome (*b* = 0.21, *SE* = 0.28, *z* = 0.77, *p* = 0.442). Although qualitatively we see a larger number of gazes in later phases from Walnut and Zebib compared to their partners Apple and Flint (Fig. 5), no statistically significant differences in gazing behavior were observed across subjects (*p* > 0.200 for all subjects).

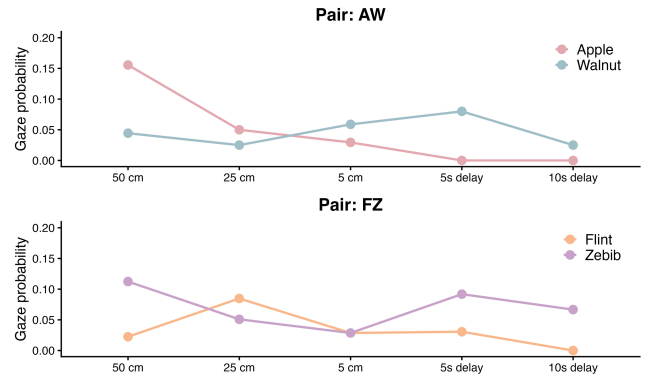

Figure 5. Gazing behavior. Probability of gazing on a trial for each subject across the non-delayed and delayed phases.

Discussion

We tested five pairs of marmosets on the rope-pulling task which requires careful temporal coordination between individuals to achieve success. The pairs were tested across phases with increasing demand for coordination, with two of five pairs progressing through all non-delayed phases and the 5s delay phase but not the 10s delay phase.

In the non-delayed phases, model results revealed that each pair seemed to have different dynamics contributing to their success which was demonstrated by variations in behavioral modification across individuals. For pair AW, Walnut significantly reduced his latency across sessions, eventually converging to Apple's consistent and short pulling latency. This is further evidenced by failed trials in the 5 cm rope phase being caused by pulls from Walnut that were too fast rather than slower pulls, demonstrating an overcorrection in response to the shorter rope. For pair FZ, while neither individual showed directional learning slopes across sessions, the gap between them narrowed across phases. Their success was likely supported by a significant reduction in latency variance on successful trials, as evidenced by the dispersion parameters in the model. This suggests that FZ may have achieved success by becoming more consistent in their latencies and that their inherent differences created a success buffer that narrowed for each successive non-delayed phase with shorter rope.

Similar asymmetries persisted in the delayed phases for pair AW, with Walnut modifying his behavior to potentially account for Apple's delay, while Apple remained consistent regardless of who was delayed. Pair FZ's success in the delayed phases seems to have been driven by Zebib's longer baseline latency, which likely acted as a buffer in the trials where she was released first. She did not need to significantly modify her behavior to accommodate Flint's delay.

While both successful pairs passed the 5s delay, they failed to meet our criterion for the 10s delay. This failure may reflect a fundamental limit in marmoset inhibitory control as 10s may exceed their capacity to suppress pulling on the rope (although marmosets have been shown to be able to delay gratification for up to 14s; Stevens et al., 2005). However, task mechanics likely also played a role. With a 5s delay and a 25 cm rope, there was enough slack for the delayed marmoset to leave the holding box and pull before the rope was fully retracted. With a 10s delay, this was less likely. To

probe this explanation, we had intended to run an additional delayed phase with a 5s delay and a shorter (5 cm) rope; however, due to pregnancies in both pairs we only completed three sessions of this condition before we had to cease further testing. While limited, preliminary results from these sessions showed consistent failures, suggesting that the observed success may be an artifact of task affordances rather than deliberate "waiting" behavior. Relatedly, while *latency to first pull* proved to be a valuable metric to measure task learning and characterize behavioral strategies, future work will also examine other behaviors (including those exhibited during the waiting period) to identify strategies that either aid in or deter successful coordination.

Contrary to expectations, social monitoring via gazing was rare and did not correlate with success, even in delayed phases where partner-tracking would be most beneficial. Gazes were more likely in early phases for both pairs, which may reflect environmental uncertainty rather than task-directed information gathering. To further probe whether marmosets rely on visual access to their partner to solve the task, the introduction of an opaque partition between the two boxes such that the two subjects cannot see each other (as in Meisner et al., 2025) may be informative. We attempted such a test (25 cm rope, no delay) with pair AW after the completion of all non-delayed phases, but abandoned testing after one session as we noted increased vocalizations (*phee* calls) from Apple indicating some signs of distress at not being able to see her partner during the task. In this single session, they were successful in 9/10 trials. Future research should introduce partitions earlier in the learning process to habituate subjects to its presence, and in turn determine if visual cues are necessary to establish strategies, even if they may not be required to maintain them.

Alongside these individual-level metrics, the composition of the pair may also affect coordination success. For example, temperament has been shown to influence task performance in marmosets (e.g., Šlipogor et al., 2022, 2026), and hence may also affect how readily a pair with a given combination of temperaments is able to establish coordination (e.g., temperament asymmetries may aid in success through role differentiation). In a similar vein, it remains to be seen whether specific combinations of individual strategies are required for effective coordination (e.g., a flexible Walnut requiring a consistent Apple), and whether such pairings only function within limited contexts, such as the non-delayed phases of this specific task. While the social constraints of marmosets make arbitrary pairings difficult, future research re-pairing individuals with established strategies would be insightful. This would reveal which strategy combinations yield success and how individuals adapt their strategies based on a partner's identity. Lastly, it is worth noting that while both successful pairs were established breeding pairs housed together, the failed pairs originated from different social contexts: one was recently paired (pair EK), and the others were siblings separated from a larger group. These housing and social dynamics likely influenced task performance. Further analysis of the limited successful trials within these failed pairs may clarify the specific barriers to their coordination.

Overall, our study suggests that coordination among marmosets can arise from individually learned behaviors in the absence of shared intentionality. This is underscored by the fact that the two pairs that were successful across 4/5 phases seemed to use different strategies to succeed. We found little evidence that marmosets understand their partner's role or consistently modify their behavior intentionally to achieve success; instead, coordination emerges through a variety of individual strategies. Rather than direct social monitoring, coordination may be driven by the marmosets simply reacting to movement of their own end of the rope as a reliable proxy for their partner's actions.

While these results, showcasing successful coordination despite strategies that ignore a partner's actions, may seem suboptimal, they ultimately lead to success in certain contexts and likely require fewer cognitive resources. This raises questions about what constitutes an "optimal" strategy for achieving coordination in social species and how this may vary depending on the structure of the problem as well as the cognitive resources available to individuals.

The rope-pulling paradigm is a standard task for probing whether individuals can account for a partner's behavior to achieve success (e.g., Hirata & Fuwa, 2007; Plotnik et al., 2011). While success in this task can in principle be used to infer higher-order cognitive abilities like shared intentionality, our results – alongside those from other non-human subjects – suggest caution. Success can emerge from a variety of distinct strategies, including those that rely on little to no understanding of the partner's role in the outcome.

Furthermore, task structure is likely to influence how coordination is achieved and interpreted. For instance, in a coordinated lever-pulling task, which lacks the direct causal transparency of seeing a shared rope move, marmosets increased social monitoring to facilitate success (Meisner et al., 2025). This suggests that the "opacity" of the task may have prompted higher rates of gazing behavior. Comparing how the same species coordinates across different task contexts is therefore essential for identifying the specific cognitive faculties that underpin successful coordination.

While marmosets are often viewed as a model species for understanding the mechanisms of coordination (e.g., Burkart et al., 2014), it is important to note that their classification as a cooperative breeder may not inherently necessitate the evolution of complex cognitive mechanisms for coordination outside of that context. Our study yields novel insights into the cognitive underpinnings of coordination, suggesting that success in this species can be driven by relatively simple, individual-level dynamics.

Acknowledgements

We thank Emily Wiler, Alexia Nelms, Enana Jacob, and Yiran (Cecilia) Zhu for their invaluable assistance with marmoset handling, data collection, and behavioral video coding. We are also grateful to Dr. Jude Mitchell for his support in designing and building the experimental equipment, as well as for his continued guidance throughout this research. Finally, we extend our gratitude to Dina Graf, the vivarium staff, and the veterinary technicians for their essential contributions to making non-human primate research possible.

Note: AI generated code, verified by researchers, was used to produce and refine the plots used in the figures.

References

Burkart, J. M., Allon, O., Amici, F., Fichtel, C., Finkenwirth, C., Heschl, A., Huber, J., Isler, K., Kosonen, Z. K., Martins, E., Meulman, E. J., Richiger, R., Rueth, K., Spillmann, B., Wiesendanger, S., & van Schaik, C. P. (2014). The evolutionary origin of human hyper-cooperation. *Nature Communications*, *5*(1), 4747. https://doi.org/10.1038/ncomms5747

Burkart, J. M., & Finkenwirth, C. (2015). Marmosets as model species in neuroscience and evolutionary anthropology. *Neuroscience Research, Marmoset Neuroscience*, *93*, 8–19. https://doi.org/10.1016/j.neures.2014.09.003

Burkart, J. M., & Heschl, A. (2006). Geometrical gaze following in common marmosets (*Callithrix jacchus*). *Journal of Comparative Psychology*, *120*, 120–130. https://doi.org/10.1037/0735-7036.120.2.120

Burkart, J. M., & van Schaik, C. P. (2020). Marmoset prosociality is intentional. *Animal Cognition*, *23*(3), 581–594. https://doi.org/10.1007/s10071-020-01363-6

Camerer, C. F. (2011). *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press.

Couzin, I. D., & Franks, N. R. (2003). Self-organized lane formation and optimized traffic flow in army ants. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1511), 139–146. https://doi.org/10.1098/rspb.2002.2210

Crawford, M. (1937). The cooperative solving of problems by young chimpanzees. *Comp Psychol Monogr*, *14*, 1–88.

Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Hawkins, R. X. D., Goodman, N. D., & Goldstone, R. L. (2019). The Emergence of Social Norms and Conventions. *Trends in Cognitive Sciences*, *23*(2), 158–169. https://doi.org/10.1016/j.tics.2018.11.003

Hirata, S., & Fuwa, K. (2007). Chimpanzees (*Pan troglodytes*) learn to act with other individuals in a cooperative task. *Primates*, *48*(1), 13–21. https://doi.org/10.1007/s10329-006-0022-1

Meisner, O. C., Shi, W., Nair, A., Nandy, G., Jadi, M. P., Nandy, A. S., & Chang, S. W. C. (2025). Diverse and flexible strategies enable successful cooperation in marmoset dyads. *Current Biology*, *35*(18), 4509-4521.e5. https://doi.org/10.1016/j.cub.2025.08.005

Melis, A. P., Hare, B., & Tomasello, M. (2006). Chimpanzees Recruit the Best Collaborators. *Science*, *311*(5765), 1297–1300. https://doi.org/10.1126/science.1123007

Miller, C. T. (2025). Animal cognition: Adaptive cooperation strategies in marmoset monkeys. *Current Biology*, *35*(18), R886–R887. https://doi.org/10.1016/j.cub.2025.07.078

Plotnik, J. M., Lair, R., Suphachoksahakun, W., & de Waal, F. B. M. (2011). Elephants know when they need a helping trunk in a cooperative task. *Proceedings of the National Academy of Sciences*, *108*(12), 5116–5121. https://doi.org/10.1073/pnas.1101765108

R Core Team. (2023). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing. https://www.r-project.org/*.

Sehner, S., Willems, E. P., Vinicus, L., Migliano, A. B., van Schaik, C. P., & Burkart, J. M. (2022). Problem-solving in groups of common marmosets (*Callithrix jacchus*): More than the sum of its parts. *PNAS Nexus*, *1*(4), pgac168. https://doi.org/10.1093/pnasnexus/pgac168

Šlipogor, V., Graf, C., Massen, J. J. M., & Bugnyar, T. (2022). Personality and social environment predict cognitive performance in common marmosets (*Callithrix jacchus*). *Scientific Reports*, *12*(1), 6702. https://doi.org/10.1038/s41598-022-10296-8

Šlipogor, V., Masilkova, M., Höflinger, A., Lang, N., Bugnyar, T., & Konečná, M. (2026). With or without you: Common marmoset, *Callithrix jacchus,* personality expression is mediated by social setting. *Animal Behaviour*, *231*, 123367. https://doi.org/10.1016/j.anbehav.2025.123367

Stevens, J. R., Hallinan, E. V., & Hauser, M. D. (2005). The ecology and evolution of patience in two New World monkeys. *Biology Letters*, *1*(2), 223–226. https://doi.org/10.1098/rsbl.2004.0285

Tomasello, M., & Moll, H. (2010). The Gap is Social: Human Shared Intentionality and Culture. In P. M. Kappeler & J. Silk (Eds.), *Mind the Gap: Tracing the Origins of Human Universals* (pp. 331–349). Springer. https://doi.org/10.1007/978-3-642-02725-3_16

Visalberghi, E., Quarantotti, B. P., & Tranchida, F. (2000). Solving a cooperation task without taking into account the partner's behavior: The case of capuchin monkeys (*Cebus apella*). *Journal of Comparative Psychology*, *114*, 297–301. https://doi.org/10.1037/0735-7036.114.3.297