BEHAVIOR

Neural basis of cooperative behavior in biological and artificial intelligence systems

Mengping Jiang†, Linfan Gu†, Mingyi Ma, Qin Li, Jonathan C. Kao*, Weizhe Hong*

INTRODUCTION: Cooperation—the process by which individuals coordinate their actions to achieve shared benefits—is fundamental to human and animal societies and increasingly critical in artificial intelligence (AI). Cooperation often requires sophisticated integration of self-monitoring, partner observation, context-dependent decision-making, and precise temporal coordination. However, the neural mechanisms and computational principles that enable such coordination remain poorly understood in both domains.

RATIONALE: We developed parallel experimental paradigms that allow direct comparison between biological and AI systems. For biological studies, we created a behavioral task in which pairs of mice must coordinate their nose poke actions within precise time windows to receive mutual rewards, combined with microendoscopic calcium imaging of the anterior cingulate cortex (ACC)—a region implicated in social cognition and decision-making. Simultaneously, we trained pairs of artificial agents using multi-agent reinforcement learning to perform an analogous coordination task, which provided complete access to their "neural" computations and enabled precise experimental manipulations that are challenging in biological systems. This comparative approach allowed us to examine whether similar computational principles govern cooperation across biological and artificial systems.

RESULTS: We found that mice can indeed learn to coordinate their actions for mutual benefits. Detailed behavioral analysis revealed that successful coordination depends on key social behavioral strategies that include approach, waiting, and interaction behaviors. These three preparatory behaviors increased substantially over the course of training and were more prominent during successful trials. Our neural recordings revealed that ACC neurons encoded different aspects of the cooperative process, including correct versus missed coordinative decisions, the three preparatory behaviors, and key decision-making processes. Furthermore, partner location was strongly represented in ACC activity, especially during moments requiring coordination decisions. Animals with stronger neural representations of behaviors and partner information showed better cooperative performance. Lastly, chemogenetic and optogenetic silencing experiments demonstrate that ACC activity causally contributes to cooperative behavior—inhibiting ACC neurons reduced cooperative success, impaired social behavioral strategies, and decreased the precision of coordinated actions.

Artificial agents trained on an analogous mutual cooperation task developed strikingly similar behavioral strategies and neural representations. Like mice, successful artificial agents exhibited waiting behavior—pausing when their partner was far away and coordinating their movements to minimize distance differences. Analysis of the agents' recurrent neural networks revealed enhanced representations of partner-related information, paralleling our observations in mouse ACC. Furthermore, artificial neurons encoding key behavioral decisions—"hold" and "proceed"—emerged during training, and selectively disrupting these neurons impaired cooperative performance.
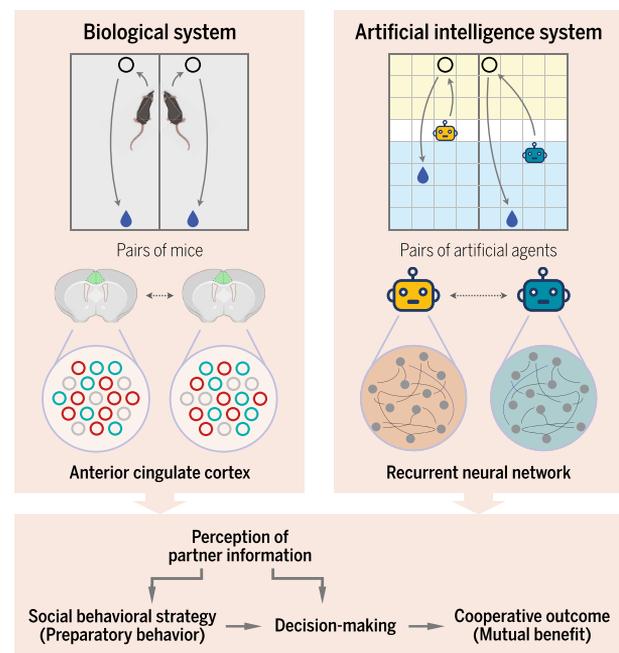
Cooperation in biological and AI systems. Both biological systems (left) and AI systems (right) engage in cooperative behavior. In biological systems, pairs of mice learned to coordinate actions within precise time windows to receive mutual rewards. In AI systems, pairs of agents navigated an artificial environment and learned to coordinate actions within specified time windows to receive mutual rewards. Through analyses of neural activities in the biological brain (the anterior cingulate cortex) and artificial brain (recurrent neural network), we reveal that both systems utilize similar computational principles underlying successful cooperation: perception of partner information, development of social behavioral strategies, and context-dependent decision-making processes that lead to cooperative outcomes with mutual benefits.

CONCLUSION: Our findings reveal that successful cooperation in mice emerges from the combination of partner information, social behavioral strategies, and context-dependent decision-making processes. The ACC serves as a critical processing hub, representing partner information, behavioral strategies, and coordination decisions. The remarkable convergence between mouse ACC activity and artificial neural network dynamics suggests that these principles represent fundamental organizational requirements for any intelligent system engaged in real-time cooperation. These findings advance our understanding of the neural basis of social behavior and demonstrate the power of comparative approaches that bridge neuroscience and AI. □

BEHAVIOR

Neural basis of cooperative behavior in biological and artificial intelligence systems

Mengping Jiang[1,2]†, Linfan Gu[1,2,3]†, Mingyi Ma[1,2], Qin Li[1,2,3], Jonathan C. Kao[4,5]*, Weizhe Hong[1,2,3]*

Cooperation, the process through which individuals work together to achieve common goals, is fundamental to human and animal societies and increasingly critical in artificial intelligence (AI). In this study, we investigated cooperation in mice and AI systems, examining how they learn to actively coordinate their actions to obtain shared rewards. We identified key social behavioral strategies and decision-making processes in mice that facilitate successful cooperation. These processes are represented in the anterior cingulate cortex (ACC), and ACC activity causally contributes to cooperative behavior. We extended our findings to AI systems by training artificial agents in a similar cooperation task. The agents developed behavioral strategies and neural representations reminiscent of those observed in the biological brain, revealing parallels between cooperative behavior in biological and artificial systems.

Cooperation is a process in which two or more individuals work together to achieve shared goals. Cooperation plays a vital role in enhancing individual fitness and promoting group survival (1, 2), whereas its breakdown often leads to detrimental social conflict and instability. Understanding the mechanisms underlying cooperation therefore carries profound societal implications. Cooperative behaviors are not exclusive to humans but are observed across a variety of animal species (3), such as chimpanzees (4, 5), elephants (6), dolphins (7), bats (8), and birds (9, 10).

An important form of cooperation is mutualism-based cooperation, which requires individuals to coordinate their actions in real time to achieve immediate (and often simultaneous) shared benefits (3). For instance, cooperative hunting demands sophisticated coordination of actions between participants to increase their likelihood of capturing prey (11–13). Previous studies have shown that humans and other species such as marmosets, shrews, and rodents are capable of engaging in mutual cooperation (14–18). Mutual cooperation involves an extraordinarily complex interplay of behavioral processes, requiring individuals to actively communicate and coordinate with each other to build shared goals and make decisions based on continuous, real-time monitoring and assessment of others' actions. Despite its importance, these intricate behavioral processes and their underlying neural mechanisms remain poorly understood.

Studies of humans and primates have implicated an important role for several brain regions in cooperative behavior, including the anterior cingulate cortex (ACC) (19–21). The ACC plays a key role in a broad range of behavioral functions, such as emotional processing, reward and punishment, negative affect, pain perception, and social behavior (22). However, how ACC neural activities play a role in the sophisticated coordination of actions during mutual cooperation remains poorly understood.

Recent advances in artificial intelligence (AI), particularly in reinforcement learning, enable artificial agents to behave proficiently in simulated environments (23). Research suggests that both artificial and biological agents can exhibit similar behavioral strategies and neural representations (24–28). This opens exciting avenues for exploring how cooperative behavior may emerge when artificial agents interact (29) and whether such interactions could be driven by neural network dynamics that resemble those observed in biological systems.

Results

Mice coordinate their actions to obtain mutualistic rewards

To determine whether mice exhibit cooperative behavior, we devised an operant task in which two mice were required to coordinate their actions to obtain mutualistic rewards. In the initial phase, we trained individual mice to nose poke and associate this action with receiving a water reward. We next paired two mice and required them to nose poke within a specified time window of each other to obtain mutualistic rewards. If only one mouse nose poked within the time window, neither received a reward. Training began with 3-s window that was subsequently reduced to 1.5 s and finally to 0.75 s (Fig. 1, A to C; movies S1 to S3; and Materials and methods).

To evaluate cooperative performance, we calculated the ratio of correct trials (in which both mice nose poked within the time window) to the total number of correct trials plus miss trials (where only one mouse nose poked within the time window). We then established the chance-level correct trial ratio by randomly pairing the two animals' nose poke event sequences from different trials (fig. S1A and Materials and methods). After training, 76% of mouse pairs exhibited an increase in the ratio of correct trials above chance (Fig. 1D and Materials and methods). Among these animals, a substantial percentage (41%) were "high-performance" pairs whose ratio of correct trials was higher than 4 standard deviations (SDs) above chance (Fig. 1, E to G; fig. S1, B to F; Materials and methods; and supplementary note S1). When analyzing the data to consider correct trials as paired nose pokes occurring within a 0.75-s window irrespective of the cooperative stage, we found that high-performance mice showed a gradual increase in correct trials and the overall cooperative index (defined as the difference between the observed correct trial ratio and the chance value) over the course of training (Fig. 1, C, E, and F, and Materials and methods; for performance defined by the original time window, see fig. S1, G to J). Further, miss trials declined over the course of training (Fig. 1G), indicating that mice not only learned to coordinate their actions but also became more proficient, making fewer errors over time. This increase in performance in high-performance pairs compared with other groups was not simply due to their overall motivations, because the total number of nose pokes was comparable between groups (fig. S1E). We examined cooperative performance in both males and females and found that they were comparable (fig. S2).

We also examined the distribution of the poke interval and found that high-performance pairs exhibited shorter poke intervals on the final day compared with the first day of the third cooperative stage (Fig. 1, H and I), suggesting that these mice achieved greater interanimal coordination over the course of training. When we examined correct trials that occurred consecutively, we found that high-performance pairs showed a higher percentage of consecutive correct trials than intermediate-performance or nonsignificant pairs (fig. S1F).

It is possible that both animals independently learn to nose poke within a fixed time window after drinking, without reciprocally coordinating their actions. If animals simply adopted a timing-based strategy, their performance would not be different from that after shuffling across trials (Materials and methods). We found that the number of

[1]Department of Neurobiology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [2]Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [3]Department of Bioengineering, Henry Samueli School of Engineering, University of California, Los Angeles, Los Angeles, CA, USA. [4]Department of Electrical and Computer Engineering, Henry Samueli School of Engineering, University of California, Los Angeles, Los Angeles, CA, USA. [5]Department of Computer Science, Henry Samueli School of Engineering, University of California, Los Angeles, Los Angeles, CA, USA. *Corresponding author. Email: whong@ucla.edu (W.H.); kao@seas.ucla.edu (J.C.K) †These authors contributed equally to this work.
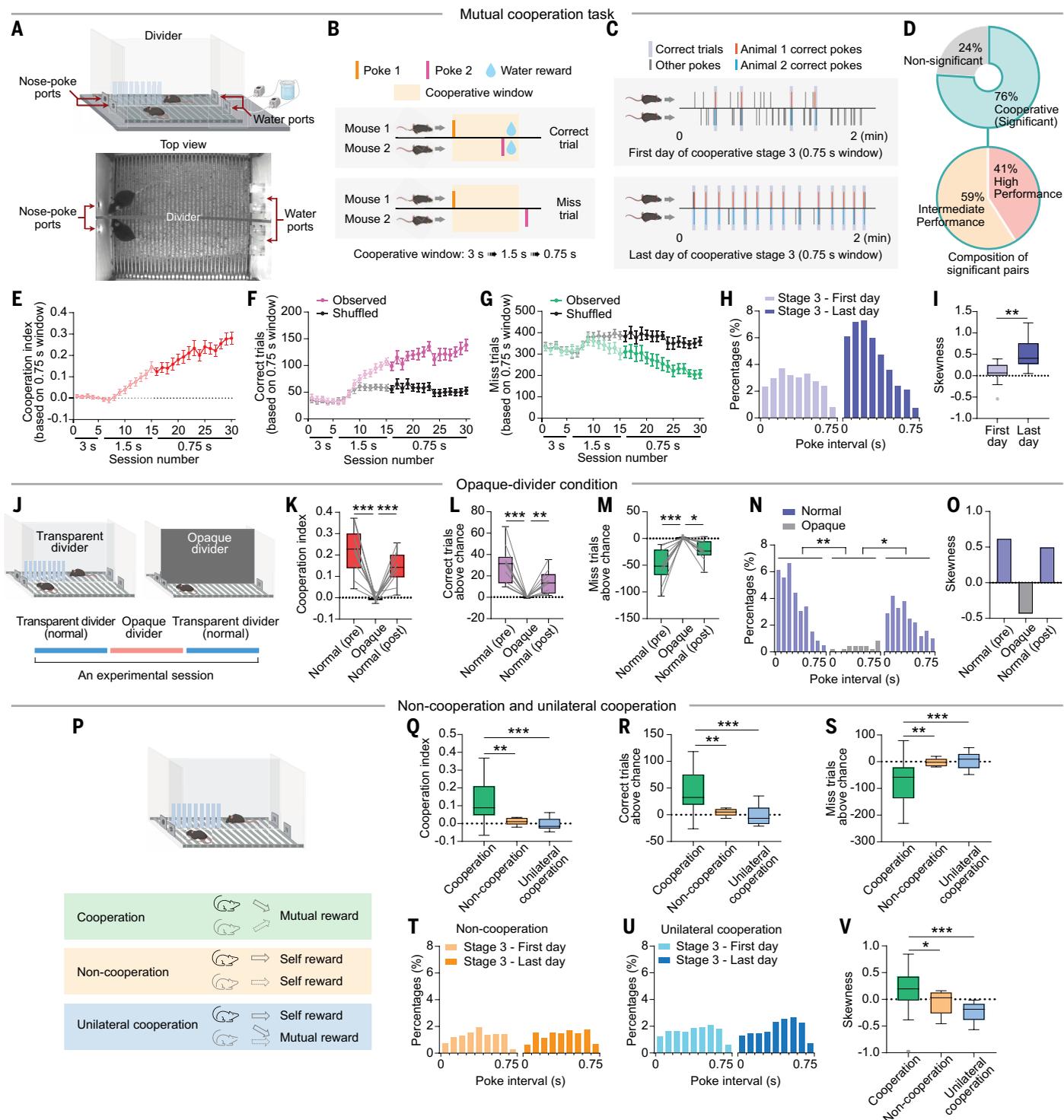
Fig. 1. Mice engage in cooperative behavior to obtain mutualistic rewards. (**A**) Schematic illustrating behavioral apparatus used for mice to perform the mutual cooperation task to actively coordinate their actions. The training chamber is divided into two compartments by a transparent, perforated partition, with each compartment equipped with a nose poke port and a water port on two opposite ends. (**B**) Schematic illustrating correct and miss trials in the cooperation task. (**C**) Representative raster plots showing nose pokes and correct trials for mice on the first and last day of the third cooperative stage (0.75-s window). (**D**) Percentages of nonsignificant, intermediate-performance, and high-performance mice. (**E** to **G**) Quantification of cooperation in high-performance mice: the cooperation index (E), correct trials in observed and shuffled data (F), and miss trials in observed and shuffled data (G) across training sessions. The cooperation index is defined as the difference between the observed ratio of correct trials and the chance value. (**H** and **I**) Poke interval distribution (H) and its skewness (I) of correct trials on the first and last day of the final cooperative stage in high-performance mouse pairs. Poke interval is defined as time between the two animals' nose pokes in a correct trial. (**J**) Schematic of the cooperation task under normal (transparent-divider) or opaque-divider conditions. (**K** to **M**) Quantification of cooperation in mice under conditions with normal (transparent-divider; before and after the opaque divider condition) or opaque-divider condition: cooperation index (K), correct trials above chance (L), and miss trials above chance (M). (**N** and **O**) Poke interval distribution (N) and skewness (O) of correct trials in normal (transparent-divider; before and after the opaque divider condition) or opaque-divider conditions. (**P**) Schematic illustrating the cooperation, noncooperation, and

unilateral cooperation tasks. (**Q** to **S**) Behavioral performance across cooperation, noncooperation, and unilateral cooperation for cooperation index (Q), correct trials above chance (R), and miss trials above chance (S). In unilateral cooperation, the behavioral performance of the cooperative (second) mice was analyzed and shown. For a fair comparison, data from the normal cooperation condition included all trained mouse pairs (nonsignificant, intermediate-performance, and high-performance pairs). (**T** and **U**) Poke interval distribution of correct trials on the first and last day of the final cooperative stage in noncooperative (T) and unilateral (U) conditions. (T), $P = 0.126$; (U), $P = 0.015$ (significant for the opposite direction). Kolmogorov-Smirnov test. (**V**) Skewness of poke interval distribution in cooperation, noncooperation, and unilateral cooperation. In (E) to (G), sessions 1 to 15 (shaded) are presented based on 0.75-s time window. In (E) to (I), $n = 12$ pairs. In (K) to (O), $n = 10$ pairs. In (Q) to (V), $n = 38$ pairs (cooperation), 10 pairs (noncooperation), and 10 pairs (unilateral cooperation). In (E) to (G), mean ± SEM; in (I), (K) to (M), (Q) to (S), and (V), box plots: center indicates median, and box indicates quartiles. In (I), Wilcoxon matched-pairs signed rank test. In (K) to (M), repeated measures one-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test. In (N), Kolmogorov-Smirnov test. In (Q) to (S), Kruskal-Wallis test with Dunn's multiple comparisons test. In (V), one-way ANOVA (Welch test) with Dunnett's T3 multiple comparisons test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

correct trials exceeded the shuffled data, whereas the number of miss trials was lower (Fig. 1, F and G, and fig. S1, C, D, G, and H). This suggests that the improved performance results from active coordination between the animals, as opposed to independent poking based on timing.

Active coordination and mutualistic benefits are required for cooperation

Cooperation is a complex social interaction that depends not only on an individual's own actions but also on their partner's actions. Although our shuffled data suggest that animals' mutualistic reward arises from active coordination between animals rather than timing-based independent poking, it is still crucial to experimentally determine whether partner information plays a role. We conducted a control experiment in which we replaced the transparent, perforated divider with a solid, opaque divider, blocking visual cues (Fig. 1J). We found that correct trials and the cooperation index decreased, whereas miss trials increased (Fig. 1, K to M). The percentage of correct trials with small nose poke intervals also substantially decreased (Fig. 1, N and O). This supports the notion that partner information is necessary for mice to coordinate their actions and that this process cannot be explained by alternative nonsocial strategies, such as timing-based independent poking.

To further determine whether coordinated nose poking indeed requires mutualistic benefits as opposed to simple behavioral mimicry or coincidental actions, we carried out an additional control experiment, in which two mice, placed side-by-side, each individually received reward after a nose poke, irrespective of whether they were cooperative (Fig. 1P). In this "noncooperation" condition, mice engaged in fewer correct trials and more miss trials, with the cooperation index being substantially lower than in the cooperation condition and not different from chance (Fig. 1, Q to S, and fig. S3, A to C).

Even if animals rely on partner information to obtain rewards, it remains possible that their nose poking behavior is driven solely by the partner's nose poke action rather than by active coordination between them. We conducted a third control experiment. In this "unilateral cooperation" condition, one mouse could obtain reward by poking alone, whereas the second mouse required coordinated nose pokes with the first animal to earn its reward (Fig. 1P). If cooperative outcomes were simply driven by copying behavior or social cues from the first animal rather than by true coordination, the second mouse could obtain reward by merely mimicking the first animal's actions or following its social cues. In this unilateral condition, the second mice engaged in fewer correct trials and more miss trials, with the cooperation index being substantially lower than in the normal cooperation condition and not different from chance (Fig. 1, Q to S, and fig. S3, E to G). This argues against the possibility that the second animal simply mimics and copies the first animal's behavior. In both noncooperation and unilateral conditions, the distributions of poke intervals did not change over the course of training (Fig. 1, T to V). The reduced performance in these control conditions was not due to a lack of task engagement, because the total number of nose pokes slightly increased in both conditions (fig. S3, D and H).

Taken together, these results suggest that this cooperative behavior does not result from independent timing-based decision-making, simple behavioral mimicry, coincidental actions, or social cue-dependent decision-making. Instead, both animals correctly follow the rule and actively coordinate their actions based on their partner's social information and mutualistic benefits in order to achieve successful cooperation.

ACC neurons encode distinct cooperative events

The ACC plays a critical role in social decision-making and represents the outcomes of self and others' behaviors ([30]–[33]). We carried out in vivo microendoscopic calcium imaging to record neural activity in the ACC when mice performed the cooperation task (Fig. 2A). Using a calcium indicator, GCaMP7f (Fig. 2, B to D), we recorded neural activities from 12,798 neurons in 17 pairs of mice. Through receiver operating characteristic (ROC) analysis, we identified individual ACC neurons that responded during specific behavioral events during the task (Fig. 2E and supplementary note S2). A fraction of neurons responded selectively to either correct or miss trials but not both, with more neurons responding to correct trials (Fig. 2, F to H).

We further imaged these neurons in the opaque-divider control condition. Although animals in this condition did not exhibit correct trials above chance, they still engaged in spontaneous nose pokes. However, neurons that responded to correct or miss trials in the cooperation condition showed minimal neural activity in the opaque-divider condition (Fig. 2, F and G, and supplementary note S3). Instead, a largely separate set of neurons responded to spontaneous poking events (Fig. 2, F to H).

By examining the activity trajectories in the principal component space, we found that ACC activities at the population level were also distinct between these three types of nose pokes (Fig. 2, J and K). Using support vector machine (SVM) decoders, we found that correct and miss trials in normal conditions can be decoded above chance and that both can be decoded from spontaneous nose pokes in the opaque-divider condition (Fig. 2, L to N; supplementary note S4; and Materials and methods).

Over the course of training, there was an increase in the fraction of correct poke-responsive neurons and decoding performance for correct versus miss trials (fig. S5, A and B), indicating that successful training leads to a stronger neural representation of coordinated actions. Indeed, we observed a positive correlation between cooperative performance and the fraction of correct poke-encoding cells, in which animals with better cooperative performance had larger populations of neurons representing successful coordination (Fig. 2I). Additionally, decoding performance for correct versus miss pokes was also correlated with cooperative performance in individual animals (Fig. 2O). This suggests that the internal representation of successful versus unsuccessful actions is a key neural signature in effective cooperators.

Behavioral strategies to achieve effective cooperation

Because the partner's social information plays a crucial role in achieving cooperation, we analyzed both animals' behavioral actions during
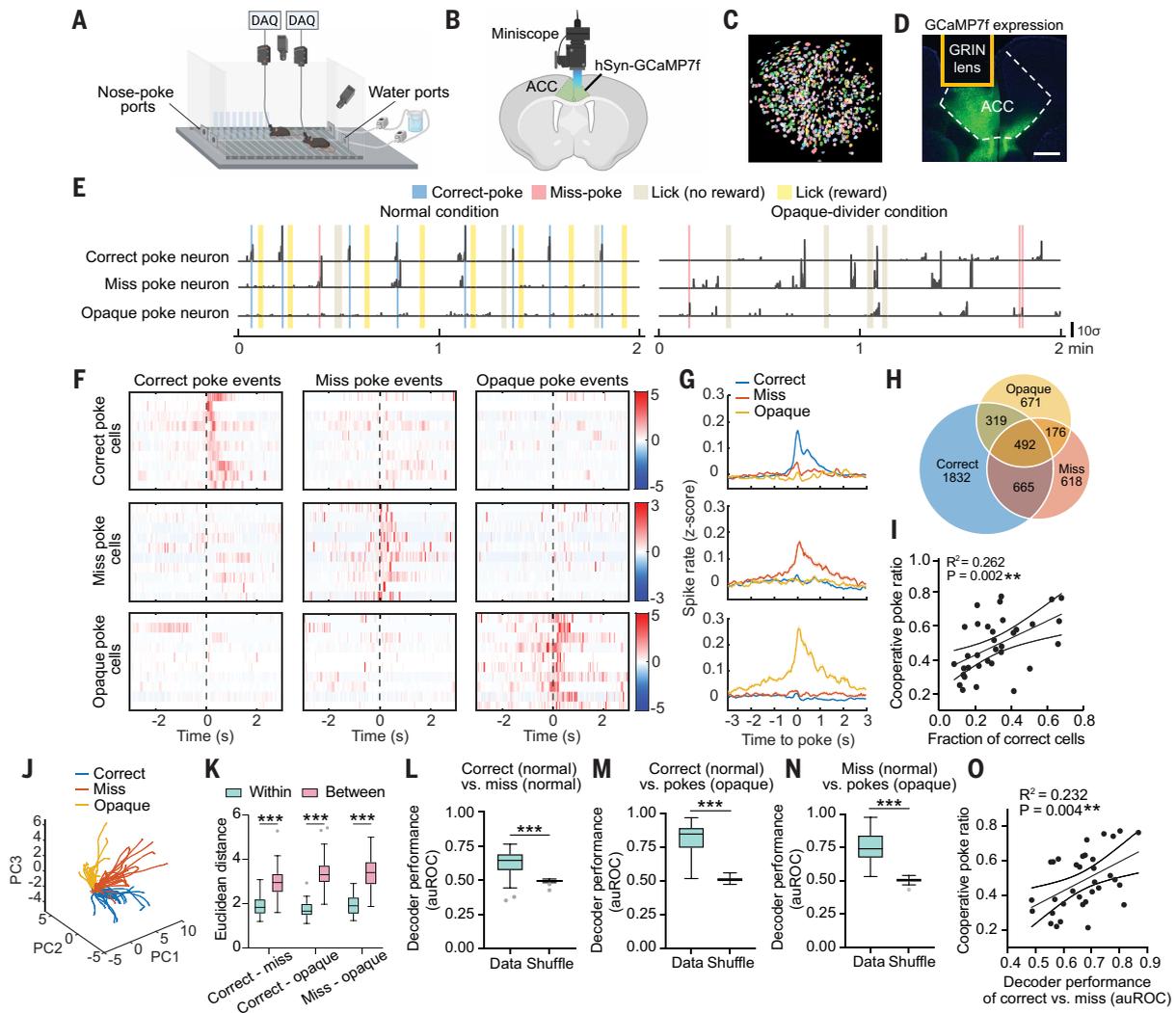
Fig. 2. Encoding of cooperative events in the ACC. (**A**) Schematic illustrating calcium imaging using miniature microendoscope. Top and front cameras were used to record animals' behavior. (**B**) Viral injection and lens implantation in the ACC. (**C** and **D**) Example ROI of neurons (C) and GCaMP7f expression (D). Scale bar, 500 μm. GRIN lens; gradient index lens. (**E**) Deconvolved calcium traces in normal (transparent) and opaque conditions, aligned with behavioral events. (**F**) Heatmaps showing average responses of example cells (each row) activated selectively during correct pokes in the normal condition, miss pokes in the normal condition, or spontaneous pokes in the opaque condition. (**G**) Event-triggered average neural activity of neurons that responded to different poke events. (**H**) Venn diagram showing neurons activated during correct pokes in the normal condition, miss pokes in the normal condition, and spontaneous pokes in the opaque condition. (**I**) Correlation between the fraction of cells encoding correct pokes and cooperative poke ratio (the fraction of correct pokes in all pokes) in the well-trained animals. Linear regression with 95% confidence intervals (CIs). (**J**) Principal component (PC) projections of population activity associated with correct pokes in the normal condition, miss pokes in the normal condition, and spontaneous pokes in the opaque condition. (**K**) Average Euclidean distance in PC space between or within correct and miss pokes in the normal condition, between or within correct pokes in the normal condition and spontaneous pokes in the opaque condition, and between or within miss pokes in normal condition and spontaneous pokes in the opaque condition. (**L** to **N**) Performance of decoders in classifying correct versus miss pokes in the normal condition (L), correct pokes in the normal condition versus spontaneous pokes in the opaque condition (M), and miss pokes in the normal condition versus spontaneous pokes in the opaque condition (N). (**O**) Correlation between decoding performance of correct versus miss pokes and animals' cooperative performance (the cooperative poke ratio). This correlation was not due to different numbers of trials across animals because down-sampling to match the number of trials across animals did not affect this correlation (Materials and methods). Linear regression with 95% CIs. $n = 34$ mice. In (G), mean $\pm$ SEM; in (K) to (N), box plots: center indicates median, and box indicates quartiles. In (K), repeated measures two-way ANOVA with Tukey's multiple comparisons test. In (L) to (N), Wilcoxon matched-pairs signed rank test. ***$P < 0.001$.

the task and the behavioral strategies that may contribute to successful cooperative outcomes (fig. S6A). Using a key point tracking algorithm, Social Leap Estimates Animal Poses (SLEAP) (*34*), we tracked the frame-by-frame poses of both animals. Over the course of training, we observed a sustained increase in animals' approach toward the partner's side of the chamber (Fig. 3, A1 and B1). This increase was not simply due to an overall increase in cooperative performance or an increase in correct trials, because within each trial, this behavior

predominantly occurred within 2 s before nose poke and increased substantially over the course of training (Fig. 3, C1, and fig. S6B). Although approach behavior was more prominent during correct trials, this increase was present in both correct and miss trials (Fig. 3, D1), suggesting that additional behaviors may be involved in facilitating successful cooperation (fig. S6A).

For mice to coordinate the timing of their actions, if one animal arrives at the nose poke port first, they would need to wait for the partner
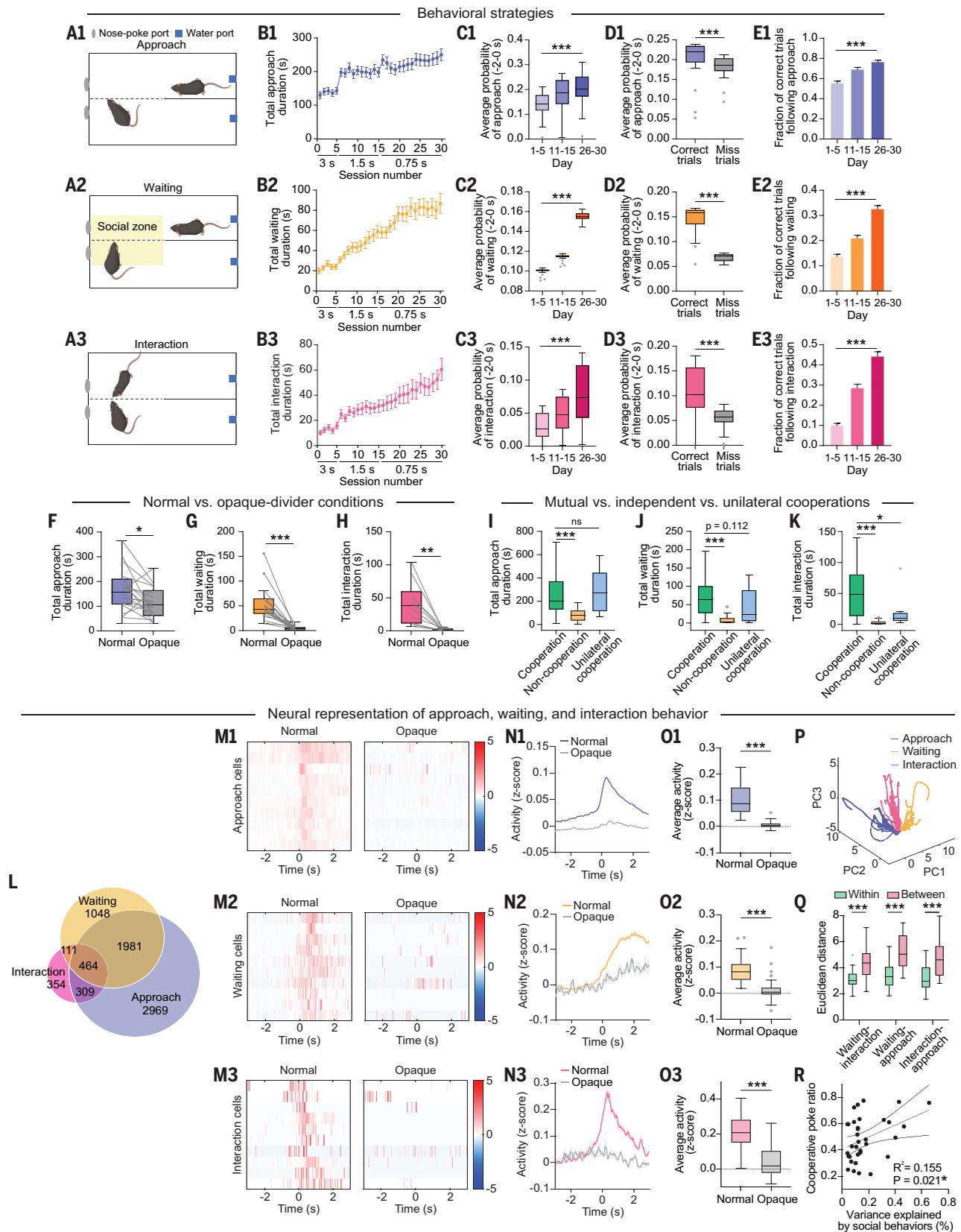
Fig. 3. Behavioral strategies to achieve active coordination and their neural representation in the ACC. (**A**) Schematic illustrating approach, waiting, and interaction behaviors. (**B**) Duration of approach, waiting, and interaction behaviors across training sessions. (**C** and **D**) Average probability of approach, waiting, and interaction behaviors that occurred within 2 s before poking in all trials across different training periods (C) or within 2 s before poking in correct and miss trials in the last five sessions (D). (**E**) Fraction of correct trials after approach, waiting, and interaction behaviors across different training periods. (**F** to **H**) Duration of approach (F), waiting (G), and interaction (H) behaviors under normal or opaque-divider conditions. For the normal condition, data represent the average of durations from the periods before and after the opaque-divider condition

(Materials and methods). (**I** to **K**) Comparison of the duration of approach (I), waiting (J), and interaction (K) behaviors in cooperative, noncooperative, and unilateral cooperative conditions. In unilateral cooperation, the behaviors of the cooperative (second) mice were analyzed and shown. (**L**) Venn diagram showing neurons activated during approach, waiting, and interaction behaviors. (**M**) Heatmaps showing average responses of example cells (each row) activated during approach, waiting, and interaction behaviors in normal and opaque conditions. (**N**) Responses of approach cells, waiting cells, and interaction cells in normal and opaque conditions. (**O**) Average responses of approach cells, waiting cells, and interaction cells during 1 s after the onset of the corresponding behavior in normal and opaque conditions. (**P**) PC projections of population activity associated with approach, waiting, and interaction behaviors. (**Q**) Average Euclidean distance in PC space between or within waiting and interaction, between and within waiting and approach, and between or within interaction and approach. (**R**) Correlation between the percentage of variance explained by social behaviors in the neural space and animals' cooperation performance (cooperative poke ratio). Linear regression with 95% CIs. In (B) to (E), $n = 38$ pairs. In (F) to (H), $n = 10$ pairs. In (I) and (J), $n = 76$ mice (38 pairs; normal cooperation), 10 pairs (noncooperation), and 10 pairs (unilateral cooperation). In (K), $n = 38$ pairs (normal cooperation), 10 pairs (noncooperation), and 10 pairs (unilateral cooperation). In (M) to (R), $n = 34$ mice. In (B), (E), and (N), mean ± SEM; in (C) and (D), (F) to (K), (O), and (Q), box plots: center indicates median, and box indicates quartiles. In (C1) to (C3), and (E1) to (E3), Friedman test. In (D1) to (D3), Mann-Whitney test. In (F), paired t test. In (G), (H), and (O), Wilcoxon matched-pairs signed rank test. In (I) to (K), Kruskal-Wallis test with Dunn's multiple comparisons. In (Q), mixed-effects model with Sidak's multiple comparisons test. Not significant (ns) indicates $P > 0.05$; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

before initiating nose poke. Indeed, we found that mice exhibited a marked increase in waiting behavior (Fig. 3, A2 and B2, and movie S2), defined as the period when a mouse entered the social zone and remained there without nose poking until its partner arrived. Within each trial, waiting behavior occurred predominantly within the 2-s window preceding nose poke and increased substantially over the course of training (Fig. 3, C2, and fig. S6C). Waiting behavior was more strongly associated with correct trials compared with miss trials (Fig. 3, D2).

In addition to waiting, animals also need to precisely coordinate their actions in order to nose poke simultaneously. We found that mice achieved this coordination by mutually poking their noses toward each other between the two sides of the divider. These mutual interactions, like waiting behavior, occurred predominantly within the 2-s window before nose poke and increased by 158.9% over the course of training (Fig. 3, A3, B3, and C3; fig. S6D; and movie S1). Mutual interaction was also more strongly associated with correct trials (Fig. 3, D3). Animals not only displayed more waiting, approach, and interaction across all trials, but even within correct trials, animals showed a substantial increase in these behaviors over the course of training (Fig. 3, E1 to E3). Mice displaying more approach also showed more waiting and interaction behaviors (fig. S6, E and F), indicating a consistency across behavioral strategies. We examined these three types of behaviors in both male and female mice and found that both showed similar patterns, durations, and training-related improvements (fig. S6, G to I).

We further analyzed the body position during interaction behavior (supplementary note S5). During initial training, mice often interact with a bigger internose distance, facing each other directly at ~180° angles (figs. S7, J to M, and fig. S8A). As training progressed, their distance to each other decreased, and the average interaction angle shifted from 180° to ~120°, with both animals facing closer toward the nose poke (figs. S7, J to M, and S8, A to C). This allows animals to maintain visual and physical contact with their partner while simultaneously positioning themselves for efficient transition to nose poking. This behavioral refinement suggests that animals interaction strategies became more precise and efficient. Indeed, animal movements became more synchronized immediately after cross-barrier interactions compared with moments before interaction or other time points (fig. S8D), suggesting that these interactions facilitate coordination.

In the opaque-divider condition, approach behavior was substantially reduced (Fig. 3F), with the remaining approach likely reflecting their motivation to seek social cues from their partner. Waiting and interaction behavior were almost entirely abolished (Fig. 3, G and H), likely owing to the lack of access to social cues. In noncooperation or unilateral conditions, waiting and interaction behaviors were similarly reduced, with interaction behavior showing the most pronounced decrease (Fig. 3, I to K). Thus, approach, waiting, and interaction—which increased substantially as animals improved their performance—may serve as key behavioral strategies for achieving successful cooperative outcomes (fig. S6A).

Encoding of behavioral strategies in the ACC

We next explored whether and how these social behaviors are represented in the ACC. Using ROC analysis, we identified distinct populations of neurons that responded differentially to approach, waiting, and interaction behavior (Fig. 3L). To rule out the possibility that these neurons simply encode the animals' motor actions and/or positions relative to the partner's side of the chamber, we analyzed moments in the opaque-divider control conditions when the animals happened to exhibit a similar motor action (Materials and methods). We found that approach, waiting, and interaction neurons were not active during these moments in the control condition (Fig. 3, M to O). At the population level, these three behaviors were associated with distinct trajectories in the top principal component space, with greater between-group distance compared with within-group distance (Fig. 3, P and Q).

We next examined the neural representation of these social behaviors in individual animals and its relationship to their cooperative performance. Using partial least squared regression (PLSR), we measured the percentage of variance in the ACC neural space that can be explained by these social behaviors and found that it was positively correlated with the cooperative poke ratio in individual animals (Fig. 3R). This indicates that animals that had a stronger encoding of these social behaviors in the ACC showed better cooperative performance.

Encoding of cooperative decision-making

As animals approach the nose poke area, they decide whether to nose poke on the basis of their partner's state: They refrain from poking when the partner is absent ("hold") and proceed to poke when the partner is also approaching the nose poke area ("proceed"). Unsuccessful cooperation occurs when an animal pokes while the partner is far away (fail-to-hold) or fails to engage in poking when both are near the nose poke location (fail-to-proceed) (Fig. 4A). The hold decision is closely related to waiting behavior; whereas "waiting" refers to the moment-by-moment actions that occur throughout the cooperation process, "hold" specifically refers to the discrete decision with respect to poking (refraining from poking) during a given trial. Over the course of training, we observed a substantial increase in trials during which animals made the correct decisions—holding when their partner was not close and proceeding when their partner was close—accompanied by a decrease in fail-to-hold and fail-to-proceed scenarios (Fig. 4, B to E). The fractions of proceed and hold trials were positively correlated with the cooperative poke ratio in individual animals (Fig. 4, F and G), suggesting that the animals learn to observe and adapt to partner actions to inform their own decisions.

To determine whether ACC activity encodes correct and incorrect cooperative decisions, we trained SVM decoders on neural activity during the period when the animals approached the nose poke location before poking. We found that both hold and proceed decisions can be decoded, both from incorrect decisions and between each other (Fig. 4, H to J). Thus, the ACC encodes the cooperative decision-making on the basis of the partner's behavioral state.
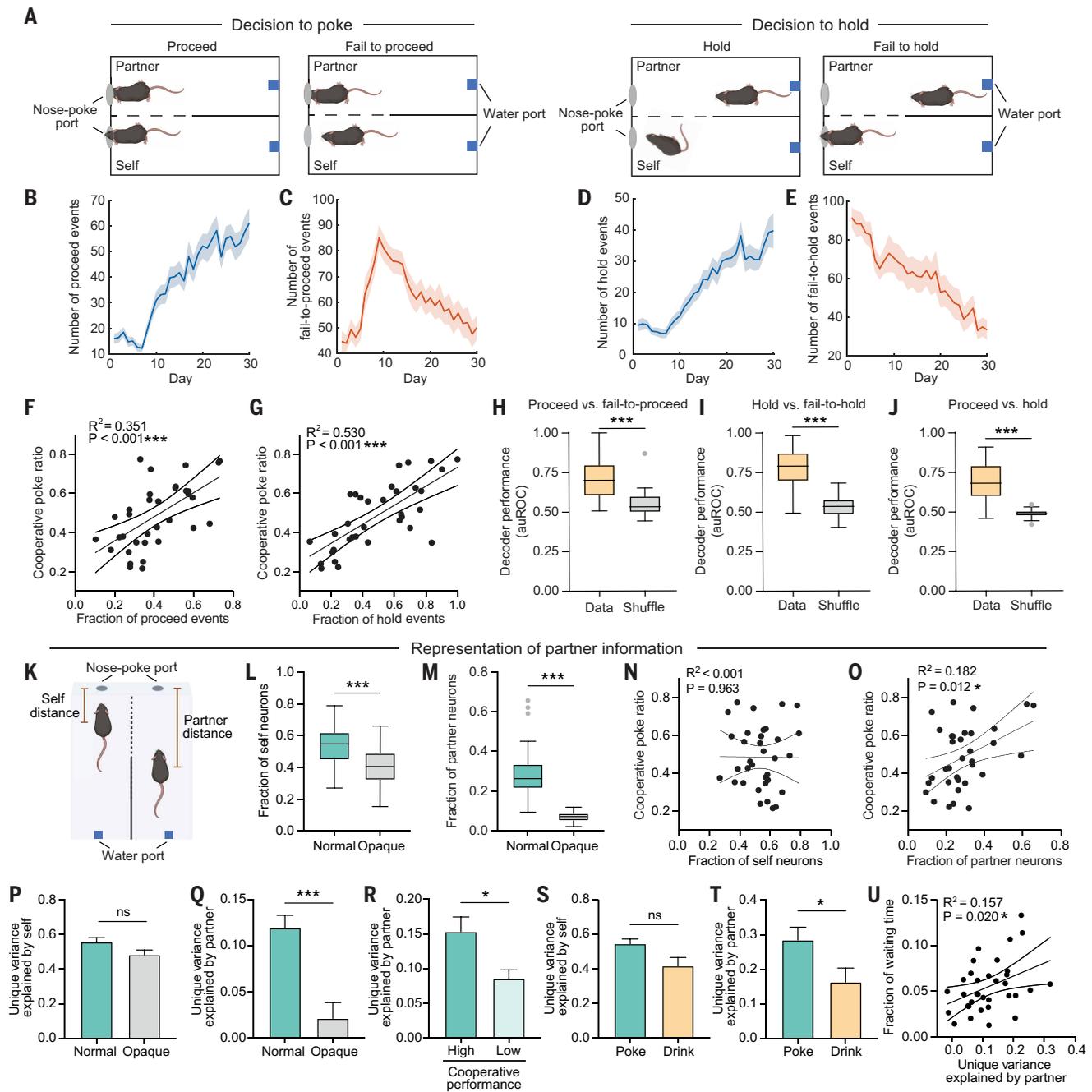
Fig. 4. Neural representation of cooperative decision-making and partner information in the ACC. (**A**) Schematic illustrating mice's decision to proceed to poke when its partner is nearby and decision to hold when its partner is far away. (**B** to **E**) The number of proceed (B), fail-to-proceed (C), hold (D), and fail-to-hold (E) events across training sessions. (**F** and **G**) Correlation between the fraction of proceed events and cooperative poke ratio (F) or between the fraction of hold events and cooperative poke ratio (G). Linear regression with 95% CIs. (**H** to **J**) Performance of decoders in classifying proceed versus fail-to-proceed (H), hold versus fail-to-hold (I), and proceed versus hold (J). (**K**) Schematic illustrating the positions of self and partner with respect to the nose poke port. (**L** and **M**) Fraction of neurons correlated with self (L) or partner (M) positions relative to the nose poke port in normal (transparent) and opaque-divider conditions. (**N** and **O**) Correlation between the fraction of self- (N) or partner (O) neurons in the normal condition and cooperative poke ratio. Linear regression with 95% CIs. (**P** and **Q**) Fraction of unique variance in neural space explained by self- (P) or partner (Q) movement among total variance explained by self and partner. (**R**) Fraction of unique variance in neural space explained by partner movement [as in (Q)] compared between animals classified as high versus low cooperative performance within each pair. High and low cooperative performance were determined by comparing the cooperative poke ratio (fraction of correct pokes among all pokes) between the two animals within each pair. (**S** and **T**) Fraction of unique variance explained by self's (S) or partner's (T) movement trajectory during the cooperative period (before poking) and reward consumption period (drinking). (**U**) Correlation between unique variance explained by partner and the fraction of trials containing waiting behavior. Linear regression with 95% CIs. In (B) to (E), $n = 24$ mice (high performance pairs in Fig. 1, E to I). In (F) to (U), $n = 34$ mice. In (B) to (E), and (P) to (T), mean ± SEM; in (H) to (J), and (L) and (M), box plots: center indicates median, and box indicates quartiles. Wilcoxon matched-pairs signed rank test, ns indicates $P > 0.05$; *$P < 0.05$, ***$P < 0.001$.

Representation of partner information correlates with cooperative performance

For animals to engage in approach, waiting, and interaction at an appropriate time, while also making correct decisions to initiate nose poking and achieving precise temporal coordination of their actions, they must be aware not only of their own location but also of their partner's location. We found that distinct populations of neurons responded to either self's or partner's position relative to nose poke location (Fig. 4, K to M). In the opaque-divider condition, the fraction of these neurons was reduced compared with the normal condition (Fig. 4, L and M). In particular, the fraction of self-position–encoding neurons was reduced by 23.33% (supplementary note S6), whereas the fraction of partner-position–encoding neurons showed a more pronounced reduction of 76.05%. Further, the fraction of neurons responding to partner position—but not self-position—in the normal condition was positively correlated with cooperative performance in individual animals (Fig. 4, N and O). Thus, although both self- and partner position are represented in the ACC, the neural representation of the partner position is more correlated with successful cooperation.

Using linear regression to model the self's or partner's movement trajectories with ACC population activities, we found that self-trajectories were represented strongly in the normal condition, with a weaker representation in the opaque condition (fig. S9A). By contrast, the partner's trajectories were only represented in the normal condition and were nearly abolished in the opaque-divider condition (fig. S9B). Because the self's and partner's location could sometimes be correlated, we determined the unique representation of either the self's or partner's location in the neural population dynamics without the correlated contribution of the other using PLSR (Materials and methods). In the normal condition, both the self's and partner's positions formed unique representations in the ACC (Fig. 4, P and Q). By contrast, in the opaque-divider condition, although the variance explained by self-position was slightly lower, the variance explained by the partner's position was substantially reduced (Fig. 4, P and Q). Moreover, the animals showing higher cooperative performance within each pair also exhibited greater representation of partner position (Fig. 4R). These suggest that partner position is represented in the ACC population activity and that this representation is associated with cooperative performance.

We hypothesized that partner position is more strongly represented in the brain immediately before correct nose pokes—when cooperative decisions are made—compared with the reward consumption period, during which animals' behaviors (drinking) do not depend on each other. Indeed, the representation of partner position was higher before correct nose pokes than during drinking (Fig. 4T). By contrast, representations of self-locations showed minimal changes between these two moments (poking versus drinking) (Fig. 4S). This result suggests that partner position is more prominently represented during cooperative decision-making than during reward consumption. Moreover, animals with stronger neural representations of partner position in the ACC exhibited more waiting behavior (Fig. 4U), suggesting that waiting behavior may be particularly reliant on recognizing partner position.

Inhibiting ACC activity impairs cooperative behavior

To examine whether the ACC plays a causal role during this process, we conducted chemogenetic inhibition of ACC neurons using the inhibitory DREADD hM4Di (Fig. 5, A and B). ACC inhibition led to an overall decrease in the cooperation index. This decrease was due to a reduction of correct trials, because the number of miss trials was not affected (Fig. 5, C to F). ACC inhibition also reduced the fraction of trials showing short intervals and increased the average poke interval of correct trials (Fig. 5, G to I), suggesting that nose pokes became less coordinated. Cooperative behaviors were not affected in control animals expressing mCherry (fig. S10, A to E), confirming that the effect was not due to clozapine-*N*-oxide (CNO) injection. Thus, inhibition of ACC neuron activity reduced the active coordination between animals, resulting in lower cooperative performance.

Furthermore, we found that the duration of approach, waiting, and interaction behaviors decreased after ACC inhibition (Fig. 5, J to L, and fig. S10, F to H), suggesting that ACC neuron activity is required for specific social behaviors critical to cooperative behavior. This reduction in social behavior was not due to increased general anxiety or decreased general sociability, because ACC inhibition does not affect the travel distance or the time spent in the center area in the open field test (fig. S10, I to M) or the general social preference in the three-chamber sociability test (fig. S10, N to R). Moreover, silencing ACC glutamatergic neurons with CaMKII-hM4Di reduced the cooperative index (fig. S10, S to W), suggesting that these neurons are required for cooperative behavior.

Lastly, to achieve precise temporal silencing of ACC neurons, we performed optogenetic experiments using a light-activated inhibitory opsin, stGtACR (Fig. 5, M and N). Brief 1-s optogenetic silencing of ACC neurons before cooperative nose poking substantially reduced the success rate of cooperative trials, whereas silencing for the same duration after cooperative nose poking did not influence the following trials (Fig. 5, O and Q). As a control, light illumination in mCherry-expressing animals did not affect cooperative performance (Fig. 5, P and R). Thus, ACC activity is specifically involved during the cooperative decision-making period leading up to cooperative actions.

Social cooperation in AI systems

We extended our investigation beyond biological systems and examined whether and how cooperation emerges in AI systems. Using multiagent reinforcement learning (MARL), we created two artificial agents with recurrent neural networks (RNNs) through RLlib (*35*) and trained them to cooperate in an artificial environment designed according to the animal task: Two agents navigated an artificial arena divided by a barrier, with each agent occupying one side and each side featuring a "nose poke" location and a "water" port (Fig. 6A and Materials and methods). The agents performed the task using rules similar to those of the animals—they obtained social information through observations and obtained mutual rewards by navigating the arena, engaging in coordinated "nose pokes" within two time steps of each other, and collecting "water" rewards (Fig. 6B and Materials and methods). Each agent implemented an actor-critic architecture with a 256-unit RNN, which had two output branches: one for the policy (actor) and another for the value function (critic) (Fig. 6A), and they were trained to maximize their reward independently using the proximal policy optimization algorithm (PPO) (*36*). The two agents did not share any network parameters; their networks were updated separately after every iteration. Similarly to training animals, artificial agents were trained in two phases to increase the probability of the reinforcement events. In the noncooperative phase, agents separately learned to associate nose poke with reward consumption without coordination (fig. S11A); in the cooperative phase, agents obtained mutualistic rewards when they nose poked within two time steps (Fig. 6B). Each phase consisted of 4000 iterations, each with 4000 environmental steps.

Over the course of training, agents learned to acquire robust cooperative performance (Fig. 6, C to E; fig. S11, B to D; and movie S4). Similarly to mice, trained agents exhibited more correct trials and fewer miss trials compared with untrained agents (those at the first iteration, Fig. 6D). This resulted in a substantial increase in the fraction of correct pokes above chance, exceeding the performance of agents trained under the noncooperative condition (Fig. 6E and fig. S11D). The nose poke interval distribution had relatively low variance in the cooperative condition, peaking around 0 (Fig. 6, F to I), indicating more precise coordination between agents. Although agents with active coordination had a smaller time interval between drinking events compared with noncooperative agents, the nose poke interval was shorter than drinking intervals (Fig. 6, H and I).
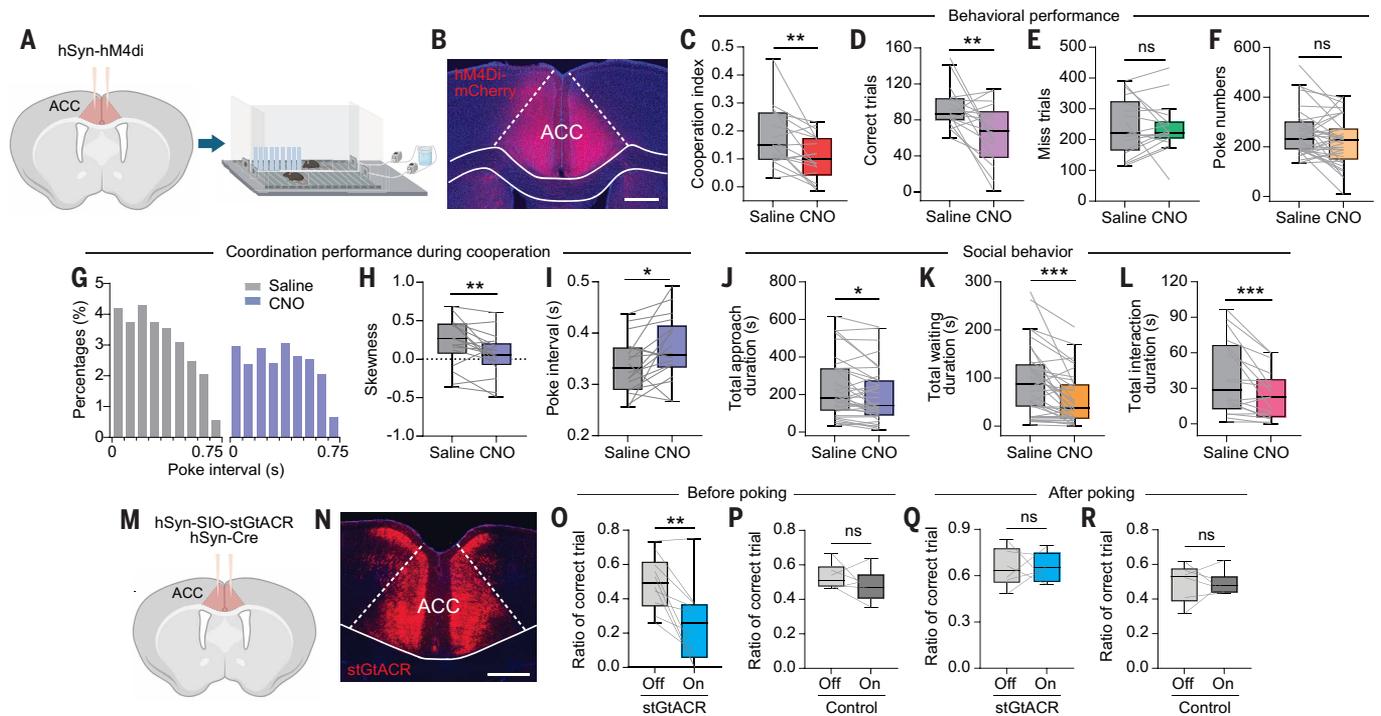
Fig. 5. Inhibition of ACC neural activity impairs cooperative behavior. (**A**) Schematic of viral injection and experimental paradigm for hSyn-hM4Di DREADD inhibition in the ACC. (**B**) Example image showing hSyn-hM4Di expression in the ACC. Scale bar, 500 μm. (**C** to **F**) Behavioral performance of mice expressing hSyn-hM4Di after saline or CNO injection: cooperation index (C), correct trials (D), miss trials (E), and poke numbers (F). (**G** and **H**) Poke interval distribution (G) and skewness (H) of correct trials for hM4Di-expressing mice after saline or CNO injection. (**I**) Average poke interval of correct trials after saline or CNO injection. (**J** to **L**) Duration of approach (J), waiting (K), and interaction (L) behaviors during the cooperation task after saline or CNO injection. (**M**) Schematic of viral injection for stGtACR2 inhibition of ACC neurons. (**N**) Example image showing stGtACR2 expression in the ACC. Scale bar, 500 μm. (**O** to **R**) Ratio of correct trial when photoinhibition (light on) and sham inhibition (light off) is applied before [(O) and (P)] or after [(Q) and (R)] poking in stGtACR2 animals and mCherry-expressing controls. In (C) to (I), and (L), $n = 18$ pairs. In (J) and (K), $n = 36$ mice. In (O), $n = 10$ mice. In (P) to (R), $n = 6$ mice. In (C) to (F), (H) to (L), and (O) to (R), box plots: Center indicates median, and box indicates quartiles. Wilcoxon matched-pairs signed rank test, ns indicates $P > 0.05$; *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

We further analyzed the position and actions of an agent when their partner engaged in a nose poke. In the cooperative condition, agents moved toward the nose poke location from an average distance of less than one step from the nose poke location, whereas in the noncooperative condition, partner position was more randomly distributed throughout the arena (Fig. 6, J and K). Although the movement of cooperative agents was strongly in the direction toward the nose poke location, this was not the case for noncooperative agents. Thus, trained agents actively coordinate their poking to obtain mutual rewards, mirroring cooperative behavior observed in mice.

Observation of partner information is required for cooperation in artificial agents

Given that social interaction is crucial for mutual cooperation in animals, we investigated whether the ability to represent partner information is also necessary for agents to achieve cooperation. Removing partner information from an agent's input substantially reduced the number of correct trials in the cooperative condition, but the same perturbation in the noncooperative condition did not affect agent performance (Fig. 6, L and M). Furthermore, when agents were trained without partner information, they learned to drink independently but did not learn to cooperate (fig. S11, E to H).

We further analyzed the agent's value estimates, which reflect the estimated sum of future rewards associated with each environment state. In both cooperative and noncooperative conditions, we observed an increase in the value output before and during self–nose poke actions, followed by an immediate decrease after that (Fig. 6, N to P).

This indicates that agents associated their own nose pokes with reward. By contrast, the value associated with the partner's nose poking actions remained flat in the noncooperative condition but increased markedly in the cooperative condition (Fig. 6, N to P). Moreover, the value increase associated with partner actions was dependent on agent's proximity to the nose poke location—it was high when the agent was one or two steps away from the nose poke location but dropped—and was indistinguishable from the noncooperative condition—when it was more than three steps away (Fig. 6Q). Thus, cooperative agents developed value representations of nose poking actions of both self and partner that increased during moments critical for successful cooperation.

Behavioral strategies and their neural representation in artificial agents

We next investigated whether agents exhibited behavioral strategies to facilitate cooperation in ways similar to those observed in animals. To address this, we analyzed each agent's actions with respect to their partner's position. Because agents can "see" each other directly, we did not expect approach or interaction like those observed in mice. However, we found that agents displayed "waiting" behavior: they paused (idled) or moved backward when their partner was farther away from the nose poke location. The movement flow field of the agents revealed that both aimed to actively coordinate their actions by minimizing the difference in their distances to the nose poke locations (Fig. 7, A to C). This active coordination occurred before correct pokes but was absent before miss pokes or during the noncooperative condition (Fig. 7D). Furthermore, this waiting behavior was positively correlated with
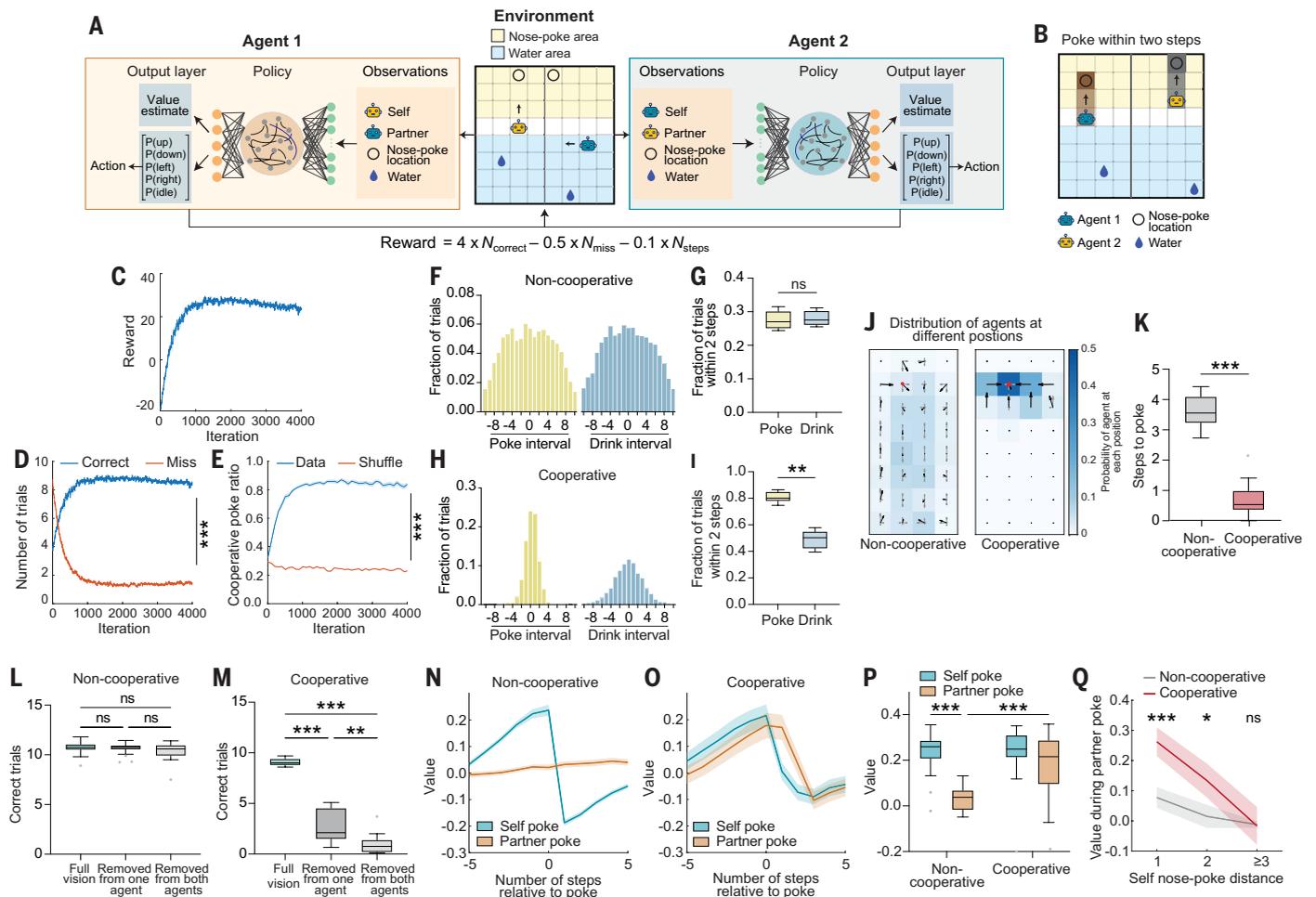
Fig. 6. Social cooperation in artificial intelligence systems. (**A**) Schematic illustrating the structure of the cooperation task in artificial agents using MARL: Two agents navigate an artificial arena divided by a barrier, observe the environment (including the other agent), and make action decisions through their policy network. In each trial, the nose poke port may appear on a random tile within the "nose poke" area, and the drink port may appear on a random tile within the "water" area. This creates a dynamic environment that requires the agents to actively coordinate their actions based on their partner's behavior, rather than relying solely on the timing or positioning of their own actions (Materials and methods). (**B**) Schematic of a successful cooperation (correct trial). Agents need to nose poke within two steps to achieve a correct trial. (**C** to **E**) Performance of artificial agents during the cooperative training phase: reward (C), number of correct and miss trials (D), and cooperative poke ratio compared with a shuffle control (E) across training iterations. (**F** to **I**) Poke intervals of agents in the noncooperative phase [(F) and (G)] or cooperative phase [(H) and (I)]. Distributions of intervals between nose poke and drink events for two agents [(F) and (H)], and fractions of poke or drink intervals within two steps [(G) and (I)]. Each pair was sorted into an "initiator" and a "follower" according to which agents initiated more trials. Poke intervals were calculated by subtracting poking time of the "follower" agent from poking time of the "initiating" agent. (**J**) Heatmap showing probabilities of agents' positions when their partner is at the nose poke area, with arrows indicating normalized moving direction. (**K**) Agents' Manhattan distance to the nose poke when their partner is at the nose poke. (**L** and **M**) Number of correct trials when observation of the partner's location is removed in the noncooperative phase (L) and cooperative phase (M). (**N** and **O**) Value output associated with self–nose pokes and partner nose pokes in the noncooperative phase (N) and cooperative phase (O). (**P**) Comparison of value output for self–nose pokes and partner nose pokes between cooperative and noncooperative phases. (**Q**) Value during partner nose pokes as a function of the Manhattan distance from self to nose poke. In (C) to (K), and (P) and (Q), $n = 20$ agents. In (L) and (M), $n = 10$ pairs. In (C) to (E), (N), (O), and (Q), mean ± SEM; in (G), (I), (K) to (M), and (P), box plots: center indicates median, and box indicates quartiles. In (D) and (E), repeated measures two-way ANOVA. In (G), (I), and (K), Wilcoxon matched-pairs signed rank test. In (L) and (M), repeated measures one-way ANOVA with Tukey's multiple comparisons test. In (P) and (Q), two-way ANOVA with Tukey's multiple comparisons test. ns indicates $P > 0.05$; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

better cooperative performance in individual agents (Fig. 7E). Thus, waiting behavior, as it did in animals, served to facilitate cooperative behavior in artificial agents.

To understand how agents' neural networks encode self- and partner-related positional information, we examined the activity of individual units in the recurrent neural networks, which reflected internal computations that transform observations into actions (Fig. 7F). Using PLSR, we measured the unique variance in neural activity attributable to either self- or partner-related information, excluding any contributions due to correlations between the two (Materials and methods). In

the cooperative condition, the unique variance explained by partner-related information was higher than in the noncooperative condition, whereas the variance explained by self-related information was lower (Fig. 7, G and H). These findings are consistent with those observed in animals, highlighting the critical role of representing partner-related information in successful cooperation. Additionally, consistent with the observations in mice, individual neurons (units) that represented the relative position of self or partner to the nose poke location increased compared with the noncooperative condition (self- or partner distance to the nose poke location) (Fig. 7, I to K).
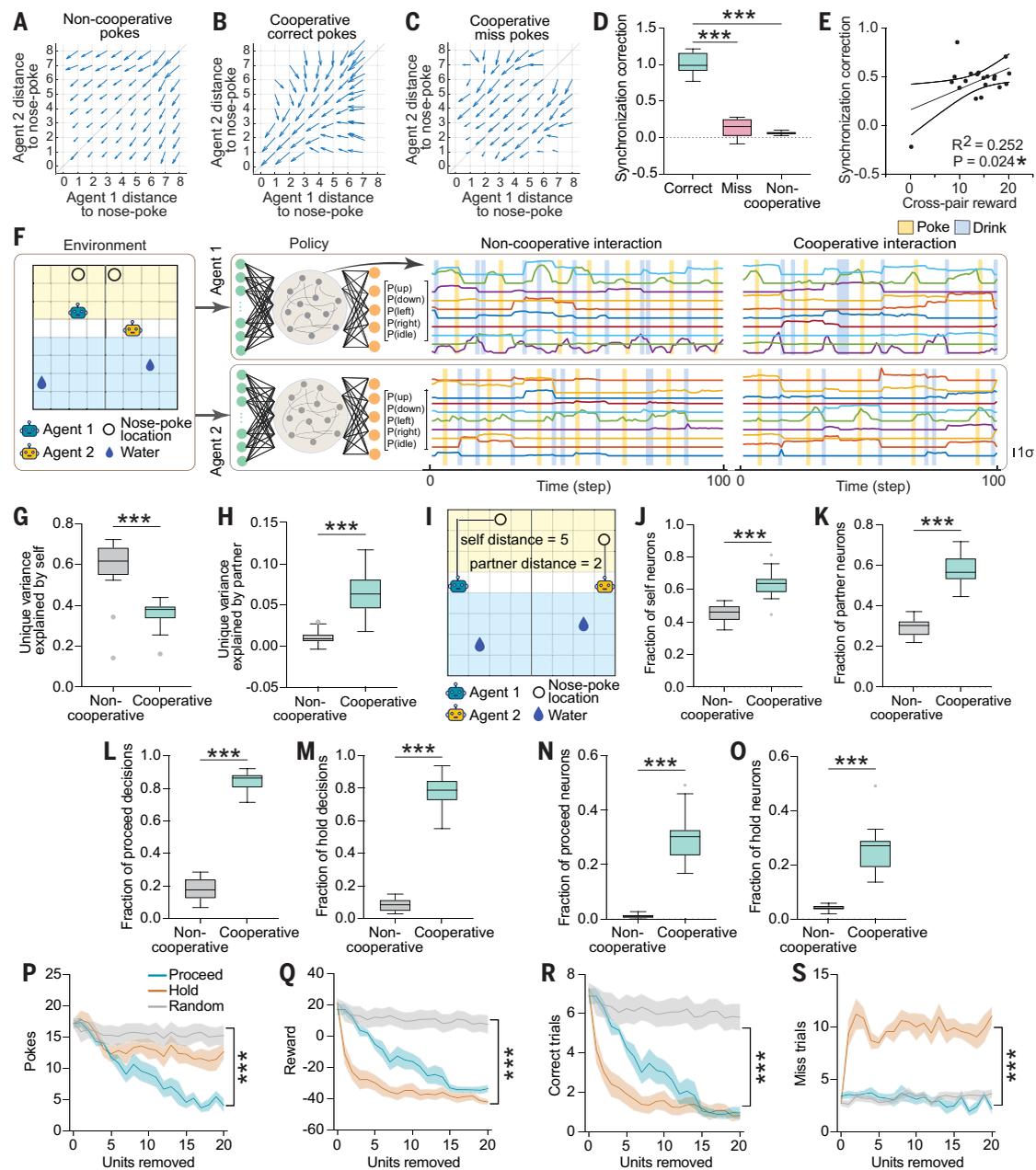
Fig. 7. Behavioral and neural dynamics of artificial agents during cooperation. (**A** to **C**) Flow field of the two agents' changes in distance to their respective nose pokes, where the diagonal line represents equal distances and perfect coordination. (A) Pokes in the noncooperative phase, (B) correct pokes in the cooperative phase, and (C) miss pokes in the cooperative phase. (**D**) Active coordination to achieve synchronization before correct or miss pokes, or pokes in the noncooperative condition. Active coordination is defined as the reduction in the difference (synchronization correction) between two agents' Manhattan distance to their corresponding nose pokes (which reflects waiting behavior), when one agent is at least two steps ahead of the other, measured per environmental step. (**E**) Correlation between active coordination to achieve synchronization in the cooperative phase and agents' cross-pair performance (reflecting their cooperative performance). Linear regression with 95% CIs. (**F**) Example traces of activations in the hidden layer of the agents' policy network. (**G** and **H**) Unique variance in the hidden layer explained by self-related information (G) and partner-related information (H) in noncooperative and cooperative conditions. (**I**) Diagram illustrating agents' Manhattan distance to the nose poke. (**J** and **K**) Fraction of units significantly correlated with self–nose poke positions (J) or partner nose poke positions (K) (based on Pearson's correlation coefficient). (**L**) Fraction of correct decisions to proceed with poking when both self- and partner agents approach the nose poke with similar distances (Materials and methods). (**M**) Fraction of correct decisions to hold from poking when the agent is close to the nose poke and the partner is far away (Materials and methods). (**N**) Fraction of neurons responding to proceed behavioral processes, defined as units significantly correlated with both self's and partner's position relative to the nose poke (showing higher activation when both agents approached their respective nose poke ports; Materials and methods). (**O**) Fraction of neurons responding to hold behavioral processes, defined as units correlated with the difference between self's and partner's positions relative to the nose poke (showing higher activation when self is close to the nose poke and the partner is far away; Materials and methods). (**P** to **S**) Effect of ablating hold, proceed, or random neurons on the total number of pokes (P), total reward (Q), number of correct trials (R), and number of miss trials (S). In (A) to (C), $n = 10$ agents. In (D) and (E), (G) and (H), and (J) to (O), $n = 20$ agents. In (P) to (S), $n = 10$ pairs. In (P) to (S), mean $\pm$ SEM; in (D), (G) and (H), and (J) to (O), box plots: center indicates median, and box indicates quartiles. In (D), repeated measures one-way ANOVA with Tukey's multiple comparisons test. In (G) and (H), and (J) to (O), Wilcoxon matched-pairs signed rank test. In (P) to (S), repeated measures two-way ANOVA. ***$P < 0.001$.

Similarly to animals, agents must make appropriate decisions based on the location of the other agent. If the partner is already at the nose poke area, the agent should proceed to poke ("proceed"); otherwise, it should refrain from poking ("hold"). Indeed, we found that these hold and proceed actions occurred substantially more often in cooperative conditions than in noncooperative ones (Fig. 7, L and M). The behavioral processes that involved hold and proceed actions were also represented at the level of individual neurons, with a higher proportion in the cooperative condition compared with the noncooperative condition (Fig. 7, N and O).

To determine whether these cooperative action-encoding neurons play a causal role in facilitating successful cooperation, we selectively abolished neurons encoding the proceed or hold actions. This led to a drastic reduction in cooperative performance, whereas removing an equivalent number of random units had little effect (Fig. 7, P to S). This observation was not due to a lower contribution of random units to total variance. Even when we increased their number of random units to match the variance of the task-encoding neurons, the effect remained the same (fig. S11, L to P). Although ablating hold and proceed neurons both disrupted cooperative coordination, they produced distinct behavioral outcomes. Both manipulations led to fewer correct trials; however, removing the proceed neurons substantially reduced the total number of pokes by 73%, whereas removing hold neurons had little impact (Fig. 7P). This difference arose because eliminating hold neurons increased miss trials, whereas removing proceed neurons had no such effect (Fig. 7S). This aligns with their functional roles—hold neurons encoded scenarios where the agent is near the nose poke while the partner is farther away, and their removal impaired the agent's ability to refrain from poking. These observations highlight the distinct functional roles of different neuronal populations in enabling social cooperation.

Discussion

In mutual cooperation, two or more individuals coordinate their actions to achieve shared goals (1–3). Using a mutual reward task, we found that both rodents and AI agents developed convergent behavioral strategies to coordinate their actions and gain rewards. In mice, neural activity in the ACC represented key aspects of coordination and played a causal role in its execution. In parallel, targeted perturbations of task-relevant neurons within artificial agents revealed functionally distinct subpopulations that drove coordination behavior in these agents. These results highlight conserved computational principles underlying social cooperation across biological and artificial systems (1–3).

To study mutual cooperation, we devised an operant task in which two mice were required to coordinate their actions to achieve a shared reward. Compared with the task environment and chamber that was previously designed for rodents and shrews (15–17), a set of improvements were made to better study cooperative interactions in mice (supplementary note S7). Using this task, we demonstrated that freely behaving mice can learn to actively coordinate their nose pokes to obtain mutual rewards. The fact that mice received mutual reward does not necessarily mean that both animals correctly understood the rule and actively coordinated their actions. Using a combination of analytical approaches and experimental controls, we demonstrated that active coordination involving perception of partner information and mutual benefits is essential for successful cooperative outcomes in mice, supporting the notion that both animals correctly followed the task rules to actively coordinate their behaviors and arguing against the possibility that the second animal simply mimics the first animal's behavior or follows their social cues.

We identified three behavioral strategies that contribute to successful cooperative outcomes: approach, waiting, and interaction behavior. In addition to these preparatory behaviors, animals also learned to observe and adapt to their partner's behavior and make appropriate decisions accordingly. As animals approached the nose poke area, they refrained from poking when the partner was absent (hold) and proceeded to poke when the partner was also approaching (proceed). These behavioral strategies and decision-making processes enabled mice to achieve successful cooperation.

Neural recordings in humans and nonhuman primates have shown that cooperation involves multiple social brain regions, including the ACC, dorsomedial prefrontal cortex (dmPFC), orbitofrontal cortex (OFC), and temporoparietal junction (19, 21, 37–39). The ACC is important for a broad array of behavioral functions, including emotional processing, reward and punishment, negative affect, and pain perception (22). In the context of social interaction, the ACC has been implicated in social cognition and decision-making, such as tracking others' actions, predicting others' decisions, and integrating information about social partners to guide behavioral choices (21, 30). Additionally, the ACC plays a key role in emotional contagion, social transfer of pain experience, and empathy-driven helping behavior (32, 40, 41). Using microendoscopic calcium imaging with single-cell resolution, we observed that the ACC not only represents individual social behavioral actions (approach, waiting, and interaction) but also plays a critical role in encoding key decision-making processes—the decision to hold or proceed—during active coordination. This process relies on closely monitoring the partner and effectively adjusting one's own actions to align with the partner. Indeed, we found a strong representation of partner location in the ACC, especially during the poking phase. These findings highlight the ACC as a hub for cooperative decision-making: It facilitates social observation, integrates self- and partner-related information, and guides appropriate cooperative responses tailored to diverse scenarios. This role is further supported by causal evidence demonstrating that silencing the ACC impairs active coordination. The ACC appears to be involved in both the expression of learned cooperative behavior and the acquisition and refinement of cooperative strategies during learning. Because our manipulations targeted general populations of ACC neurons, the causal role of specific neural representations of behavioral strategies and decision-making processes remains to be tested. Additionally, because the ACC is functionally connected with other cooperation-related brain regions, including the dmPFC and OFC (19), our study opens up directions for future studies to explore the interactions between ACC and other brain regions during cooperative behavior. It is possible that the integrated information in the ACC is being used by other interconnected brain regions toward coordinated movement, and this possibility remains to be tested.

Computational models using AI have shown promise in modeling behaviors of single individuals, but their application to interindividual social interactions remains underexplored (23, 42). In this study, we directly compared the acquisition of cooperative behaviors in both biological and artificial systems, finding similarities between rodents' social interaction strategies and those of AI agents during cooperation tasks. Both biological brains and artificial networks organize into functional groups that enhance their response to stimuli (42–44). We identified select subsets of neurons in both biological brains (ACC) and artificial networks that encode the key cooperative decisions. Paralleling our observations in mice, artificial agents developed enhanced representations of both their own and their partner's information as they learned coordinative actions, with partner location becoming increasingly associated with reward outcomes. These results underscore the critical role of real-time social information integration and suggest that principles derived from biological systems can potentially inform the design of more sophisticated collaborative AI architectures. Although artificial neural networks cannot fully capture the complexities of neuronal subtypes found in the biological brain (supplementary note S8), our findings suggest that artificial agents can serve as useful models to interrogate the computational principles of social cognition.

Conversely, the artificial system provides a powerful platform for testing hypotheses that are technically challenging in biological systems. A key benefit of our multiagent environment is the complete access to each agent's neural network, enabling precise manipulations. We demonstrated that selectively perturbing task-relevant units in artificial agents impaired distinct aspects of cooperative behavior. Thus, AI systems may serve as tractable models for testing mechanistic hypotheses that are technically difficult or impossible in animal models. Beyond targeted perturbations, artificial frameworks also enable systematic scaling to more sophisticated network architectures and more complex social environments.

Multiagent AI systems are becoming increasingly important for real-world scenarios requiring coordination among multiple autonomous entities, such as distributed robotics systems (*45*). Our findings provide valuable strategies for probing the plausibility of biological mechanisms in AI agents and exploring more naturalistic, self-motivated forms of social interaction in artificial agents and their alignment with neural processes in the biological brain during human-AI collaborations.

Materials and methods

Animals

To study the cooperative behavior in mice, C57BL/6J male and female mice (aged 8 weeks) were purchased from Jackson Laboratories (000664). Animals were housed under a 12-hour light/dark cycle, with food available ad libitum. Water was restricted based on behavioral performance and experimental requirements during training. The housing facility maintained a temperature of 21° to 23°C and humidity of 30 to 70%. Animal care and experimental procedures adhered to the NIH Guide for the Care and Use of Laboratory Animals and were approved by UCLA IACUC.

Stereotaxic surgeries

Surgery for calcium imaging: We performed calcium imaging after training in the mutual cooperation task. Viral injections and GRIN lens implantations were performed as previously described (*32*). Briefly, 300 nl AAV5-hSyn-GCaMP7f-mCherry (Addgene) was injected into the ACC (AP: +1.0 mm, ML: +0.5 mm, DV: −1.8 mm, relative to bregma) at a rate of 30 nl/min using a fine glass capillary (WPI). Following a 5-day recovery, a GRIN lens (Inscopix; 1.0-mm diameter × 4.0-mm length) was implanted above the viral injection site (AP: +1.0 mm, ML: +0.5 mm, DV: −1.6 mm, relative to bregma).

Surgery for chemogenetic manipulation: For chemogenetic inhibition of ACC neurons, AAV5-hSyn-hM4Di-mCherry, AAV5-CamKIIa-hM4Di-Cherry, or AAV5-hSyn-mCherry as the control (Addgene) were injected. For hSyn-hM4Di and mCherry-expressing animals, we injected 300 nl AAV bilaterally at different sites along the AP axis (AP 0.0 and 1.0, ML ±0.25, DV −1.8; and AP 2.0, ML ±0.25, DV −1.5) to ensure effective coverage of the ACC. For CamKIIa-hM4Di-expressing animal, we injected 300 nl AAV bilaterally into the ACC (AP 0.0 and 1.0, ML ±0.25, DV −1.8).

Surgery for optogenetics: For optogenetic inhibition experiments in the ACC, 400 nl AAV2-hSyn-Cre combined with AAV1-hSyn-SIO-stGtACR2-FusionRed or AAV2-hSyn-DIO-mCherry (as a control) was bilaterally injected into the ACC (AP 1.0, ML ±0.25, DV −1.8). Ferrule fiber-optic cannulas (200-μm core diameter, 0.37 numerical aperture; Inper) were implanted with 0.40 mm above the virus injection sites with a 15° angle on both hemispheres. All mice were allowed to recover for at least 7 days before behavioral testing.

Histology

Eight weeks after the imaging or DREADD experiments, and 1 week after the optogenetics experiments, mice were transcardially perfused with 4% paraformaldehyde (PFA) and post-fixed in the same solution for 24 hours. Coronal sections with 60-mm thickness were obtained using a vibratome. These sections were stained with DAPI (SouthernBiotech)

and mounted on slides. Images were acquired using a Leica microscope to confirm the position of lens implantation and expression of GCaMP7f, hSyn-hM4Di-mCherry, CamKIIa-hM4Di-Cherry, stGtACR2-FusionRed or mCherry.

Behavioral assays

Mutual cooperation task: We devised a mutual cooperation task that improved upon previous methods (supplementary note S7). Mice were housed in pairs before and throughout the training period. Behavioral experiments were conducted in a chamber divided into two compartments by a transparent divider that featured a perforated segment near the nose poke port, allowing visual, olfactory, and physical contact between mice (Fig. 1A). Each compartment contained a nose poke port (1.5-cm diameter) and a water port, both equipped with an infrared detector (Fig. 1A). When mice engaged with either port, the detector triggers a signal that was recorded by an Arduino microcontroller. The behavior was recorded using a top-mounted camera.

During the pretraining stage, each mouse was independently trained to nose poke for a water reward. We removed animals that exhibited very low motivation to nose poke (showing less than 30 pokes per 30-min session), which accounted for less than 10% of the total animals we tested. Training began in a short chamber (25 cm by 30 cm by 15 cm) for three to seven daily sessions, followed by a longer chamber (25 cm by 30 cm by 30 cm) for three to five daily sessions (at least 100 trials per session for three sessions). During the cooperative stages, mice were trained (side-by-side in the long chamber) to nose poke simultaneously within progressively shorter time windows (3, 1.5, and 0.75 s) to receive mutual rewards (Fig. 1B). In the first two stages (3 and 1.5 s), successful cooperation was signaled by a sound cue, which was removed in the final stage (0.75 s). The three cooperative stages comprised 5, 10, and 15 daily sessions, respectively. Training sessions were conducted once daily, lasting 30 min each. Plain water was used as the reward throughout all behavioral sessions. The training system was controlled by an Arduino microcontroller, with outputs recorded via a custom MATLAB script.

Throughout all experiments, mice were maintained on a controlled water restriction schedule where they received water primarily during behavioral sessions. Specifically, mice were water-restricted every day outside each training session, during which time they received water rewards upon successful task performance. This protocol applied to both the pretraining and cooperative stages. To ensure animal welfare, all mice were weighed daily throughout the experiment. Any mouse whose body weight fell more than 15% below their baseline weight received supplemental water access for 2 to 3 min at 1.5 hours post-training.

To enable a direct comparison of cooperation metrics across training sessions, sessions 1 to 15 (the first two cooperative stages with 3- and 1.5-s windows) in Fig. 1, E to G; fig. S1, B to D; fig. S2; and fig. S3, A to C, and E to G were analyzed by defining correct trials as paired nose pokes occurring within a 0.75-s window, regardless of the cooperative stage. These data points are shown in a shaded color. Sessions in Fig. 1, E to G were also analyzed using the original time window, with the corresponding results presented in fig. S1, G to J. A detailed comparison between males and females were shown in fig. S2 and in fig. S6, G to I.

Noncooperation and unilateral cooperation conditions: In the noncooperation condition, mice poked independently to obtain rewards for another 30 sessions after pretraining stage. In the unilateral cooperation condition, following pretraining, the noncooperative mice (the first mouse) poked independently for rewards, whereas the cooperative mice (the second mouse) were required to poke within the same time window as their partners, similar to normal cooperation, to receive rewards. The behavioral performance of the cooperative mice (the second mouse) was analyzed and shown in Fig. 1, Q to S, and V, and Fig. 3, I and J. In addition, data from noncooperation and unilateral conditions

in Fig. 1, Q to S, and V, and Fig. 3, I and J was compared to all pairs of mice trained in the normal cooperation condition, which included nonsignificant, intermediate-performance, and high-performance pairs.

Cooperation tasks with transparent and opaque dividers: After cooperative training, mice underwent additional testing with both transparent (normal) and opaque dividers within a session. We implemented a within-session design where each pair of mice underwent three consecutive periods (Fig. 1J): (i) 16- to ~20-min transparent-divider condition, (ii) 25- to ~30-min opaque-divider condition, and (iii) 16- to ~20-min transparent-divider condition. This design allows us to directly compare performance within the same animals while controlling for potential time-of-session effects. To ensure consistent comparison, we analyzed the same duration (16 min) of each period for all comparisons. A solid, opaque divider was used to prevent visual communications of social cues. To ensure robust behavioral comparisons, the average durations from the first and second transparent-divider condition (before and after the opaque-divider condition) were compared to those from the opaque-divider condition. Although it is possible that animals may hear or smell the other animal and infer their general location, this is unlikely to allow animals to precisely determine each other's location.

Cooperation task without water deprivation: On the day prior to testing, one animal in each pair was given free access to water for 2 hours before being reunited with its partner in the home cage following cooperative training. On the testing day, the same animal was again given free access to water for more than 1 hour prior to the cooperation task. The hydrated animal was assigned to either the left or right chamber in a counterbalanced manner across pairs. We found that when one animal lacks motivation due to not being thirsty, this animal would substantially reduce nose poking behavior (fig. S3, I to L), which would prevent both animals from achieving successful coordination and receiving rewards. Water deprivation was a standard practice to encourage task engagement, as is commonly used in behavioral neuroscience studies. We anticipate that natural rewards without deprivation would lead to similar cooperative outcomes, as the fundamental neural mechanisms underlying temporal coordination and social information processing are likely independent of the specific motivational context, although the strength and persistence of cooperative behavior might differ. This remains an interesting question for future study.

Home cage behavioral test: To observe the mice's normal behavior in their home cage, we placed a water bottle in the home cage after the task. The behavior of well-trained animals in their home cage was measured on the day following the final training session. We did not observe differences in general social behaviors such as sniffing, social grooming, and huddling between high-performance pairs and other mice in their home cage environment (fig. S1, M to S). Composite score in fig. S1S was calculated by first z-scoring each of the three behaviors (sniffing, social grooming, and huddling) across animals, then computing the mean of these z-scores for each animal. This suggests that the higher cooperative performance of high-performance pairs is not a reflection of generally increased social behavior. Instead, the behavioral differences appear to be specific to the cooperation task context, where high-performance pairs exhibited more precise coordination.

Estimating chance level of poke synchronization: To estimate the chance level of poke synchronization, we randomly shuffled the mice's poke event sequences (fig. S1A). Within each pair of mice, we kept the poke time of one mouse the same and shuffled the poke events of the second mouse. Specifically, we divided the whole session into behavior periods, each defined as the interval between two correct pokes, and then randomly shuffled the order of these periods to reconstruct the entire behavioral sequence. To determine the chance-level correct and miss trials, we compared the real poke times of one mouse to the shuffled

behavioral sequence of its partner. The chance level was calculated as the average across 1000 shuffled iterations.

To determine if mice showed cooperative performance significantly above chance, we computed the ratio of correct trials in the last five sessions in the real data and compared it to the shuffled data. Pairs showing significantly higher correct trial ratio were defined as cooperative pairs. We also subdivided the cooperative pairs into high-performance pairs (more than 4 SD above chance) and intermediate-performance pairs (less than 4 SD above chance). Specifically, we computed the standard deviation of chance-level correct trial ratio across all pairs (based on the average of the last two sessions of the shuffled data). Pairs with an average ratio of correct trials (real data minus shuffled data) exceeding four times the standard deviation were designated as high-performance pairs, and the remaining pairs were categorized as intermediate-performance pairs.

Animal pose tracking and identification: We used SLEAP (https://sleap.ai/) to track animals' positions and poses. SLEAP is a deep-learning-based multianimal pose tracking algorithm (*34*). A model was developed by training on at least 6000 manually labeled frames in videos, with each frame containing two instances. We tracked six nodes (nose, head, left ear, right ear, body, tail base) and five edges (nose-head, head-left ear, head-right ear, head-body, body-tail) for each instance. A custom Python script was applied to identify inaccurate predictions and tracks, and manual examinations and corrections were applied to all behavioral videos to further ensure prediction accuracy.

"Approach behavior" was defined as the moment (i) when the distance between the nose and the divider was less than 6 pixels or (ii) when the distance between the nose and the divider was between 6 and 12 pixels and the angle of the nose-head extension with the divider was between 30° and 150°. The moments that temporally overlapped with poking or with a nose-head angle change greater than 90° within 300 ms were excluded. "Waiting behavior" was defined as the moment when one animal had entered the social zone when the other animal was not in the social zone, approached at least once, and stayed in this area for more than 1.5 s without poking. The social area is defined as the half of the compartment closer the nose poke port. The partner entered the social zone at the end (last 0.5 s) of the waiting behavior period. Interaction behavior was defined as moments when (i) both animals approached the divider, (ii) the distance between their noses was less than 60 pixels, and (iii) their head orientations were facing toward the divider (head angle within 60° of the divider's perpendicular axis). Additionally, we excluded instances when animals were rapidly turning their bodies around after nose poke and facing each other by chance (turning more than 90° within 300 ms). These three behaviors were also defined in the opaque-divider condition using the same criteria.

t-SNE embedding of tracking data: To obtain an unsupervised low-dimensional embedding and clustering of behavioral features, we performed t-SNE (*46*). For each animal pair, we analyzed 5-min video segments from each training stage (day 2: 3-s stage, day 10: 1.5-s stage, day 30: 0.75-s stage). Using SLEAP tracking data, we extracted 12 behavioral features per frame, including each animal's head angle, velocity, nose distance to the divider and poke ports, and head angle relative to the divider, as well as interanimal nose distance and relative velocity. Because key behavioral events were unequally distributed (for example, fewer frames during poking than during traveling or drinking), we applied importance sampling to avoid bias from overrepresented behaviors. This approach relied on an initial embedding to guide sampling (*46*). We first randomly sampled 30,000 frames and embedded them into a two-dimensional (2D) t-SNE space, identifying 11 clusters using watershed segmentation. We then sampled 30 frames per cluster from each session, yielding a final set of 36,220 frames. A final t-SNE embedding was computed using these frames. To generalize the embedding, we

trained a multilayer perceptron regressor to map the full set of behavioral features onto the 2D t-SNE space. We projected all video frames for analysis and visualization.

Chemogenetic inhibition of ACC

Mutual cooperation task: We performed chemogenetic inhibition in the mutual cooperation task. Mice were tested in two counterbalanced sessions. In the first session, half of the pairs received 1% dimethyl sulfoxide (DMSO) in saline (control), and the other half received CNO (5 mg/kg body weight for animals injected with AAV5-hSyn-hM4Di-mCherry or 1.6 mg/kg body weight for animals injected with AAV5-CamKIIa-hM4Di-mCherry; Enzo, catalog number BML-NS105) in saline with 1% DMSO, administered 30 min prior to testing. In the second session, the treatment groups were switched: mice previously receiving saline were administered CNO, and those initially given CNO received saline. Between the first and second manipulation sessions, mice were trained to perform the cooperation task as usual, without injections of saline or CNO. Behavioral performance from the first and second sessions was analyzed statistically.

Open field test: To examine whether inhibition of ACC neural activity influences mice's anxiety, we performed an open field test in an open arena (40 cm by 40 cm) 30 min after injections of CNO (5 mg/kg) or saline. The injections were conducted on day 1 and 3. Mice were allowed to explore the open field for 30 min and their behaviors were recorded using a top-view camera. We used SLEAP to track the mice's positions. The center area comprised 50% of the total space. The movement distance and time spent in the center and outer areas were quantified.

Three-chamber social test: To examine whether inhibition of ACC neural activity influences mice's general sociability, we performed a three-chamber social test in a three-chamber apparatus. The apparatus consists of two side chambers (25 cm by 25 cm) and a center chamber (12.5 cm by 25 cm). Mice were habituated to the apparatus with an empty cup in each of the side chambers for 3 days. On the testing day, the subject mice were injected with CNO (5 mg/kg) or saline 30 min before testing in a counterbalanced manner. Initially, mice were allowed to explore the three chambers freely for 10 min with an empty cup in each side chamber. A same-sex stranger mouse was subsequently introduced to one of the cups (social chamber), and the subject mouse was allowed to explore the chambers for another 10 min. The social chamber was randomly assigned in a counterbalanced manner across subject mice. Behaviors were recorded using a top-view camera and tracked using SLEAP after the experiments. The social and nonsocial areas were defined as the quarters of the side chambers containing the cup with the mouse inside and the empty cup, respectively. Social preference score was defined as (time spent in the social area − time spent in the nonsocial area)/(time spent in the social area + time spent in the nonsocial area).

Optogenetic inhibition of the ACC

Following stereotaxic surgeries, animals recovered in their home cages for at least 7 days before resuming cooperative training. Animals were trained with optic fibers. On test days, animals performed the cooperation task as usual. Blue light (473 nm, CNI Laser) was manually delivered for 1 s at an irradiance of ~3 to 10 mW/mm^2 in the target region. For manipulations prior to nose pokes, light was delivered when both animals entered the social zone. For manipulations after nose pokes, light was delivered following correct trials. As a control, sham stimulation (no light delivery) was performed in the same animals under identical behavioral conditions, with real and sham stimulations delivered in an interleaved manner. Additionally, light stimulation was applied to mCherry-expressing control animals under the same behavioral conditions and parameters. Behaviors were manually annotated frame-by-frame using custom-written Python software (Behavior Annotator).

Microendoscopic calcium imaging

Behavioral assay: We performed calcium imaging during the cooperation task. The imaging experiment consisted of four consecutive sessions. A transparent, perforated divider was used in the first and third sessions, and a solid, opaque divider was used in the second and fourth sessions. Each session lasted at least 10 min and could extend up to 60 min to ensure that all sessions included at least 20 to 70 total trials. The calcium fluorescence signals were recorded using UCLA Miniscope V4 equipped with a data acquisition board (Open Ephys) and the behaviors were recorded using an infrared camera (FLIR) simultaneously. The miniscope was controlled by the UCLA miniscope DAQ. Calcium fluorescence signals were recorded at 30 fps, and behaviors were recorded at 15 fps. Behaviors were manually annotated frame by frame using a custom Python software (Behavior Annotator) to identify behavioral events.

We observed that, at times, both animals appeared to poke but did not insert their noses far enough into the nose poke port to trigger the sensor detecting poking events. In these cases, both animals likely perceived that they had completed a correct trial, distinguishing these instances from those where one animal was clearly not near the nose poke area. As a result, the animals may internally represent these as correct trials, even though the behavior is categorized as a miss trial. Therefore, whereas all miss trials were included in the behavioral analysis across all figures, for neural analyses involving miss trials, we excluded these ambiguous cases and only considered miss trials in which one animal was not at the nose poke location. This approach ensures that the neural representation of miss trials accurately reflects the animals' perception.

In addition to well-trained animals, we performed additional recording sessions during the initial noncoordinated stage and in early training stage of poke coordination (3-s and 1.5-s stages). Although the extended duration of training (30 days) prevents us from tracking individual cell identities over the entire course, we were able to analyze the fraction of neurons responsive to correct pokes across different training stages.

Extraction of calcium signals: Calcium fluorescence videos of both animals were simultaneously recorded at 30 fps through a miniaturized microendoscope (UCLA Miniscope v4). Raw videos were first fed into the motion-correction algorithm NoRMCorre to eliminate motion artifact (*47*). We then used the bandpass filter function in ImageJ (filterLarge = 40, filterSmall = 3, percentage of the image size) to remove the fluorescent background from the corrected videos. We next applied CNMF-E (constrained nonnegative matrix factorization) to the filtered videos to automatically detect and extract regions of interest (ROIs) (*48*). All ROIs were manually inspected to remove duplicated ROIs and the ROIs that did not represent cell bodies. A total of 12,798 single neurons were identified.

Unless otherwise specified, the CNMF-E denoised $\Delta F/F$ traces were used in the analyses. For analyses that require higher temporal resolution (ROC analysis and decoder analyses for different pokes), we used the deconvolved spike activities from the denoised calcium traces provided by the CNMF-E software package. The OASIS algorithm, which performs fast online deconvolution on the calcium traces, was used to obtain deconvolved spike trains (*49*).

Single-cell analysis for poking: The relationship between a single neuron's deconvolved spike responses and cooperative behaviors was quantified using a ROC analysis. Prior to downstream analysis, all spike traces were *z*-scored and presented throughout in units of standard deviation. We applied ROC analysis to identify neurons that significantly responded to each cooperative behavior. For poking behavior, a window from 150 ms before to 150 ms after poking onset was used as the positive class. Time periods with no behavior were used as the negative class. True positive rates and false positive rates were computed

over a range of binary decision thresholds that spanned the full range of the neural signal. These rates were used to construct an ROC curve, which depicts the detection capability of the neural signal at various thresholds. The area under the ROC curve (auROC) was then determined to quantify how strongly neural activity was influenced by each event. To evaluate significance, the observed auROC was compared against a null distribution, generated by circularly permuting the deconvolved calcium signals with random circular time shifts 1000 times. A neuron was deemed significantly responsive ($\alpha < 0.05$) if its auROC exceeded the 97.5th percentile (indicating activation) or fell below the 2.5th percentile (indicating suppression) of the null distribution.

Clustering of task-encoding neurons: To investigate how task-encoding neurons evolve throughout the reward sequence (from poke to drink), we clustered neurons that were activated during poking or drinking, as identified through ROC analysis across all animals. For each responsive neuron, we extracted trial-averaged activity within a time window of –2.5 to +2.5 s aligned to both poking and drinking events. The activity was binned into 0.5-s intervals and concatenated to form a single response vector per neuron. Clustering was performed using MATLAB's linkage function (agglomerative hierarchical clustering) to group neurons based on the similarity of their temporal response profiles.

Principal components analysis: To visualize population responses during cooperative behaviors, we applied principal components analysis (PCA) to obtain components that maximize the variance of the neural population activity during behavior events. For poking events (correct, miss, and opaque pokes), we computed trial-averaged neural responses within a time window of –2 to 2 s relative to poke onset. Baseline activity was defined as the average response from –2 to –1 s relative to poke onset. For social events (waiting, interaction, and approach), trial-averaged responses were computed from the previous entry into the social zone to the next exit (all activities within the social zone including the poke). Baseline activity was set as the average response during the period of 1 s following entry. To align trials, timelines were interpolated so that entry, poke, and exit times matched across events. Neural population activities were concatenated across event types for PCA. For each behavior bout, population activities were projected onto the first three principal components for visualization. For comparison of population responses, we calculated the Euclidean distances between PC-projected populations (using the first three principal components) within or across behaviors.

Population decoding of poking events: A SVM decoder was trained to decode pairwise between correct pokes under normal condition (with transparent divider), miss pokes under normal condition (with transparent divider), and spontaneous pokes under opaque-divider condition using deconvolved spike traces. Average activity over the window of –150 to 150 ms was taken for each bout. For each comparison, bouts of the two behavioral classes were balanced by randomly drawing from the class with more bouts, such that the number of bouts was equal. Performance of the decoder was tested using a leave-one-out cross-validation (LOOCV) procedure, where one bout served as the test set and the rest as the training set, repeatedly applied for all bouts. To eliminate contamination between training and test sets, the training samples that were within 5 s from the test sample were eliminated from the training set. The test samples' prediction scores were compared against the true labels to produce auROC values. To generate shuffled performance, deconvolved calcium activities were circularly shifted with random time lags relative to the behaviors for 500 times, and an auROC value was calculated for each shuffle. The average auROC value across the 500 shuffles was compared to the averaged auROC of 50 runs from the experiment data. To correlate auROC values of the correct versus miss pokes with each animal's performance, 25 randomly sampled bouts per event were taken from each animal for a fair comparison.

Within individual sessions, cooperative performance showed a non-significant trend of decrease at the end of training sessions, likely reflecting reduced motivation (fig. S5C). Similarly, whereas neural activity in correct-poke-responsive neurons decreases over time, decoding performance shows no significant differences (fig. S5, D and E).

Population decoding of interaction distance and angle: A support vector regression model was trained to decode animals' interaction distance and angle. For each interaction bout, the average distance between noses, or the average head angle between animals were predicted using the average population activity. Performance of the decoder is tested using an LOOCV procedure as described in "Population decoding of poking events." To assess model performance, the predicted distance and angle was correlated against the true values, and the correlation coefficient was compared to that generated using randomly shuffled data.

Single-cell analysis of social behaviors: Using ROC analysis, we compared single neuron responses to approach, waiting, and interaction behavior between normal (transparent) and opaque-divider conditions. These three behaviors were defined in "Animal pose tracking and identification" based on SLEAP tracking and manual behavioral annotation, with the same criteria applied to normal (transparent) and opaque-divider conditions. In opaque-divider conditions, we identified similar body positions corresponding to each of the three behaviors observed in the normal cooperative condition. ROC analysis was performed as described in "Single-cell analysis for poking."

Variance explained by social behaviors: To determine the percentage of variance in the neural activities that is explained by the three social behaviors, a PLSR analysis was performed with the binary matrix of the three social behaviors as $\mathbf{X}$ and population activity as $\mathbf{Y}$. Through singular value decomposition of $\mathbf{X}$ and $\mathbf{Y}$, PLSR decomposes components from both matrices that maximize the covariance between them. The total variance explained by the three PLS components in the neural space was used to identify correlations with individual animal's cooperative performance.

Characterization of cooperative decision-making categories: Among correct trials, we classified behaviors indicative of cooperative decision-making into two categories: hold and proceed trials. "Hold" trials were defined as instances where a mouse entered the social zone and waited for its partner, holding for at least 1.5 s before initiating a poke. "Failure to hold" would result in a miss trial initiated by the self. This reflects a mouse performing a poke while its partner was not present in the social zone, suggesting a lack of social attention. "Proceed" trials were defined as instances where both mice entered the social zone nearly simultaneously, with a time difference of no more than 500 ms, and performed the poke together. "Failure to proceed" resulted in a miss trial initiated by the partner, where the self-mouse failed to reciprocate its partner's poke despite both animals being in the social zone and adjacent to the poke site.

Decoding trial subtypes: We decoded hold versus failure to hold, proceed versus failure to proceed, and hold versus proceed from the deconvolved neural population activity using an SVM. To exclude confounding effects related to poking execution and trial outcomes, the decoder was trained on activities during the decision-making phase between 1.5 and 0.5 s prior to poke onset. Decoder performance was evaluated using a LOOCV procedure, as described in "Population decoding of poking events."

Predicting self- and partner positions with ACC population activity: A linear regression model was trained on ACC population activity to predict the longitudinal movements of either the subject animal or their partner between the nose port and the water port. The dataset

was split evenly, with the model trained on one half and tested on the other. The total number of frames used for training and testing was balanced across normal (transparent) and opaque-divider conditions.

Unique variance of self and partner in the neural space: We performed PLSR between the movement trajectories of both mice and the neural population activity to obtain the total variance of the PLSR components in the neural space. To determine the unique contribution of self-movement to neural population variance, self-trajectories were circularly shifted 1000 times and partner trajectories were kept intact. For each shuffle, a PLSR was performed between the shuffled trajectories and the population activity. The unique variance of self-movement is defined as the total variance subtracted by the average variance explained in the shuffled data. Similarly, the unique variance of partner movement was obtained by circularly shuffling the partner's movement trajectories. As a control, both self- and partner trajectories were shuffled to compute the chance-level explained variance. The fraction of partner variance explained was obtained by dividing the partner's unique variance by the total nonrandom variance (total variance minus chance variance). To obtain unique variance for self and partner during specific poke and drinking periods, the time window 2 s before poke and 2 s after drinking were concatenated respectively, and PLSR was performed on the concatenated trace using the same method.

Single-cell response to distances to the nose poke port: We calculated the Pearson correlation between each cell's calcium traces and the relative position of self or partner to their corresponding nose poke port. Self- and partner positions were circularly shifted 1000 times with random time lags and correlated with each neuron to obtain a null distribution. A neuron was deemed positively correlated if its auROC exceeded the 97.5th percentile and negatively correlated if it fell below the 2.5th percentile of the null distribution ($\alpha = 0.05$).

Multiagent reinforcement learning
Task design: To study cooperative behavior in artificial agents, we employed a MARL framework. We created an artificial environment in an 8-by-8 grid world with two agents, each moving in a 4-by-8 area. Within each agent's area, we designated a "nose poke port" location and a "water port" location. To prevent agents from simply reinforcing repetitive movement trajectories and to incorporate variance in task goals, the nose poke location was assigned to a random tile within a 4-by-3 region in the top half of the environment and the water port location was assigned to a random tile within a 4-by-4 region in the bottom half of the environment at the start of each trial (after each successful cooperation). Nose poke and drinking actions were defined as the agents entering the nose poke and water port locations, respectively. Training consisted of cooperative and noncooperative phases. In the noncooperative phase, each agent independently nose poked and drank water, with rewards for each poke-drink action sequence. The reward was defined to be

$$\text{Reward} = 4N_{\text{trial}} - 0.1N_{\text{step}}$$

where N_{trial} is the number of complete trials of nose poking and drinking, N_{step} is the number of steps excluding nose poking and drinking.

In the cooperative stage, agents were rewarded at the nose poke port (+2) and water port (+2) only when they nose poked within a two-step window, and were negatively rewarded when any agent poked outside this window. We referred to nose pokes within a two-step window as a "synchronized nose poke."

$$\text{Reward} = 4N_{\text{correct}} - 0.5N_{\text{miss}} - 0.1N_{\text{step}}$$

where N_{correct} is the number of synchronized nose pokes followed by drinking, N_{miss} is the number of individual pokes, and N_{step} is the number of steps excluding nose poking and drinking. Note that poking and drinking were coupled such that each correct trial yielded a total

reward of 4 points (+2 for poking and +2 for drinking); however, if a session ended before drinking could be completed, the final correct trial received only 2 points (+2 for poking alone).

Agent architecture: Each agent was represented by a vanilla RNN. The input to the RNN was a flattened vector containing the x and y coordinates of both agents, both nose poke port locations, and the self–water port location. The current hidden state was computed as the sum of affine transformations of the input and previous hidden state, followed by the ReLU activation function:

$$\mathbf{h}_t = \text{ReLU}\left(\mathbf{W}_{\text{input}}\mathbf{x}_t + \mathbf{b}_{\text{input}} + \mathbf{W}_{\text{rec}}\mathbf{h}_{t-1} + \mathbf{b}_{\text{rec}}\right)$$

where

$$\mathbf{W}_{\text{input}} \in \mathbb{R}^{256 \times 200}, \mathbf{x}_t \in \mathbb{R}^{200 \times 1}, \mathbf{W}_{\text{rec}} \in \mathbb{R}^{256 \times 256}, \mathbf{h}_t \in \mathbb{R}^{256 \times 1},$$
$$\mathbf{b}_{\text{input}} \in \mathbb{R}^{256 \times 1}, \mathbf{b}_{\text{rec}} \in \mathbb{R}^{256 \times 1}$$

The RNN hidden state was mapped to two parallel linear layers to compute the action logits (actor) and the state value (critic). There were five possible actions: move up, left, down, right, and idle. The action logits were used to compute softmax probabilities for each action. The action logits were computed as

$$\mathbf{a}_t = \mathbf{W}_{\text{action}}\mathbf{h}_t + \mathbf{b}_{\text{action}}$$

where

$$\mathbf{W}_{\text{action}} \in \mathbb{R}^{5 \times 256}, \mathbf{h}_t \in \mathbb{R}^{256 \times 1}, \mathbf{b}_{\text{action}} \in \mathbb{R}^{5 \times 1}, \mathbf{a}_t \in \mathbb{R}^{5 \times 1}$$

The value of a state was computed as

$$v_t = \mathbf{w}_{\text{value}}^T \mathbf{h}_t + b_{\text{value}}$$

where

$$\mathbf{w}_{\text{value}} \in \mathbb{R}^{256 \times 1}, \mathbf{h}_t \in \mathbb{R}^{256 \times 1}, b_{\text{value}} \in \mathbb{R}, v_t \in \mathbb{R}$$

We also used RLlib's curiosity module to encourage exploration at the early stage of training (*50*). The curiosity module defines an intrinsic reward based on the uncertainty of future states of the environment. The intrinsic reward is to the extrinsic environment reward to encourage agents to explore the environment.

Training and evaluation: Agents were trained using PPO for 4000 iterations in each phase. Each iteration consisted of 20 episodes, with each episode being 200 time steps. We applied L2 regularization to $\mathbf{W}_{\text{rec}}$ with loss function $\lambda = 0.1$. Ten pairs of agents were trained through the noncooperative and cooperative stages. After the cooperative stage, to enhance the generalizability of cooperation to new agents, we added an additional training stage where agents were trained in a cross-pair fashion with a pool of other agents. Within a pool, agents were interleaved every episode, meaning agents were exposed to many different partners. After cross-pair training, agents showed generalizable cooperative performance to never-before-seen partner agents. We rolled out 10 episodes of 500 time steps for each pair post-training to evaluate performance and perform analyses.

Training without partner observation or under low motivation: To model the effects of limited social information, we designed a training condition in which agents could not observe their partner, mirroring the opaque condition used in animal experiments. In this setting, partner-related observations were removed from the agent's inputs, while information about self-location, nose poke locations, and water port locations remained available. Agents were trained using the same curriculum as in the standard condition.

In the low motivation condition, we trained agents where one agent received rewards for coordinated nose pokes while the other received

only random environment rewards that were unrelated to coordinated behavior. Specifically, one agent did not receive rewards for poking or drinking. Instead, it was given random rewards at each time step. The other agent's rewards remained unchanged. Training followed the same curriculum as in the standard condition. We found that whereas the reward-receiving agent was able to learn individual poking, agents failed to develop coordinated poking behavior (fig. S11, I to K). This result is consistent with our observations in mice and supports the conclusion that mutual motivation is essential for the emergence of coordination.

Action probability during partner nose pokes: To analyze an agent's actions when its partner nose poked, we rolled out episodes where the nose poke port was fixed. This enabled us to visualize the agent's locations and actions, averaging over a fixed nose poke location. For each agent, we recorded their location and action distributions when their partner nose poked, excluding situations when the agent had already nose poked within the last five time steps. We plotted the normalized histogram of the agent's location as a heatmap within the 4-by-8 arena. For each tile, the actions in each direction were normalized and plotted as an arrow in gray, and the vector sum of all actions was plotted as an arrow in black. The arrows were not shown if the agent visited a tile less than 0.1% of the time.

Unique variance: The unique variance of self and partner in the neural space was calculated in the same way as for the mice. We performed PLSR between self–agent location and partner agent location and self-neural activities. Unique variance is defined as the reduction of variance explained in the neural space when either the self-location or the partner location was shuffled.

Characterization of waiting behavior: Since the nose poke port location was randomized after each trial and asymmetric between agents, we used the distance to the nose poke port to measure the synchronization between the agents during cooperation. For the 15 steps leading up to a nose poke, we generated a plot where the x axis represented one agent's distance to nose poke and the y axis represented the other agent's distance to nose poke. We analyzed the dynamics of the nose poke distances by calculating the average change in distance at each subsequent step. Perfect synchronization is represented by the central line, where the distances of the two agents to their respective nose poke port are equal. A correct trial could occur when the difference between the agent's distances is less than or equal to two steps. If the difference exceeds two steps, the agent with the smaller distance must wait to potentially receive reward. To quantify waiting behavior, we calculated synchronization correction, defined as the change in the absolute difference between the two agent's distances to the nose poke port when the difference exceeded two steps. Waiting behavior corresponds to a decrease in this distance. We compared synchronization correction in correct trials, miss trials, and noncooperative trials.

Single unit response to nose poke distance: To evaluate how neurons respond to the distance of both the self and the partner from the nose poke port, we calculated the correlation between each neuron's activation and self- and partner distance to the nose poke port. Consistent with the approach used in previous animal analyses, we circularly shuffled activations 500 times and compared the true correlation with the null distribution to determine if a neuron was significantly correlated with self- or partner distance.

Hold and proceed trials: Similar to animals, agents must learn to hold (refrain from nose poking when the partner was far away) and proceed (nose poke when the partner was nearby). Hold trials were defined as correct trials where an agent moved backward or remained idle when its partner was more than two steps away from the nose poke port. Proceed trials were defined as correct trials where the agents

were synchronized before poking (when the difference in their distances to the nose poke port was less than two steps). The fraction of correct hold decisions was the number of hold trials divided by the sum of the number of hold trials and the number of miss trials when the agent poked while the partner was more than two steps away. The fraction of correct proceed decisions was the number of proceed trials divided by the sum of the number of proceed trials and the number of miss trials when the agent, within two steps of its partner, failed to poke after its partner.

Single-unit response to hold and proceed behavior processes: Similarly to our procedure for animals, we calculated Pearson correlation between each artificial unit's activation and self–nose poke distance, partner nose poke distance, and the difference between self–nose poke distance and partner nose poke distance. We evaluated if a unit is significantly correlated with these variables by generating a null distribution of chance level correlation, as described in "Single-cell response to distances to the nose poke port." We classified units as responding to hold or proceed behavioral processes based on their correlation with self and partner Manhattan distance to the nose poke port. Hold units were defined as units whose activity significantly correlated with the difference between self- and partner distances to the nose poke port (active when the agent was close to the nose poke port while its partner was far away). Proceed units were defined as units whose activity significantly correlated with both self- and partner distances to the nose poke port (active when both agents approached their respective nose poke ports).

Network manipulation: Hold and proceed units were ranked by their correlation coefficients with the respective variables. We ablated the top 1 to 40 units in each agent neural network and quantified the change in number of nose pokes, correct trials, and miss trials. Specifically, for 10 episodes of 200 steps each, we set these ablated unit activations to zero. As a control, we randomly ablated up to 40 nonsignificant neurons to match the variance seen in the task-coding neuron populations.

Quantification and statistical analysis

All statistical analyses were conducted using Prism (GraphPad) or MATLAB (MathWorks). Details about the types of statistical tests used and sample sizes are provided in figure legends. Types of statistical tests were determined based on data distribution. P values were corrected for multiple comparisons when necessary. Sample sizes were not predetermined using statistical methods. Our sample sizes are similar to those used in previous publications in the field, and are deemed appropriate based on the size and significance of the effects and consistency across animals. Animals of appropriate genotype, sex, age, and weight were randomly assigned to experimental or control group. Test order was counterbalanced across animals whenever necessary. Experimenters were not blind to group allocation during data acquisition or analysis. All behavioral, imaging, and manipulation experiments were replicated in multiple animals with similar results (see figure legends for exact numbers of animals and/or trials for each experiment). Example micrographs were based on at least three independent biological samples (animals) showing similar results. The center line in the box plots indicates the median, the box limits indicate the upper and lower quartiles, and the whiskers indicate data within 1.5× interquartile range.

REFERENCES AND NOTES

1. D. G. Rand, M. A. Nowak, Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013). doi: 10.1016/j.tics.2013.06.003; pmid: 23856025
2. M. Stallen, A. G. Sanfey, The cooperative brain. *Neurosci.* **19**, 292–303 (2013). doi: 10.1177/1073858412469728; pmid: 23300215
3. T. Clutton-Brock, Cooperation between non-kin in animal societies. *Nature* **462**, 51–57 (2009). doi: 10.1038/nature08366; pmid: 19890322
4. S. Hirata, K. Fuwa, Chimpanzees (Pan troglodytes) learn to act with other individuals in a cooperative task. *Primates* **48**, 13–21 (2007). doi: 10.1007/s10329-006-0022-1; pmid: 17103081

5. R. Chalmeau, Do chimpanzees cooperate in a learning task? *Primates* **35**, 385–392 (1994). doi: 10.1007/BF02382735

6. J. M. Plotnik, R. Lair, W. Suphachoksahakun, F. B. M. de Waal, Elephants know when they need a helping trunk in a cooperative task. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5116–5121 (2011). doi: 10.1073/pnas.1101765108; pmid: 21383191

7. K. Jaakkola, E. Guarino, K. Donegan, S. L. King, Bottlenose dolphins can understand their partner's role in a cooperative task. *Proc. Biol. Sci.* **285**, 20180948 (2018). doi: 10.1098/rspb.2018.0948; pmid: 30232161

8. G. S. Wilkinson, Reciprocal food sharing in the vampire bat. *Nature* **308**, 181–184 (1984). doi: 10.1038/308181a0

9. F. Péron, L. Rat-Fischer, M. Lalot, L. Nagle, D. Bovet, Cooperative problem solving in African grey parrots (*Psittacus erithacus*). *Anim. Cogn.* **14**, 545–553 (2011). doi: 10.1007/s10071-011-0389-2; pmid: 21384141

10. S. Torres Ortiz, A. Corregidor Castro, T. J. S. Balsby, O. N. Larsen, Problem-solving in a cooperative task in peach-fronted conures (*Eupsittula aurea*). *Anim. Cogn.* **23**, 265–275 (2020). doi: 10.1007/s10071-019-01331-9; pmid: 31760558

11. R. D. Estes, J. Goddard, Prey Selection and Hunting Behavior of the African Wild Dog. *J. Wildl. Manage.* **31**, 52 (1967). doi: 10.2307/3798360

12. C. Boesch, H. Boesch, Hunting behavior of wild chimpanzees in the Taï National Park. *Am. J. Phys. Anthropol.* **78**, 547–573 (1989). doi: 10.1002/ajpa.1330780410; pmid: 2540662

13. C. Muro, R. Escobedo, L. Spector, R. P. Coppinger, Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations. *Behav. Processes* **88**, 192–197 (2011). doi: 10.1016/j.beproc.2011.09.006; pmid: 21963347

14. A. R. Conde-Moro, F. Rocha-Almeida, R. Sánchez-Campusano, J. M. Delgado-García, A. Gruart, The activity of the prelimbic cortex in rats is enhanced during the cooperative acquisition of an instrumental learning task. *Prog. Neurobiol.* **183**, 101692 (2019). doi: 10.1016/j.pneurobio.2019.101692; pmid: 31521703

15. M. Jiang *et al.*, Evolution and neural representation of mammalian cooperative behavior. *Cell Rep.* **37**, 110029 (2021). doi: 10.1016/j.celrep.2021.110029; pmid: 34788618

16. K. A. Han, T. H. Yoon, J. Shin, J. W. Um, J. Ko, Differentially altered social dominance- and cooperative-like behaviors in Shank2- and Shank3-mutant mice. *Mol. Autism* **11**, 87 (2020). doi: 10.1186/s13229-020-00392-9; pmid: 33126897

17. K.-M. Zhang *et al.*, A new paradigm of learned cooperation reveals extensive social coordination and specific cortical activation in mice. *Mol. Brain* **16**, 40 (2023). doi: 10.1186/s13041-023-01032-y; pmid: 37170174

18. O. C. Meisner *et al.*, Development of a Marmoset Apparatus for Automated Pulling to study cooperative behaviors. *eLife* **13**, RP97088 (2024). doi: 10.7554/eLife.97088; pmid: 39466838

19. Y. Zoh, S. W. C. Chang, M. J. Crockett, The prefrontal cortex and (uniquely) human cooperation: A comparative perspective. *Neuropsychopharmacology* **47**, 119–133 (2022). doi: 10.1038/s41386-021-01092-5; pmid: 34413478

20. S. W. C. Chang, J.-F. Gariépy, M. L. Platt, Neuronal reference frames for social decisions in primate frontal cortex. *Nat. Neurosci.* **16**, 243–250 (2013). doi: 10.1038/nn.3287; pmid: 23263442

21. K. Haroush, Z. M. Williams, Neuronal prediction of opponent's behavior during cooperative social interchange in primates. *Cell* **160**, 1233–1245 (2015). doi: 10.1016/j.cell.2015.01.045; pmid: 25728667

22. S. van Heukelum *et al.*, Where is Cingulate Cortex? A Cross-Species View. *Trends Neurosci.* **43**, 285–299 (2020). doi: 10.1016/j.tins.2020.03.007; pmid: 32353333

23. M. Botvinick, J. X. Wang, W. Dabney, K. J. Miller, Z. Kurth-Nelson, Deep Reinforcement Learning and Its Neuroscientific Implications. *Neuron* **107**, 603–616 (2020). doi: 10.1016/j.neuron.2020.06.014; pmid: 32663439

24. A. Banino *et al.*, Vector-based navigation using grid-like representations in artificial agents. *Nature* **557**, 429–433 (2018). doi: 10.1038/s41586-018-0102-6; pmid: 29743670

25. H. Makino, Arithmetic value representation for hierarchical behavior composition. *Nat. Neurosci.* **26**, 140–149 (2023). doi: 10.1038/s41593-022-01211-5; pmid: 36550292

26. A. J. E. Kell, D. L. K. Yamins, E. N. Shook, S. V. Norman-Haignere, J. H. McDermott, A Task-Optimized Neural Network Replicates Human Auditory Behavior, Predicts Brain Responses, and Reveals a Cortical Processing Hierarchy. *Neuron* **98**, 630–644.e16 (2018). doi: 10.1016/j.neuron.2018.03.044; pmid: 29681533

27. R. Hattori *et al.*, Meta-reinforcement learning via orbitofrontal cortex. *Nat. Neurosci.* **26**, 2182–2191 (2023). doi: 10.1038/s41593-023-01485-3; pmid: 37957318

28. S. H. Singh, F. van Breugel, R. P. N. Rao, B. W. Brunton, Emergent behaviour and neural dynamics in artificial agents tracking odour plumes. *Nat. Mach. Intell.* **5**, 58–70 (2023). doi: 10.1038/s42256-022-00599-w; pmid: 37886259

29. A. Dafoe *et al.*, Open Problems in Cooperative AI. arxiv:2012.08630 [cs.AI] (2020).

30. P. Gangopadhyay, M. Chawla, O. Dal Monte, S. W. C. Chang, Prefrontal-amygdala circuits in social decision-making. *Nat. Neurosci.* **24**, 5–18 (2021). doi: 10.1038/s41593-020-00738-9; pmid: 33169032

31. M. A. J. Apps, M. F. S. Rushworth, S. W. C. Chang, The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron* **90**, 692–707 (2016). doi: 10.1016/j.neuron.2016.04.018; pmid: 27196973

32. M. Zhang, Y. E. Wu, M. Jiang, W. Hong, Cortical regulation of helping behaviour towards others in pain. *Nature* **626**, 136–144 (2024). doi: 10.1038/s41586-023-06973-x; pmid: 38267578

33. M. Carrillo *et al.*, Emotional Mirror Neurons in the Rat's Anterior Cingulate Cortex. *Curr. Biol.* **29**, 1301–1312.e6 (2019). doi: 10.1016/j.cub.2019.03.024; pmid: 30982647

34. T. D. Pereira *et al.*, SLEAP: A deep learning system for multi-animal pose tracking. *Nat. Methods* **19**, 486–495 (2022). doi: 10.1038/s41592-022-01426-1; pmid: 35379947

35. E. Liang *et al.*, "RLlib: Abstractions for Distributed Reinforcement Learning" in *Proceedings of the 35th International Conference on Machine Learning, ICML 2018*, Stockholm, Sweden, 10 to 15 July, 2018 (PMLR, vol. 80, 2018), pp. 3053–3062.

36. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal Policy Optimization Algorithms. arxiv:1707.06347 [cs.LG] (2017).

37. G. G. Knyazev, A. N. Savostyanov, A. V. Bocharov, P. D. Rudych, A. E. Saprigyn, Multivariate pattern analysis of cooperation and competition in constructive action. *Neuropsychologia* **202**, 108956 (2024). doi: 10.1016/j.neuropsychologia.2024.108956; pmid: 39002772

38. L. Tsoi, J. Dungan, A. Waytz, L. Young, Distinct neural patterns of social cognition for cooperation versus competition. *Neuroimage* **137**, 86–96 (2016). doi: 10.1016/j.neuroimage.2016.04.069; pmid: 27165762

39. J. Decety, P. L. Jackson, J. A. Sommerville, T. Chaminade, A. N. Meltzoff, The neural bases of cooperation and competition: An fMRI investigation. *Neuroimage* **23**, 744–751 (2004). doi: 10.1016/j.neuroimage.2004.05.025; pmid: 15488424

40. J. P. Burkett *et al.*, Oxytocin-dependent consolation behavior in rodents. *Science* **351**, 375–378 (2016). doi: 10.1126/science.aac4785; pmid: 26798013

41. M. L. Smith, N. Asada, R. C. Malenka, Anterior cingulate inputs to nucleus accumbens control the social transfer of pain and analgesia. *Science* **371**, 153–159 (2021). doi: 10.1126/science.abe3040; pmid: 33414216

42. X. Zhang *et al.*, Inter-brain neural dynamics in biological and artificial intelligence systems. *Nature* **2025**, 1–11 (2025). doi: 10.1038/s41586-025-09196-4; pmid: 40604289

43. M. Rigotti, D. Ben Dayan Rubin, X.-J. Wang, S. Fusi, Internal representation of task rules by recurrent dynamics: The importance of the diversity of neural responses. *Front. Comput. Neurosci.* **4**, 24 (2010). doi: 10.3389/fncom.2010.00024; pmid: 21048899

44. D. Bau *et al.*, Understanding the role of individual units in a deep neural network. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30071–30078 (2020). doi: 10.1073/pnas.1907375117; pmid: 32873639

45. Z. Ning, L. Xie, A survey on multi-agent reinforcement learning and its application. *J. Autom. Intell.* **3**, 73–91 (2024). doi: 10.1016/j.jai.2024.02.003

46. L. Willmore, C. Cameron, J. Yang, I. B. Witten, A. L. Falkner, Behavioural and dopaminergic signatures of resilience. *Nature* **611**, 124–132 (2022). doi: 10.1038/s41586-022-05328-2; pmid: 36261520

47. E. A. Pnevmatikakis, A. Giovannucci, NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* **291**, 83–94 (2017). doi: 10.1016/j.jneumeth.2017.07.031; pmid: 28782629

48. P. Zhou *et al.*, Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *eLife* **7**, e28728 (2018). doi: 10.7554/eLife.28728; pmid: 29469809

49. J. Friedrich, P. Zhou, L. Paninski, Fast online deconvolution of calcium imaging data. *PLOS Comput. Biol.* **13**, e1005423 (2017). doi: 10.1371/journal.pcbi.1005423; pmid: 28291787

50. D. Pathak, P. Agrawal, A. A. Efros, T. Darrell, "Curiosity-driven exploration by self-supervised prediction" in *Proceedings of the 34th International Conference on Machine Learning, ICML 2017*, Sydney, Australia, 6 to 11 August, 2017 (PMLR, vol. 70, 2017), pp. 2778–2787.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

Neural basis of cooperative behavior in biological and artificial intelligence systems

Mengping Jiang, Linfan Gu, Mingyi Ma, Qin Li, Jonathan C. Kao, and Weizhe Hong

Editor's summary
Cooperation between individuals is a fundamental behavior in many animals. However, the neural correlates of cooperative behavior remain to be fully elucidated. Jiang *et al.* developed a behavioral paradigm to investigate cooperative behavior in mice and identified key behavioral elements, approach, waiting, and interaction, in the decision-making process during cooperation. In vivo imaging showed that the anterior cingulate cortex is a central hub for cooperative decision making. Finally, the authors modeled cooperation between artificial intelligence systems and showed that they developed behavioral strategies and neural representations reminiscent of those observed in mice. The study provides valuable insights into the mechanisms involved in the formation of groups and societies in the animal kingdom. —Mattia Maroso

View the article online
https://www.science.org/doi/10.1126/science.adw8151
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service