Journal Pre-proofs

Article

Oxytocin-mediated empathy internally facilitates cooperative behaviors in rats

Miaoyaoxin Wang, Qianqian Shi, Yukai Shao, Mengping Jiang, Ao Fu, Yanwang Huang, Qingxiu Wang, Lei Wei, Ziye Zhang, Jingze Xu, Tianming Yang, Yulong Li, Zuoren Wang

PII: S2095-9273(25)00873-4
DOI: https://doi.org/10.1016/j.scib.2025.08.042
Reference: SCIB 3549

To appear in: *Science Bulletin*

Received Date: 10 July 2025
Revised Date: 29 July 2025
Accepted Date: 18 August 2025

Please cite this article as: M. Wang, Q. Shi, Y. Shao, M. Jiang, A. Fu, Y. Huang, Q. Wang, L. Wei, Z. Zhang, J. Xu, T. Yang, Y. Li, Z. Wang, Oxytocin-mediated empathy internally facilitates cooperative behaviors in rats, *Science Bulletin* (2025), doi: https://doi.org/10.1016/j.scib.2025.08.042

Oxytocin-mediated empathy internally facilitates cooperative behaviors in rats

Miaoyaoxin Wang[a,b], Qianqian Shi[c], Yukai Shao[a,b], Mengping Jiang[d,e], Ao Fu[a,b], Yanwang Huang[a,h,i], Qingxiu Wang[a], Lei Wei[a], Ziye Zhang[a,b], Jingze Xu[a,b], Tianming Yang[a,b], Yulong Li[f,g,*], Zuoren Wang[a,b,*]

[a] Institute of Neuroscience, State Key Laboratory of Neuroscience, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[b] University of Chinese Academy of Sciences, Beijing 100049, China

[c] Center for Brain-Inspired Computing Research (CBICR), Department of Precision Instrument, Tsinghua University, Beijing 100084, China

[d] Department of Neurobiology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA

[e] Department of Biological Chemistry, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA 90095, USA

[f] State Key Laboratory of Membrane Biology, School of Life Sciences, Peking University; PKU-IDG/McGovern Institute for Brain Research, Beijing 100871, China

[g] Peking-Tsinghua Center for Life Sciences, New Cornerstone Science Laboratory, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China

[h] Department of Neurosurgery, Huashan Hospital, MOE Frontiers Center for Brain Science, Fudan University, Shanghai 200040, China

[i] National Center for Neurological Disorders, Shanghai Key Laboratory of Brain Function and Restoration and Neural Regeneration, Huashan Hospital, Fudan University, Shanghai 200040, China

* Corresponding author.

Email addresses: zuorenwang@ion.ac.cn (Z. Wang), yulongli@pku.edu.cn (Y. Li).

Article History:

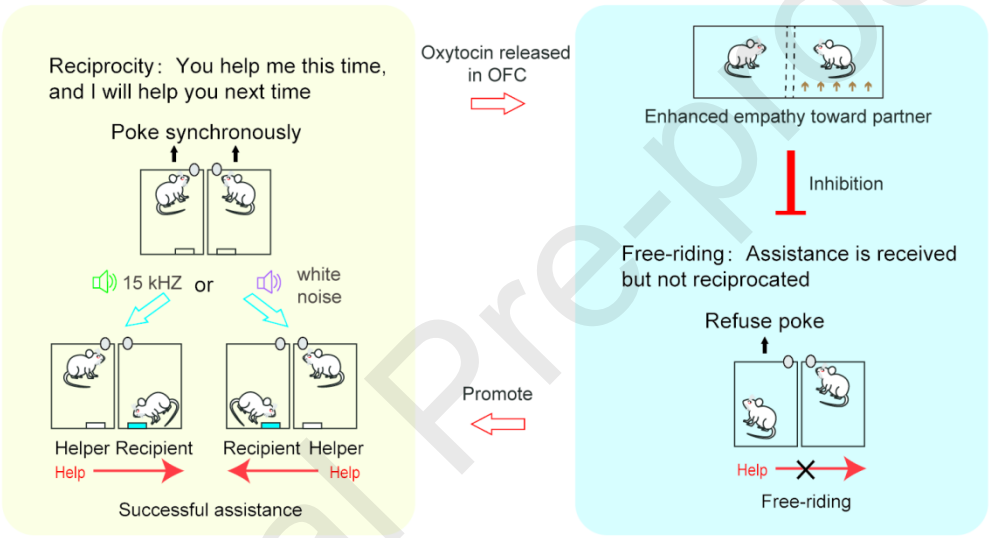
Received 10 July 2025

Received in revised form 29 July 2025

Accepted 18 August 2025

Available online

Graphical abstract



Abstract

Reciprocity is considered one of the vital mechanisms that sustain the evolution of cooperative behavior. However, free-riding, where assistance is received but not reciprocated, poses a serious threat to reciprocity behavior, which relies on future payback. Previous theories proposed that third-party punishment plays a vital role in preventing free-riding behavior. However, this external mechanism has inherent limitations, particularly in situations where third parties are absent. Empathy, the ability to perceive and share the emotional states of others, has long been considered a driving force behind prosocial behavior, yet its role in cooperative behavior remains underexplored. In this study, we have designed a new reciprocity paradigm, and

demonstrate that rats' reciprocity behavior can stably establish even in the absence of the external mechanisms. Additionally, reciprocity experiences can enhance the empathy of wild type rats, but not oxytocin-deficient rats, towards their partners. Furthermore, oxytocin-deficient rats exhibit more free-riding behaviors. Through fiber photometry recording of oxytocin probe, we found that oxytocin remarkably is released in the orbitofrontal cortex during the reciprocity task, significantly exceeding levels observed in both mutualism and individual tasks. Based on our results, we suggest that oxytocin-mediated empathy enhancement reduces rats' free-riding behavior towards their partners, thereby making reciprocity behavior more stable. This empathy-mediated internal driving force complements the previously proposed external mechanisms, providing new theories and perspectives for understanding the evolution of cooperative behavior.

Key Words: Cooperation; Reciprocity; Mutualism; Free-riding; Empathy; Oxytocin

1. Introduction

Cooperative behavior is essential for the survival and reproduction of animal species, particularly in humans, who are known as super cooperators [1, 2]. Cooperation is a cornerstone of human society [3], evident in every facet of life, from major international organizations like the United Nations and the World Trade Organization to initiatives like the Human Genome Project. It underpins our society, culture, and economy.

Understanding the evolution of cooperation is crucial for comprehending the social structures of both humans and animals. It aids sociologists in exploring the mechanisms that foster social cohesion, trust, and collective action [4]. Despite its significance, questions about "How Did Cooperative Behavior Evolve?" remain, highlighted by *Science* as 16th of the 125 critical unsolved most important questions in its 125th anniversary issue in 2005 [5].

Traditionally, it is believed that several mechanisms have driven the evolution of cooperative behavior, including kin selection, group selection, direct reciprocity, indirect reciprocity, and network reciprocity [6]. The later three forms--collectively referred to as reciprocity behavior, along with mutualism, represent the most common types of cooperative behaviors [7]. While mutualism involves immediate rewards for all participants, reciprocity emphasizes delayed rewards and enhances the complexity and flexibility of cooperation [8, 9].

Reciprocity behavior is particularly crucial in scenarios that value long-term gains over immediate returns. However, the delay between the cost incurred and the benefit received by an actor in reciprocal actions can easily lead to free-riding, where individuals fail to reciprocate help, posing a significant threat to the stability and evolution of cooperation [8]. Research over the past fifty years has shown that mechanisms such as second-party punishment [10] and third-party punishment [11-14] are effective in mitigating free-riding and enhancing cooperation in humans. Third-party punishment, also known as altruism punishment, where someone who is not directly involved in the cooperation incurs a cost to punish a free-rider without direct benefits, is especially critical [11]. However, this approach has evident limitations, notably the need for a third-party intervener (large-scale). Moreover, the presence and effectiveness of such mechanisms in animals are less clear. Processes resembling third-party punishment has been noted in several species [15-17]. However, a recent study showed that in chimpanzees—our closest evolutionary relatives, have not found evidence of third-party punishment for violations of cooperation [18].

Nonetheless, the widespread occurrence of cooperative behavior in animals suggests that additional mechanisms might support its evolution. In pursuit of these mechanisms, we have developed a behavioral paradigm that transitions from mutualism to reciprocity and have begun exploring its neural underpinnings.

2. Materials and methods

2.1 Animals

Wild-type female Sprague-Dawley rats weighing 200–250 g (8–10 weeks old) were purchased from Shanghai SLAC (China), Oxytocin-deficient transgenic rats were generated by Beijing Biocytogen Co., Ltd. Rats were pair-housed and maintained on a standard 12-h light-dark cycle (lights on at 07:00). Water deprivation (20–25 mL, per day per animal) was applied prior to experiments to motivate the animals to earn a water reward. Experiments on rats were conducted at the Center for Excellence on Brain Science. All of the experimental protocols followed the institutional guidelines and were approved by the Animal Care and Use Committee of the Center for Excellence on Brain Science and Intelligence Technology (Institute of Neuroscience), CAS. The ethics approval number is NA-002-2022.

2.2 Device employed in reciprocity paradigm

The apparatus consisted of two chambers (22×40×50 cm, WHD) made of aluminum plates with black oxidation. The two neighboring chambers were separated by a grille partition of 3 cm allowing for visual and physical contact. Each chamber contained a

nose port (diameter of 3.5 cm, depth of 2.5 cm) and a water port with an infrared beam inside for detection of nose pokes or water-licking. The locations of the nose ports were fixed and set to 10 cm above the bottom of the chamber (Fig. S1b online).

2.3 Training protocol

There are three stages in a complete reciprocity task training (Fig. S1a online). All rats were paired randomly at the beginning of training and went through all stages. All stages have 30 s dark time to allow rats to habitue to chamber and partner, which means a house-cage light will be open after this 30 s dart time. The first stage, termed the "Individual Stage", requires rats to independently perform a nose-poke to obtain water, thereby learning the association between the action and the reward. To progress, a rat must complete at least 100 trials in a 30 min session and finish five such sessions. The subsequent "Mutualism Stage" introduces cooperative requirements, where paired rats must simultaneously nose-poke within a 3 to 1 s window to earn their reward. Advancement from this stage requires completing 150 trials per 30 min session, over three sessions. The final "Reciprocity Stage" incorporates two paradigms: trial-by-trial and block-by-block. In the trial-by-trial paradigm, each trial initiated with a randomly selected reminder cue indicating which rat might receive a reward. A 15 kHz sound and white noise designated potential rewards for the right and left-sided rats, respectively. Successful coordination of nose-pokes within a 1 s window halted the reminder cue, subsequently triggering a "go" cue to signal reward availability. It's important to note that water rewards were dispensed only when rats accessed the water port. Meanwhile, in the block-by-block paradigm (Fig. S3a online), each 2 min block consistently used the same reminder cue. Apart from this difference, the task requirements remained identical to those in the trial-by-trial paradigm. Rats were defined as well-trained when they exhibited a significant difference in licking latency between the two cues on three cumulative days.

2.4 Modulation Index (MI):

$$\mathrm{MI} = \frac{(\mathrm{X_{self}} - \mathrm{X_{partner}})}{(\mathrm{X_{self}} + \mathrm{X_{partner}})}, -1 \leq \mathrm{MI} \leq +1, \qquad (1)$$

where X is the mean behavioral response to each cue.

2.5 Both Get Reward (Mutualism) and No-reminder-cue (NRC) control tests.

In the mutualism test, reminder cues were still randomly initiated at the start of each trial. However, the modification allowed both rats to access water within the same trial, provided they coordinated their pokes within a 1 s window. This approach aimed to

evaluate their cooperative behavior when both had an equal chance of receiving a reward. Conversely, in the No-reminder-cue (NRC) test, while we still classified trial types as either 15 kHz or white noise, the reminder cue was notably absent at the beginning of each trial. This meant the rats lacked prior knowledge of who could potentially access the water reward. In this scenario, water was dispensed randomly to one side, maintaining a 50% probability for each rat. This test was designed to assess the rats' cooperative behavior under conditions of uncertainty regarding reward distribution.

2.6 Empty cage control test

The test includes two scenarios: in the first, the focal rat and its reciprocity training partner are placed together in a training box. In the second scenario, only the focal rat is placed in its side of the box, while the other side remains empty. All test rats will randomly experience both scenarios. In each scenario, a partner cue for focal rat will continuously play for 2 min, allowing the focal rat to only offer help unilaterally.

2.7 Multiple partner test

The initial experiment encompasses three distinct scenarios, each comprising a helper and a recipient randomly assigned to a training box. In this configuration, the helper possesses the unilateral capability to assist the recipient. The scenarios vary solely based on the recipient's identity: in the first scenario, the recipient is a familiar partner from prior training in reciprocal tasks. In the second, the recipient is an unfamiliar rat of the same sex, not previously engaged in help exchanges but trained in object nose-touching. In the third scenario, the recipient, also an unfamiliar rat of the same sex, has neither previously assisted the test rat nor received any task training, thereby precluding the possibility of receiving help from the helper.

The second test is divided into two phases. In the initial receive phase, focal rats, serving as recipients, either receive assistance from a cooperator or do not get help from a defector over three consecutive days, each session lasting 5 min. On the fourth day, during the return phase, roles are reversed between helpers and recipients to examine the focal rats' propensity to assist various designated partners.

2.8 Tail suspension test

The rat will be suspended by its tail for 5 to 10 s, during which it may exhibit anxious struggling behaviors. Following each suspension, there will be a rest period of 10 to 30

s before the next suspension is initiated. This process will be repeated a total of five times.

2.9 Video acquisition

Video recordings were conducted at a frame rate of 25 frames per second, with a resolution of 1280 × 960 pixels. The camera was strategically placed 75 cm above the midpoint between two chambers, a position that ensured comprehensive coverage of both areas.

2.10 Behavior analysis using DeepLabCut

We utilized DeepLabCut (DLC, version 2.2) to track unmarked body parts of animals in both chambers simultaneously. This tracking was performed offline for all videos from the first and twentieth days of the reciprocity stage. The video analysis' s temporal resolution aligned with the camera' s frame rate of 16.7 ms, while the spatial resolution was determined by measuring a known distance in pixel values at the level of the animals' movement in the chamber, calculated as 0.1 cm per pixel.

A cropped image (1280 × 640 pixels), representing the chamber, was designated as the region of interest (ROI) for training the model. DLC training utilized videos from 16 different animal pairs, with 30 frames per video, resulting in a total of 480 frames. These frames were used to train a ResNet-50 neural network over 1,120,000 iterations, specifically tailored for multi-animal models.

Frame data from DLC with a likelihood below 0.9 was removed. Given the linear movement patterns of rats, we employed regression analysis using data from 10 frames before and after each gap to interpolate missing data. Furthermore, we calculated the orientation of one rat towards another by measuring the angle between two vectors: one extending from the midpoint of its head (situated halfway between the ears) to its nose, and the other from the same midpoint to the other rat' s nose.

2.11 Definition of partition-crossing and mutual approach via DeepLabCut

For each video frame in the recordings, we obtained the coordinate of two rat' s noses in frame to compute their real location in chamber and their inter-nose distance. The *X*-coordinate of each rat's nose was specifically used to identify partition-crossing behavior. Given that the left barrier's coordinate is precisely at 320 pixels and the right barrier at 350 pixels, we defined partition-crossing behavior as occurring when a rat's

nose crossed over these pixel coordinates on its assigned side.

Mutual approach was defined as two criteria (1) Inter-nose distance should be less than 60 pixels (6cm), and (2) Mutual head orientation should be under 30°.

2.12 Criteria of different zone in chamber

We divided the two chambers into specific zones based on their coordinates on the *X* and *Y* axes. The middle coordinates along the *X*-axis for dividing the two chambers are set at 210 and 470 pixels. These coordinates serve as the demarcation points to differentiate the areas proximal to the barrier (and distal to the barrier zone in each chamber. Along the *Y*-axis, the middle coordinate is established at 320 pixels. This coordinate is used to differentiate between the nose-poke zone and the port-licking zone in each chamber.

2.13 Granger causality analysis using PDC

To assess the influence of one rat's serial behavior on another during the reciprocity tasks, we employed Partial Directed Coherence (PDC) following the work [19], PDC is a frequency decomposition of Granger causality analysis[52] which can investigate frequency-specific directed connections between multiple time series. Firstly, we preprocess the data for each condition. In detail, for calculating the Granger causality between the within-trial positions, across numbers (the number of cross partition) and the orientation of two rats, we discarded the trials containing less than 100 frames (60 frames/s) of interaction for analysis (51 out of 620 total trials discarded), and averaged every N/100 frames to obtain aligned time series (Conditions within trials have the same time length, thus aligned). Following the work [19], the aligned data were then subsampled with a sampling rate of 10 samples per second, and were sorted to two groups (helper's and recipient's locations) under different conditions (correct or failed trials), and normalized. Next, we use the methods as follows to implement the Granger causality:

We used a single vector autoregressive (VAR) model to fit the specific data of time series of interest (e.g., positions data during a period of time from beginning to success);From the VAR model, we calculated the information PDC (iPDC) to quantify the directed connectivity from one rat's behavior (time series) to another (e.g., from helper to recipient, and vice versa);We integrated each iPDC to obtain information flow (I_{flow}) which represents the causality from one rat group (helper or recipient) to another rat in units of information transfer (bits):

$$I_{flow} = \frac{1}{fs}\int_0^{fs/2} \log_2(1 - iPDC(f))df \qquad (2)$$

where the fs is the sampling rate.

For testing significance of information flow (I_{flow}), we shuffled the trials in each condition as the surrogate datasets, thus there is no interaction between two rats in the surrogate datasets. We calculated a *P*-value following:

$$P = \frac{(\#\text{ of surrogate } I_{flow} > \text{real } I_{flow}) + 1}{\#\text{ of surrogate} + 1} \qquad (3)$$

Furthermore, for testing the significance of information flow difference between different conditions, we implemented bootstrap methods by selecting n trials randomly 1000 times, and obtaining iPDC and I_{flow} as before. By computing a 2-side finite-bias-corrected *P*-value, we can see the significance difference between two conditions

$$P = 2\text{mininem}\left\{\frac{(\#\text{ of surrogate } I_{flow}\text{differences} > 0) + 1}{\#\text{of bootstrap} + 1}, \frac{(\#\text{of bootstrap } I_{flow}\text{differences} < 0) + 1}{\#\text{of bootstrap} + 1}\right\} \qquad (4)$$

2.14 Viruses

The virus and its titer were as follows:

rAAV2/9-hSys-OXT1.0-WPRE-hGH-polyA = 5×10^{12} ,all virus purchased on (Brain VTA (Wuhan) Co., Ltd).

2.15 Surgery

Rats were anesthetized with isoflurane, and in photometry recording, the positional coordinates used to locate the orbitofrontal cortex (OFC) were anterior-posterior (AP) 3.24 mm, medial-lateral (ML) ± 2.52 mm, and dorsal-ventral (DV) 5.0 mm. The injection volume for OT1.0 was 250 nL. For optical fiber implantation, the fiber (diameter: 200 μm; numerical aperture: 0.37, Nanjing Qianao Technology) was implanted 0.2 mm above the injection site immediately following the virus injection. Dental cement was used to secure the optical fiber to the skull. Post-surgery, rats were allowed a recovery period of 2 weeks prior to recording. For additional procedural details, more details of procedures see [20, 21]

2.16 Fiber photometry recording

A 470 nm laser was released to activate the green fluorescent signal, while an A405 laser was simultaneously used to control for motion noise as describe in [22]. The fluorescence data recorded by the 470 channel was processed for motion noise removal using the data from the 405 channels before final analysis. The ΔF/F values were calculated using the formula (Fsignal-F_0)/ Fsignal. During the mutualism and reciprocity phase, F_0 is the average of the Fsignal during the waiting stage. In the individual phase, F_0 is the average of the Fsignal from -1 to 0 s before the go cue playback. Trials were divided into four stages based on the go cue playback, the rat entering the water port (port entry), and the rat leaving the water port (port off). Only data where the signal in the mutualism or reciprocity task significantly exceeds its baseline F_0 ($P < 0.05$) and where the peak ΔF/F values during the mutualism or reciprocity task exceed 1 were used for further analysis.

2.17 Immunohistology and imaging

First, brain slices were soaked in phosphate buffer solution (PBS; 5 min, twice), and then DAPI (1:1000; CST, #4083S) was added. After 5 min, slices were washed with PBS to remove DAPI (5 min, twice) and then mounted with 75% glycerin. Brain slices were then imaged using an Olympus VS.120 high-through fluorescence microscopic imaging system (10 × objective).

2.18 Oxytocin receptor (Oxtr) mRNA fluorescent in s*itu* hybridization

For Oxtr mRNA fluorescence in *situ* hybridization in OFC, animals were euthanized under isoflurane anesthesia and subjected to transcardial perfusion with DEPC-treated phosphate-buffered saline followed by 4% DEPC-treated paraformaldehyde (PFA) to preserve RNA integrity and fix tissues. Brains were postfixed in 4% DEPC-PFA at 4°C for 4 h, then cryoprotected by immersion in 30% DEPC-treated sucrose solution until fully submerged. Coronal sections (20 μm thick) were prepared using a Leica CM1950 cryostat, and collected on glass slides for further processing.

For the multiplex fluorescent RNAscope assay (323100, ACD Biotech), brain sections were hybridized with commercially available antisense Oxtr mRNA probes (rat-specific; 483671-C3, ACD Biotech) following the manufacturer's protocol. After signal amplification and staining with TSA-570, slides were coverslipped with Fluomount-G medium containing DAPI (0100-20, Southern Biotech) to visualize nuclei and seal samples. Images were acquired using an Olympus VS120 slide-scanning

microscope under 20 × objective (Olympus Corporation, Japan).

2.19 Empathy test

Two rats were placed in a compartmentalized box, separated by a fence. An electric shock plate was installed at the bottom of one side of the box, referred to as the demonstrator side. Ultrasonic microphones were strategically positioned at the top of both sides of the box to capture sound data. The duration of the test was 8 min. For the initial 3 min, no electric shock was administered, allowing for the assessment of the rats' baseline freezing behavior. During the final 5 min of the test, an electric shock of 0.8 mA, lasting for 1 s, was administered every 29 s. The empathy levels of the rats were quantitatively measured using Noldus EthoVision (Noldus Information Technology, Netherlands).

2.20 Behavioral tests

Before behavioral tests include open filed test, elevated plus maze test and three chamber social preference test. Each rat was handled for 5–10 min at a time, twice a day, for a total of two days. The behavioral tests in our experiment were performed sequentially with an interval of one day. For all behavioral tests, we used Noldus EthoVision (Noldus Information Technology, Netherlands) to track the center of gravity of rats to acquire the specific parameters of each behavioral test. Details see [21]

3 Results

3.1 Design of new reciprocity paradigm

To quantitatively study the evolution of cooperative behavior, we previously established a mutualism cooperative paradigm applicable to mice, rats, and tree shrews, in which the two animals poke together and then drink together [20]. To establish a behavioral paradigm that meets the definition of reciprocity, we have modified the previous task into a new version, in which the two rats poke together, but they drink alternately (Fig. 1a and Fig. S1a–c online).

In detail, during a trial of mutualism behavior, two rats need to poke nose port synchronously in a 1 s cooperative time window to get reward for both themselves. The modification we made in our new paradigm is that after the rats learned mutualism behavior, we introduced a reciprocity phase. In this phase, the rats still need to poke the nose port simultaneously within a 1 s cooperative time window, but in each trial, only

one rat receives a water reward (Fig. S1d online). The information about which rat will receive the reward is determined at the start of the trial by a random reminder cue. There are two types of reminder cues, a 15 kHz sound and white noise, which indicate that only the rat on the right or left, respectively, will receive the reward. A trial where the rat itself receives a reward is defined as the 'self-trial' and a trial, while the trial where its partner receives a reward is defined as the 'partner-trial.' In each trial, the rat that receives the water reward is defined as the "recipient", and the rat that helps its partner receive the reward is defined as the 'helper' (Fig. S1e online).

During partner-reward trials, the helper incurs a cost by performing the nose-poke without an immediate reward. The actor can only benefit from a subsequent self-reward trial, which relies on the partner's reciprocating. The time delay between the partner-reward and self-reward trials introduces a dynamic consistent with the fundamental structure of reciprocity: "You help me this time, and I will help you next time." In different trials, the two rats switch roles between recipient and helper randomly with equal probability, receiving help from or assisting their partner to obtain rewards, thus achieving reciprocity.

3.2 Rats show robust direct reciprocity behavior

We initially trained co-housed rats on the task (Fig. S2 online). Only those that completed at least 150 successful trials per day for three consecutive days during the mutualism phase were included in the reciprocity phase (Fig. S2a online). At the beginning of training, rats showed no significant difference in their responses to the two reminder cues. However, as training progressed, their responses to the two cues gradually diverged (Fig. S2b–d online). By day 14, rats exhibited significantly different behaviors in response to the two cues. Specifically, during self-trials—when rats had the opportunity to obtain a reward—they engaged in nose-poking more frequently, exhibited shorter latencies, and transitioned to the water port more rapidly compared to partner-trials (Fig. S2e online). In our task, water is only provided after the rats enter the water port, making this step necessary for receiving a reward. If rats understand that the partner-trial mean helping their partner receive a reward rather than receiving one themselves, they should reduce entries into the water port during partner-trials. Our data confirmed this: after training, there was a significant increase in the number of trials where rats did not enter their own water port during partner-trial (Fig. 1b). This increase was not due to their partners drinking faster, as the time of drinking by the partners did not change significantly before and after training (Fig. 1c). Rats were defined as well-trained when they exhibited a significant difference in licking latency between the two cues for three cumulative days. Based on this criterion, 87.5% of the rats (14 out of 16) successfully met the criteria (Fig. S2f online). These results suggest that the rats understood that they could not receive a reward during partner-trials. Interestingly,

however, they continued to poke the nose port during partner-trials (Fig. S2e online).

To further explore the motivations behind rats' actions during partner-trial, we conducted additional control experiments. Considering the task structure of trial-by trial, if rats were not inclined to assist their partners, they would need to wait 30 s for a 50% chance of entering a self-trial to potentially receive a reward. To discount the hypothesis that rats were merely rushing through partner-trial to reach self-trial more quickly, we introduced a block-by-block training paradigm. In this setup, each block lasted a fixed duration of 2 min and featured a consistent type of reminder cue throughout (Fig. S3a online), thereby eliminating the possibility of rushing within the block. Our findings revealed that rat behavior in the block-by-block structure closely resembled that observed in the trial-by-trial paradigm (Fig. S3b–g online). Notably, all rats (16 out of 16) successfully met the well-trained criteria (Fig. S3e online). After learning to differentiate between the two types of reminder cues, rats consistently engaged in behaviors that facilitated their partner's access to rewards during partner-trial (Fig. S3d online).

To ensure that the rats' performance during the reciprocity phase was not simply driven by mutualistic reinforcement, we designed a control experiment in which the reciprocity phase is replaced with a "both-get-reward" (mutualism) phase, where two reminder cues were still assigned and played in each trial. However, in this phase, both rats could receive a reward in each trial, regardless of the type of reminder cue. Unlike the reciprocity task, rats in the mutualism phase showed no differentiation between the two reminder cues after 14 d of training (Fig. S4a, b online). These results indicate that rats in the reciprocity phase applied a different behavioral rule compared to those in the mutualism phase. Further analysis of no-reminder-cue task (NRC) confirmed that these actions were not the result of associative learning based on a 50% chance but were influenced by the design of our task (Fig. S4c, d online).

If a rat's poke in the partner-trials is intended to help the partner, this behavior should be influenced by the presence of a recipient. To test this, we exposed focal rats to two conditions in a randomized sequence: one with a recipient in the opposite cage and one without (Fig. S4e online). The results showed a significant reduction in nose-poke frequency when no recipient was present during the partner-trials (Fig. S4f online), suggesting that the helper's nose-pokes in the partner-trials are influenced by the presence of the partner.

Rats, which are highly social animals, demonstrate both prosocial motivation and the ability to evaluate the outcomes of others [23], making them an ideal model organism for studying the reciprocity behavior [24]. They have been shown to exhibit direct and generalized reciprocity behavior, and adjust their reciprocity tendencies flexibly based on various conditions [25-32]. To further validate our reciprocity paradigm, we replicated the experimental protocol established by Kettler N et al. [34], by pairing rats with

partners possessing different prior social experiences. Our results confirmed that rats preferentially provided assistance to partners who had previously helped them (Fig. S5a–f online), aligning with findings from other studies[33, 34].

Based on our comprehensive data, we demonstrated that after a period of training, rats successfully transitioned from mutualistic to reciprocal behavior and reliably exhibited stable, direct reciprocity.

3.3 Social interactions are essential to reciprocity behavior

Social interactions, which transmit social information between individuals, play a critical role in establishing and maintaining social behaviors [35]. The importance of social interaction for cooperation has been noted in several previous studies, yet they frequently lack detailed analysis [20, 36, 37]. DeepLabCut, a powerful method based on deep neural networks, allowed us to precisely extract the positions of unmarked body parts of interacting animals with high spatial and temporal resolution [38]. Using this technology, we identified two types of social interactions between rats: "partition-crossing" (Fig. 1d–f) and "mutual approach behavior" (Fig. 1g–i). We observed that these social interactions occurred more frequently in the task execution area and were more prevalent in reciprocity tasks that required information exchange, compared to individual tasks, where a single poker could get reward without such information exchange required (Fig. S6a–h online). Furthermore, recipients engage in partition-crossing faster (Fig. 1e) and initiate more mutual approach behaviors than helpers (Fig. 1h), suggesting that these interactions might serve to convey the need for assistance. Additionally, a higher proportion of these social interactions within a trial correlated with faster completion of that trial (Fig. 1f, i), indicating that social interactions facilitate the execution of reciprocity tasks.

To further explore the relationship between social interactions and the execution of reciprocity tasks, we placed a transparent barrier between the two cages, specifically suppressed the aforementioned types of social interactions (Fig. 1j). We found that after adding the transparent barrier, the frequency of both types of social interactions significantly decreased, accompanied by a significant reduction in the number of reciprocity trials (Fig. 1k–l). These results above further support the notion that social interactions are essential to reciprocity behavior.

To investigate the interactions and mutual influences between two rats, we initially measured the distance between their noses (Fig. S7a online). We found that the execution time of reciprocity tasks is positively related to the distance between the noses of two rats, suggesting that closer social distances correspond to better execution of reciprocal behaviors (Fig. S7b, c online). For effective social interactions, maintaining attention towards a partner's actions is crucial [39]. According to previous

study[19],we defined the angle between the line extending from a rat's nose to head and the line connecting the rat's nose to its partner's nose as head orientation, which serves as an index of social attention (Fig. S7d online). Our results indicate that a low head orientation (high social attention) towards a partner corresponds to shorter execution time of task (Fig. S7e–f online), supporting that high social attention enhances the execution of reciprocity. Using Granger causality analysis, we further found that in successful reciprocity trials (SRT), the spatial positions of both rats influenced each other, indicating a mutual following behavior. Conversely, in failed reciprocity trials (FRT), the position of the helper influenced the recipient, but not vice versa (Fig. S8a–g online), suggesting that helpers might lose interest in the recipients, failing to follow them, which results in a failed assistant. Additionally, social interactions in SRT appeared to elicit social attention from the partner, a phenomenon absent in FRT (Fig. S8h–p online).

Taken together, out results support that close social distance and high social attention foster reciprocity, paralleling with the findings in prosocial behavior [19].

3.4 Free-riding behavior doesn't increase in the absence of second- and third-party punishment mechanisms

Free-riding behavior severely disrupts reciprocity interactions and affects the stability and sustainability of cooperative behaviors [40]. We then ask whether this is the scenario in rats. We define the trials in which the helpers failed to assist their partners as free-riding behavior, and found that free-riding also negatively impacted the reciprocity in our task (Fig. S9a, b online). In detail, our observations showed that rats were less likely to assist their partners after experiencing free-riding compared to when they were helped in the previous trial, in both co-housed and separately-housed conditions (Fig. S9a, b online).

Second- and third-party punishments have been proposed to play an important role in inhibiting free-riding and enhancing cooperation [10, 11]. However, these studies are predominantly from the results of human studies. To determine whether similar mechanisms operate in rats, we first examined free-riding behavior under co-housed condition, where second- and third-party punishment could theoretically occur. To better analyze this behavior, we employed two metrics to characterize free-riding. The first metric, the ratio of failed to all partner-trials, focuses on the global outcome and indexes no matter whether helper of recipient is responsible for the failure. The second metric, the ratio of those partner-trials in which helper rat even did not engage with the nose-poke at all to all partner-trials, highlights the motivation for helping, i.e., if helper rat never poked the nose port in a partner trial, it clearly indicates a lack of intention to assist the recipient. Our data show no significant difference in these metrics before and after training among co-housed rats undergoing reciprocity task (Fig. S9c, d online).

This suggests that free-riding did not become prevalent in co-housed condition, which permits rats to interact with their partners and third parties and potentially helps inhibit free-riding through second- and third-party punishments. To control for the potential influence of second- and third-party punishments on inhibiting free-riding behavior, we trained rats in a separately-housed condition, where second- and third-party punishments were theoretically excluded (i.e., the rats never encountered their training partner outside of training, and no third party was aware of the results of their reciprocity interactions). Our data show that the numbers of two metrics that represent free-riding didn't increase among reciprocity training in separately-housed rats (Fig. S9e, f online), who show similar behaviors and a comparable ability to perform the reciprocity task to those of co-housed rats (Fig. S10a–c online). Thus, our results indicate that even in the absence of the external mechanisms punishments, free-riding still did not become prevalent. This implies the existence of additional mechanisms that inhibit free-riding and promote the evolution of cooperation.

3.5 Reciprocity training enhances empathy in rats towards their partners

Empathy, often defined as the ability to perceive and share the emotional states of others, is fundamental to social behavior [41]. In rodents, empathy manifests in diverse behaviors, with emotional contagion being a classical form [42, 43]. Previous studies have demonstrated that rodents can empathize with the emotional states of their peers and initiate prosocial behaviors in response [44-46]. Moreover, higher levels of empathy correlate with increased prosocial behavior [47]. Empathy is considered as a driving force behind prosocial actions, offering both ultimate and proximate explanations for such behaviors [48]. While reciprocity is a form of bidirectional prosocial behavior [49], whether empathy plays a role in reciprocity remains unexplored. Thus, we hypothesize that empathy serves as additional mechanism for inhibiting free-riding behavior, thereby contributing to reciprocity.

To test this hypothesis, we first conducted a classic empathy test—the observational fear learning test [50] —on two groups of rats: one trained in reciprocity tasks and another in individual tasks—where activating the nose port alone yielded a water reward. In this test, one rat in each pair, the demonstrator (demo), received 8shocks to induce fear, while the observer (obs) did not receive shocks but could empathize with the demo's fear, thus experiencing fear itself. We quantified the rats' fear levels by measuring the duration of freezing. Rats trained in reciprocity tasks, both co- and separately-housed (Fig. 2a, b), exhibited significantly higher freezing levels when their partners were shocked compared to those trained in individual tasks (Fig. 2c, d), suggesting they have higher emotional contagion levels toward their partner. However, such an increase was not observed in rats trained with mutualism tasks (Fig. S11a–c online), ruling out the possibility that general cooperative (mutualism)

experience alone enhances emotional contagion.

To identify the factors that might lead to higher emotional contagion levels, we first measured the freezing levels of the demo rats during the test, and found no significant differences between the two groups (Fig. 2e). Social attention, as measured by the head orientation of the observers towards the demonstrators during empathy, also showed no significant differences (Fig. 2f). Additionally, ultrasonic vocalizations at 22 kHz, previously reported to be linked to fear transmission [51], revealed no significant differences between the two groups during the tests (Fig. 2g). Further analyses of rats' motor abilities, anxiety levels, and social preferences revealed no significant differences (Fig. S12a–c online), supporting that reciprocity training specifically enhances rats' emotional contagion levers towards their partners. Although the precise mechanisms remain unclear, these findings suggest that the social experiences associated with reciprocity may contribute to increased emotional contagion levels.

In combination with previous studies showed that higher levels of empathy correlating with increased prosocial behavior [47], our results suggest that enhanced empathy by reciprocity training may serve as the candidate mechanism for free-riding inhibition as mentioned above.

3.6 Oxytocin release in the OFC was observed during the reciprocity task

To explore the neural mechanisms by which training for reciprocity enhances rats' empathy levels towards their partners, our research focused on oxytocin. Oxytocin, a cyclic peptide composed of nine amino acids [52], was initially studied for its role in childbirth and lactation but has since been linked to a range of social behaviors including emotional recognition [53], helping behavior [54], consolation actions [55], and pair bonding behaviors [56]. Oxytocin has also been implicated in modulating empathy across species from mice and prairie voles to humans [57-59].

We previously demonstrated that orbitofrontal cortex (OFC) encodes cooperation [20] and here confirmed the expression of the oxytocin receptor (Oxtr) in this region (Fig. S13a online). To examine whether oxytocin is released in OFC during cooperation, we used a genetically encoded, fluorescent oxytocin sensor, namely $GRAB_{OXT1.0}$ [60], to optically record (Fig. 3a) the signal of oxytocin. We initially examined the reliability of the sensor through the tail suspension experiment, which elevates anxiety levels and triggers oxytocin release known for its anti-anxiety effects [61]. We observed a significant increase in the $GRAB_{OXT1.0}$ signal during tail suspension, confirming the reliability of our $GRAB_{OXT1.0}$ (Fig. S13a–d online).

As shown in Fig. 3b, we recorded the oxytocin signals in OFC when rats undergo

individual (left), mutualism (middle) and reciprocity (right) tasks. As a control, rats undergoing individual task (Fig. 3c, black) did not show a marked increase oxytocin signals in the periods from GoCue to PortExit. For rat undergoing mutualism tasks (Fig. 3c, blue), oxytocin signal slightly increased in the same period. However, during reciprocity tasks, oxytocin release was markedly elevated (Fig. 3c, red). To further compared the dynamics of oxytocin signals in rats of different tasks, we statistically analyzed the peak of the oxytocin signal from GoCue to PortEntry point, the peak (Fig. 3d) and average (Fig. 3e) of oxytocin signals were significantly higher in rats undergoing reciprocity tasks (red) compared to those in mutualism (blue) or individual (black) tasks. These findings indicate that oxytocin is released more robustly in the OFC during reciprocity tasks than during mutualism or individual tasks.

3.7 Oxytocin-deficient rats exhibit more free-riding behaviors

To further explore how training in reciprocity tasks enhances empathy and its association with oxytocin, we trained oxytocin-deficient (OXT-KO, homozygous) rats performing the same reciprocity task. OXT-KO rats learned the task more slowly than wild-type (WT) rats, but at the end of training, they achieved the comparable competency at a group level (Fig. S14a–c online). However, OXT-KO rats exhibited a higher incidence of free-riding behavior (Fig. 4a, b) and fewer reciprocity trials lasting over 15 s (Fig. S15a, b online) compared to WT rats. Moreover, they were more sensitive to free-riding, evidenced by a lower likelihood of reciprocating after experiencing free-riding (Fig. 4c). We excluded the possibility that the increased free-riding behavior in OXT-KO rats stems from their inability to synchronize actions within the 1-s cooperative time window (Fig. S15c online).

3.8 Reciprocity training fails to enhance empathy of OXT-KO rats

Our collective results above indicate that in OXT-KO rats, the mechanisms that typically inhibit free-riding behavior may not be fully effective due to the absence of oxytocin. Thus, we conducted empathy tests on OXT-KO rats after their training in reciprocity tasks. Unlike WT rats, OXT-KO rats did not exhibit an increase in emotional contagion levels towards their partners even after undergoing reciprocity training (Fig. 4d). To ensure that this outcome was not due to differences in task mastery, we further categorized OXT-KO rats into those who had learned the task and those who had not, and confirmed there was no difference in empathy between these two groups (Fig. S15d online).

Overall, these findings demonstrate that the enhancement of emotional contagion levels towards partners resulting from training in reciprocity tasks is dependent on the

presence of oxytocin. In OXT-KO rats, the absence of oxytocin led to no increase in emotional contagion levels towards their partners, contributing to more frequent free-riding behavior.

4 Discussion and conclusion

Reciprocity behavior is considered one of the fundamental drivers of the evolution of cooperative behavior. It fosters the development of social structures and complexity, facilitates efficient resource allocation and defense mechanisms, reduces internal competition, and resolves social conflicts. However, due to the significant time delay between the initiation of a helping action and its reciprocation, free-riding behavior can easily occur within reciprocity interactions. Free-riding, where one receives help but does not reciprocate, poses a serious threat to the occurrence and maintenance of reciprocity behavior. In humans, although various mechanisms can prevent the occurrence of free-riding [10, 11, 62] these mechanisms have limitations, for instance, these mechanisms often require third-party intervention and it is uncertain if they are applicable to animals. Currently, no third-party punishment has been observed in chimpanzees [18], our closest evolutionary relatives, yet cooperative behavior is widely present in chimpanzees and even more evolutionarily ancient animals. This suggests that other mechanisms must exist in animals to inhibit free-riding and maintain reciprocity behavior, thereby aiding the evolution of cooperation.

Compared to multiple paradigms of reciprocity have been established [25, 63], our paradigm represents the fully automated implementation of reciprocity in rats. In our work, we designed trial-by-trial and block-by-block reciprocity paradigms where rats consistently displayed stable direct reciprocity behavior (Fig. 1 and Figs. S2 and S3). We also found that social interactions positively facilitate the execution of reciprocity tasks, and isolation from social interactions reduces the efficiency of reciprocity behavior among rats (Fig. 1 d–n). Interestingly, compared to those performing tasks alone, WT rats trained in reciprocity tasks showed higher levels of empathy toward both co- and separately-housed partners, and this was not due to visual or auditory stimuli (Fig. 2). Through optical fiber recordings of oxytocin probes, we discovered that oxytocin is released the highest in the OFC brain area during reciprocity tasks compared to mutualism or individual tasks, demonstrating that oxytocin also participates in reciprocity tasks (Fig. 3). Furthermore, even with reciprocity training, OXT-KO rats did not show increased levels of empathy toward their partner (Fig. 4), and compared to WT rats, which showed enhanced empathy toward their peers, oxytocin-deficient rats exhibited higher levels of free-riding (Fig. 4).

Recent mouse studies have documented striking one-shot rescue behaviors: a free

helper vigorously licks and tongue-drags an unconscious cage-mate to accelerate recovery [64, 65], a manoeuvre acting through dedicated arousal circuits [66] and coordinated by parallel oxytocin pathways that couple empathic arousal to vigorous grooming [67]. Each paradigm captures a single, unidirectional act of aid that ends once the distressed partner revives; because the beneficiary is never required—or able—to reciprocate, these tasks sidestep the classic free-rider dilemma. Our reciprocity paradigm targets a fundamentally different social mechanism: mutualistic cooperation with delayed return. Two rats must synchronize nose-pokes within a 1 s window; only one drinks in each trial, and roles switch randomly across hundreds of trials. The helper therefore incurs an immediate cost, banking on the partner's willingness to return the favor later in the session. This structure introduces (i) repeated, bidirectional interactions, (ii) explicit temporal separation between giving and receiving, (iii) the strategic temptation to defect, and (iv) a requirement for partner-specific memory.

Mirror neurons, originally discovered in the premotor cortex of macaques, are defined as neurons that discharge both when an individual performs an action and when they observe the same action performed by another [68]. These neurons are proposed to underlie social cognition, empathy, and learning. In rodents, mirror-like responses have been reported in the anterior cingulate cortex for both self-experienced and observed pain [69] in the somatosensory pathways for contagious itch [70], and recently in the hypothalamus for aggressive behaviors, where neurons respond similarly to fighting and witnessing aggression [71]. Together, these findings suggest that rats possess a primitive but functional mirror neuron system that can process both negative affective states and social actions. While mirror neuron systems have been identified for aversive states such as pain, itch, and aggression in rodents, no direct evidence has yet demonstrated mirror-like responses for positive social experiences. Nevertheless, positive social reward clearly exists in animals: behaviors such as grooming and social touch activate dedicated affective pathways, and disruption of these pathways reduces social bonding and affiliative behaviors [72]. Investigating whether positive social states can be vicariously represented—analogous to mirroring—may shed light on the motivational mechanisms underlying reciprocity, especially in cases where helping behavior occurs without immediate personal gain.

Oxytocin has been extensively demonstrated in rodents to enhance social salience and affiliative motivation, facilitate both social recognition memory and associative learning, and promote cognitive flexibility. These domain-general functions may, in principle, support reciprocity by increasing attention to social partners, reinforcing memory of past cooperative acts, and supporting the behavioral flexibility required to alternate between helping and receiving roles across time. Resolving whether oxytocin facilitates reciprocity via general social functions or dedicated circuitry will require causal, circuit-level interventions targeting oxytocin pathways within regions mediating partner monitoring, memory, and behavioral updating, while carefully controlling for potential off-target effects on the broader social, cognitive, and

motivational processes previously implicated in oxytocin function. Despite this, empirical evidence points to its necessity: OXT-KO rats exhibit increased free-riding behavior and fail to display the reciprocity-induced enhancement in empathic responses observed in wild-type controls. These findings suggest that an intact oxytocin system is essential for the emergence and maintenance of reciprocity.

Integrating all our data, we propose that oxytocin-mediated enhancement of empathy reduces free-riding behavior among rats, making reciprocity behavior more stable and possibly sustaining the evolution of cooperation. This internally mediated direct drive complements the previously second- and third-party punishment mechanisms. All of these mechanisms work together to inhibit the occurrence of free-riding behavior, contributing to maintenance of reciprocity which promotes evolution of cooperation. In conclusion, our findings provide new perspectives to our fundamental understanding of the general mechanisms underlying evolution of cooperative behavior.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the STI2030-Major Projects (2022ZD0205100), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB32010300), and the Shanghai Municipal Science and Technology Major Project (2018SHZDZX05). We are grateful to the technical support from the optical imaging facility and animal facility of the Institute of Neuroscience, Center for Brain Science and Intelligence Technology, CAS.

Author contributions

Miaoyaoxin Wang and Zuoren Wang conceived this project, designed the experiments and wrote the paper. Miaoyaoxin Wang, Yukai Shao, Ao Fu, Ziye Zhang, and Jingze Xu performed the experiments. Miaoyaoxin Wang, and Qianqian Shi analyzed the data. Qingxiu Wang and Lei Wei provide the OXT-KO rats. Yanwang Huang performed experiment of Oxtr mRNA FISH. Yulong Li provided the fluorescent oxytocin probe. Mengping Jiang and Tianming Yang. provide guidance on data analysis.

Appendix A. Supplementary material

Supplementary data to this article can be found online.

Reference

[1] Axelrod R, Hamilton WD. The evolution of cooperation. Science 1981; 211: 1390-1396

[2] Nowak M, Highfield R. Supercooperators: Altruism, evolution, and why we need each other to succeed. Simon and Schuster, 2012

[3] Boyd R, Richerson PJ. Culture and the evolution of human cooperation. Philos Trans R Soc Lond B Biol Sci 2009; 364: 3281-3288

[4] Sterelny K. The Pleistocene social contract: Culture and cooperation in human evolution. Oxford University Press, 2021

[5] Pennisi E. How did cooperative behavior evolve? Science 2005; 309: 93-93

[6] Nowak MA. Five rules for the evolution of cooperation. Science 2006; 314: 1560-1563

[7] Clutton-Brock T. Cooperation between non-kin in animal societies. Nature 2009; 462: 51-57

[8] Trivers RL. The evolution of reciprocal altruism. Q Rev Biol 1971; 46: 35-37

[9] Taborsky M, Frommen JG, Riehl C. Correlated pay-offs are key to cooperation. Philos Trans R Soc Lond B Biol Sci 2016; 371: 20150084

[10] Fehr E, Gächter S. Cooperation and punishment in public goods experiments. Am Econ Rev 2000; 90: 980-994

[11] Fehr E, Gächter S. Altruistic punishment in humans. Nature 2002; 415: 137-140

[12] Fehr E, Fischbacher U. Third-party punishment and social norms. Evol Hum Behav 2004; 25: 63-87

[13] Zoh Y, Chang SWC, Crockett MJ. The prefrontal cortex and (uniquely) human cooperation: A comparative perspective. Neuropsychopharmacology 2022; 47: 119-133

[14] Jordan JJ, Hoffman M, Bloom P, et al. Third-party punishment as a costly signal of trustworthiness. Nature 2016; 530: 473-476

[15] Raihani NJ, Grutter AS, Bshary R. Punishers benefit from third-party punishment in fish. Science 2010; 327: 171-171

[16] Ratnieks FL, Wenseleers T. Policing insect societies. Science 2005; 307: 54-56

[17] Monnin T, Ratnieks FL, Jones GR, et al. Pretender punishment induced by chemical signalling in a queenless ant. Nature 2002; 419: 61-65

[18] Riedl K, Jensen K, Call J, et al. No third-party punishment in chimpanzees. Proc Natl Acad Sci USA, 2012, 109: 14824-14829

[19] Gachomba MJM, Esteve-Agraz J, Caref K, et al. Multimodal cues displayed by submissive rats promote prosocial choices by dominants. Curr Biol 2022; 32: 3288-3301.

[20] Jiang M, Wang M, Shi Q, et al. Evolution and neural representation of mammalian cooperative behavior. Cell Rep 2021; 37: 110029

[21] Zheng C, Wei L, Liu B, et al. Dorsal BNST Drd2^{+} neurons mediate sex-specific anxiety-like behavior induced by chronic social isolation. Cell Rep 2023; 42: 112799

[22] Krausz TA, Comrie AE, Kahn AE, et al. Dual credit assignment processes underlie dopamine signals in a complex spatial environment. Neuron 2023; 111: 3465-3478.

[23] Ben-Ami Bartal I. The complex affective and cognitive capacities of rats. Science 2024; 385: 1298-1305

[24] Schweinfurth MK. The social life of Norway rats (*Rattus norvegicus*). Elife 2020; 9: e54020

[25] Rutte C, Taborsky M. Generalized reciprocity in rats. PLoS Biol 2007; 5: e196

[26] Schweinfurth MK, Taborsky M. Reciprocal trading of different commodities in norway rats (*Rattus norvegicus*). Curr Biol 2018; 28: 594-599.

[27] Schneeberger K, Roder G, Taborsky M. The smell of hunger: norway rats (*Rattus norvegicus*) provision social partners based on odour cues of need. PLoS Biol 2020; 18: e3000628

[28] Dolivo V, Taborsky M, Herberstein M. Cooperation among norway rats (*Rattus norvegicus*): The importance of visual cues for reciprocal cooperation, and the role of coercion. Ethology 2015; 121: 1071-1080

[29] Gerber N, Schweinfurth MK, Taborsky M. The smell of cooperation: Rats increase helpful behaviour when receiving odour cues of a conspecific performing a cooperative task. Proc Biol Sci 2020; 287: 20202327

[30] Schmid R, Schneeberger K, Taborsky M. Feel good, do good? Disentangling reciprocity from unconditional prosociality. Ethology 2017; 123: 640-647

[31] Schweinfurth MK. Cooperative intentions and their implications on reciprocal cooperation in norway rats (*Rattus norvegicus*). Ethology 2021; 127: 865-871

[32] Dolivo V, Taborsky M. Cooperation among Norway rats (*Rattus norvegicus*): The importance of visual cues for reciprocal cooperation, and the role of coercion. Ethology 2015; 121: 1071-1080

[33] Rutte C, Taborsky M. The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. Behav Ecol Sociobiol 2007; 62: 499-505

[34] Kettler N, Schweinfurth MK, Taborsky M. Rats show direct reciprocity when interacting with multiple partners. Sci Rep 2021; 11: 3228

[35] Alexander RD. The evolution of social behavior. Annu Rev Ecol Syst 1974; 5: 325-383

[36] Paulsson NI, Taborsky M. Reaching out for inaccessible food is a potential begging signal in cooperating wild-type norway rats (*Rattus norvegicus*). Front Psychol 2021; 12: 712333

[37] Łopuch S, Popik P. Cooperative behavior of laboratory rats (*Rattus norvegicus*) in an instrumental task. J Comp Psychol 2011; 125: 250-253

[38] Mathis A, Mamidanna P, Cury KM, et al. DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 2018; 21: 1281-1289

[39] Ohtsubo Y, Tamada S. Social attention promotes partner intimacy. Lett Evol Behav Sci 2016; 7: 21-24

[40] Apicella CL, Silk JB. The evolution of human cooperation. Curr Biol 2019; 29: R447-R450

[41] De Waal FB, Preston SD. Mammalian empathy: behavioural manifestations and neural basis. Nat Rev Neurosci 2017; 18: 498-509

[42] Kim A, Keum S, Shin HS. Observational fear behavior in rodents as a model for empathy. Genes Brain Behav 2019; 18: e12521

[43] Preston SD, De Waal FB. Empathy: Its ultimate and proximate bases. Behav Brain Sci 2002; 25: 1-20

[44] Wu YE, Dang J, Kingsbury L, et al. Neural control of affiliative touch in prosocial interaction. Nature 2021; 599: 262-267

[45] Wu YE, Hong W. Neural basis of prosocial behavior. Trends Neurosci 2022; 45: 749-762

[46] Jeon D, Kim S, Chetana M, et al. Observational fear learning involves affective pain system and Cav1.2 Ca^{2+} channels in ACC. Nat Neurosci 2010; 13: 482-488

[47] Scheggia D, La Greca F, Maltese F, et al. Reciprocal cortico-amygdala connections regulate prosocial and selfish choices in mice. Nat Neurosci 2022; 25: 1505-1518

[48] Ben-Ami Bartal I, Decety J, Mason P. Empathy and pro-social behavior in rats. Science 2011; 334: 1427-1430

[49] Erreygers S, Vandebosch H, Vranjes I, et al. Positive or negative spirals of online behavior? Exploring reciprocal associations between being the actor and the recipient of prosocial and antisocial behavior online. New Media Soc 2017; 20: 3437-3456

[50] Kim SW, Kim M, Shin HS. Affective empathy and prosocial behavior in rodents. Curr Opin Neurobiol 2021; 68: 181-189

[51] Bonauto SM, Greuel OM, Honeycutt JA. Playback of rat 22-kHz ultrasonic vocalizations as a translational assay of negative affective states: an analysis of evoked behavior and brain activity. Neurosci Biobehav Rev 2023; 153: 105-118

[52] Froemke RC, Young LJ. Oxytocin, neural plasticity, and social behavior. Annu Rev Neurosci 2021; 44: 359-381

[53] Ferretti V, Maltese F, Contarini G, et al. Oxytocin signaling in the central amygdala modulates emotion discrimination in mice. Curr Biol 2019; 29: 1938-1953.

[54] Yamagishi A, Lee J, Sato N. Oxytocin in the anterior cingulate cortex is involved in helping behaviour. Behav Brain Res 2020; 393: 112814

[55] Burkett JP, Andari E, Johnson ZV, et al. Oxytocin-dependent consolation behavior in rodents. Science 2016; 351: 375-378

[56] Young LJ, Wang Z. The neurobiology of pair bonding. Nat Neurosci 2004; 7: 1048-1054

[57] Pisansky MT, Hanson LR, Gottesman II, et al. Oxytocin enhances observational fear in mice. Nat Commun 2017; 8: 2102

[58] Kitano K, Yamagishi A, Horie K, et al. Helping behavior in prairie voles: A model of empathy and the importance of oxytocin. iScience 2022; 25: 103991

[59] Barchi-Ferreira AM, Osório FL. Associations between oxytocin and empathy in humans: A systematic literature review. Psychoneuroendocrinology 2021; 129: 105268

[60] Qian T, Wang H, Wang P, et al. A genetically encoded sensor measures temporal oxytocin release from different neuronal compartments. Nat Biotechnol 2023; 41: 944-957

[61] Yoon S, Kim YK. The role of the oxytocin system in anxiety disorders. Adv Exp Med Biol 2020; 1191: 103-120

[62] Fehr E, Fischbacher U, Gächter S. Strong reciprocity, human cooperation, and the enforcement of social norms. Hum Nat 2002; 13: 1-25

[63] Festucci F, Buccheri C, Cerniglia L, et al. A new paradigm for prosocial behavior and reciprocity, assessed in wt and het rats for the DAT gene. Behav Brain Res 2020; 393: 112746

[64] Sun F, Wu YE, Hong W. A neural basis for prosocial behavior toward unresponsive individuals. Science 2025; 387: eadq2679

[65] Sun W, Zhang G-W, Huang JJ, et al. Reviving-like prosocial behavior in response to unconscious or dead conspecifics in rodents. Science 2025; 387: eadq2677

[66] Cao P, Liu Y, Ni Z, et al. Rescue-like behavior in a bystander mouse toward anesthetized conspecifics promotes arousal via a tongue-brain connection. Sci Adv 2025; 11: eadq3874

[67] Zhang F-R, Liu J, Wen J, et al. Distinct oxytocin signaling pathways synergistically mediate rescue-like behavior in mice. Proc Natl Acad Sci USA 2025; 122: e2423374122

[68] Rizzolatti G, Fadiga L, Gallese V, et al. Premotor cortex and the recognition of motor actions. Cogn Brain Res 1996; 3: 131-141

[69] Carrillo M, Han Y, Migliorati F, et al. Emotional mirror neurons in the rat's anterior cingulate cortex. Curr Biol 2019; 29: 1301-1312.

[70] Yu Y-Q, Barry DM, Hao Y, et al. Molecular and neural basis of contagious itch behavior in mice. Science 2017; 355: 1072-1076

[71] Yang T, Bayless DW, Wei Y, et al. Hypothalamic neurons that mirror aggression. Cell 2023; 186: 1195-1211.

[72] Liu B, Qiao L, Liu K, et al. Molecular and neural basis of pleasant touch sensation. Science 2022; 376: 483-491

Figures

Fig. 1 Social interactions promote reciprocity behavior.

(a) Schematic diagram of mutualism and reciprocity task.

(b) Between Day 1 and Day 14, there was a significant increase in the number of trials that rats (n =16) refrained from entering the water port after successful partner-reward trials (Wilcoxon test, P <0.0001).

(c) Between Day 1 and Day 14, a comparison of drinking duration yielded no significant difference (n = 16) (Wilcoxon test, P = 0.124).

(d) Example of partition-crossing behavior.

(e) Comparison the latency of first-time partition-crossing (Wilcoxon test, P = 0.025) between the recipient (n = 16) and helper (n = 16) in SRT.

(f) Correlation between trial duration (Y axis) and the ratio of partition-crossing behavior time to trial duration (X axis) for helper (left panel, P = 0.0017) and recipient (right panel, P <0.0001) in SRT. Data were fitted using linear regression.

(g) Example of mutual approach behavior.

(h) Comparison the number of initial approach behaviors between the helper (n = 16) and recipient (n = 16) in SRT (Wilcoxon test, P = 0.0315).

(i) Correlation between trial duration (Y axis) and the ratio of mutual approach behavior time to trial duration (X axis) in SRT (P <0.0001). Data were fitted using linear regression.

(j) Task with a transparent barrier added into the middle of two cages.

(k) Comparison of the number of mutual approach (n = 16) (Paired Student's t-test, P <0.001) under conditions with and without the transparent barrier.

(l) Comparison of the number of partition-crossing behaviors (n = 16) (Paired Student's t-test, P <0.0001) under conditions with and without the transparent barrier.

(m) Comparison of the number of trial that rats (n = 16) help partner under conditions with and without the transparent barrier (Paired Student's t-test, P = 0.0009). Data are shown as mean ± SEM * P <0.05; *** P <0.001; **** P <0.0001. ns., no significant difference.Abbreviations: #, number.

Fig. 2 Rats after reciprocity training show higher emotional contagion toward both co-housed and separately-housed partners.

(a) Schematic diagram of training co-housed (left) and separately-housed (right) rats to perform either reciprocity or individual tasks.

(b) Schematic diagram of emotional contagion test. The emotional contagion level is indexed by the immobility time of the observer.

(c) Comparison of immobility time of observer rats from reciprocity task (RT) (n = 8) and individual task (IT) (n = 8) groups (co-housed with demonstrator) (two-way RM ANOVA, training task (RT vs. IT) × min interaction effects: F7,248 = 1.928, P = 0.0581; min effects: F7,248 = 9.726, P <0.0001; training task effects: F1,248 = 0.1271, P = 0.0005).

(d) Comparison of immobility time of observer rats from RT (n = 8) and IT (n = 8) groups (separately-housed with demonstrator) (two-way RM ANOVA, training task (RT vs. IT) × min interaction effects: F7,96 = 0.5381, P = 0.8036; min effects: F7,96 = 9.361, P <0.0001; training task: F1,96 = 5.067, P = 0.0267).

(e) Comparison of immobility time of demonstrator rats from RT (n = 8) and IT (n = 8) groups (co-housed with observer) (two-way RM ANOVA, training task (RT vs. IT) × min interaction effects: F7,248 = 05372, P = 0.8059; min effects: F7,248 = 45.18, P < 0.0001; training task effect: F1,248 = 0.3783, P = 0.5391).

(f) Comparison of head orientation of observer rats from RT (n = 8) and IT (n = 8) groups toward their partner during emotional contagion task.

(g) Comparison of ultrasound (22 kHz) emitted from demonstrator from RT (n = 8) and IT (n = 8) groups during emotional contagion test. Data are shown as mean ± SEM * P <0.05; *** P <0.001. ns, no significant difference.

Fig. 3 Oxytocin released in OFC specifically during reciprocity task.

(a) Schematic diagram of virus injection.

(b) Diagram depicting the protocol of fiber photometry recording experiment.

(c) $GPAB_{OXT1.0}$ signals in OFC of rats performing reciprocity task (n = 19), mutualism task (n = 16) and individual task.

(d-e) Comparison of Peak and Average Z-scored $\Delta F/F_0$ of $GPAB_{OXT1.0}$ signals in OFC of rats (n = 19) in reciprocity, mutualism (n = 16) and individual task. (d) Peak Z-scored $\Delta F/F_0$ in reciprocity task is significantly higher than other tasks. No significant change found in mutualism task between individual task (one-way ANOVA dissecting the interaction [F (3,101) = 14.48, P <0.0001)]. (e) Mean Z-scored $\Delta F/F_0$ in reciprocity task is significantly higher than other tasks. No significant change found in mutualism task between individual task (one-way ANOVA dissecting the interaction [F (3,101) = 12.23, P <0.0001)]. Data are shown as mean ± SEM * P < 0.05, ** P <0.01; ****P <0.0001. ns, no significant difference.

Fig. 4 Oxytocin-deficient (OXT-KO) rats show higher free-riding, and their empathy level failed to be enhanced by reciprocity training.

(a) Ratio of failed partner-trials to total partner-trials of WT (n = 16) and OXT-KO (n = 16) rats (two-way RM ANOVA, groups (WT vs. OXT-KO) × min interaction effects: F13,420 = 1.302, P = 0.2084; min effects: F13,420 = 4.356, P <0.0001; groups effects: F1,420 =23.82, P <0.0001).

(b) Ratio of failed partner-trials in which helper never engage in nose port to total partner-trials of WT (n = 16) and OXT-KO (n = 16) rats. (two-way RM ANOVA, groups (WT vs. OXT-KO) × min interaction effects: F13,420 = 0.9441, P = 0.5069; min effects: F13,420 = 4.173, P <0.0001; groups effects: F1,420 =18.16, P <0.0001).

(c) Probability of WT (n = 16) and OXT-KO (n = 16) rats to assist partner after being betrayed in the last trial. (two-way RM ANOVA, groups (WT vs. OXT-KO) × min interaction effects: F13,297 = 1.826, P = 0.0387; min effects: F13,297 =2.401, P = 0.0044; groups effects: F1,297 =18.90, P <0.0001).

(d) Comparison of immobility time of OXT-KO observer rats from RT (n = 8) and IT (n = 8) groups (co-housed with demonstrator) (two-way RM ANOVA, training task (RT vs. IT) × min interaction effects: F7,304 = 0.09953, P = 0.998; min effects: F3,04

= 9.744, P <0.0001; training task effects: F1,304 = 0.05798, P = 0.8099). Data are shown as mean ± SEM *** P <0.001; ****P <0.0001. ns, no significant difference.