Oxytocin modulates the neurocomputational mechanisms engaged in learning social hierarchy Authors

Jiawei Liu, Chen Qu, Rémi Phillipe, Siying Li, Edmund Derrington, Brice Corgnet, Jean-Claude Dreher

Oxytocin modulates the neurocomputational mechanisms engaged in learning social hierarchy

Authors

Jiawei Liu[1,2,3]†, Chen Qu[3]†, Rémi Phillipe [1,2], Siying Li[3,4], Edmund Derrington[1,2], Brice Corgnet[5], Jean-Claude Dreher[1,2]*

[1] Laboratory of Neuroeconomics, Institut des Sciences Cognitives Marc Jeannerod, CNRS, Lyon, France.

[2] Université Claude Bernard Lyon 1, Lyon, France.

[3] Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, China; School of Psychology, Center for Studies of Psychological Application, and Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, China.

[4] Faculty of Education, Northeast Normal University, Changchun, China.

[5] EM Lyon Business School, and CNRS, GATE Lab, UMR 5824, Ecully, 69130, France.

† These authors contributed equally to this work.

* Lead contact, Correspondence: dreherjeanclaude@gmail.com

Abstract

It is unknown whether oxytocin modulates the neurocomputational mechanisms engaged in learning social ranks and in making transitive inferences based on social memory. In a new fMRI study combined with computational modeling, participants learned two non-linear social hierarchies by trial and error, one in which they were embedded and the other not. During a subsequent test phase, they inferred hierarchical relationships between face pairs never encountered during training. During training, oxytocin modulated computations of the action selection policy from a Reinforcement Learning-ELO model and this effect was associated with increased activity of the medial prefrontal cortex. During test phase, oxytocin modulated the balance between self-other social memory, boosting correct performance for the other-related hierarchy and increasing amygdala response. These findings demonstrate that oxytocin modulates the brain computations underlying both choice selection related to social ranks and social memory needed to make transitive inferences.

Highlight sentences

- Oxytocin modulates brain computations underlying choice selection of social ranks
- OXT increases the activity of the ventromedial prefrontal cortex and the posterior cingulate cortex, which compute the action selection policy at the time of choice.
- OXT modulates the memory of social hierarchies differently depending on whether one is embedded in the hierarchy or not.
- This effect occurs both the behavioral and the brain system levels in the amygdala and parahippocampal regions

Keywords

Reinforcement learning; model-based fMRI; Social hierarchy; Self-other distinction; Oxytocin.

1. Introduction

Dominance hierarchies, characterized by ranking systems based on the relative social power of individuals within a group, are a major evolutionary force that drive dyadic asymmetries. Dominance hierarchies may be learned incrementally, by accumulating feedback on relationships to specific members of a social group. The learning of hierarchical status can be framed as a reinforcement learning (RL) problem, inspired by neurocomputational approaches traditionally applied to non-social cognition, both when learning dominance relationships by observation of other interacting individuals (Kumaran et al., 2016), or by direct dyadic competitive interactions between oneself and another individual (Ligneul et al., 2016; Qu et al., 2017). The medial prefrontal cortex (mPFC) and the amygdala/hippocampus may play distinct computational roles during and after learning 'linear' ($r_1<r_2<\ldots<r_k$) social hierarchies. During learning, the mPFC (both dorsal and ventral parts) has been shown to compute the monitoring and updating of dominance status. These are fundamental signals needed to make choices between two individuals when learning their social dominance relationship by trial and error (Janet et al., 2022; Kumaran et al., 2016; Ligneul et al., 2016). In contrast, when required to infer the relationship between two non-adjacent members in a linear hierarchy, while only adjacent relationships have been learned, the amygdala-hippocampal complex is necessary to retrieve learned relationships between other members of the social structure from social memory (Kumaran et al., 2012, 2016).

Previous studies have considered social hierarchies as a series of relationships whereby individuals are ranked successively in a linear fashion (Jensen et al., 2019; Kumaran et al., 2012, 2016). However, social hierarchy in real life is often not a simple linear structure. Real hierarchies often include individuals of equal social ranks. The first question we addressed is how such non-linear 'diamond-shaped' hierarchies might be learned and represented in the brain (Fig. 1). We characterize the computations and brain systems engaged when making transitive inferences, after learning a diamond-shaped hierarchy by repeated learning trials between pairs of people with feedback. These transitive inferences were performed in the absence of corrective feedback, while participants viewed pairs not encountered during training. Such transitive inferences require social memory of the previously learned paired members, to infer the relative ranks of members, based on partial information provided during the learning trials.

Oxytocin (OXT) is an evolutionarily conserved neurohypophysial hormone, engaged in lactation and parenting. The function of this hormone might have been sculpted during evolution to facilitate the expression of adaptive social behaviors, as required for forming hierarchical relationships (Lee et al., 2019). There is evidence in non-human primates and rodents that OXT modulates social hierarchy. For example, OXT suppresses vigilance towards faces of dominant males in the rhesus macaque (Chang & Platt, 2014; Ebitz et al., 2013) and relaxes social interactions between monkeys, thereby flattening the social hierarchy (Jiang & Platt, 2018). In mice, socially dominant males have higher OXT receptor binding compared to subordinates, in brain regions such as the amygdala (Lee et al., 2019). This effect may reflect that the different social environments experienced by dominant and subordinate animals may alter receptor expression, and potentially facilitate adaptive social behaviors.

Here, we focused on the neurocomputational processes that take place while learning non-linear (diamond-shaped) social hierarchies by observation. We used model-based functional magnetic resonance imaging (fMRI), and an extension of a two-phase social hierarchy learning paradigm. Participants learned hierarchical ranks by trial and error during a training phase (Fig. 1). Then, during a test phase, they made transitive inferences of relative hierarchical rank between pairs of individuals who were never encountered before. We developed a new reinforcement learning model to account for social hierarchy

learning when it is necessary to learn both equal ranks and different ranks and investigated how OXT administration might impinge on the model parameters.

We tested whether OXT plays a specific role in either phase of our new social hierarchy learning paradigm while: (a) learning social ranks by reinforcement during the training phase and (b) encoding and retrieval of social memories, required to make transitive inferences between learned members of the hierarchical structure during the test phase. We also compared how OXT administration might affect these processes with respect to self- and other-orientated hierarchies (i.e., when oneself is included or not in the hierarchy) (Fig. 1C).

During training, we hypothesized that learning non-linear social hierarchies engaged computational mechanisms and brain representations similar to those engaged in learning one-dimensional linear social hierarchies by observation. Specifically, we hypothesized that the mPFC, engaged in learning linear social hierarchy (Janet et al., 2022; Kumaran et al., 2016; Ligneul et al., 2016), will also be engaged when learning diamond-shaped social hierarchies. Preliminary evidence suggests that OXT may modulate distinct neurocomputational mechanisms engaged in reinforcement learning, including selection between different options during choice behavior (Piva & Chang, 2018) or prediction error, as previously suggested in the domain of prosocial behavior (Clark-Elford et al., 2014; Hu et al., 2015; Hung et al., 2017; Hurlemann et al., 2010; Ide et al., 2018; Kruppa et al., 2019; Liao et al., 2020; Martins et al., 2022; Zhuang et al., 2021). Therefore, we hypothesized that OXT could modulate both action selection policy, reflecting the probability of choosing the correct rank of a given individual and updating of social ranks at the time of outcome.

During the test phase of the task, social memory processes are engaged when making transitive inferences between members of the learned structure (pairs of individuals who were never encountered together before). Previous studies in mice indicate that OXT regulates social memory (Baram et al., 2021; Ferguson et al., 2000; H. J. Lee et al., 2008; MacBeth et al., 2009; Raam et al., 2017; Wirth et al., 2021). We thus hypothesized that OXT affects social memory also in humans and could modulate responses of the amygdala/hippocampal region, which are known to be engaged when making transitive inferences between items (Kumaran et al., 2016). Finally, we tested whether OXT

modulates social memory and its neural correlates differently for social hierarchy structures in which participants were embedded or not (Self vs Other-orientated hierarchy). This is because well-functioning social interactions may require differential other-related and self-related information processing (Bang et al., 2022; Ereira et al., 2020; Wittmann et al., 2016) and because OXT differently modulates value computation of self–other allocations (Chang, Barter, et al., 2012; Chang & Platt, 2014; Liao et al., 2020) and in-group/out-group decisions (De Dreu et al., 2010).

2. Results

2.1 Behavioral results

We fitted two generalized linear mixed-effects models (GLMM) to analyze how performance of learning social ranks during the training phase (GLMM1) or making transitive inferences during the test phase (GLMM2) varied according to the Status of the members (Superior, Intermediate and Inferior), Hierarchy Orientation (Self: participant at the center of the hierarchy, labelled 'Self' *versus* Other: another participant at the center of the hierarchy, labelled 'Other'), 9 Blocks, Group (Oxytocin *versus* Placebo) and pertinent interaction factors.

During the training trials, participants performed better on trials comparing hierarchy members with Superior (odds ratio = 3.392, $p < 0.001$) and Inferior (odds ratio = 3.156, $p < 0.001$) status, compared to trials involving hierarchy members of equal status (Fig. 2A). A *post-hoc* test showed that participants also performed better on trials involving Superior than Inferior hierarchy members (odds ratio = 1.331, SE = 0.06, $p < 0.001$, $z = 6.316$). As the number of blocks increased, participants improved their performance (odds ratio = 1.433, $p < 0.001$). An interaction effect of Group × Block further showed that participants in the Placebo group were more likely than those in the Oxytocin group to improve performance over consecutive blocks (odds ratio = 0.966, $p = 0.042$, Fig. 2B). We also found participants learned better in the Self- than Other-orientated hierarchy (odds ratio = 1.237, $p < 0.001$, Fig. 2C). There was a significant interaction effect of Hierarchy Orientation (Self/Other) x Status (Superior: odds ratio = 0.543, $p < 0.001$; Inferior: odds ratio = 0.616, $p < 0.001$). A *post-hoc* test revealed that participants performed better in

Self-orientated Hierarchy than the Other-orientated Hierarchy condition only for trials involving Intermediate status (odds ratio = 1.783, SE = 0.183, $p < 0.001$, $z = 5.646$). With regard to results of response time (RT) analyzed by LMM1 during the training trials, RTs were longer for trials involving both Superior ($b = 0.032$, $p = 0.008$) and Inferior ($b = 0.128$, $p < 0.001$) hierarchy elements than trials involving the elements of Intermediate status (Fig. 2A, right). A *post-hoc* test further revealed that RTs was shorter for trials involving Superior elements than Inferior ($b = -0.096$, SE = 0.009, $p < 0.001$, $z = -10.529$). These results indicate that participants' RTs are consistent with a speed-accuracy tradeoff. Moreover, RTs gradually reduced over consecutive trial blocks, indicating faster responses as the hierarchies were learned better ($b = -0.045$, $p < 0.001$, Fig. S1A). The Oxytocin group also showed quicker reduction of RTs than the Placebo group over blocks ($b = -0.022$, $p < 0.001$, Fig. S1A). In addition, RTs reduction across blocks was faster in the Self than the Other-orientated condition ($b = -0.007$, $p = 0.034$, Fig. S1B), suggesting people learn faster when they are included in the hierarchy.

In test trials, participants improved performance over consecutive blocks (GLMM2, odds ratio = 1.373, $p < 0.001$). The Oxytocin group improved accuracy faster across blocks than the Placebo group (odds ratio = 1.04, $p = 0.023$, Fig. 2D). The interaction effect of Group × Hierarchy Orientation showed that the Oxytocin group outperformed the Placebo group in the Self-orientated hierarchy compared to Other-orientated hierarchy (odds ratio = 1.262, $p = 0.005$, Fig. 2E). Analysis of RT by LMM2 revealed that participants were slower in the Self-orientated hierarchy condition compared to Other-orientated hierarchy ($b = 0.09$, $p < 0.001$). Moreover, participants decreased their RT faster across blocks in the Self-orientated hierarchy compared to the Other-orientated condition ($b = -0.014$, $p < 0.001$). The Oxytocin group also decreased RT faster over blocks than the Placebo group ($b = -0.015$, $p < 0.001$). See supplementary Table S1 & S2 for complete behavioral results.

2.2 Computational model results

We tested 5 models accounting for how learning the item values in the hierarchy structure. Model 1 was the RL-ELO model (Model 1) which has been shown to be successful to account for learning linear hierarchies (Kumaran et al., 2016). Model 2 was a modified RL-ELO model with two different learning rates for hierarchically superior and inferior members of the hierarchy. Model 3 was the classical Rescorla-Wagner (RW) model and

model 4 was the Value Transfer model, which are known to fail in learning linear hierarchies as they associate equally with positive and negative outcomes (Kumaran et al., 2016). We also developed a new RL-ELO model, called the Combined RL-ELO-based associative learning model (Model 5), which enables to learn hierarchical structures involving equal ranks (i.e., non-linear 'diamond shape'). Because learning equal ranks may involve different strategies than learning unequal ranks, we assumed that learning equal ranks was equivalent to memorizing matching pairs. For this reason, we propose a model in which learning occurs in two steps. First, an associative learning rule is updated according to the fact that ranks are equal or unequal. Second, the value of each item in the presented pair is updated according to the past choice and reward. That is, if ranks between items are unequal, the value of the highest (or the lowest) item in the hierarchy is increased (or decreased respectively). If items in the pair were equal, their values became closer according to a proportion of their initial distance (See also computational modeling section in methods). All models were fitted separately for choices by each individual using Matlab VBA-toolbox (Daunizeau et al., 2014). Protected exceedance probability (PXP) was used for model selection. It accounts for the probability that the considered model is more frequent in the studied population, which can be used as a quantitative measure of the amount of evidence for the best model (Daunizeau et al., 2014). The Self- and the Other-orientated conditions were fitted separately to allow participants to learn differently between conditions.

Comparisons between models showed that the best model was the RL-ELO-based associative learning model (Model 5: PXP = 1; Fig. 3D), with a single alpha learning rate to update item values for each of the two hierarchy orientations (Self and Other). The fit of this model with the performance results are plotted over blocks for correctly selected items during training (Fig. 3A) and test blocks (Fig. 3B). The mean evolution of each item power ($Q_{(i)t}$) across all subjects is shown in Fig. 3C. Finally, we ran a 2-way mixed ANOVA analysis on the parameters of the RL-ELO-based associative learning model according to hierarchy orientation (Self/Other: within-subject factor) and treatment group (Oxytocin/Placebo: between-subject factor) (See supplementary Table S7). We observed significant effects only for the learning rate α_{eq} used to update choices of members of equal rank. That is, for this learning rate α_{eq}, we observed a main effect of hierarchy orientation ($F_{(1,56)} = 7.42$, $p = 0.009$, $\eta^2 = 0.04$; Self: $Mean_{(\alpha_{eq})} = 0.10$, Other: $Mean_{(\alpha_{eq})}$=

0.18; Fig. 3E) and a marginally significant Group effect ($F_{(1,56)} = 3.53$, $p = 0.065$, $\eta^2 = 0.003$; OXT: $Mean_{(\alpha_{eq})} = 0.10$, Placebo: $Mean_{(\alpha_{eq})} = 0.18$), but no interaction effect ($F_{(1,56)} = 0.02$, $p = 0.89$, $\eta^2 = 0.001$).

2.3 Brain regions engaged in learning social hierarchy

As the Combined RL-ELO-based associative learning model can predict the participants' behavior during social hierarchy learning, we fitted parametric trial-by-trial regressors of the model-estimated probability that a participant chooses the correct response as a proxy of the knowledge of the hierarchical structure (GLM1). This probability of choosing the correct answer was added as a parametric regressor at the time of presentation of the two faces. When collapsing over the Placebo and Oxytocin groups, a whole brain analysis indicated that the ventromedial prefrontal cortex (vmPFC), posterior cingulate cortex (PCC) as well as the temporal gyrus, parietal lobule and occipital cortex encoded the probability of choosing the correct answer during training (Fig. 4A; also see Supplementary Table S4 for details of activated regions). Subsequently, we carried out region-of-interest (ROI) analyses in vmPFC and PCC, based on independent coordinates from a meta-analysis related to social learning and self-referential processing (Bartra et al., 2013). Beta estimate values were extracted from the interaction between hierarchical orientation (Self/Others) and Group treatment during the learning trials, and a follow-up mixed ANOVA was performed. Results showed a significantly stronger neural response in the Oxytocin than the Placebo group in both ROIs (vmPFC: $F_{(1,56)} = 4.06$, $p = 0.0488$, $\eta^2 = 0.033$, Fig. 4B; PCC: $F_{(1,56)} = 4.66$, $p = 0.035$, $\eta^2 = 0.023$, Fig. 4C).

In addition, the prediction error (PE), which reflects the difference between expected and actual rewards, was encoded at the outcome feedback in training trials by robust activity in the vmPFC and the ventral putamen bilaterally (GLM2, Fig. 5A). However, none of these regions showed significant differences between Oxytocin and the Placebo groups (illustrated in Fig. 5B, 5C & 5D).

2.4 Brain regions engaged in social memory when making transitive inferences

During the test phase, a whole brain fMRI data analysis (GLM3, using the probability of choosing the correct response computed by the RL-ELO associative model) showed

engagement of a brain network including the mPFC, the anterior and posterior cingulate cortex and the temporal cortex (Fig. 6A). Moreover, comparison in the test phase between hierarchy orientation (Other > Self) revealed stronger activity in the dorsal anterior cingulate cortex (dACC, Fig. 6B & 6C), the right inferior parietal lobule (IPL) and in the precentral gyrus. We also examined putative differences between groups in vmPFC and PCC ROIs. However, no significant difference between placebo and Oxytocin groups were observed (vmPFC: $t = 2$, $df = 56$, $p = 0.1$; PCC: $t = -1$, $df = 56$, $p = 0.3$).

Finally, because the behavioral analyses indicated an interaction effect between Group and Hierarchy Orientation during the test phase (Fig. 2E), we investigated the brain correlates of this behavioral effect (GLM4). This analysis revealed an interaction with the same pattern between Group and Hierarchy Orientation (Fig. 7A), observed in the left amygdala (Fig. 7B) and in the right parahippocampal region (Fig. 7C). BOLD response in these regions was stronger during the Self- than Other-orientated hierarchy trials in the Placebo group, whereas in the Oxytocin group activity was stronger for the Other- than for the Self-orientated hierarchy.

3. Discussion

The present study investigated how intranasal OXT modulates the neurocomputational mechanisms subserving learning and remembering diamond-shaped social hierarchies in which participants were or not embedded. Our results provide mechanistic insights into how OXT modulates the neurocomputational signals determining action selection policy during learning and social memory recall when making transitive inferences in the test phase. OXT increased activity of the vmPFC and PCC related to action selection policy, reflecting the probability of identifying the correct rank relationship between two individuals (i.e., equal or unequal ranks in the non-linear 'diamond shape' structure). During the test phase, participants had to decide, by making transitive inferences, whether two items were of equal or unequal ranks. In this test phase, OXT led to improved performance for the Other-orientated hierarchy, and reduced performance for the Self-orientated hierarchy. This interaction between group (OXT, Placebo) and Self/Other-hierarchy orientation was associated with higher activity in the amygdala and the parahippocampal region. Together, these findings characterize how OXT modulates

neurocomputational signals engaged in learning social hierarchies and in remembering learned social relationships.

Social dominance hierarchies are a common form of social organization that requires individuals to learn and remember their status relative to others. Once relative status in a given structure has been learned, individuals maintain consistent patterns of social behavior (e.g. competitive or submissive) based on their past social interactions (Chase & Seitz, 2011; Williamson et al., 2016). Unlike previous studies on linear social hierarchy structure (Kumaran et al., 2012; Vasconcelos, 2008), we developed a diamond-shaped social hierarchy task in which relative ranks could be equal. Our combined RL-ELO-based associative learning model captured the learning of such hierarchies. This model considers learning social hierarchy as an incremental process accumulating feedback concerning members of the social group, in which one may or may not be embedded. Our RL-ELO-based model reflects dynamic learning of social relationships. This was done in a two-stage process. Participants updated the associative value representing the learned association between the pairs and the action of choosing equality between members, given past exposure to this pair. Then, participants updated their representations of the hierarchy structure by correcting the value of each item of the pair according to past choices and rewards. This model allowed us to identify the medial PFC and PCC as implementing the action selection policy. Moreover, the PE engaged the ventral putamen and the vmPFC during the feedback outcome, extending the classical PE signal to learning of social relationships (30, 38, 39). Recent computational models and experimental studies have proposed that the same brain representations that map space can be extended to a broad range of non-spatial problems in abstract cognitive space (Kurth-Nelson et al., 2015). These studies support that the mPFC and the hippocampus are involved in non-spatial relational memory tasks, allowing them to make transitive inferences (Boorman et al., 2021; Dayan, 1993; Sharpe et al., 2017) and to organize the relationships between entities (e.g., places, objects, people, events, or abstract task states) to afford generalization and flexible decision making. Our results extend these findings and explain how abstract structural relationships such diamond-shaped social hierarchies are learned at the computational level.

Here, we characterize the effect of OXT in dynamic learning of social hierarchy at the neurocomputational level. Our findings demonstrate distinct computational roles of OXT during the training and test phases of the task. During training, OXT modulates the

neurocomputational signals needed for implementing an action policy to choose the correct answer. OXT did not modulate learning rate *per se*, nor PE-related brain activity, but modulated the brain regions computing a choice policy during training. There is mixed evidence for a role of OXT on reinforcement learning in humans. Some studies report that OXT may facilitate reinforcement learning (Hu et al., 2015; Hurlemann et al., 2010; Zhuang et al., 2021) and modulate computational signals underlying prosocial reinforcement learning (Martins et al., 2022). However, other studies report that OXT reduces the encoding of PE during in an iterative trust game (Ide et al., 2018) and during social reward learning tasks (Clark-Elford et al., 2014; Kruppa et al., 2019). These mixed findings are consistent with the fact that the social effects of OXT are highly dependent on the behavioral context (Bartz et al., 2011).

It has been difficult to separate learning from choice selection processes because value representations drive choosing, as expected values of different available options are compared, and the best is selected. However, they also drive learning, as the expected value of an outcome is compared to the reward that is actually received, and future expectations are updated. Yet, value-based learning and value-based choosing may depend on separate neural representations (Miller et al., 2022; Rudebeck et al., 2017). Our findings demonstrate that OXT modulated the neurocomputational signals engaged in choice selection, increasing engagement of the vmPFC and PCC that encode the action selection policy during learning of the social hierarchy. At the computational level, these results validate the proposal that OXT may modulate the option selection stage, shifting the choice probability curve (Piva & Chang, 2018), and increasing sensitivity to the relative value during a two-choice economic task (Chang et al., 2011; Piva et al., 2017).

During the test phase, OXT differentially modulated social memory needed for transitive inferences according to whether the participant was embedded or not in the hierarchy. Thus, OXT enhanced performance for the making of transitive inferences with respect to the other-related hierarchy, and reduced performance for the hierarchy in which the participant themself was embedded. At the brain system level, OXT modulated amygdala and parahippocampal activities when making transitive inferences. These brain regions showed interactions in the same direction between group treatment and hierarchy orientation factors. The amygdala and parahippocampal regions have previously been found to encode the status of individuals in linear hierarchies (Kumaran et al., 2016). Both of these regions are rich in OXT receptors and receive OXT from the axonal projections of

parvocellular neurons in the hypothalamus (Hawthorn et al., 1985; Jurek & Neumann, 2018; Thijssen & Wiegerink, 1978). Thus, they are in a critical position to be regulated by OXT. The amygdala directly mediates OXT influences on social behavior in rodents (Ferguson et al., 2001; Huber et al., 2005) and non-human primates (Chang et al., 2015). OXT regulates social memory inferences in rodents and non-human primates (Ferguson et al., 2000; H. J. Lee et al., 2008; MacBeth et al., 2009). For example, OXT receptors are causally involved in social memory in rodents. Knockout of the OXT gene results in the failure of male mice to develop social memory but treatment with OXT rescued this (Ferguson et al., 2000). In rhesus macaques, OXT delivered directly into the basolateral amygdala enhances prosocial tendencies and attention to the recipients of prosocial decisions (Chang et al., 2015). Moreover, OXT modulates value computation of self–other allocations, amplifying both vicarious reinforcement and self-reinforcement in rhesus macaques (Chang, Barter, et al., 2012; Chang & Platt, 2013)). In humans, OXT influences the balance between self and other mental processes (e.g., trust and prosocial behavior) and their underlying neural mechanisms (De Dreu, 2012; Feng et al., 2020; Liao et al., 2020; Tomova et al., 2019; Zhao et al., 2016). Together, these studies provide behavioral and neuroscientific evidence for OXT effects on self-other referential processing. Our findings extend these previous studies in an important way to self-other social memory when making transitive inferences. They indicate distinct patterns of OXT modulation of brain activities engaged in social memory, when one is oneself (or not) embedded in the hierarchy. Overall, our findings show that the effects of OXT on social memory are modulated by self-other regarding considerations, and that these effects are mediated in the amygdalo-hippocampal complex.

We also observed a modulation for Other- vs Self-orientated hierarchy in the dACC, related to action selection policy during the test phase (Fig. 6). This brain region was more robustly engaged for Other- than Self-orientated social hierarchy during the test trials. Efficient social interactions require processing of both Other-related and Self-related information (Bang et al., 2022; Ereira et al., 2020; Wittmann et al., 2016). The anterior cingulate gyrus (ACCg) is implicated in empathy, social learning (Camille et al., 2011; Hadland et al., 2003; Hayden & Platt, 2010), and computations of 'other-orientated' information (Apps & Ramnani, 2014; Behrens et al., 2008; Chang, Gariépy, et al., 2012). Both the anterior and posterior portions of the ACCg, which overlap with the mPFC and dACC regions in our study, play a key role in facilitating cooperative and competitive

interactions, by tracking parameters such as the value and cost of options to others (Apps et al., 2016). As such, our results suggest that the ACCg/mPFC tracks the motivation of others through representations that not only code for the values and costs of reward to other individuals (Apps et al., 2016), but also incorporates status information, particularly within one's own social group.

Limitations and clinical implications

One strength of our study is that we controlled for participants' age and gender (Meyer-Lindenberg et al., 2011; Wang et al., 2017). We decided to test only male participants because the effects of OXT on the brain valuation system are likely to be different in men and women, as OXT increases activation in the ventral striatum in men, but decreases activation in the same region in women during the canonical prisoner's dilemma game (Feng et al., 2015). Another reason for testing only men is that the hierarchy was composed only of male faces, and that men's and women's attitudes towards male faces may differ when learning hierarchies. Further studies will be needed to determine whether the current findings depend upon sex, by comparing hierarchies of women only or mixed samples.

Intranasal OXT spray in humans increases OXT levels in the cerebrospinal fluid. The pathways through which OXT reaches the brain include direct nose-to-brain transport, blood-to-brain transport through the blood-brain barrier or a combination of both. A recent study in primates showed that when administered intranasally, OXT can reach the brain (Lee et al., 2020). However, it is also possible that absorption of OXT into the blood might impact the brain indirectly. For instance, peripheral administration of OXT in rats requires vagal signaling to reduce methamphetamine self-administration and reinstatement of methamphetamine-seeking behaviors (Everett et al., 2020). Therefore, we cannot exclude that our behavioral and neural findings of intranasal OXT on option selection during the training stage and on self-other social memory during the test phase, might reflect indirect peripheral effects or a possible interaction between peripheral and central actions. Future studies with concomitant administration of non-brain penetrant antagonists will help to dissect these effects further.

Our findings may have clinical implications because OXT may increase social functioning in clinical populations, by modulating the option selection stage of social decision making

and/or social memory. Although there have been mixed findings (Cid-jofr et al., 2021; Wang et al., 2017), recent reports have suggested a potential role of OXT to enhance social learning in autism spectrum disorder (ASD) (Kruppa et al., 2019). Single-dose administration studies have generally shown positive effects in ASD (Domes et al., 2013; Gordon et al., 2013), which might be related to an overall enhancement of the motivational system in social contexts. However, longer-term treatment studies, using multiple doses per day, often failed to demonstrate treatment effects (Guastella & Hickie, 2016). One reason might be that these were not designed to provide specific social learning contexts around the times of administration (Aoki et al., 2014), or may have interfered with psychotropic medication (Yatawara et al., 2015). Future studies performed in ASD are needed to investigate how OXT modulates social learning and social memory.

4. Materials and Methods

4.1 Participants

We recruited 60 healthy male undergraduate or graduate students (mean age 21.1 years, SD = 2.10, range 18-29 years). Participants were not taking any prescribed drugs and had no history of drug abuse or smoking. We instructed participants to abstain from alcohol and heavy exercise for 24 hours and from food or any beverage other than water for at least 2 hours before scanning. Participants gave written informed consent. Human Research Ethics Committee for Non-Clinical Faculties in South China Normal University approved the study. Two subjects were excluded because of too low performance in the last session (blocks 7~9) both in training and test trials, based on the scale Z-score method, with z scores lower than -2.0, as there was less than a 2.27% probability that these two participants would have a lower average accuracy than our sample. Also, we went through all participants' decisions and found the learning rates of these two fluctuated irregularly. Finally, we included 29 subjects (mean age 20.72 years, SD = 1. 87) in the OXT group and 29 subjects (mean age 21.41 years, SD = 2.23) in the Placebo group for analysis.

4.2 Stimulus and Task

The two diamond-shaped hierarchy structures consisted of 16 face pictures (8 for Self-orientated condition, 8 for Other-orientated condition) and two shaded pictures referring to Self and Other (letter 'X' as its name). Real face pictures were selected from a large Chinese face database: the CAS-PEAL face database (Wen Gao et al., 2008). Images with neutral facial expression were cropped congruously and converted to grayscale. For each hierarchy structure, position 5 was always Self or Other, the face images were randomly distributed to the other locations.

The fMRI task proceeded in 3 separate scanning sessions of 6 blocks (Self: 3 blocks, Others: 3 blocks) and 150 trials each on average and took around 60 mins in total (Fig. 1). Each block combined 12 learning trials followed with 13 test trials. Each learning trial block included 5 superior paired trial types (red line: P1 vs P2, P2 vs P3, P1 vs P5, P2 vs P5, P3 vs P5), 5 inferior paired trial types (green line: P7 vs P8, P8 vs P9, P7 vs P5, P8 vs P5, P9 vs P5) and 2 intermediate paired trial types (orange line: P4 vs P5, P6 vs P5). The Test trial block comprised (pink line: P1 vs P3, P2 vs P4, P3 vs P6, P1 vs P6, P4 vs P8, P6 vs P7, P7 vs P9, P4 vs P9, P2 vs P8, P3 vs P7, P1 vs P8, P2 vs P9, P4 vs P6). See Fig 1 for a display of these pairs. The Self- and Other-orientated conditions occurred consecutively with a random starting order that was indicated by a cue of text description on the screen.

To learn these novel diamond-shaped hierarchy structures, participants performed a three-choice social hierarchy learning task. During the learning trials, they press ''Left (>)'', ''Equal (=)'' or ''Right (<)'' within 3 s when they were shown two-paired face items on the screen to indicate the relationship between the two. To confirm the reaction, the pictures were highlighted with yellow or blue borders for 0.2 s, indicating whether it was a Self or Other-oriented hierarchy structure respectively. The order of Self/Other trial blocks and their color indications were balanced across different subjects. Then, participants received a feedback screen (2 seconds) indicating Correct (+20) or Wrong (-20). In the test trials, participants had to infer the relationships between the two faces that were presented (pairs of non-adjacent faces from the same hierarchy) from what they had learnt before. In the test phase they received no immediate feedback, however at the end of each test trial block they were informed of their global score of correct responses.

4.3 Procedure

When participants arrived at the lab, they were first asked to self-administer an intranasal dose of 24 IU (Oxytocin Spray— Sichuan Meike Pharmacy Co. Ltd, Sichuan, China; 3 puffs of 4 IU per nostril with 30 s between each puff) of either oxytocin or placebo under the experimenters' supervision. Our placebo contained the same solvent (sodium chloride and glycerin) but no oxytocin and was packaged in the same manufactured bottle. Then, participants were required to fill out questionnaires assessing psychological measures: The State-Trait Anxiety Inventory (STAI, subscales: SAI and TAI) to evaluate state and trait anxiety, the Beck Depression Inventory (BDI) to assess depression, and the Positive and Negative Affect Schedule (PANAS, subscales: PAS and NAS) to rate current affective state.

Prior to the scanning session, to motivate learning of the hierarchy task, subjects were given a cover story. They role-played as a candidate applying for a job at their ideal company. Concurrently, they imagined another person X (a stranger) also applying for this job. Subsequently they completed 10 low difficulty multiple choice questions, representing personality, math and literature (e.g., general knowledge: "How do you use a knife and fork?"; math tests: "According to the net of a cube below, select a valid way to reshape it."). These had to be answered by subjects within 5 mins. Then, participants were told that both he and the X person would be assigned to corresponding positions in two different departments of the company based on their scores.

Once the quiz was completed, the participants received experimental instructions, that they and X were employed by department A (B) and B (A) separately. As beginners, they were required to learn the relationship between each person's level of competence in each of the two departments, to better integrate into the organization. Instructions were also included with respect to the learning and test trials in the hierarchy task. We briefly repeated these instructions at the beginning of pretest program which comprised three trials in the Self- and Other-orientated hierarchy conditions, to confirm they thoroughly understood the task. Afterwards, 16 pictures were presented in a random order three times, each for 2 seconds to reduce the potential effects of novel stimuli during the scanning procedure. 30 minutes after the administration of the spray participants began to perform the formal task which took about 60 minutes in the MRI scanner. When the fMRI experiment finished, participants completed a self-reported questionnaire to investigate their attitude towards the study, as well as the State Anxiety Inventory (STAI-SAI), the Positive and Negative Affect Schedule (PANAS), Social Dominance Orientation Social

(SDO), Achievement Motivation Scale (AMS) and Iowa-Netherlands Comparison Orientation Measure (INCOM), to control the emotion and personality variables (Supplementary Table S5).

4.4 Behavioral analysis

Performance was modeled using a generalized linear mixed-effects model (GLMM) with a logit link function for the binomial (correct/ wrong) answer by the glmer function in 'lme4' package on R, allowing random intercepts per subjects. The fixed effects included 4 variables (Group (OXT/Placebo), Hierarchy-orientation condition (Self/Other), Status (Superior, Intermediate, Inferior), Block (1 to 9)), and their 2-way interaction factors. Non-significant interactions would be excluded from final models. For response time (RT), we used a linear mixed model (LMM) to analyze data with the lmer function in 'lmertest' package on R. Follow-up tests of significant main effects or interactions were conducted using the 'emmeans' package. Detailed results tables for all models are reported in Supplementary files.

4.5 Computational modeling

The social hierarchy structure concerned both equal and different ranks, which makes it two dimensional and hence harder to learn than linear hierarchies. Consequently, we developed a new RL-ELO model, termed Combined RL-ELO-based associative learning models (Model 5, detailed below), enabling the model to learn hierarchical structures involving equal ranks. Because learning equal ranks may involve different strategies than learning unequal ranks, we assumed that learning equal ranks was equivalent to memorizing matching pairs. For this reason, we propose a model in which learning happens in two steps. Firstly, the Model 5 increased equality pair trial weights, when updating value processes, by judging the paired trial as equal or unequal response p_{eq} in the first step, with learning rate α_{asso}, using an associative learning model (Equation 1~2). Secondly, participants updated their representation of the structure of the hierarchy by updating the value of each item in the pair according to the past choice and reward. Each item value was evaluated within the internal hierarchy structure, with learning rate α in unequal trials, and the chosen item and learning rate α_{eq} in equal trials by the RL-ELO model (Equation 3~10). That is, the model increased (or decreased respectively) the value

of the highest (or the lowest) item in the hierarchy. If both items were equal in the hierarchy, their values became closer according to a proportion (α_{eq}) of their initial distance. The softmax choice rule (Equation 11~14) converted action values (e.g., V_L) to action probabilities (e.g., p_L). The inverse temperature β, controled the steepness of the softmax function, which decides how decisively the response alternatives contrast with each other.

4.6 RL-ELO model combining associative and structure learning (Model 5)

Free parameters: α_{asso} = learning rate of associative learning; α = learning rate in unequal trial chosen item; α_{eq} = proportion of the distance items' value are brought closer after equal feedback; δ_1 = weight on value from associative learning; δ_2 = weight on values from global structure learning; γ = bias representing preference towards selecting the equal response; β = temperature representing the stochasticity of choices.

Associative learning (distinguish '=' and '<>' choices):

if choice at t-1 is '=', then Eq = 1

if choice at t-1 is either '<' or '>', then Eq = -1

$$p_{eq,t-1} = p_{(action_{t-1} = equal)} \quad (1)$$

$$Q_{eq,t} = Q_{eq,t-1} + \alpha_{asso} * (Eq - p_{eq,t-1}) \quad (2)$$

Structure learning (distinguish relationship between a pair of items):

if Left item is correct choice:

$$p_{win,t-1} = p_{L,t-1} \quad (3)$$

$$I_{L,t-1} = 1,\ I_{R,t-1} = -1 \quad (4)$$

if Right item is correct choice:

$$p_{win,t-1} = p_{R,t-1} \quad (5)$$

$$I_{R,t-1} = 1,\ I_{L,t-1} = -1 \quad (6)$$

Therefore, if previous response is unequal ('<' or '>'):

$$Q_{L,t} = Q_{L,t-1} + I_{L,t-1} * \alpha * (1 - p_{win,t-1}) \quad (7)$$

$$Q_{R,t} = Q_{R,t-1} + I_{R,t-1} * \alpha * (1 - p_{win,t-1}) \quad (8)$$

if previous response is equal ('='):

$$Q_{L,t} = Q_{L,t-1} + \alpha_{eq} * (Q_{R,t-1} - Q_{L,t-1}) \quad (9)$$

$$Q_{R,t} = Q_{R,t-1} - \alpha_{eq} * (Q_{L,t-1} - Q_{R,t-1}) \quad (10)$$

Probability of estimated action:

$$V_L = Q_L - Q_R \quad (11)$$

$$V_E = \delta_1 * Q_{eq} + \frac{\delta_2}{|Q_L - Q_R|} * N_{trial} + \gamma \quad (12)$$

$$V_R = Q_R - Q_L \quad (13)$$

Action policy (softmax function):

$$p_{(choose\ option\ i)} = \frac{e^{\beta * V_i}}{e^{\beta * V_L} + e^{\beta * V_E} + e^{\beta * V_R}} \quad (14)$$

Lastly, several models were tested to update the item values in the hierarchy structure. For example, the Rescorla-Wagner (RW) model (Model 3) and the Value Transfer model (Model 4), which are known to fail in learning linear hierarchies as they associate equally with positive and negative outcomes. The classical RL-ELO model (Model 1) which functions successfully in learning linear hierarchy structures was also compared, as well as its form with two different learning rates for hierarchically superior and inferior members of the hierarchy (Model 2).

The Combined RL-ELO-based associative learning model increased equality pair trial weights when updating the value process by judging the paired trial to be equal or unequal response p_{eq} in the first step with learning rate α_{asso}, using the associative learning model (Equation 1~2); the next step evaluated each item's value within the internal hierarchy structure, with learning rate α in unequal trial chosen items and learning rate α_{eq} *i*n equal trial chosen items, according to the RL-ELO model (Equation 3~10).

Model 1: RL-ELO model with 2 alphas

Classical RL models assume only a single learning rate α, ranging from 0 to 1. The higher α, the stronger the weighting of the prediction error (PE) for updating value. However, due to the three-level hierarchies (i.e., superior, intermediate, inferior), we developed two RL-ELO models. The first RL-ELO model with 2 alphas (Model 1), with α representing the learning rate for unequal choice $\alpha_{unequal}$, and the other is equal choice α_{equal}. The second RL-ELO model (Model 2), further subdivided the learning rate $\alpha_{unequal}$ into $\alpha_{superior}$ and $\alpha_{inferior}$.

RL-ELO model with 2 alphas is focused on the difference value between pairwise items in each trial. The implementation of this model is approximately similar to the process of Equation 3 to 12. What differs from these segments is the estimated equal action probability within Equation 12 which will be calculated as Equation 15:

$$V_E = \frac{\delta_2}{|Q_L - Q_R|} \qquad (15)$$

Model 2: RL-ELO model with 3 alphas

This model assumes that the following three various hierarchies: superior, intermediate and inferior have their own learning rate. Therefore, the update value of unequal trials (Equation 7 to 8) is split, as when a previous pair is related to superior hierarchy, then conducted with $\alpha_{superior}$, corresponding to inferior hierarchy performed with $\alpha_{inferior}$.

Model 3 & 4: Rescorla-Wagner (RW) model & Value transfer model

The classical RW model is inspired by Pavlovian conditioning — animal learning from conditioned stimulus (CS) paired with an unconditioned stimulus (US). According to mechanism of RW model, the value transfer model contains another component, the accrued value θ from associative strength of the indirect (wrong) item paired with CS. While RW model set θ as zero.

if Left item is correct choice, trial outcome vector (O) & indirect item vector(W):

$$O_{L,t-1} = 1,\ O_{R,t-1} = -1 \qquad (16)$$

$$W_{L,t-1} = 0,\ W_{R,t-1} = 1 \qquad (17)$$

if Right item is correct choice, trial outcome vector & indirect item vector:

$$O_{L,t-1} = -1,\ O_{R,t-1} = 1 \quad (18)$$

$$W_{L,t-1} = 1,\ W_{R,t-1} = 0 \quad (19)$$

Update values of previous direct item:

$$V\_direct_{L,t-1} = (O_{L,t-1} - V_{L,t-1}) * \alpha \quad (20)$$

$$V\_direct_{R,t-1} = (O_{R,t-1} - V_{R,t-1}) * \alpha \quad (21)$$

Update values of previous indirect item:

$$V\_indirect_{L,t-1} = V_{R,t-1} * W_{L,t-1} * \theta \quad (22)$$

$$V\_indirect_{R,t-1} = V_{L,t-1} * W_{R,t-1} * \theta \quad (23)$$

Total update:

$$V_{L,t} = V\_direct_{L,t-1} + V\_indirect_{L,t-1} \quad (24)$$

$$V_{R,t} = V\_direct_{R,t-1} + V\_indirect_{R,t-1} \quad (25)$$

4.7 fMRI data acquisition and analyses

Brain images were collected at the Imaging Center of South China Normal University on a Siemens 3T Tim Trio scanner equipped with a 32-channel head-coil. Functional images were collected by using an echo-planar imaging (EPI) sequence (repetition time = 2000 ms, echo time = 30 ms; flip angle = 90°; slice thickness = 3.5 mm, field of view = 224 mm, in-plane resolution = 3 × 3 mm). In addition to functional images, a high-resolution anatomical scan was also collected (MDEFT, 1 × 1 × 1 mm resolution). Stimulations and response recordings were controlled by E-prime software on a Windows laptop located in the scanner control room. Choice responses were collected using a 4- button response pad.

The following pre-processing was applied: slice timing correction, spatial realignment, co-registered to the mean functional images, segmentation of structural T1 image, normalization motion correction, spatial smoothing using a Gaussian kernel of full-width-at-half-maximum 8.0 mm.

We ran four independent general linear models (GLMs) with a 2-way flexible factorial ANOVA, with Self- and Other-orientated hierarchy as a within subject factor and Group (Oxytocin/Placebo) as a between subject factor. In GLM1 and GLM3, onset regressor

began at the presentation of the 2 paired items with null duration, modulated by the parametric regressor: action probability of the choice for correct responses by the Combined RL-ELO-based associative learning model (Model 5). GLM1 was applied to training trials and GLM3 to test trials. To improve overall model fit and account for potential confounds, the design matrix included the following additional regressors which were not analyzed: two regressors that modelled the onset of the highlighted response screen and the onset of the feedback screen. Both were performed separately for training trials and test trials. The GLM2 model employed a regressor at the onset of the feedback screen with 2 s duration, modulated by two parametric regressors: an unsigned PE of judging equal or unequal trial PE_{eq}, and an unsigned PE of choosing response action PE_i by the winning Model 5. Parametric modulators were not orthogonalised with respect to each other. The GLM4 was a basic GLM analysis without model parameters, which set the regressor with whole test trial while the two items were displayed and the highlighted response screen with actual duration.

For identification of brain (de)activation patterns we first applied whole-brain FWE-correction with a cluster-wise threshold of $p<0.05$, combined with an uncorrected voxel-level threshold of $P<0.001$. Next, referencing prior neuroimaging work on subjective value of choice alternatives, we extracted BOLD activity from 10 mm radius spherical ROI in the vmPFC (x = -2, y = 46, z = -8) and PCC (x = -4, y = -30, z = 36) according to a meta-analysis of value representation (Bartra et al., 2013). We also added the left amygdala as ROI, identified through the WFU PickAtlas 3 toolbox. The results were visualized using the Mango image viewer software (http://rii.uthscsa.edu/mango/mango.html). Brain regions were labeled according to the automated anatomical labeling template via the xjView toolbox (http://www.alivelearn.net/xjview8/). All significantly activated brain regions are reported in Supplementary Table S5.

References

Aoki, Y., Yahata, N., Watanabe, T., Takano, Y., Kawakubo, Y., Kuwabara, H., Iwashiro, N., Natsubori, T., Inoue, H., Suga, M., Takao, H., Sasaki, H., Gonoi, W., Kunimatsu, A., Kasai, K., & Yamasue, H. (2014). Oxytocin improves behavioural and neural deficits in inferring others' social emotions in autism. *Brain*, *137*(11), 3073–3086. https://doi.org/10.1093/BRAIN/AWU231

Apps, M. A. J., & Ramnani, N. (2014). The Anterior Cingulate Gyrus Signals the Net Value of Others' Rewards. *Journal of Neuroscience*, *34*(18), 6190–6200. https://doi.org/10.1523/JNEUROSCI.2701-13.2014

Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron*, *90*(4), 692–707. https://doi.org/10.1016/J.NEURON.2016.04.018

Bang, D., Moran, R., Daw, N. D., & Fleming, S. M. (2022). Neurocomputational mechanisms of confidence in self and others. *Nature Communications 2022 13:1*, *13*(1), 1–14. https://doi.org/10.1038/s41467-022-31674-w

Baram, A. B., Muller, T. H., Nili, H., Garvert, M. M., & Behrens, T. E. J. (2021). Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. *Neuron*, *109*(4), 713-723.e7. https://doi.org/10.1016/j.neuron.2020.11.024

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/J.NEUROIMAGE.2013.02.063

Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: context and person matter. *Trends in Cognitive Sciences*, *15*(7), 301–309. https://doi.org/10.1016/J.TICS.2011.05.002

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., & Rushworth, M. F. S. (2008). Associative learning of social value. *Nature 2008 456:7219*, *456*(7219), 245–249. https://doi.org/10.1038/NATURE07538

Boorman, E. D., Sweigart, S. C., & Park, S. A. (2021). Cognitive maps and novel inferences: a flexibility hierarchy. *Current Opinion in Behavioral Sciences*, *38*, 141–149. https://doi.org/10.1016/J.COBEHA.2021.02.017

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double Dissociation of Stimulus-Value and Action-Value Learning in Humans with Orbitofrontal or Anterior Cingulate Cortex Damage. *Journal of Neuroscience*, *31*(42), 15048–15052. https://doi.org/10.1523/JNEUROSCI.3164-11.2011

Chang, S. W. C., Barter, J. W., Ebitz, R. B., Watson, K. K., & Platt, M. L. (2012). Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (Macaca mulatta). *Proceedings of the National Academy of Sciences of the United States of America*, *109*(3), 959–964. https://doi.org/10.1073/PNAS.1114621109

Chang, S. W. C., Fagan, N. A., Toda, K., Utevsky, A. V., Pearson, J. M., Platt, M. L., & Gazzaniga, M. S. (2015). Neural mechanisms of social decision-making in the primate amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(52), 16012–16017.

Chang, S. W. C., Gariépy, J. F., & Platt, M. L. (2012). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience 2012 16:2*, *16*(2), 243–250. https://doi.org/10.1038/NN.3287

Chang, S. W. C., & Platt, M. L. (2013). *Oxytocin and social cognition in rhesus macaques: Implications for understanding and treating human psychopathology*. https://doi.org/10.1016/j.brainres.2013.11.006

Chang, S. W. C., & Platt, M. L. (2014). Oxytocin and social cognition in rhesus macaques: Implications for understanding and treating human psychopathology. *Brain Research*, *1580*, 57–68. https://doi.org/10.1016/J.BRAINRES.2013.11.006

Chang, S. W. C., Winecoff, A. A., & Platt, M. L. (2011). Vicarious reinforcement in rhesus macaques (Macaca mulatta). *Frontiers in Neuroscience*, *0*(MAR), 27. https://doi.org/10.3389/FNINS.2011.00027/BIBTEX

Chase, I. D., & Seitz, K. (2011). Self-Structuring Properties of Dominance Hierarchies: A New Perspective. *Advances in Genetics*, *75*, 51–81. https://doi.org/10.1016/B978-0-12-380858-5.00001-0

Cid-jofr, V., Moreno, M., Reyes-parada, M., & Renard, G. M. (2021). *Role of Oxytocin and Vasopressin in Neuropsychiatric Disorders : Therapeutic Potential of Agonists and Antagonists*.

Clark-Elford, R., Nathan, P. J., Auyeung, B., Voon, V., Sule, A., Müller, U., Dudas, R., Sahakian, B. J., Phan, K. L., & Baron-Cohen, S. (2014). The effects of oxytocin on social reward learning in humans. *International Journal of Neuropsychopharmacology*, *17*(2), 199–209. https://doi.org/10.1017/S1461145713001120

Daunizeau, J., Adam, V., & Rigoux, L. (2014). VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Computational Biology*, *10*(1), e1003441.

Dayan, P. (1993). Improving Generalization for Temporal Difference Learning: The Successor Representation. *Neural Computation*, *5*(4), 613–624. https://doi.org/10.1162/NECO.1993.5.4.613

De Dreu, C. K. W. (2012). Oxytocin modulates cooperation within and competition between groups: An integrative review and research agenda. *Hormones and Behavior*, *61*(3), 419–428. https://doi.org/10.1016/j.yhbeh.2011.12.009

De Dreu, C. K. W., Greer, L. L., Handgraaf, M. J. J., Shalvi, S., Van Kleef, G. A., Baas, M., Ten Velden, F. S., Van Dijk, E., & Feith, S. W. W. (2010). The Neuropeptide Oxytocin Regulates Parochial Altruism in Intergroup Conflict Among Humans. *Science*, *328*(5984), 1408–1411. https://doi.org/10.1126/SCIENCE.1189047

Domes, G., Heinrichs, M., Kumbier, E., Grossmann, A., Hauenstein, K., & Herpertz, S. C. (2013). Effects of intranasal oxytocin on the neural basis of face processing in autism spectrum disorder. *Biological Psychiatry*, *74*(3), 164–171. https://doi.org/10.1016/j.biopsych.2013.02.007

Ebitz, R. B., Watson, K. K., & Platt, M. L. (2013). Oxytocin blunts social vigilance in the rhesus macaque. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(28), 11630–11635. https://doi.org/10.1073/PNAS.1305230110

Ereira, S., Hauser, T. U., Moran, R., Story, G. W., Dolan, R. J., & Kurth-Nelson, Z. (2020). Social training reconfigures prediction errors to shape Self-Other boundaries. *Nature Communications 2020 11:1*, *11*(1), 1–14. https://doi.org/10.1038/S41467-020-16856-8

Everett, N. A., Turner, A. J., Costa, P. A., Baracz, S. J., & Cornish, J. L. (2020). The vagus nerve mediates the suppressing effects of peripherally administered oxytocin on methamphetamine self-administration and seeking in rats. *Neuropsychopharmacology 2020 46:2*, *46*(2), 297–304. https://doi.org/10.1038/S41386-020-0719-7

Feng, C., Hackett, P. D., DeMarco, A. C., Chen, X., Stair, S., Haroon, E., Ditzen, B., Pagnoni, G., & Rilling, J. K. (2015). Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans. *Brain Imaging and Behavior*, *9*(4), 754–764. https://doi.org/10.1007/S11682-014-9333-9

Feng, C., Zhou, X., Zhu, X., Zhu, R., Han, S., & Luo, Y.-J. (2020). *Effect of intranasal oxytocin administration on self-other distinction: Modulations by psychological distance and gender*. https://doi.org/10.1016/j.psyneuen.2020.104804

753 Ferguson, J. N., Aldag, J. M., Insel, T. R., & Young, L. J. (2001). Oxytocin in the Medial Amygdala is Essential for
754 Social Recognition in the Mouse. *Journal of Neuroscience*, *21*(20), 8278–8285.
755 https://doi.org/10.1523/JNEUROSCI.21-20-08278.2001
756 Ferguson, J. N., Young, L. J., Hearn, E. F., Matzuk, M. M., Insel, T. R., & Winslow, J. T. (2000). Social amnesia in
757 mice lacking the oxytocin gene. *Nature Genetics*, *25*(3), 284–288. https://doi.org/10.1038/77040
758 Gordon, I., Vander Wyk, B. C., Bennett, R. H., Cordeaux, C., Lucas, M. V., Eilbott, J. A., Zagoory-Sharon, O.,
759 Leckman, J. F., Feldman, R., & Pelphrey, K. A. (2013). Oxytocin enhances brain function in children with
760 autism. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 20953–
761 20958. https://doi.org/10.1073/PNAS.1312857110
762 Guastella, A. J., & Hickie, I. B. (2016). Oxytocin Treatment, Circuitry, and Autism: A Critical Review of the
763 Literature Placing Oxytocin into the Autism Context. *Biological Psychiatry*, *79*(3), 234–242.
764 https://doi.org/10.1016/j.biopsych.2015.06.028
765 Hadland, K. A., Rushworth, M. F. S., Gaffan, D., & Passingham, R. E. (2003). The anterior cingulate and reward-
766 guided selection of actions. *Journal of Neurophysiology*, *89*(2), 1161–1164.
767 https://doi.org/10.1152/JN.00634.2002/ASSET/IMAGES/LARGE/9K0232906003.JPEG
768 Hawthorn, J., Ang, V. T. Y., & Jenkins, J. S. (1985). Effects of lesions in the hypothalamic paraventricular,
769 supraoptic and suprachiasmatic nuclei on vasopressin and oxytocin in rat brain and spinal cord. *Brain*
770 *Research*, *346*(1), 51–57. https://doi.org/10.1016/0006-8993(85)91093-5
771 Hayden, B. Y., & Platt, M. L. (2010). Neurons in Anterior Cingulate Cortex Multiplex Information about Reward and
772 Action. *Journal of Neuroscience*, *30*(9), 3339–3346. https://doi.org/10.1523/JNEUROSCI.4874-09.2010
773 Hu, J., Qi, S., Becker, B., Luo, L., Gao, S., Gong, Q., Hurlemann, R., & Kendrick, K. M. (2015). Oxytocin
774 selectively facilitates learning with social feedback and increases activity and functional connectivity in
775 emotional memory and reward processing regions. *Human Brain Mapping*, *36*(6), 2132–2146.
776 https://doi.org/10.1002/HBM.22760
777 Huber, D., Veinante, P., & Stoop, R. (2005). Vasopressin and oxytocin excite distinct neuronal populations in the
778 central amygdala. *Science*, *308*(5719), 245–248.
779 https://doi.org/10.1126/SCIENCE.1105636/SUPPL_FILE/HUBER-SOM.PDF
780 Hung, L. W., Neuner, S., Polepalli, J. S., Beier, K. T., Wright, M., Walsh, J. J., Lewis, E. M., Luo, L., Deisseroth, K.,
781 Dölen, G., & Malenka, R. C. (2017). Gating of social reward by oxytocin in the ventral tegmental area. *Science*
782 *(New York, N.Y.)*, *357*(6358), 1406–1411. https://doi.org/10.1126/SCIENCE.AAN4994
783 Hurlemann, R., Patin, A., Onur, O. A., Cohen, M. X., Baumgartner, T., Metzler, S., Dziobek, I., Gallinat, J., Wagner,
784 M., Maier, W., & Kendrick, K. M. (2010). Oxytocin enhances amygdala-dependent, socially reinforced
785 learning and emotional empathy in humans. *The Journal of Neuroscience : The Official Journal of the Society*
786 *for Neuroscience*, *30*(14), 4999–5007. https://doi.org/10.1523/JNEUROSCI.5538-09.2010
787 Ide, J. S., Nedic, S., Wong, K. F., Strey, S. L., Lawson, E. A., Dickerson, B. C., Wald, L. L., La Camera, G., &
788 Mujica-Parodi, L. R. (2018). Oxytocin attenuates trust as a subset of more general reinforcement learning, with
789 altered reward circuit functional connectivity in males. *NeuroImage*, *174*, 35–43.
790 https://doi.org/10.1016/J.NEUROIMAGE.2018.02.035

Janet, R., Ligneul, R., Losecaat-Vermeer, A. B., Philippe, R., Bellucci, G., Derrington, E., Park, S. Q., & Dreher, J.-C. (2022). Regulation of social hierarchy learning by serotonin transporter availability. *Neuropsychopharmacology 2022*, 1–8. https://doi.org/10.1038/S41386-022-01378-2

Jensen, G., Terrace, H. S., & Ferrera, V. P. (2019). Discovering Implied Serial Order Through Model-Free and Model-Based Learning. In *Frontiers in Neuroscience* (Vol. 13). https://doi.org/10.3389/fnins.2019.00878

Jiang, Y., & Platt, M. L. (2018). Oxytocin and vasopressin flatten dominance hierarchy and enhance behavioral synchrony in part via anterior cingulate cortex. *Scientific Reports*, *8*(1), 8201. https://doi.org/10.1038/s41598-018-25607-1

Joiner, J., Piva, M., Turrin, C., & Chang, S. W. C. (2017). Social learning through prediction error in the brain. *Npj Science of Learning*, *2*(8). https://doi.org/10.1038/s41539-017-0009-2

Jurek, B., & Neumann, I. D. (2018). The oxytocin receptor: From intracellular signaling to behavior. *Physiological Reviews*, *98*(3), 1805–1908. https://doi.org/10.1152/PHYSREV.00031.2017/ASSET/IMAGES/LARGE/Z9J0021828400013.JPEG

Kruppa, J. A., Gossen, A., Oberwelland Weiß, E., Kohls, G., Großheinrich, N., Cholemkery, H., Freitag, C. M., Karges, W., Wölfle, E., Sinzig, J., Fink, G. R., Herpertz-Dahlmann, B., Konrad, K., & Schulte-Rüther, M. (2019). Neural modulation of social reinforcement learning by intranasal oxytocin in male adults with high-functioning autism spectrum disorder: a randomized trial. *Neuropsychopharmacology*, *44*(4), 749–756. https://doi.org/10.1038/S41386-018-0258-7

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., & Dayan, P. (2016). Computations Underlying Social Hierarchy Learning: Distinct Neural Mechanisms for Updating and Representing Self-Relevant Information. *Neuron*, *92*(5), 1135–1147. https://doi.org/10.1016/j.neuron.2016.10.052

Kumaran, D., Melo, H. L., & Duzel, E. (2012). The Emergence and Representation of Knowledge about Social and Nonsocial Hierarchies. *Neuron*, *76*(3), 653–666. https://doi.org/10.1016/j.neuron.2012.09.035

Kurth-Nelson, Z., Barnes, G., Sejdinovic, D., Dolan, R., & Dayan, P. (2015). Temporal structure in associative retrieval. *Elifesciences.Org*, *4*, 4919. https://doi.org/10.7554/eLife.04919

Lee, H. J., Caldwell, H. K., Macbeth, A. H., Tolu, S. G., & Young, W. S. (2008). A Conditional Knockout Mouse Line of the Oxytocin Receptor. *Endocrinology*, *149*(7), 3256–3263. https://doi.org/10.1210/EN.2007-1710

Lee, M. R., Shnitko, T. A., Blue, S. W., Kaucher, A. V, Winchell, A. J., Erikson, D. W., Grant, K. A., & Leggio, L. (2020). Labeled oxytocin administered via the intranasal route reaches the brain in rhesus macaques. *Nature Communications*, *11*(1), 1–10.

Lee, W., Hiura, L. C., Yang, E., Broekman, K. A., Ophir, A. G., & Curley, J. P. (2019). Social status in mouse social hierarchies is associated with variation in oxytocin and vasopressin 1a receptor densities. *Hormones and Behavior*, *114*, 104551. https://doi.org/10.1016/J.YHBEH.2019.06.015

Liao, Z., Huang, L., & Luo, S. (2020). Intranasal oxytocin decreases self-oriented learning. *Psychopharmacology 2020 238:2*, *238*(2), 461–474. https://doi.org/10.1007/S00213-020-05694-7

Ligneul, R., Obeso, I., Ruff, C. C., & Dreher, J. C. (2016). Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex. *Current Biology*, *26*(23), 3107–3115. https://doi.org/10.1016/j.cub.2016.09.015

MacBeth, A. H., Lee, H. J., Edds, J., & Young, W. S. (2009). Oxytocin and the oxytocin receptor underlie intrastrain, but not interstrain, social recognition. *Genes, Brain and Behavior*, *8*(5), 558–567. https://doi.org/10.1111/J.1601-183X.2009.00506.X

Martins, D., Lockwood, P., Cutler, J., Moran, R., & Paloyelis, Y. (2022). Oxytocin modulates neurocomputational mechanisms underlying prosocial reinforcement learning. *Progress in Neurobiology*, *213*, 102253. https://doi.org/10.1016/j.pneurobio.2022.102253

Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature Reviews Neuroscience 2011 12:9*, *12*(9), 524–538. https://doi.org/10.1038/NRN3044

Miller, K. J., Botvinick, M. M., & Brody, C. D. (2022). Value representations in the rodent orbitofrontal cortex drive learning, not choice. *ELife*, *11*. https://doi.org/10.7554/eLife.64575

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature 2006 442:7106*, *442*(7106), 1042–1045. https://doi.org/10.1038/NATURE05051

Piva, M., & Chang, S. W. C. (2018). An integrated framework for the role of oxytocin in multistage social decision-making. *American Journal of Primatology*, *80*(10), e22735. https://doi.org/10.1002/AJP.22735

Piva, M., Zhang, X., Noah, J. A., Chang, S. W. C., & Hirsch, J. (2017). Distributed neural activity patterns during human-to-human competition. *Frontiers in Human Neuroscience*, *11*, 571. https://doi.org/10.3389/FNHUM.2017.00571/BIBTEX

Qu, C., Ligneul, R., Van der Henst, J.-B., & Dreher, J.-C. (2017). An Integrative Interdisciplinary Perspective on Social Dominance Hierarchies. *Trends in Cognitive Sciences*, *21*(11), 893–908. https://doi.org/10.1016/j.tics.2017.08.004

Raam, T., McAvoy, K. M., Besnard, A., Veenema, A., & Sahay, A. (2017). Hippocampal oxytocin receptors are necessary for discrimination of social stimuli. *Nature Communications 2017 8:1*, *8*(1), 1–14. https://doi.org/10.1038/S41467-017-02173-0

Rudebeck, P. H., Saunders, R. C., Lundgren, D. A., & Murray, E. A. (2017). Specialized Representations of Value in the Orbital and Ventrolateral Prefrontal Cortex: Desirability versus Availability of Outcomes. *Neuron*, *95*(5), 1208-1220.e5. https://doi.org/10.1016/j.neuron.2017.07.042

Sharpe, M. J., Chang, C. Y., Liu, M. A., Batchelor, H. M., Mueller, L. E., Jones, J. L., Niv, Y., & Schoenbaum, G. (2017). Dopamine transients are sufficient and necessary for acquisition of model-based associations. *Nature Neuroscience 2017 20:5*, *20*(5), 735–742. https://doi.org/10.1038/NN.4538

Thijssen, J. H. H., & Wiegerink, M. A. H. M. (1978). Intra- and extrahypothalamic vasopressin and oxytocin pathways in the rat. *Cell and Tissue Research 1978 192:3*, *192*(3), 423–435. https://doi.org/10.1007/BF00212323

Tomova, L., Heinrichs, M., & Lamm, C. (2019). The Other and Me: Effects of oxytocin on self-other distinction. *International Journal of Psychophysiology*, *136*(January), 49–53. https://doi.org/10.1016/j.ijpsycho.2018.03.008

Vasconcelos, M. (2008). Mini-review Transitive inference in non-human animals: An empirical and theoretical analysis. *Behavioural Processes*, *78*, 313–334. https://doi.org/10.1016/j.beproc.2008.02.017

Wang, D., Yan, X., Li, M., & Ma, Y. (2017). Neural substrates underlying the effects of oxytocin: A quantitative meta-analysis of pharmaco-imaging studies. *Social Cognitive and Affective Neuroscience*, *12*(10), 1565–1573. https://doi.org/10.1093/scan/nsx085

Williamson, C. M., Lee, W., & Curley, J. P. (2016). Temporal dynamics of social hierarchy formation and maintenance in male mice. *Animal Behaviour*, *115*, 259–272. https://doi.org/10.1016/J.ANBEHAV.2016.03.004

Wirth, S., Soumier, A., Eliava, M., Derdikman, D., Wagner, S., Grinevich, V., & Sirigu, A. (2021). Territorial blueprint in the hippocampal system. *Trends in Cognitive Sciences*. https://doi.org/10.1016/J.TICS.2021.06.005

Wittmann, M. K., Kolling, N., Faber, N. S., Scholl, J., Nelissen, N., & Rushworth, M. F. S. (2016). Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. *Neuron*, *91*(2), 482–493. https://doi.org/10.1016/J.NEURON.2016.06.022

Yatawara, C. J., Einfeld, S. L., Hickie, I. B., Davenport, T. A., & Guastella, A. J. (2015). The effect of oxytocin nasal spray on social interaction deficits observed in young children with autism: a randomized clinical crossover trial. *Molecular Psychiatry 2016 21:9*, *21*(9), 1225–1231. https://doi.org/10.1038/MP.2015.162

Zhao, W., Yao, S., Li, Q., Geng, Y., Ma, X., Luo, L., Xu, L., & Kendrick, K. M. (2016). Oxytocin blurs the self-other distinction during trait judgments and reduces medial prefrontal cortex responses. *Human Brain Mapping*, *37*(7), 2512–2527. https://doi.org/10.1002/HBM.23190

Zhuang, Q., Zhu, S., Yang, X., Zhou, X., Xu, X., Chen, Z., Lan, C., Zhao, W., Becker, B., Yao, S., & Kendrick, K. M. (2021). Oxytocin-induced facilitation of learning in a probabilistic task is associated with reduced feedback- and error-related negativity potentials. *Journal of Psychopharmacology (Oxford, England)*, *35*(1), 40–49. https://doi.org/10.1177/0269881120972347

Acknowledgments

We thank Y. Huang for the help in participant recruitment and scanning. We also thank all of the participants who participated in the study.

Funding:

This research has benefited from the financial support of IDEXLYON from Université de Lyon (project INDEPTH) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005), and of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by the French National Research Agency, and the grant from the Agence Nationale pour la Recherche to JCD (ANR-21-CE37-0032). This work was also supported by the National Natural Science Foundation of China (No. 31970982 and No. 32171019). J.L was supported by the China Scholarship Council.

Author contributions:

Conceptualization: J.L, C.Q, S.L, J.D

Methodology: J.L, R.P, S.L,

Investigation: J.L, S.L,

Visualization: J.L

Supervision: C.Q, J.D

Writing—original draft: J.L, R.P,

Writing—review & editing: J.L, E.D, B.C, J.D

Declarations of interests:

Authors declare that they have no competing interests.

Data and materials availability:

All data needed to evaluate the conclusions in the paper are present in the paper, the Supplementary Materials, and the OSF repository: https://osf.io/wjbpz/?view_only=d9ed0a26788b45a98e51e367059e0588

Figures

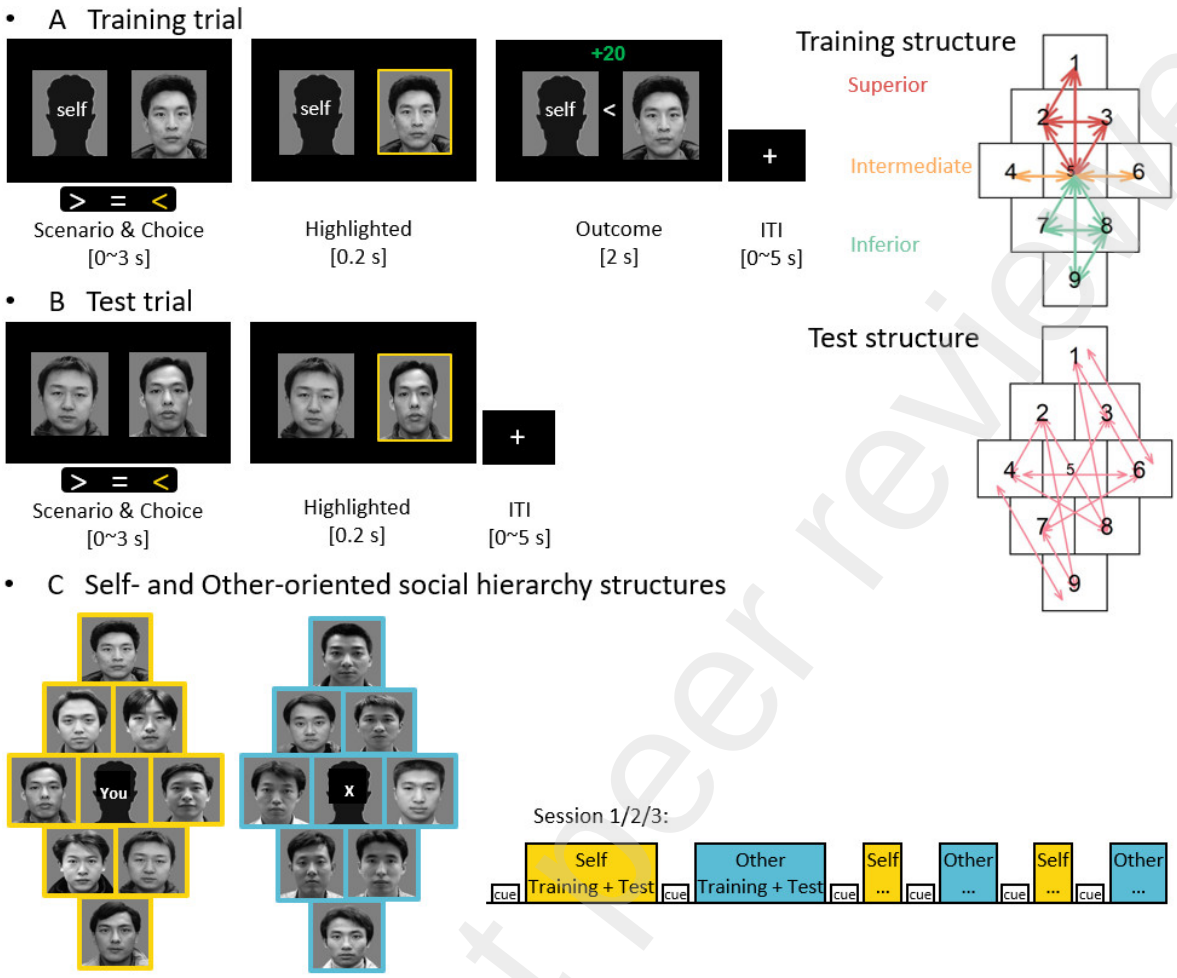

Fig. 1. Task structure. (A) On each training trial, two members of the social hierarchy were presented. Participants indicated their hierarchical relationship from the three available options: Left is superior (>), equal status (=), or right is superior (<) within 3 s. After the decision, the higher rank item's border was highlighted for 0.2 s (if "=" was pressed, both borders lighted up at the same time), and the outcome feedback was presented for 2 seconds subsequently: right (+20) or wrong (-20). Each trial finished with an ITI showing a jittered fixation (1–5 s). The right panel shows the relationships between members of the social hierarchy structure which were presented together, composing 12 training trials: 5 superior pairs (red line: P1 vs P2, P2 vs P3, P1 vs P5, P2 vs P5, P3 vs P5), 5 inferior pairs (green line: P7 vs P8, P8 vs P9, P7 vs P5, P8 vs P5, P9 vs P5) and 2 intermediate pairs (orange line: P4 vs P5, P6 vs P5). (B) Test trial blocks were presented after training trial blocks. New pairs of items, that were not in the training trial blocks were presented in a similar way, and participants attempted to infer the hierarchical relationships, however, there was no feedback. The right panel shows the social hierarchy structure contained 13 test trials (pink line: P1 vs P3, P2 vs P4, P3 vs P6, P1 vs P6, P4 vs P8, P6 vs P7, P7 vs P9, P4 vs P9, P2 vs P8, P3 vs P7, P1 vs P8, P2 vs P9, P4 vs P6). (C) The fMRI task included Self-orientated and Other-orientated social

hierarchy structures in which the central position (P5) is labelled as "You" or "X" respectively, along with the corresponding color (yellow or blue). However, these structures were never shown to the participants, and are presented here only to explain the experiment. The right panel illustrates one fMRI session, which comprised 3 blocks (integrated Self- and Other-orientated blocks, Self: 3 mini blocks, Other: 3 mini blocks). It was indicated in text on the cue screen whether the mini block was Self- or Other-orientated at the start of the block. (See Methods for a detailed description).

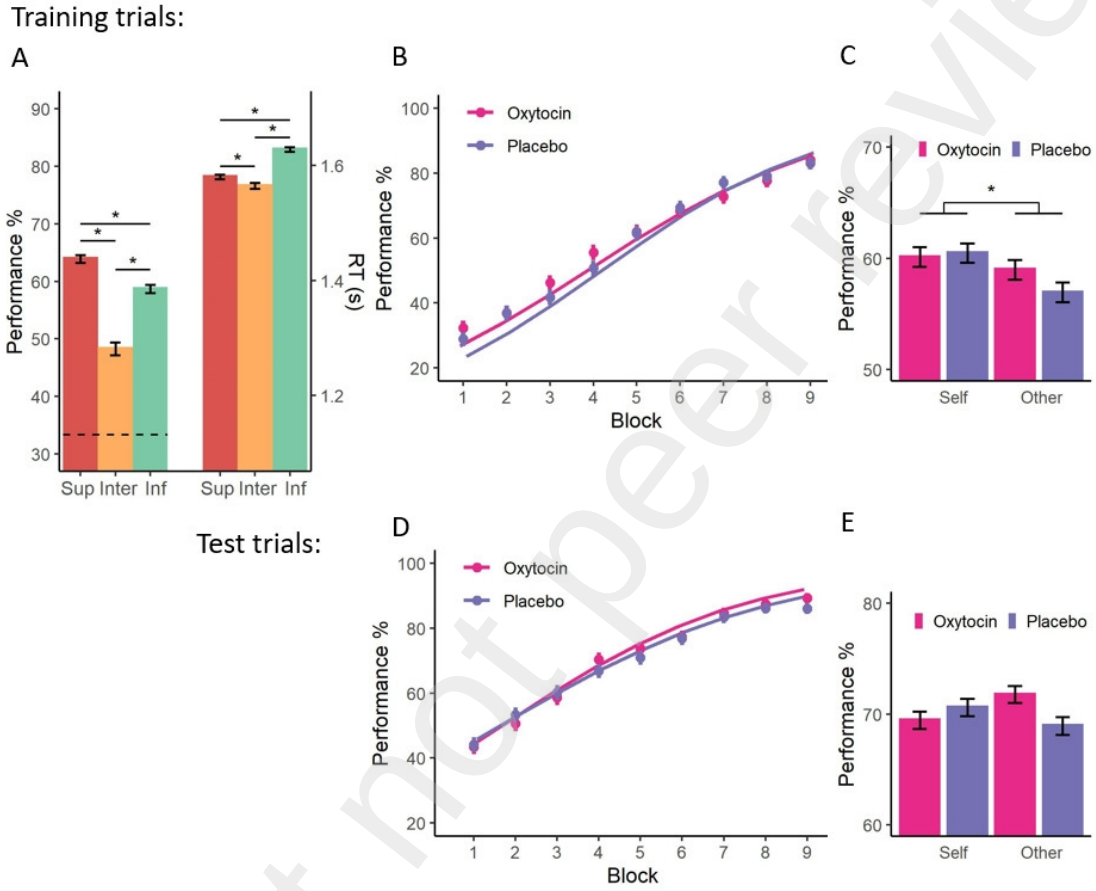

Fig. 2. Behavioral results. (A) The graphs show significant main effect for hierarchy in performance accuracy and response times (RT) during training trials by GLMM1 (performance) and LMM1 (RT) for superior, intermediate and inferior paired items. The dotted line indicates the performance level of random choice at 1/3. (B) Learning curve showing the mean performance of Oxytocin and Placebo groups in training trials, averaged across within-group stimuli. Participants in Placebo group were more likely learned better than Oxytocin group across time (GLMM1). (C) For illustration purposes we indicate mean performance of interaction between Groups and Hierarchy orientation, but only the Hierarchy orientation effect is significant. Performance was more likely better for the Self- than for the Other-orientated conditions (GLMM1). (D) Learning curve showing the mean performance for the Oxytocin and Placebo groups in test trials, averaging across within-group stimuli. Participants in the Oxytocin group were more likely to be

perceived as learning inferences better across time than the Placebo group (GLMM2). (E) The interaction effect of Group × Hierarchy Orientation in test trials showed that Placebo-treated participants performed better in the Self- than the Other-orientated hierarchical inference learning condition. In contrast, performance for the Other- was better than for the Self-orientated hierarchy in the Oxytocin treated group (GLMM2). Error bars correspond to SEM, lines refer to estimated performance rate.

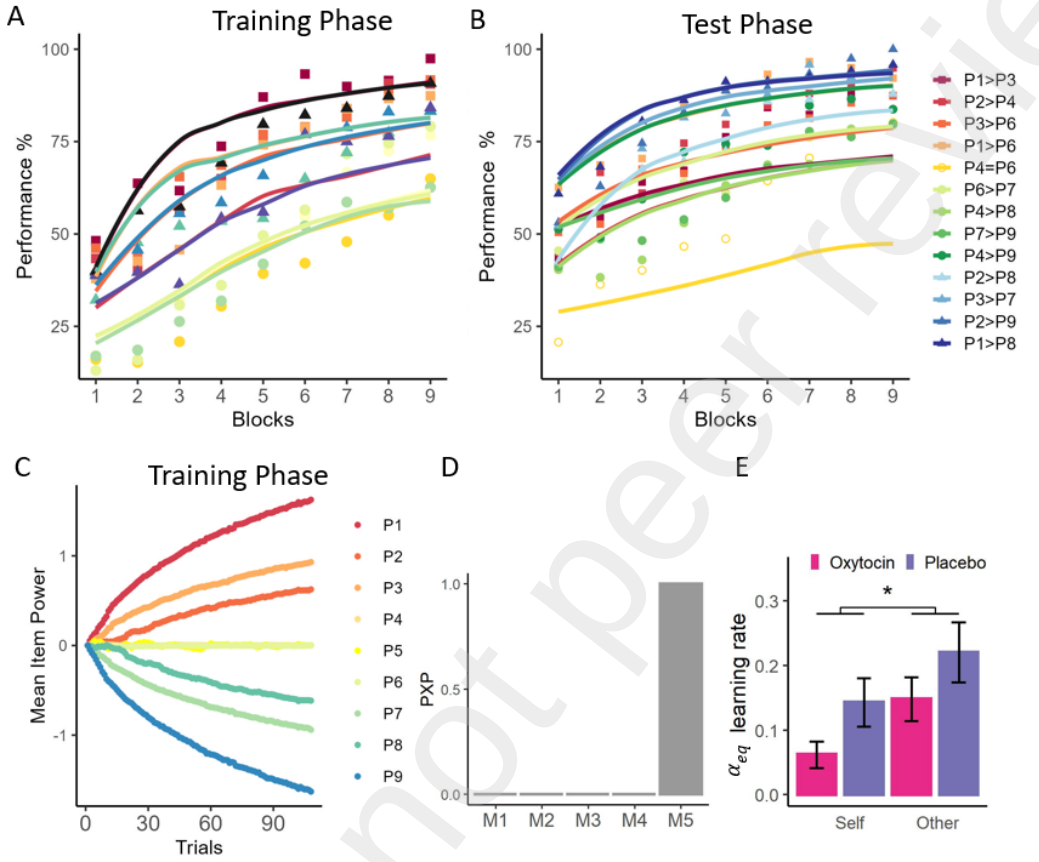

Fig. 3. Model validation on RL-ELO-based associative learning model and model comparison. (A) Learning curves, pooled over all factors estimated by the computational model with the probability of items being correct during training trials. Dots represent mean performance results of participants on that specific trial. (B) The plot shows the mean item power changed trial-by-trial, captured by the computational model, which indicates the difference between the expectation over the estimated power of the presented item. (C) Similar learning curves analysis but estimated for test trials. (D) Bayesian model selection in terms of the protected exceedance probabilities (PXP) with all subjects, which indicates that the combinatorial models based on RL-ELO and associative learning (M5) estimates decisions in hierarchy relationship better than other models: RL-ELO model with 1 alpha (M1); RL-ELO model with 3 alphas (M2); Rescorla-Wagner model (M3); Value transfer model (M4). (E) The learning rate $\boldsymbol{\alpha_{eq}}$ of equal trials, modeled by the RL-ELO associative learning model, is higher for the Other- than Self- oriented

hierarchy, and marginally significantly different with respect to the treatment group effect (Oxytocin or Placebo).

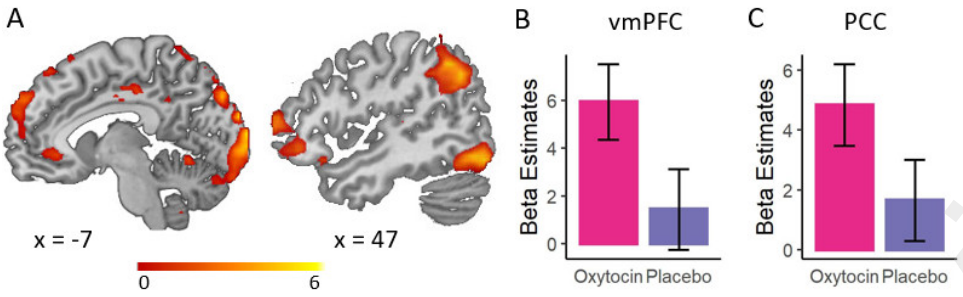

Fig. 4. Probability to choose the correct answer in training trials. (A) Whole-brain analysis revealed a network of brain regions encoding probability to choose the correct answer regardless of group (GLM 1). The regions of interest (ROI) of vmPFC and PCC were defined based on a meta-analysis (see methods). (B and C) Illustration of beta estimates in vmPFC and PCC and show a significant difference between Oxytocin and Placebo groups. Display threshold: $P < 0.001$ uncorrected at the voxel-level with $k >= 40$. Error bars show SEM. The activation figures hereinafter were made in the same way unless stated otherwise. (vmPFC = ventromedial prefrontal cortex, PCC = posterior cingulate cortex).

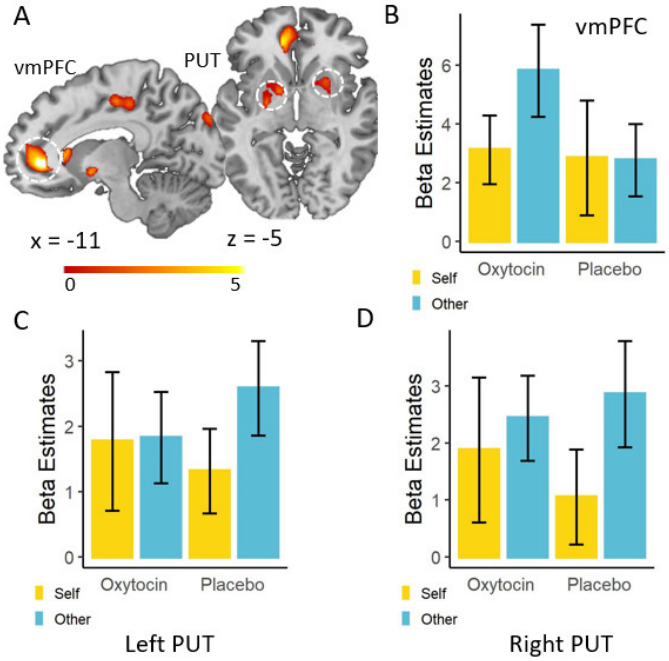

Fig. 5. Prediction error value for the winning model in training trials. (A) Whole-brain analysis revealed a network of brain regions encoding prediction error regardless of group (GLM 2). (B, C & D) Illustration of beta estimate values in vmPFC, and left and right putamen. These regions showed significant activity, but no significant differences between conditions (vmPFC: ventromedial prefrontal cortex, PUT: putamen).

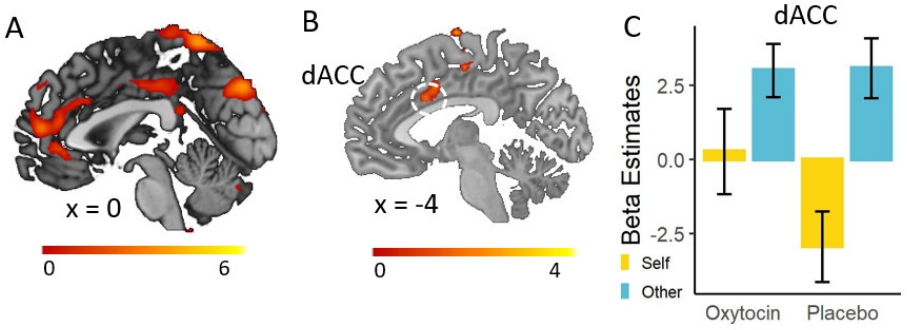

Fig. 6. Probability to choose the correct answer of the winning model in test trials. (A) Whole-brain analysis revealed a network of brain regions encoding the probability to choose the correct answer regardless of group by GLM 3. (B &C) Neural responses are significantly stronger to the Other- than to the Self-orientated hierarchy in the dACC (dorsal anterior cingulate cortex).

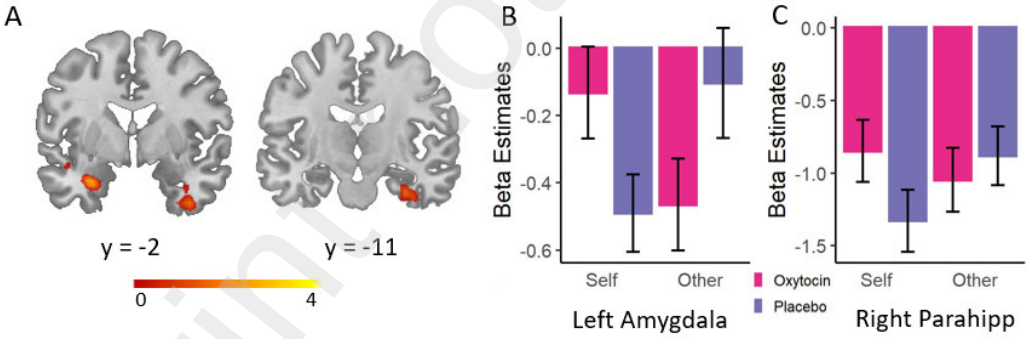

Fig. 7. Brain areas tracking the transitivity process in test trials. (A) Interaction effect of Group × Hierarchy Orientation in test trials showed that BOLD deactivations in the left amygdala and right inferior parahippocampal gyrus were stronger during the Self- than Other-orientated hierarchy trials in the Placebo group, while in Oxytocin group the activity was stronger for the Other- than for the Self-orientated hierarchy when tracking the transitive inference process. (B & C) Illustration of activities in amygdala and Parahippocampus showing significant differences between Oxytocin and Placebo groups.

Supplementary results

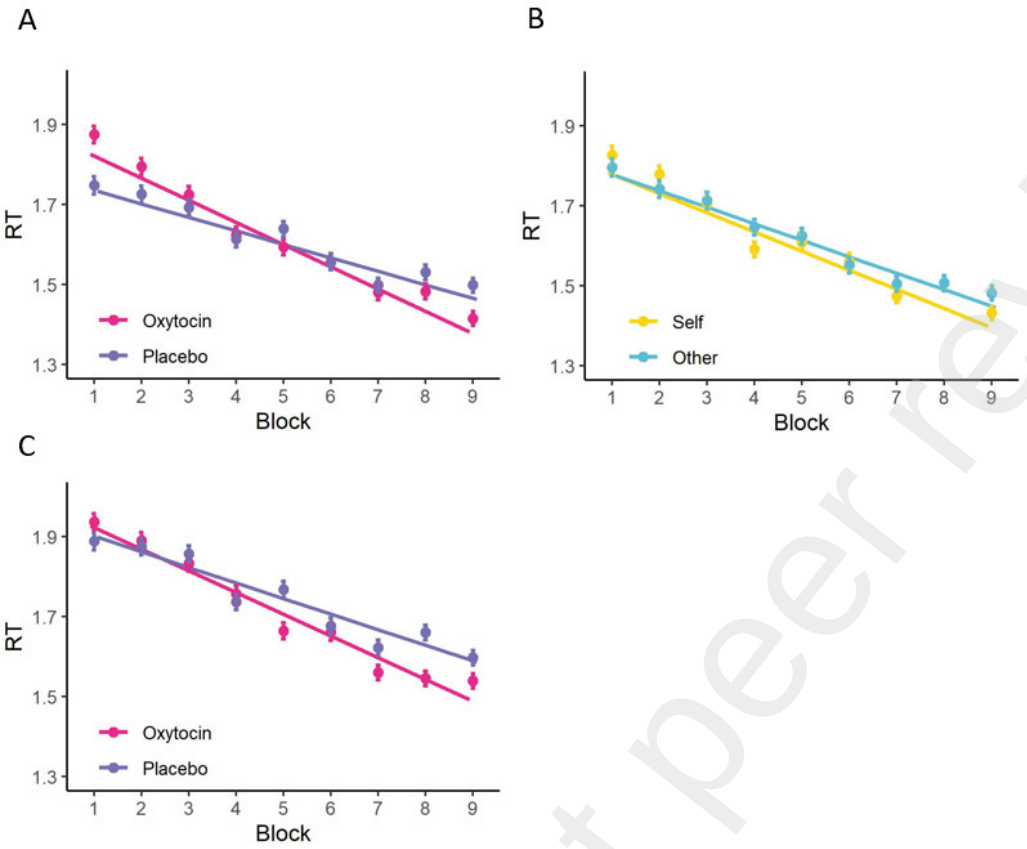

Fig. S1.

Behavior results of response time (RT).

(A) RT in oxytocin group declines faster than in placebo group during learning trials.

(B) RT in Self-oriented hierarchy structure declines faster than in Other-oriented hierarchy structure during learning trials.

(C) RT in oxytocin group declines faster than in placebo group during test trials.

Table S1.

Performance results of generalized linear mixed-effects model and response time (RT) results of linear mixed effect model in training trials.

	Performance			RT		
Predictor	Odds ratios	SE	z	Beta	SE	t
(Intercept)	0.234 ***	0.022	-15.472	1.824 ***	0.03	60.595
Group (PB)	1.226	0.221	1.135	0.114	0.06	1.905
S_O (Other)	1.237 ***	0.056	4.711	0.014	0.019	0.762
Block	1.433 ***	0.014	37.503	-0.045 ***	0.002	-27.667
Hierarchy (Superior)	3.392 ***	0.471	8.79	0.032 **	0.012	2.637
Hierarchy (Inferior)	3.156 ***	0.437	8.29	0.128 ***	0.012	10.592
Group × Block	0.966 *	0.016	-2.038	-0.022 ***	0.003	-6.763
S_O × Hierarchy (Superior)	0.543 ***	0.066	-5.06	0.082 **	0.024	3.401
S_O × Hierarchy (Inferior)	0.616 ***	0.074	-4.05	0.052 *	0.024	2.154
Block × Hierarchy (Superior)	0.92 **	0.023	-3.259			
Block × Hierarchy (Inferior)	0.882 ***	0.022	-5.01			
S_O × Block				-0.007 *	0.003	-2.125

Notes: contrast coding of Group: 0.5 = Oxytocin, -0.5 = Placebo (reference level); contrast coding of S_O: 0.5 = Self, -0.5 = Others (reference level); contrast coding of rank: $\frac{1}{3}$ = Superior, $\frac{1}{3}$ = Intermediate (reference level), $\frac{1}{3}$ = Inferior. Block was coded as ordinal levels. Significance: $*$ p < 0.05, ** p < 0.01, *** p < 0.001.

Table S2.

Performance results of generalized linear mixed-effects model and response time (RT) results of linear mixed effect model in test trials.

	Performance			RT		
Predictor	Odds ratios	SE	z	Beta	SE	t
(Intercept)	0.59 ***	0.057	-5.416	1.959 ***	0.032	61.594
Group (PB)	0.881	0.172	-0.649	0.045	0.064	0.705
S_O (Other)	1.021	0.042	0.511	0.09 ***	0.019	4.821
Block	1.373 ***	0.012	36.518	-0.047 ***	0.002	-28.17
Group × Block	1.04 *	0.018	2.268	-0.015 ***	0.003	-4.672
Group × S_O	1.262 **	0.104	2.82			
S_O × Block				-0.014 ***	0.003	-4.077

Notes: Reference level of Group: 0.5 = Oxytocin, -0.5 = Placebo; Reference level of S_O: 0.5 = Self, -0.5 = Others. Significance: ∗ $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table S3.

Parameters of combined RL-ELO-based associative learning model.

Mean (SEM)								
		α_{asso}	α	α_{eq}	β	δ_1	δ_2	γ
Oxytocin	Self	0.55	0.35	0.06	0.36	0.59	0.00	-0.55
		(0.01)	(0.03)	(0.02)	(0.08)	(0.06)	(0.00)	(0.10)
	Other	0.54	0.37	0.15	0.38	0.50	0.00	-0.57
		(0.01)	(0.02)	(0.03)	(0.09)	(0.08)	(0.00)	(0.10)
Placebo	Self	0.55	0.41	0.14	0.24	0.63	0.00	-0.53
		(0.01)	(0.02)	(0.04)	(0.09)	(0.06)	(0.00)	(0.08)
	Other	0.55	0.34	0.22	0.16	0.58	0.00	-0.64
		(0.01)	(0.02)	(0.05)	(0.11)	(0.07)	(0.00)	(0.09)

Note: SEM means standard error of the mean.

Table S4.

Brain activity related to GLMs 1-4.

Brain Region	Hemisphere	Cluster Size	T-value	MNI			BA
				x	y	z	
Regions encoding the value of chosen correct option in training phase							
(pooled all conditions; GLM1)							
Positive modulation in all subjects:							
CB/OG* **	L	4311	7.34	-33	-75	-21	7/17/19
MTG/IPL* **	L	1397	7	-66	-39	0	21/39/40
SFG/MFG* **	L	1830	5.77	-27	27	54	8
MFG/IFG* **	L	492	5.65	-39	60	-3	10/47
TG* **	R	368	4.78	60	-24	-21	21/22
vmPFC **	B	88	4.58	-12	36	-9	10
MCC **	B	131	4.32	6	-30	42	31/24
Negative modulation in all subjects:							
LG*	L	62	4.34	-24	-69	-6	19
OXT > PB:							
PCC (p=0.052)	R	56	4.41	15	-24	42	31
Positive modulation in Self-orientated condition:							
CB/OG* **	L	1342	6.39	-30	-75	-21	16/18
VS/Hipp* **	L	186	5.67	-24	-27	9	27/13
PCUN* **	B	537	5.58	3	-75	57	7/13
CB/OG* **	R	287	5.1	45	-78	-18	18/19
MFG* **	R	366	5.06	48	51	9	8/16
MFG* **	L	161	4.96	-30	24	54	8
MTG*	L	61	4.7	-66	-39	0	21
AG* **	R	181	4.69	48	-57	42	40
AG* **	L	397	4.64	-42	-51	36	41
IFG*	L	71	4.46	-30	30	-15	47
Positive modulation in Other-orientated condition:							
TG* **	L	236	5.23	-66	-39	0	21
LG* **	R	224	4.86	9	-96	-3	17

IFG**	L	105	4.82	-39	36	-9	47
SFG***	L	251	4.47	-15	42	45	9
CB/OG***	L	206	4.35	-39	-72	-24	19/37
AG***	L	168	4.12	-57	-57	30	39/40
AG*	R	82	4.08	57	-57	24	39/40
vmPFC*	B	80	3.92	3	30	-12	11
OG/PCUN*	L	76	3.74	-12	-87	33	7
Positive modulation in Oxytocin group:							
AG/TPJ/TG***	L	1620	6.21	-54	-33	0	40/22/21
AG/TPJ/TG***	R	829	4.91	51	-30	3	40/22/21
PCC***	B	683	4.84	9	-33	51	31
MFG/IFG***	R	239	4.71	48	51	9	46/47
SFG***	L	319	4.56	-15	42	51	9
TG**	R	135	4.49	63	-9	-12	21/38
MFG/SFG**	R	122	4.47	36	39	45	8
SFG*	R	57	4.43	12	33	54	6
PCUN*	R	63	4.38	6	-96	0	17
ACC**	L	86	4.1	-9	39	0	24
vmPFC**	B	90	4	-3	57	12	10
OG*	L	79	3.72	-39	-84	-12	37
Positive modulation in Placebo group:							
CB/PCUN***	L	2156	7.66	-30	-78	-21	7/19
SFG**	L	87	4.38	-12	63	30	10
AG*	R	59	4.14	45	-63	30	40
OFC**	L	98	4.13	-39	54	-3	46
Negative modulation in Placebo group:							
IFG*	L	76	5.32	-45	3	24	9
LG*	L	80	5.04	-24	-69	-3	19
LG*	R	60	4.41	27	-51	-9	19
Regions encoding the value of prediction error in training phase							
(all conditions pooled, GLM2)							
Positive modulation in all subjects:							
vmPFC**	B	181	5.74	-12	42	0	32
Cau*	R	106	5.2	18	-6	27	

Put*	L	101	4.91	-24	-12	-3	
Put**	R	131	4.84	30	3	15	
OG*	B	124	4.51	-3	-90	33	19
Negative modulation in all subjects:							
MFG/SFG***	R	410	5.92	42	21	30	8/9
MFG/IFG***	L	458	5.82	-39	21	27	46/9
SFG/SMA***	L	253	5.76	-6	9	57	6/8
IFG/Ins*	R	94	5.33	33	21	0	47
IPL***	L	243	5.04	-36	-45	45	40
Negative modulation in Self-orientated condition:							
Thalamus*	R	117	4.46	9	-24	0	
MFG***	L	265	4.41	-39	24	27	9
MFG*	R	118	4.24	42	21	27	9
Positive modulation in Other-orientated condition:							
vmPFC*	B	108	5.15	-12	39	-3	32
Putamen*	R	93	4.11	27	9	0	
Negative modulation in Other-orientated condition:							
MFG*	R	129	4.7	42	24	27	8
MFG/IFG*	L	122	4.3	-39	21	27	
Negative modulation in Oxytocin group:							
IPL*	L	92	4.1	-36	-45	45	40
Negative modulation in Placebo group:							
MFG**	R	183	5.05	42	21	30	8
SFG/SMA***	R	260	4.9	9	24	48	6
MFG/IFG**	L	195	4.52	-42	15	30	9
IFG/Ins*	R	87	4.29	39	24	0	47
Regions encoding the value of chosen correct option in test phase							
(all conditions pooled, GLM3)							
Positive modulation in all subjects:							
IFG/MFG**	L	125	6.11	-51	30	0	47
PCUN***	R	542	5.85	9	-84	39	19

VS* **	R	270	5.76	30	-21	3	13
MTG* **	R	357	5.42	60	-42	-3	21
Ins **	L	108	5.27	-30	-6	3	13
SFG* **	R	608	5.25	9	51	42	9
TPJ* **	L	495	5.06	-54	-42	24	40
TPJ* **	R	248	5.01	51	-51	30	40
IFG/MFG* **	R	242	4.79	48	39	-3	46/10
CB* **	L	287	4.73	-33	-75	-24	
CB/OG* **	R	341	4.53	21	-93	-12	18
PCC* **	B	278	4.13	12	-42	42	31
Negative modulation in all subjects:							
LG **	L	132	5.76	-15	-78	-3	18
MedFG* **	B	271	5.08	-9	12	51	6
LG*	R	82	4.59	18	-75	-3	18
Other > Self:							
dACC*	R	94	4.29	15	6	45	33
PreCG*	R	78	4.21	18	-15	57	6
IPL*	R	104	3.76	57	-27	33	40
Positive modulation in Self-orientated condition:							
IFG **	L	109	5.85	-48	30	0	47
MFG*	R	99	5.18	48	39	-3	47
SFG **	R	114	4.5	9	54	42	9
AG*	R	88	4.35	48	-51	30	39
TG **	R	122	4.29	66	-33	-6	20/21
MTG **	L	107	4.17	-48	-39	0	21
SFG*	L	96	4.17	-15	27	60	9
Negative modulation in Self-orientated condition:							
Ins*	R	82	6.18	33	18	12	47
MCC/SFG* **	L	383	5.38	-24	0	54	24/6
Positive modulation in Other-orientated condition:							
VS/Amy* **	R	447	6.01	30	-18	6	13
OG* **	R	321	5.85	36	-81	-21	19
PCUN* **	R	1342	5.77	9	-84	39	19
LG* **	L	254	5.57	-15	-105	-9	18

TPJ* **	L	662	5.13	-51	-39	24	40/41
PoCG* **	R	327	5.13	60	-24	33	2/40
TG **	R	132	4.76	63	-42	0	21
TG/OG*	L	99	4.41	-54	-57	6	37
Negative modulation in Other-orientated condition:							
LG*	L	73	4.7	-12	-75	-3	18
Positive modulation in Oxytocin group:							
TG **	L	110	5.24	-51	-36	6	22
PCUN*	R	100	4.83	6	-78	36	7
TPJ/Ins*	L	89	4.67	-54	-42	24	40/42
VS*	R	83	4.64	30	-6	6	
PCC* **	B	321	4.45	3	-33	33	31
Positive modulation in Placebo group:							
TG **	R	170	4.96	63	-42	-3	21
SFG **	R	171	4.84	12	54	42	9
OG* **	R	246	4.79	15	-87	36	19
OG* **	L	314	4.63	-42	-84	-12	37
ACC **	B	106	4.2	3	36	-24	11
CB*	R	104	4.06	42	-75	-24	
Negative modulation in Placebo group:							
LG*	L	87	5.39	-15	-75	-3	18
Others - Self in Placebo group:							
SFG/PreCG* **	B	436	4.62	6	-3	78	6/31
LG/OG*	R	84	4.29	24	-96	-9	18
CB*	L	73	4.26	-27	-81	-39	
Regions encoding the value of inferring social hierarchy in test phase (GLM4)							
Oxytocin_(Self – Other) > Placebo_(Self – Other):							
ITG **	L	131	4.27	30	12	-30	38

Note: Regions shown here met the uncorrected voxel-level threshold of $p < 0.001$ with $k = 40$. With the above uncorrected threshold as the cluster-defining threshold, clusters met the Family-

Wise Error corrected cluster-level (cl-FWE) threshold were marked: *p < 0.05, **p < 0.01, ***p < 0.001.

Coordinates shown here were based on Montreal Neurological Institute (MNI) coordinate system.

Abbreviations: B: bilateral, L: left, R: right, BA: Brodmann Area; ACC: anterior cingulate cortex, AG: angular gyrus, Amy: amygdala, Cau: caudate, CB: cerebellum, dACC: dorsal anterior cingulate cortex, FG: fusiform gyrus, IFG: inferior frontal gyrus, Ins: insula, IPL: inferior parietal lobule, ITG: inferior temporal gyrus, LG: lingual gyrus, MCC: middle cingulate cortex, MedFG: medial frontal gyrus, MFG: middle frontal gyrus, ,MTG: middle temporal gyrus, OFC: orbitofrontal cortex, OG: occipital gyrus, PCC: posterior cingulate cortex, PCUN: precuneus, PoCG: post-central gyrus, PreCG: precentral gyrus, Put: putamen, SFG: superior frontal gyrus, SmG: supramarginal gyrus, SMA: supplementary motor area, TG: temporal gyrus, TPJ: temporal parietal junction, vmPFC: ventromedial prefrontal cortex, VS: ventral striatum.

Table S5.

Various scales' scores on Pre - & Post - scanning task.

Scales	Oxytocin Mean	Placebo Mean	t	p
Before the fMRI task				
BDI	29.48 (6.07)	29.62 (7.30)	-0.078	0.938
STAI	77.62 (13.77)	75.55 (16.82)	0.513	0.61
STAI _SAI	36.38 (7.22)	34.76 (9.37)	0.738	0.464
STAI _TAI	41.24 (7.98)	40.79 (8.78)	0.204	0.839
PANAS	52.83 (7.47)	50.21 (9.26)	1.186	0.241
PANAS _PAS	32.97 (5.52)	30.90 (5.77)	1.396	0.168
PANAS _NAS	19.86 (5.41)	19.31 (7.52)	0.321	0.75
After the fMRI task				
STAI _SAI	35.97 (9.45)	33.28 (8.77)	1.124	0.266
PANAS	49.24 (9.07)	44.69 (9.27)	1.89	0.064
PANAS _PAS	31.76 (7.77)	29.10 (7.53)	1.321	0.192
PANAS _NAS	17.48 (5.51)	15.59 (6.27)	1.224	0.226
SDO	3.48 (1.58)	3.60 (1.79)	-1.053	0.292
SDO_GBD	3.85 (1.66)	4.22 (1.74)	-2.298	0.022 *
SDO_OEQ	3.11 (1.40)	2.99 (1.62)	0.89	0.374
INCOM	38.03 (-4.34)	38.31 (-5.53)	-0.211	0.833
INCOM _Abilities	21.03 (-2.74)	20.45 (-2.8)	0.805	0.424
INCOM _Opinions	17 (-2.24)	17.86 (-3.44)	-1.131	0.264
AMS	5.9 (17.07)	2.79 (12.01)	0.801	0.427

Table S6.

Post-experimental questionnaire.

Questions	Oxytocin Mean (SD)	Placebo Mean (SD)
Please rate other's willingness to partner with someone in another Department (from your perspective) *	0.43 (1.65)	0.33 (1.78)
Please rate your willingness to partner with someone in another Department	0.48 (1.76)	0.43 (1.86)
Please rate perceived social distance with each colleague in your company	47.59 (43.71)	44.44 (25.30)
Guess whether you received oxytocin ① or placebo ②	1.38 (0.49)	1.48 (0.51)
Do the characters in your department match the ones you imagine?	2.07 (0.70)	2.10 (0.77)
Do the characters in other's department match the ones you imagine?	1.93 (0.65)	1.97 (0.68)
How would you grade your relationship with those at the top of your department?	4.24 (0.64)	4.31 (0.60)
How would you grade your relationship with those at the top of the other's department?	4.14 (0.95)	4.17 (0.80)
How would you grade your relationship with those at the lowest level of your department?	1.76 (0.91)	1.38 (0.68)
How would you grade your relationship with those at the lowest level of the other's department?	1.62 (0.82)	1.41 (0.68)
How satisfied do you feel intuitively about your company?	3.31 (0.93)	3.24 (0.83)

Table S7.

Model parametric in RL-Elo associative learning model using repeated- measures ANOVAs.

Parametric	Factor	$F_{(1,56)}$	p	η^2
α_{asso}	Group	0.25	0.62	0.00293
α_{asso}	S_O	0.00	0.95	0.00003
α_{asso}	Group×S_O	0.13	0.72	0.00085
α_{eq}	Group	3.53	0.07	0.03946
α_{eq}	S_O	7.42	0.01	0.04412
α_{eq}	Group×S_O	0.02	0.88	0.00013
α	Group	0.30	0.58	0.00341
α	S_O	1.69	0.20	0.01093
α	Group×S_O	4.34	0.04	0.02766
β	Group	2.31	0.13	0.03025
β	S_O	0.21	0.65	0.00091
β	Group×S_O	0.69	0.41	0.00301
δ_1	Group	0.49	0.49	0.00592
δ_1	S_O	1.59	0.21	0.00900
δ_1	Group×S_O	0.08	0.78	0.00047
δ_2	Group	3.01	0.09	0.02588
δ_2	S_O	1.13	0.29	0.01015
δ_2	Group×S_O	0.03	0.86	0.00030
γ	Group	0.05	0.82	0.00075
γ	S_O	1.06	0.31	0.00428
γ	Group×S_O	0.65	0.42	0.00262

Note: We examined differences between Self/Other hierarchy conditions (S_O) in the model parametric at the group level using repeated-measures ANOVAs.

Oxytocin modulates the neurocomputational mechanisms engaged in learning social hierarchy

Authors

Jiawei Liu[1,2,3]†, Chen Qu[3]†, Rémi Phillipe [1,2], Siying Li[3,4], Edmund Derrington[1,2], Brice Corgnet[5], Jean-Claude Dreher[1,2]*

[1] Laboratory of Neuroeconomics, Institut des Sciences Cognitives Marc Jeannerod, CNRS, Lyon, France.

[2] Université Claude Bernard Lyon 1, Lyon, France.

[3] Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, China; School of Psychology, Center for Studies of Psychological Application, and Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, China.

[4] Faculty of Education, Northeast Normal University, Changchun, China.

[5] EM Lyon Business School, and CNRS, GATE Lab, UMR 5824, Ecully, 69130, France.

† These authors contributed equally to this work.

* Lead contact, Correspondence: dreherjeanclaude@gmail.com