Neural Response Effects of Oxytocin and Vasopressin on Human Learning for Social Cooperation

João Simões, joaomiguelsimoes7@tecnico.ulisboa.pt[1]

[1]Instituto Superior Técnico, University of Lisbon, Portugal

ABSTRACT

Background: Oxytocin (OT) and vasopressin (AVP) are neuropeptides thought to have essential roles during social interactions, while interacting with brain regions of the dopaminergic system, which influences reinforcement learning (RL) mechanisms.

Objective: Investigate the roles of OT and AVP on social RL and its neural correlates, in different social contexts, and examine the added effect of participant's sex.

Methods: Participants (148 men; 144 women) randomly received intranasal OT, AVP or placebo (PB) and played the prisoner's dilemma game. Behavioural data was modelled using computational RL models, model parameters were analyzed, and trial-by-trial Reward Prediction Error (RPE) signals were correlated with whole-brain and region-of-interest (ROI) brain activation.

Results: OT increased the α_C (learning rate) parameter in women playing with human, compared with computer partners. Generally, OT promoted a higher $Q0$ parameter (initial bias) in games with human than computer partners. AVP increased $Q0$ in men playing with humans compared to OT, and when playing with computers compared to PB. Both amygdala and caudate regions revealed a higher RPE-brain activation correlation in women under OT playing with computer than human partners. Only in the amygdala ROI, AVP increased the RPE correlation in males playing with human compared to computer partners.

Conclusions: OT may promote a pro-self bias at the beginning of a social dilemma interaction, while increasing impulsive behaviours in women after their cooperation. Contrarily, AVP might promote a pro-social bias in males. Additionally, OT may enhance social learning in women during non-social contexts; and AVP might enhance social learning in males during social contexts.

***Index Terms*—**Reinforcement learning; reward prediction error; prisoner's dilemma; neuropeptides; striatum; amygdala.

I. INTRODUCTION

OXYTOCIN (OT) is a neuropeptide produced by the human brain, which plays an essential role during social interactions. Early studies have reported that this neuropeptide enhances the human prosocial behaviour [1], promoting generosity [2], cooperation [3], stress reduction and communication during conflict situations [4], and increasing trust [5], (even after betrayal [6]), emotion recognition [7], attention and memory of positive social stimuli [8]. However, recent studies have revealed a different role of OT, with its effects being context-dependent, increasing aggression [9], envy, gloating [10] and ethnocentrism, leading to in-group favoritism [11]. In fact, these behavioural effects are related to OT neuronal modulation [6]. OT is produced by the hypothalamus, specifically by the neurons of the paraventricular and supraoptic nuclei [12], being later secreted to other brain regions, namely the hippocampus, the brainstem, the amygdala, the striatum, and many others [13], not only associated with emotional and behavioural functions, but also with reward [14]. These regions also belong to the dopaminergic system, which has important roles in motivation, reward, reproductive, maternal behaviours, and reinforcement learning (RL) [15]. Thus, those two systems influence and interact with each other, with previous studies showing that OT has multiple binding sites within the dopaminergic system [16].

Due to these findings, new hypotheses started to arise, regarding the possibility of OT having a role in the RL process in social contexts. The striatum, a region common to both neuropeptide systems, is known to signal reward and prediction, influencing future behaviour [17]. Striatal activation is also associated with the processing of reward prediction error (RPE) signals, which are produced through phasic activation of dopamine (DA) neurons [18], with RPE being the difference between a reward and its prediction. In fact, only two studies have started to investigate this role: *Ide et al.* [19], by using computational modelling and trial-by-trial RPEs, found that OT attenuated the RPEs encoding during a social interaction; *Kruppa et al.* [20], by using similar methods, found that intranasal OT enhanced social RL in patients with autism spectrum disorder (ASD).

Another neuropeptide, vasopressin (AVP), has high similarities with OT, not only being synthesized in the hypothalamus and secreted to the same brain areas [13], but also presenting similar affinity for the same receptors [21]. Socially, AVP promotes mutual cooperation [22], increases awareness and memory for social behaviours [23], conciliatory gestures in women [24] and the likelihood of reciprocating cooperation among men [25]. As OT's, AVP's effects are context-dependent, offsetting male-aggression in affiliative contexts [26] and increasing agnostic facial motor patterns in men when unfamiliar male faces were shown, reducing the friendliness recognition for those faces, while increasing the friendliness recognition in women when female faces were shown [24]. Furthermore, AVP is also reported to interact with the DA system [27], so this neuropeptide might also have a role in the RL process in social contexts [28], with no study having researched it yet.

As during interactions, the effects of OT and AVP on the RL process might be conditioned by the social context itself,

i.e., environmental (*e.g.*, person or group who is acting with the individual) and interindividual factors (*e.g.*, the sex of the participant) may have an important role in this process, which has not been studied yet. To study those effects, the prisoner's dilemma (PD) game can be employed. In this game, two players play with each other, in order to elicit relationships established on reciprocal altruism [25]. In each trial, the two players can choose to cooperate or defect independently, and, by the end of it, each player receives a payoff based on the two decisions, with several emotions being evoked according to the four possible outcomes [25].

Previous findings showed that L-DOPA, a DA precursor, increased the positive correlation between RPEs and the nucleus accumbens (NAcc), a region of the striatum [29], and that amisulpride, a DA receptor antagonist at a high dose, and memantine, an antagonist of NMDA receptors (which also regulate DA neuronal activity [30]), decreased the positive correlation between RPEs and the striatum [31]. Thus, by assuming that L-DOPA and OT provide a similar effect on the RL circuit, while amisulpride and memantine would have opposite effects in comparison to OT, it was expected that OT increased the RPE-striatal activation correlation. Regarding the social context, *Kruppa et al.* [20] found that OT increased the positive correlation between RPEs and the left NAcc during tasks with social feedback representation in individuals with ASD. Even though these results were obtained in ASD patients, previous evidence not only showed that the striatum region has a major importance on the RL process [32], but also that social information facilitates learning [33]. Herein, it was expected that the social context (*i.e.*, playing with a human partner) increased the previous hypothetical effect.

With this in mind, the present study aims to investigate the role of OT and AVP on social RL, and its neural correlates, in different contexts (*i.e.*, with different partner types, human or computer) and examine the added effect of participant's sex, using an existing behavioural and neuroimaging dataset in an RL model-based approach (commonly employed in DA research). This dataset was acquired from a placebo-controlled, randomized and double-blind study using intranasal OT and AVP during the PD game.

II. Methods

A. Participants and Drug Administration

The present study is based on a sample of 153 men and 151 women from the Emory University community, with ages between 18 and 22 years, acquired from previous reports following the same procedures [25] [34] [35] [36]. From that sample, only the data of 292 participants could be used (148 men with a mean age of 20.2 and a standard deviation of 1.3 years, and 144 women with a mean age of 20.2 and a standard deviation of 1.3 years), due to acquisition problems and missing data.

The 292 participants were randomized to be treated with either placebo (PB) ($n = 52$ for men and $n = 49$ for women), intranasal OT ($n = 48$ for men and $n = 47$ for women) or intranasal AVP ($n = 48$ for both men and women). According to studies reporting social cognitive behavioural effects previous to the data of collection, a dose of 24 IU of OT (Syntocinon-Spray, Novartis, Basel, Switzerland) for the OT group and a dose of 20 IU of AVP (American Reagents Laboratory, Shirley, NY, USA) for the AVP group were self-administered. Furthermore, subjects were informed they would receive OT or AVP, and provided written informed consent; the study was approved by the Emory University Institutional Review Board and the U.S. Food and Drug Administration.

B. Prisoner's Dilemma Task

The iterated PD is a model where two players play with each other, in order to elicit relationships established on reciprocal altruism [25]. In each trial, the two players can choose to cooperate or defect independently, and, by the end of it, each player receives a payoff based on the two decisions [25].

A specific version of this game is used in multiple studies, the sequential-choice PD game, where Player 1 decides, and Player 2 is then allowed to see Player 1's decision before determining their choice [25]. This implies that Player 1 must choose to trust Player 2 (by cooperating) or not, and Player 2 must choose to either reciprocate cooperation (or defection) or not [34]. The four possible outcomes are assigned with a different payoff. Player 1's cooperation, accompanied by Player's 2 cooperation (CC), pays \$2 to both. However, if Player 1 cooperates, and Player 2 defects (CD), it pays \$0 to Player 1 and \$3 to Player 2. The opposite payoff occurs if Player 1 defects, and Player 2 cooperates (DC). If both players defect (DD), it pays \$1 to both [25].

For the present study, only the games where the participant played as Player 1 were analyzed. As Player 1, participants were informed they would play two separate game runs, one with a human partner and another with a computer partner. However, although they thought they were playing with the same-sex human introduced before the experiment for one game, they were always playing with a pre-programmed computer algorithm, made to simulate human strategies.

Participants were rewarded with 2/3 of the total amount earned across the games.

C. fMRI Data Acquisition and Pre-Processing

All functional magnetic resonance imaging (fMRI) data acquisition and pre-processing were performed in previous studies [25] [34] [35] [36], which the reader is recommended to look through, as they explain these processes in detail.

D. Computational Reinforcement Learning Models

With the aim of achieving the best model fit to the data, ten RL models were created and used to fit the behavioural data. All models resulted from the adaptation of Rescorla-Wagner models to the present task.

Model estimation was performed using the VBA toolbox [37] of MATLAB (MathWorks). Due to parameter similarities, each model was grouped into one of two major families (the Simple family and the tit-for-tat (TT) family).

1) Simple Family - Simple Model:

The Simple model is the basic RL model, represented by equation 1.

$$V_t(A) = V_{t-1}(A) + \alpha \times \delta_t, \quad t > 0 \tag{1}$$

Where V_t represents the predicted outcome (or action value) for the trial t and action A (either cooperate, C, or defect, D), V_{t-1} is the predicted outcome of the previous trial, $t-1$, α is the learning rate, δ represents the RPE and consists of the difference between the real outcome (R) and the predicted one ($\delta = R - V_{t-1}$).

The α parameter will be estimated through model fitting and is constrained to $0 < \alpha < 1$, adjusting the impact of the δ on the next prediction [38]. Specifically, when $\alpha = 0$, the value of the selected option is not updated, while when $\alpha = 1$, the full RPE is used to update the value of the selected option. In other words, higher learning rates lead to faster learning. However, they cause considerable increases in action values, after positive outcomes, and considerable decreases, after one negative outcome, leading to oversensitivity [38].

As each outcome prediction is based on the previous one, a value for $V_0(A)$ is required, in order to calculate the outcome prediction of trial one. In fact, $V_0(A)$ simulates the initial tendency that the participant had towards cooperation or defection. In this case, the value of $V_0(A)$ was set to 0, represented by equation 2, assuming no initial tendency.

$$V_t(C) = 0 \ \ and \ \ V_t(D) = 0, \quad t = 0 \tag{2}$$

After updating the predicted outcomes (Equation 1), the following phase aims to perform a new decision on the next trial, by using those values. This is performed by using the Softmax choice rule [39], with the probability of cooperating on the trial t, $p_t(C)$, being calculated according to the following equation:

$$p_t(C) = \frac{e^{\beta \times V_t(C)}}{e^{\beta \times V_t(C)} + e^{\beta \times V_t(D)}} = \frac{e^{\beta \times [V_t(C) - V_t(D)]}}{e^{\beta \times [V_t(C) - V_t(D)]} + 1} \tag{3}$$

While the probability of defecting on the trial t is defined by:

$$p_t(D) = 1 - p_t(C) \tag{4}$$

Where $p_t(D)$ is the probability of defecting, and β is the inverse temperature parameter ($\beta > 0$), which quantifies the consistency of choices [38]. In other words, the higher the β, the higher the higher consistency of choices [38].

Furthermore, all models in the Simple family will use equations 3 and 4 to perform new decisions.

2) Simple Family - Q0 Model:

By analyzing the averaged cooperating probability from all participants in the initial trials, it was possible to conclude that, for most participants, there was a higher tendency to cooperate.

Thus, a new parameter $Q0$ was added to the previous model, in order to modulate this tendency (equation 5), similar to what has been done in the study [40]. Furthermore, the predicted outcome was still calculated by equation 1.

$$\begin{cases} V_t(C) = Q_0 \ \ and \ \ V_t(D) = 0, \ \ t = 0 \ \ and \ \ Q_0 > 0 \\ V_t(D) = |Q_0| \ \ and \ \ V_t(C) = 0, \ t = 0 \ \ and \ \ Q_0 \leq 0 \end{cases} \tag{5}$$

The lower the $Q0$, the higher the participants' tendency to defect in the initial trials, while the higher the $Q0$, the higher the initial tendency to cooperate. Furthermore, since the rewards for the cooperation choice and defecting were unequal ($R = \{0, 2\}$\$ for cooperation and $R = \{1, 3\}$\$ for defection), and the maximum reward while cooperating is 2\$ (and not 3\$), the $Q0$ parameter was restricted to $-2 < Q0 < 2$.

3) Simple Family - 2LR Model:

Another behavioural hypothesis is the fact that participants might have asymmetries in learning from trials when they cooperated and from trials when they defected. In order to model this, the 2LR model required two learning rate parameters, leading to equation 6, similar to what has been done in the study [41]. Thus, as before, the initial tendency modelled by the $Q0$, shown in equation 5, was applied.

$$\begin{cases} V_t(C) = V_{t-1}(C) + \alpha_C \times [R - V_{t-1}(C)], \quad t > 0 \\ V_t(D) = V_{t-1}(D) + \alpha_D \times [R - V_{t-1}(D)], \quad t > 0 \end{cases} \tag{6}$$

Where α_C is the learning rate from trials when the participant cooperated, and α_D is the learning rate from trials when the participant defected. As before, both parameters were restricted between 0 and 1.

4) Simple Family - 2LR Partner Model:

Similar to the previous model, it is also possible to hypothesize that the participant might have asymmetries in learning from trials when the partner cooperated and from trials when the partner defected. Again, two learning rate parameters were implemented, leading to equation 7, similar to what has been done in the study [42]. The initial tendency modelled by the $Q0$, shown in equation 5, was still applied.

$$\begin{cases} V_t(A) = V_{t-1}(A) + \alpha_{pn(C)} \times [R - V_{t-1}(A)], \ t > 0 \ and \ A(pn)_t = C \\ V_t(A) = V_{t-1}(A) + \alpha_{pn(D)} \times [R - V_{t-1}(A)], \ t > 0 \ and \ A(pn)_t = D \end{cases} \tag{7}$$

Where $\alpha_{pn(C)}$ is the learning rate from the trials when the partner cooperated, $A(pn)_t = C$, and $\alpha_{pn(D)}$ is the learning rate from the trials when the partner defected, $A(pn)_t = D$. As before, both parameters were restricted between 0 and 1.

5) Simple Family - 4LR Model:

The 4LR model translates the hypothesis of the participant having asymmetries in learning from different outcomes. For example, trials with the CC outcome might have a different importance and influence on the decision of the next trial, consequently allowing a different learning rate than trials with the DC outcome. To model it, four learning rates were added to the Q0 model, leading to equation 8. Once again, the initial tendency modelled by the $Q0$, shown in equation 5, was still applied.

$$\begin{cases} V_t(C) = V_{t-1}(C) + \alpha_{CC} \times [R - V_{t-1}(C)], \ t > 0 \ and \ \ A(pn)_t = C \\ V_t(C) = V_{t-1}(C) + \alpha_{CD} \times [R - V_{t-1}(C)], \ t > 0 \ and \ \ A(pn)_t = D \\ V_t(D) = V_{t-1}(D) + \alpha_{DC} \times [R - V_{t-1}(D)], \ t > 0 \ and \ \ A(pn)_t = C \\ V_t(D) = V_{t-1}(D) + \alpha_{DD} \times [R - V_{t-1}(D)], \ t > 0 \ and \ \ A(pn)_t = D \end{cases} \tag{8}$$

Where α_{CC} is the learning rate from the trials with the CC outcome, α_{CD} from the CD trials, α_{DC} from the DC trials and α_{DD} from the DD trials. As before, all learning rates were restricted between 0 and 1.

6) TT Family:

The other family of models is based on the tit-for-tat strategy, in which the participant chooses what the partner chose in the previous trial [43]. This can be translated to the model by adding the parameter TT to the Softmax choice rule (represented in equation 3), depending on the partner's decision, originating the equation 9.

$$\begin{cases} p_t(C)=\dfrac{e^{\beta\times V_t(C)+TT}}{e^{\beta\times V_t(C)+TT}+e^{\beta\times V_t(D)}}, & A(pn)_{t-1}=C \\ p_t(C)=\dfrac{e^{\beta\times V_t(C)}}{e^{\beta\times V_t(C)}+e^{\beta\times V_t(D)+TT}}, & A(pn)_{t-1}=D \end{cases} \tag{9}$$

If the partner cooperated in the previous trial, the TT is added to the cooperation exponential ($e^{\beta\times V_t(C)}$), increasing the participant's probability of cooperating in the next trial. Similarly, when the partner defected in the previous trial, the TT is added to the defection exponential. The TT parameter is estimated by being restricted to $TT > 0$. As before, the probability of the participant defecting in the trial t, $p_t(D)$, is achieved by equation 4.

Thus, five new models were created (similar to the ones from the Simple family), being fitted using the changed Softmax rule.

7) Model Selection and Validation:

Model comparison and selection were performed using the VBA toolbox, according to their estimated frequencies (which evaluates the likelihood of selecting one model for any subject randomly chosen [44]) and exceedance probabilities (which evaluates the likelihood of a model being more suitable than any other [44]). The selected model was then validated through a PPC procedure and a parameter recovery analysis.

E. Behavioural Analysis

The behavioural analyses were performed using R (R Foundation for Statistical Computing, Vienna, Austria). Using the winning model parameters as dependent variables, a mixed design repeated-measures analysis of variance (ANOVA) was performed, with the within-subject partner factor (computer, human) and the sex (male, female) and drug (PB, OT and AVP) as between-subject factors. The main and interaction effects that resulted from this analysis were considered statistically significant if the p-value was below 0.05. Furthermore, regarding the *post hoc* tests, Bonferroni's correction was performed.

F. fMRI Data Analysis

After fMRI scans' pre-processing (performed and described by [36]), data analysis was performed using Statistical Parametric Mapping software (SPM12; Wellcome Department of Imaging Neuroscience, Institute of Neurology, London UK), and the Sandwich Estimator (SwE) [45], including custom code written in MATLAB.

Of the 292 participants, only 253 (121 men; 132 women) played two games, one with a human partner and the other with a computer partner. Thus, only the data of these participants was used for the subject and group-level fMRI analysis.

1) Subject-Level Analysis:

A GLM was created on SPM to analyze the brain BOLD data, which included the following four regressors: (1) the outcome epoch when the reward was received for the game with the human partner; (2) the RPE parametric regressor for the game with the human partner (hRPE), obtained through parametric modulation with the previous regressor, with a polynomial expansion of first order; (3) the outcome epoch when the reward was received for the game with the computer partner; (4) the RPE parametric regressor for the game with the computer partner (cRPE), obtained through parametric modulation with the previous regressor, with a polynomial expansion of first order. Furthermore, no orthogonalization was performed between parametric regressors, in order to not attribute the shared variance to either of them, and the parametric values were mean centered.

Afterwards, two contrasts were defined to integrate the group-level analysis: the response due to RPEs when playing with the human partner being greater than 0 ($\beta_{hRPE} > 0$, or $c1 = [0, 1, 0, 0]$), and the response due to RPEs when playing with the computer partner being greater than 0 ($\beta_{cRPE} > 0$, or $c2 = [0, 0, 0, 1]$).

2) Group-Level Analysis:

A GLM was created using the SwE toolbox, with the two subject-level contrasts as scan inputs. The main goal of this analysis was to estimate the effect of three factors and their interaction on the correlation between RPEs' and brain activations: 1) between-subject factor "Drug" (OT, AVP or PB); 2) between-subject factor participant's "Sex" (male or female); and 3) within-subject factor "Partner type" with whom they played the game (human or computer), as each player performed two games (as Player 1), one with the human and another with the computer.

The model was set up using the "Modified" SwE, which assumes that subjects that belong to the same group can share a common covariance matrix, the "C2" small-sample adjustment, since, according to SwE [46], it is the most optimal correction, allowing to remove the bias in multiple scenarios correctly, and the "approx III" degrees of freedom type, being recommended by default [46]. Furthermore, since our main goal is to analyze the differences in brain activations to RPEs between the different factors, the following twelve explanatory variables (EVs) were created to build the design matrix: (1) EV selecting the games played under the effect of PB; (2) EV selecting the games played under the effect of OT; (3) EV selecting the games played under the effect of AVP; (4) EV comparing the games played with a human and with a computer under the effect of PB; (5) EV comparing the games played with a human and with a computer under the effect of OT; (6) EV comparing the games played with a human and with a computer under the effect of AVP; (7) EV comparing the games played by females and males under the

effect of PB; (8) EV comparing the games played by females and males under the effect of OT; (9) EV comparing the games played by females and males under the effect of AVP; (10) EV comparing the games played with a human and with a computer and by females and males under the effect of PB; (11) EV comparing the games played with a human and with a computer and by females and males under the effect of OT; (12) EV comparing the games played with a human and with a computer and by females and males under the effect of AVP.

In order to obtain family-wise error (FWE) corrected results in SwE, a non-parametric wild bootstrap was performed. Thus, the procedure was set up using the "C2" small-sample adjustment for WB resampling, as before, using 999 bootstraps, the "U-SwE" type of SwE, since it would allow a less biased estimator [46] and a Voxelwise Inference. All main effects and interactions were considered statistically significant if the result was FWE corrected and the p-value was below 0.05.

3) Region of Interest Analysis:

A region of interest (ROI) analysis was performed, besides a whole-brain analysis. Specifically, based on previous literature stating that both striatum and amygdala play central roles in the RL process [18], a striatum mask and an amygdala mask were acquired, separately, from the probabilistic "Harvard Oxford cortical and subcortical structural atlases" provided by the Harvard Center for Morphometric Analysis (with a threshold of 25%) [47], using the MRIcron software.

Two other separated masks, the right and the left caudate, were also used. Those masks were created by specific studies ([35] and [48]) using the same sample as the one applied on the present analysis, and were acquired from an activation map, with a FWE correction of $p < 0.001$, through contrasting (OT>PB) in male - (OT>PB) in female for CC trials while playing with a human partner.

As in the whole-brain analysis, all main effects and interactions were considered statistically significant if the result was FWE corrected and the p-value was below 0.05.

III. Results

A. Computational Reinforcement Learning Models

After fitting the model to each game data, one model must be selected to be used in the following analysis. Thus, a free-energy value was acquired per model fit (*i.e.*, one value per game). First, to measure which family best fitted the data, the estimated frequencies of each family were calculated. The estimated frequency of the Simple family was $0,5907$, while for the TT family was $0,4093$. Therefore, the Simple family was selected.

Following a similar approach, a specific model from the Simple family was chosen. Again, to measure the model that best fitted the data, the estimated frequencies were calculated, with the Simple model having an estimated frequency of $0,066$, the Q0 model of 0, the 2LR model of $0,894$, the 2LR Partner of 0, and the 4LR model of $0,039$. Furthermore, the exceedance probabilities were also determined to increase the veracity of the previous method, with the Simple model having an exceedance probability of 0, the Q0 model of 0, the 2LR model of 1, the 2LR Partner of 0, and the 4LR model of 0. Therefore, the 2LR model was selected. Figure 1 compares the 2LR with the real data, again, using the trial-by-trial averaged cooperating probability from all participants.

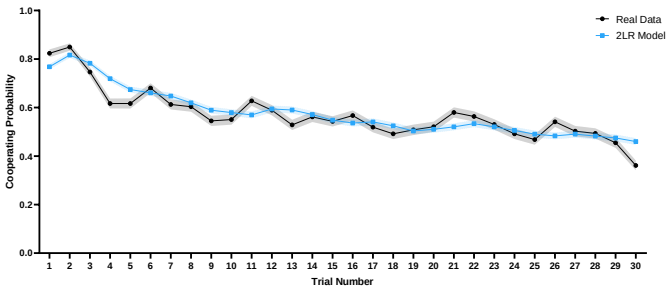


Fig. 1: Trial-by-trial averaged cooperating probability from all participants obtained from the real data and from the 2LR model. Solid lines indicate the mean, while the shaded error bars represent the standard error.

To validate the chosen model, the PPC procedure was performed. Hence, using the 2LR model's parameters, new artificial participant choices were simulated, and a similar comparison using the trial-by-trial averaged cooperating probability was performed, shown in Figure 2. In order to assess the individual variation of the 2LR model, the cooperation probability of the artificial data was averaged across trials (Figure 3(a)). Thus, the closer the mean values are to the identity line, the better the model.

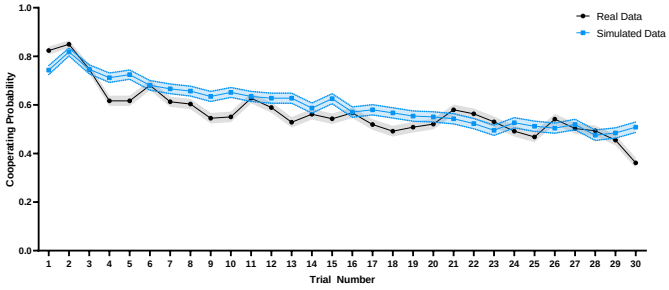


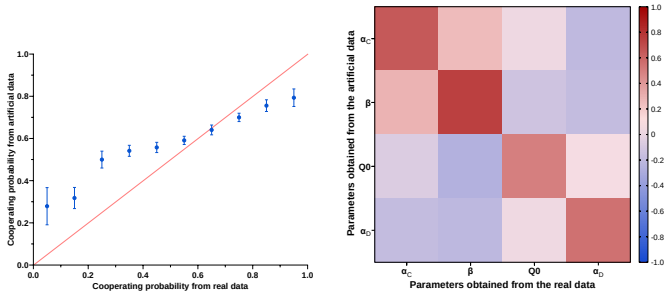
Fig. 2: Model validation: Trial-by-trial averaged cooperating probability from real data and from artificial data acquired using the 2LR model. Solid lines indicate the mean, while the shaded error bars represent the standard error.

Furthermore, since RL models might not accurately and selectively identify their parameters, a parameter recovery analysis was performed. Thus, Spearman's correlation coefficients were calculated per pair of parameters (2LR model's parameters from real data and the ones acquired from artificial simulated data). Figure 3(b) shows the correlation matrix.

B. Behavioural Analysis

1) α_C Analysis:

A significant three-way interaction (drug $\times$ sex $\times$ partner) was found for the mean α_C $[F(2,247) = 4.25, p = 0.015, \eta^2 = 0.014]$. In the OT group, female participants had a higher α_C when they played with a human than with a computer partner ($p = 0.020$) (Figure 4), while the same effect was not present for male players ($p = 0.132$) (Figure



(a) (b)

Fig. 3: Model validation: (a) individual's artificial cooperating probability acquired using the 2LR model (in blue, mean ± standard deviation) compared with real data, in relation to the identity line (in red); (b) correlation matrix of the 4 parameters of the 2LR model.

4). In the other drug groups, no significant effects were found ($p > 0.212$) (Figure 4).

2) α_D and β Analysis:

No significant main effects or interactions were found for mean α_D ($p > 0.058$) and mean β ($p > 0.180$).

3) $Q0$ Analysis:

Regarding the $Q0$ parameter, a significant main effect of partner was found $[F(1, 247) = 3.95, p = 0.048, \eta^2 = 0.007]$, with participants playing with human partners having a higher $Q0$ ($M = 0.84, SD = 0.58$) than participants playing with computer partners ($M = 0.73, SD = 0.74$).

A significant drug × sex × partner interaction was also found $[F(2, 247) = 5.63, p = 0.004, \eta^2 = 0.02]$. For male players, a higher $Q0$ was found when they played with computer partners, under AVP than PB ($p = 0.039$) (Figure 4). In the OT group versus the PB and in the OT group versus the AVP, no significant differences were found ($p > 0.647$) (Figure 4). Also for males, a higher $Q0$ was found when they played with human partners under AVP than OT ($p = 0.035$) (Figure 4), and under PB than OT ($p = 0.002$) (Figure 4), with non-significant differences being shown between AVP and PB ($p = 1$) (Figure 4). In fact, the latter significant effect was not exclusive to males, with participants under PB having a higher $Q0$ when playing with human partners, in comparison to OT ($p = 0.032$). For females, specifically in the OT group, a higher $Q0$ was found when they played with human than computer partners ($p = 0.034$) (Figure 4).

C. fMRI Data Analysis

1) Whole-Brain Analysis:

The whole-brain analysis revealed that, in the PB group, a higher correlation of the RPE signal with the superior temporal gyrus (STG) was found when playing with a human partner, in comparison to a computer partner ($p = 0.039$) (Table I).

2) Striatum ROI Analysis:

Using the striatum ROI, no significant effects were found ($p > 0.05$).

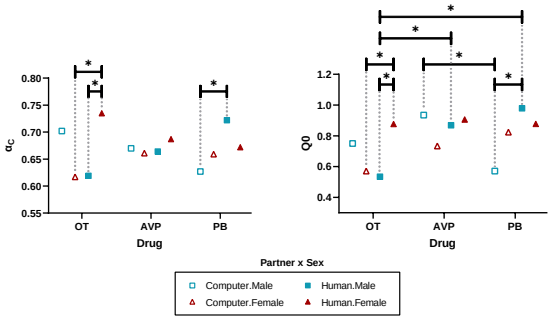


Fig. 4: Mean values of the α_C (left plot) and $Q0$ (right plot) parameters for the different combination of three factors (drug, partner and sex). OT: oxytocin; AVP: vasopressin; PB: placebo.
⋆ All statistically significant at $p < 0.05$ (on the text, only the most important results are mentioned)

TABLE I: Whole Brain fMRI Results (FWE-correct, $p < 0.05$). HUM: human partner; CPU: computer partner.

Contrast	Cluster Size	p-value (FWE-corr)	Z-score	x	y	z	Brain Region
Drug x Partner: HUM>CPU in PB	2	0.039	4.44	62	-44	24	Right superior temporal gyrus

3) Amygdala ROI Analysis:

Using an amygdala ROI, two significant three-way interactions were found (Table II, Figure 5).

Note that, the present study's design matrix did not allow to compare the three drug groups simultaneously during a three-way interaction. Thus, a three-way interaction drug (OT vs PB) × sex × partner ($p < 0.038$) was found. For female players, there was not only a higher RPE-amygdala activation correlation under PB when they played with human than computer partners ($p = 0.043$) (Figure 6(a)), but also under OT when they played with computer than human partners ($p < 0.045$) (Figure 6(a)). For male players, there were no significant effects.

Additionally, a three-way interaction drug (AVP vs PB) × sex × partner ($p < 0.031$) was found. For male players under AVP, there was a higher RPE-amygdala activation correlation when they played with human than computer partners ($p = 0.007$), with no significant effects being seen under PB. For females, no significant effects were found (Figure 6(b)).

TABLE II: Amygdala ROI fMRI Results (FWE-correct, $p < 0.05$). In the Contrast column, the four possible combinations of each three-way interaction are detailed, since all of them are produced by the same mathematical contrast. Note that, each three-way interaction has more than one cluster. OT: oxytocin; AVP: vasopressin; PB: placebo; HUM: human partner; CPU: computer partner.

Contrast	Cluster Size	p-value (FWE-corr)	Z-score	x	y	z
Three-way: OT>PB in HUM>CPU in Male>Female; OT>PB in CPU>HUM in Female>Male; PB>OT in CPU>HUM in Male>Female; PB>OT in HUM>CPU in Female>Male	8	0.015	3.37	-28	0	-16
	2	0.031	3.16	28	-10	-12
	5	0.038	3.07	30	-4	-20
	1	0.038	3.06	26	-10	-16
Three-way: AVP>PB in HUM>CPU in Male>Female; AVP>PB in CPU>HUM in Female>Male; PB>AVP in CPU>HUM in Male>Female; PB>AVP in HUM>CPU in Female>Male	1	0.020	3.27	28	-6	-14
	1	0.031	3.06	26	-8	-12

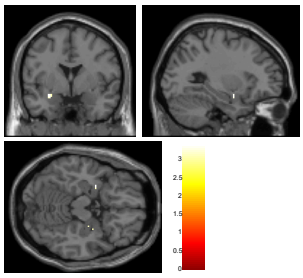


Fig. 5: Neural representation of RPE encoded in the Amygdala ($x = -28, y = 0, z = -16$). The most representative coronal (on the upper left), sagittal (on the upper right), and transverse (on the lower left) slices are shown, all obtained with the MRIcron software. Display thresholded at $p < 0.05$, family-wise error (FWE) corrected.

TABLE III: Right Caudate ROI fMRI Results (FWE-correct, $p < 0.05$). In the Contrast column, the four possible combinations of the three-way interaction are detailed, since all of them are produced by the same mathematical contrast. OT: oxytocin; AVP: vasopressin; PB: placebo; HUM: human partner; CPU: computer partner.

Contrast	Cluster Size	p-value (FWE-corr)	Z-score	x	y	z
Three-way: OT>AVP in HUM>CPU in Male>Female; OT>AVP in CPU>HUM in Female>Male; AVP>OT in CPU>HUM in Male>Female; AVP>OT in HUM>CPU in Female>Male	1	0.036	2.92	18	22	-4

4) Right Caudate ROI Analysis:

Using the right caudate ROI, a significant three-way interaction (drug (OT vs AVP) $\times$ sex $\times$ partner) was found ($p = 0.036$) (Table III). However, no significant two-way interactions or simple effects were obtained (Figure 6(c)).

5) Left Caudate ROI Analysis:

Using the left caudate ROI analysis, two significant three-way interactions were found (Table IV).

A three-way interaction drug (OT vs PB) $\times$ sex $\times$ partner interaction was found. For female players, not only there was a higher RPE-left caudate activation correlation under PB when they played with human than computer partners ($p = 0.008$) (Figure 6(d)), but also under OT when they played with computer than human partners ($p = 0.013$) (Figure 6(d)). For male players, there were no significant effects.

Furthermore, a three-way interaction drug (OT vs AVP) $\times$ sex $\times$ partner interaction was found. In the OT group, female players had a higher RPE-left caudate activation correlation when playing with a computer than with a human partner ($p < 0.013$) (Figure 6(d)), with no such effects being seen for male players. In the AVP group, no significant effects were found (Figure 6(d)).

IV. Discussion

The present study aimed to investigate the roles of OT and AVP on social RL, and its neural correlates, in different social contexts (*i.e.*, with different partner types, human or computer), and examine the added effect of participant's sex.

A. Computational Reinforcement Learning Models

Ten different RL models were created and used to fit the behavioural data, with the 2LR model from the Simple family providing the best fit. Afterwards, a PPC analysis was performed, revealing reduced divergences between the real and the artificial data acquired using the winning model, both at the trial level (Figure 2), and at the participant level (Figure 3(b)). The parameter recovery analysis also showed that the winning model was able to accurately recover its parameters (Figure 3(b)).

TABLE IV: Left Caudate ROI fMRI Results (FWE-correct, $p < 0.05$). In the Contrast column, the four possible combinations of the three-way interaction are detailed, since all of them are produced by the same mathematical contrast. OT: oxytocin; AVP: vasopressin; PB: placebo; HUM: human partner; CPU: computer partner.

Contrast	Cluster Size	p-value (FWE-corr)	Z-score	x	y	z
Three-way: OT>PB in HUM>CPU in Male>Female; OT>PB in CPU>HUM in Female>Male; PB>OT in CPU>HUM in Male>Female; PB>OT in HUM>CPU in Female>Male	3	0.001	4.10	-28	2	-16
Three-way: OT>AVP in HUM>CPU in Male>Female; OT>AVP in CPU>HUM in Female>Male; AVP>OT in CPU>HUM in Male>Female; AVP>OT in HUM>CPU in Female>Male	1	0.040	2.87	-28	2	-16

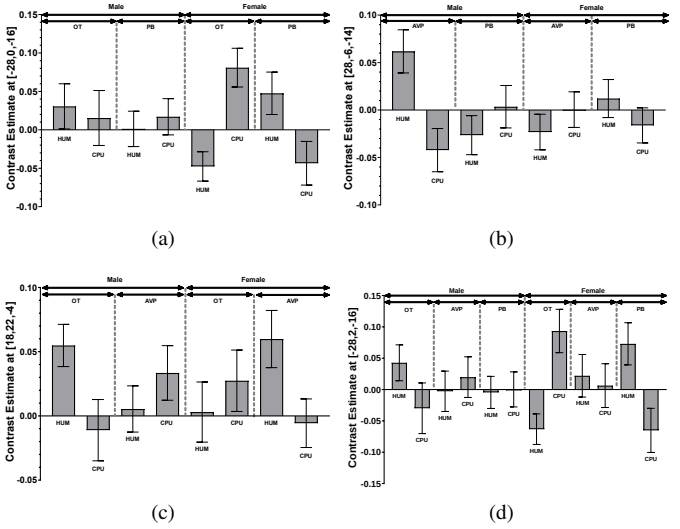


Fig. 6: Neural correlates at the corresponding peak voxel for the three-way interaction drug $\times$ sex $\times$ partner: (a) contrast estimate for the amygdala ROI ($x = -28, y = 0, z = -16$); (b) contrast estimate for the amygdala ROI ($x = 28, y = -6, z = -14$); (c) contrast estimate for the right caudate ROI ($x = 18, y = 22, z = -4$); (d) contrast estimate for the left caudate ROI ($x = -28, y = 2, z = -16$). The β-values shown in the vertical axis represent contrast estimates for the degree of the correlation between the brain activation and the RPE. OT: oxytocin; AVP: vasopressin; PB: placebo; HUM: human partner; CPU: computer partner.

B. Behavioural Analysis

The behavioural analysis performed in this study aimed to investigate the impact of intranasal OT, AVP, partner type and participant sex on the RL model parameters.

1) α_C Analysis:

As previously described, the α_C parameter represents the learning rate from trials when the participant cooperated. No main effects of OT in comparison to PB or a two-way interaction between drug and partner were found for α_C.

Nevertheless, intranasal OT increased the α_C among women, when they played with human partners than with computer partners, but not under PB. Previous studies reported

that OT is associated with defensive aggression focused on protecting and negating threats induced by out-groups [49], also reported in a Social Orientation Paradigm [9] and in women during a social task [50]. Thus, aggression may lead to impulsive actions, since the previous interaction has higher importance on the next decision. A social context (*i.e.*, the partner being human) might also augment the previous effect, leading to increased defensive aggression and impulsive behaviours.

The fact that there were significant results for α_C reveals that the cooperation decision had high importance on the subsequent trial's decision. Since cooperation implies trusting the other person, every time participants performed this decision, they were more exposed to their partner's will (*i.e.*, they might receive the worst possible outcome), which might raise impulsiveness on the subsequent trial's decision. Herein, the present results suggest that women under OT, after cooperating, are more prone to rapidly change their perception when in social contexts, compared to non-social ones.

2) $Q0$ Analysis:

As previously referred, the $Q0$ parameter represents the tendency that the participant has, at the beginning of the task, to cooperate or defect.

The main and simple effects revealed an increased $Q0$ when participants played with a human partner than with a computer partner, meaning that participants had a higher tendency to cooperate at the beginning of the game when playing with human partners. A study using a PD task with human and computer partners analyzed the impact of the two partner types, while manipulating the participant's knowledge of it, *i.e.*, participants played with human partners while assuming they were playing with computer partners and vice-versa [51], and showed that participants cooperated more when assuming that they were playing with the human partner, even if they were not. Thus, considering the present results and previous evidence, the cognitive representation of the partner may have an important role in the human-human and human-computer interactions, specifically with strangers or acquaintances, not only during the game, but also at the beginning of it, leading to a higher cooperation bias when interacting with humans.

Furthermore, although some studies [34] report that OT may promote anthropomorphism of computer partners in women, the present results showed a significantly higher $Q0$ when female participants under OT played with a human partner than with a computer partner. Although these results might seem contradictory, the present study shows that subjects have a higher cooperating bias towards human partners at the beginning of the game. Nevertheless, throughout the game, OT might increase the number of cooperating choices when playing with a computer partner, leading to anthropomorphism.

Significant drug effects were also found, with participants having a higher $Q0$ under the effect of PB than under OT, when they played with human partners. These findings are in line with the OT's antisocial or pro-self behaviours, promoting in-group favouritism and intergroup bias [11], and may also reveal that a pro-self bias and threat identification might occur at the beginning of an interaction.

A higher $Q0$ was also found for male participants under the effect of AVP than OT when playing with human partners, and under the effect of AVP than PB when playing with computer partners. Although AVP's effects are context-dependent, the present results show that AVP might induce a pro-social bias in males at the beginning of an interaction in a non-social context. Previous evidence [22] showed that AVP increased the willingness to cooperate in males using the Stag Hunt task, through increasing the desire to take risks, compared to PB. Herein, by enhancing the willingness to take risks, AVP might promote a pro-social bias in males, inducing cooperation, at the beginning of an interaction in a non-social context.

C. fMRI Analysis

The fMRI analysis performed in this study aimed to investigate the impact of intranasal OT, AVP, partner type and participant sex on the RPE-brain area activation correlations.

1) Whole-Brain Analysis:

In the whole-brain analysis, contrary to the hypotheses of this study, no main effects of OT or a two-way interaction between OT and the partner type on the RPE-striatal activation correlation were found. However, a statistically significant simple effect was found, with the PB group expressing a higher positive correlation between the RPE signal and the STG when participants played with a human than with a computer partner. Although the STG is traditionally associated with language and auditory processing [52], a study [53] has reported that it has an important role in processing social stimuli. Specifically, a previous study [54] reported that STG has an essential role in behavioural monitoring and reappraisal and another [55] studied the reinforcement and decision making in patients with psychopathy, revealing decision making deficits due to STG dysfunction. Hence, the findings reported here agree with some previous evidence, suggesting an additional role of the STG in the social RL process.

2) ROI Analyses:

Additionally, four different ROI analyses were performed, each using a separated mask, namely the striatum mask, the left caudate mask, the right caudate mask and the bilateral amygdala mask.

As previously stated, the striatum is a brain region that plays an essential role in RL, comprising a prominent dopaminergic neuronal projection that codes RPEs [18]. However, contrary to our expectations, no main effects or interactions were found in this ROI. On the other hand, the analyses using both the left and the right caudate ROIs (two brain areas that are components of the dorsal striatum) revealed three significant three-way interactions. These results might seem counter-intuitive, however, each of these ROIs is narrower than the striatum ROI. Moreover, the caudate ROI masks were derived from an activation map (acquired from the studies [35] and [48]) using the same neuronal data as the present study.

Regarding the left caudate, the present study's results suggest that OT enhances social learning in females when playing with a computer partner, compared to a human partner, while,

under PB, playing with a human partner enhances social learning in females, in comparison to computer partners. Similar results were found in the NAcc, a striatum region, in males by *Kruppa et al.* [20], while also using computational RL modelling of behavioural data and trial-by-trial RPE signals. Since humans are more used to learning from social contexts (for example, language learning requires social interactions [56]), this result might suggest that OT compensates and reinforces learning from non-social contexts. Another hypothesis might be that OT increases learning from social partners to a point where it also increases the learning from non-social partners in females. Moreover, the present results might help to corroborate and explain, with model-based fMRI, the findings of *Neto et al.* [43], a study using the same behavioural data as the present analysis, which suggested a female anthropomorphization of computer partners facilitated by OT, *i.e.*, females with increased levels of OT treated the computer partners as humans. In agreement with their findings, the present results indicate that OT might enhance females' learning of how cooperating is the best decision to increase their gains throughout the game ("taught" by a computed tit-for-tat algorithm). Under PB, the social context (*i.e.*, playing with a human partner) facilitates learning, as reported by a previous study [33], also using RL models and trial-by-trial RPE.

Regarding the right caudate, even though no significant two-way interactions or simple main effects were found, a similar trend as the one in the left caudate was found.

An exploratory analysis was also performed using an amygdala ROI, and two significant three-way interactions were found. The amygdala is a brain area that plays important roles in the emotional learning [57] and processing of emotional information [58], while recognizing the stimulus for the needs and goals of the organism [58]. Furthermore, previous studies also reported that the amygdala applies social attention, information and emotions in decision-making [59], and also has an important role in the RL process [18].

In fact, similar significant results to the ones found in the caudate were found in the amygdala. A previous study [19] using computational RL models and trial-by-trial RPE reported similar results, with males under PB having an enhanced RPE-amygdala activation correlation when playing with human partners, compared to OT. Multiple studies have revealed that the striatum (which includes the caudate region) and the amygdala work in series [18], with both structures receiving multiple DA projections [60]. Physiologically, studies have shown that the stimulation of the basolateral amygdala may increase DA release in the ventral striatum due to glutamatergic input signals [61] and that DA delivery to the ventral striatum was reduced due to inactivation of the basolateral amygdala, while maintaining the DA release to the ventral tegmental area, using a reward predicting cue [62]. Although these findings were related to the ventral striatum, one might hypothesize that similar effects would occur in the dorsal striatum and, together with the present results, it suggests that the previous caudate hypotheses also apply to the amygdala region, leading to similar activation correlations.

Additionally, an AVP simple effect was found in males, with AVP enhancing the RPE-amygdala correlation in males when playing with a human partner, compared to a computer partner. In agreement with what was previously described in the amygdala, one might hypothesize that AVP, in social contexts, might increase social learning in males. In fact, although previous evidence has reported that AVP might be involved in the learning process [63], this hypothesis is relatively unexplored. Herein, further research should be conducted to study the role of AVP in the RL process.

V. Conclusion

The present study suggests new specific roles for OT and AVP in the social RL process, which is consistent with the implication they are currently believed to have in general social cognition.

Through the parameter behavioural analysis (of both α_C and $Q0$ parameters), two different behavioural mechanisms of OT were suggested. Firstly, present study results revealed that OT may promote a pro-self (non-social) bias and threat identification at the beginning of a social interaction. Secondly, throughout a social interaction, after cooperating, women under OT might be more prone to impulsive behaviours and rapidly change their perception of the partner (*i.e.*, whether they are a threat or not) based on defensive aggression, in comparison to a non-social context. On the other hand, AVP might promote a pro-social bias at the beginning of an interaction in males, which may be caused by an enhanced willingness to take risks.

Furthermore, new neurological mechanisms of OT and AVP on the social RL process were also suggested. The whole-brain analysis revealed that the STG might have an important role in the social RL process, being positively correlated with the RPE. The caudate and amygdala ROI results suggest that OT enhances social learning in females in non-social contexts, compared to social ones. Additionally, the amygdala ROI results revealed that AVP might increase the social learning of males, during social contexts.

As there are novel findings, it is essential to further replicate this evidence. Such is a promising research avenue as these neuropeptides may prove to be important allies in the treatment of disorders associated with social deficits.

Acknowledgment

This document was written and made publically available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biomedical Engineering of the author at Instituto Superior Técnico. The work previously described was performed at Institute of Biophysics and Biomedical Engineering of the University of Lisbon (Lisbon, Portugal), during the period March-October 2021, under the supervision of Dr. Diana Prata. The thesis was co-supervised at Instituto Superior Técnico by Prof. Manuel Lopes.

REFERENCES

[1] M. Piva et al. "An integrated framework for the role of oxytocin in multistage social decision-making". In: *American Journal of Primatology* 80.10 (2018), e22735.

[2] J. Barraza et al. "Oxytocin infusion increases charitable donations regardless of monetary resources". In: *Hormones and Behavior* 60.2 (2011), pp. 148–151.

[3] C. Declerck et al. "The effect of oxytocin on cooperation in a prisoner's dilemma depends on the social context and a person's social value orientation". In: *Social Cognitive and Affective Neuroscience* 9.6 (2014), pp. 802–809.

[4] B. Ditzen et al. "Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict". In: *Biological Psychiatry* 65.9 (2009), pp. 728–731.

[5] M. Kosfeld et al. "Oxytocin increases trust in humans". In: *Nature* 435.7042 (2005), pp. 673–676.

[6] T. Baumgartner et al. "Oxytocin shapes the neural circuitry of trust and trust adaptation in humans". In: *Neuron* 58.4 (2008), pp. 639–650.

[7] G. Domes et al. "Oxytocin improves "mind-reading" in humans". In: *Biological Psychiatry* 61.6 (2007), pp. 731–733.

[8] A. Guastella et al. "Oxytocin enhances the encoding of positive social memories in humans". In: *Biological Psychiatry* 64.3 (2008), pp. 256–258.

[9] R. Ne'eman et al. "Intranasal administration of oxytocin increases human aggressive behavior". In: *Hormones and Behavior* 80 (2016), pp. 125–131.

[10] S. Shamay-Tsoory et al. "Intranasal administration of oxytocin increases envy and schadenfreude (gloating)". In: *Biological Psychiatry* 66.9 (2009), pp. 864–870.

[11] C. De Dreu et al. "Oxytocin promotes human ethnocentrism". In: *Proceedings of the National Academy of Sciences* 108.4 (2011), pp. 1262–1266.

[12] L. Peled-Avron et al. "Exogenous effects of oxytocin in five psychiatric disorders: a systematic review, meta-analyses and a personalized approach through the lens of the social salience hypothesis". In: *Neuroscience & Biobehavioral Reviews* 114 (2020), pp. 70–95.

[13] A. Meyer-Lindenberg et al. "Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine". In: *Nature Reviews Neuroscience* 12.9 (2011), pp. 524–538.

[14] K. MacDonald et al. "The peptide that binds: a systematic review of oxytocin and its prosocial effects in humans". In: *Harvard Review of Psychiatry* 18.1 (2010), pp. 1–21.

[15] M. Klein et al. "Dopamine: functions, signaling, and association with neurological diseases". In: *Cellular and Molecular Neurobiology* 39.1 (2019), pp. 31–59.

[16] R. Stoop. "Neuromodulation by oxytocin and vasopressin". In: *Neuron* 76.1 (2012), pp. 142–159.

[17] S. Hyman et al. "Neural mechanisms of addiction: the role of reward-related learning and memory". In: *Annual Review of Neuroscience* 29 (2006), pp. 565–598.

[18] B. Averbeck et al. "Motivational neural circuits underlying reinforcement learning". In: *Nature Neuroscience* 20.4 (2017), pp. 505–512.

[19] J. Ide et al. "Oxytocin attenuates trust as a subset of more general reinforcement learning, with altered reward circuit functional connectivity in males". In: *Neuroimage* 174 (2018), pp. 35–43.

[20] J. Kruppa et al. "Neural modulation of social reinforcement learning by intranasal oxytocin in male adults with high-functioning autism spectrum disorder: a randomized trial". In: *Neuropsychopharmacology* 44.4 (2019), pp. 749–756.

[21] M. Heinrichs et al. "Oxytocin, vasopressin, and human social behavior". In: *Frontiers in Neuroendocrinology* 30.4 (2009), pp. 548–557.

[22] C. Brunnlieb et al. "Vasopressin increases human risky cooperative behavior". In: *Proceedings of the National Academy of Sciences* 113.8 (2016), pp. 2051–2056.

[23] A. Guastella et al. "Intranasal arginine vasopressin enhances the encoding of happy and angry faces in humans". In: *Biological Psychiatry* 67.12 (2010), pp. 1220–1222.

[24] R. Thompson et al. "Sex-specific influences of vasopressin on human social communication". In: *Proceedings of the National Academy of Sciences* 103.20 (2006), pp. 7889–7894.

[25] J. Rilling et al. "Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men". In: *Psychoneuroendocrinology* 37.4 (2012), pp. 447–461.

[26] J. Goodson. "Deconstructing sociality, social evolution and relevant nonapeptide functions". In: *Psychoneuroendocrinology* 38.4 (2013), pp. 465–478.

[27] E. Keverne et al. "Vasopressin, oxytocin and social behaviour". In: *Current Opinion in Neurobiology* 14.6 (2004), pp. 777–783.

[28] M. Lim et al. "Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene". In: *Nature* 429.6993 (2004), pp. 754–757.

[29] R. Chowdhury et al. "Dopamine restores reward prediction errors in old age". In: *Nature Neuroscience* 16.5 (2013), pp. 648–653.

[30] L. Wang et al. "NMDA receptors in dopaminergic neurons are crucial for habit learning". In: *Neuron* 72.6 (2011), pp. 1055–1066.

[31] G. Jocham et al. "Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism". In: *Journal of Neuroscience* 34.39 (2014), pp. 13151–13162.

[32] M. Liljeholm et al. "Contributions of the striatum to learning, motivation, and performance: an associative account". In: *Trends in Cognitive Sciences* 16.9 (2012), pp. 467–475.

[33] L. Zhang et al. "A brain network supporting social influences in human decision-making". In: *Science Advances* 6.34 (2020), eabb4159.

[34] J. Rilling et al. "Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction". In: *Psychoneuroendocrinology* 39 (2014), pp. 237–248.

[35] C. Feng et al. "Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans". In: *Brain Imaging and Behavior* 9.4 (2015), pp. 754–764.

[36] X. Chen et al. "Effects of oxytocin and vasopressin on the neural response to unreciprocated cooperation within brain regions involved in stress and anxiety in men and women". In: *Brain Imaging and Behavior* 10.2 (2016), pp. 581–593.

[37] J. Daunizeau et al. "VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data". In: *PLoS Computational Biology* 10.1 (2014), e1003441.

[38] L. Zhang et al. "Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices". In: *Social Cognitive and Affective Neuroscience* 15.6 (2020), pp. 695–707.

[39] R. Sutton et al. "Reinforcement learning: An introduction". In: 2nd ed. MIT press, 2018, pp. 25–45.

[40] B. Lindström et al. "Social threat learning transfers to decision making in humans". In: *Proceedings of the National Academy of Sciences* 116.10 (2019), pp. 4732–4737.

[41] T. Katthagen et al. "Striatal Dopamine and Reward Prediction Error Signaling in Unmedicated Schizophrenia Patients". In: *Schizophrenia Bulletin* 46.6 (2020), pp. 1535–1546.

[42] S. McDougle et al. "Neural signatures of prediction errors in a decision-making task are modulated by action execution failures". In: *Current Biology* 29.10 (2019), pp. 1606–1613.

[43] M. Neto et al. "Oxytocin and vasopressin modulation of prisoner's dilemma strategies". In: *Journal of Psychopharmacology* 34.8 (2020), pp. 891–900.

[44] K. Stephan et al. "Bayesian model selection for group studies". In: *Neuroimage* 46.4 (2009), pp. 1004–1017.

[45] B. Guillaume et al. "Fast and accurate modelling of longitudinal and repeated measures neuroimaging data". In: *NeuroImage* 94 (2014), pp. 287–302.

[46] B. Guillaume. "Accurate non-iterative modelling and inference of longitudinal neuroimaging data". PhD thesis. Université de Liège, Liège, Belgique, 2015.

[47] D. Kennedy et al. *Harvard Oxford cortical and subcortical structural atlases*. Last accessed 30 September 2021. 2016. URL: https://ftp.nmr.mgh.harvard.edu/pub/dist/freesurfer/tutorial_packages/centos6/fsl_507/doc/wiki/Atlases.html.

[48] X. Chen et al. "Within vs. between-subject effects of intranasal oxytocin on the neural response to cooperative and non-cooperative social interactions". In: *Psychoneuroendocrinology* 78 (2017), pp. 22–30.

[49] C. De Dreu et al. "Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense". In: *Biological Psychiatry* 79.3 (2016), pp. 165–173.

[50] L. Kubzansky et al. "A heartfelt response: oxytocin effects on response to social stress in men and women". In: *Biological Psychology* 90.1 (2012), pp. 1–9.

[51] K. Miwa et al. "Impact of two types of partner, perceived or actual, in human–human and human–agent interaction". In: *Computers in Human Behavior* 28.4 (2012), pp. 1286–1297.

[52] N. Mesgarani et al. "Phonetic feature encoding in human superior temporal gyrus". In: *Science* 343.6174 (2014), pp. 1006–1010.

[53] E. Bigler et al. "Superior temporal gyrus, language function, and autism". In: *Developmental Neuropsychology* 31.2 (2007), pp. 217–238.

[54] R. Adolphs. "Cognitive neuroscience of human social behaviour". In: *Nature Reviews Neuroscience* 4.3 (2003), pp. 165–178.

[55] S. Gregory et al. "Punishment and psychopathy: a case-control functional MRI investigation of reinforcement learning in violent antisocial personality disordered men". In: *The Lancet Psychiatry* 2.2 (2015), pp. 153–160.

[56] H. Feldman. "How young children learn language and speech: Implications of theory and evidence for clinical pediatric practice". In: *Pediatrics in Review* 40.8 (2019), p. 398.

[57] M. Gallagher et al. "The amygdala and emotion". In: *Current Opinion in Neurobiology* 6.2 (1996), pp. 221–227.

[58] T. Brosch et al. "The Impact of Emotion on Perception, Attention, Memory, and Decision-making". In: *Swiss Medical Weekly* 143 (2013), w13786.

[59] C. Mosher et al. "Neurons in the monkey amygdala detect eye contact during naturalistic social interactions". In: *Current Biology* 24.20 (2014), pp. 2459–2464.

[60] S. Haber et al. "Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum". In: *Journal of Neuroscience* 20.6 (2000), pp. 2369–2382.

[61] S. Floresco et al. "Association basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat". In: *European Journal of Neuroscience* 10.4 (1998), pp. 1241–1251.

[62] J. Jones et al. "Basolateral amygdala modulates terminal dopamine release in the nucleus accumbens and conditioned responding". In: *Biological Psychiatry* 67.8 (2010), pp. 737–744.

[63] N. Cilz et al. "Oxytocin and vasopressin in the rodent hippocampus". In: *Genes, Brain and Behavior* 18.1 (2019), e12535.