TÉCNICO LISBOA

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

João Miguel Baptista Simões

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisors: Prof. Dr. Diana Maria Pinto Prata
Prof. Dr. Manuel Fernando Cabido Peres Lopes

Examination Committee

Chairperson: Prof. Dr. João Miguel Raposo Sanches
Supervisor: Prof. Dr. Diana Maria Pinto Prata
Members of the Committee: Prof. Dr. Nuno Miguel de Pinto Lobo e Matela
MD. Ângelo Rodrigo Neto Dias

October 2021

I declare that this document is an original work of my own authorship and that it fulfils all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Preface

The work presented in this thesis 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.

Acknowledgments

I started this journey without knowing its path, and with the help, support and guidance of so many people, I have reached this point.

Firstly, I would like to express my gratitude to my supervisor Dr. Diana Prata. The “glimpse” of the human brain provided by your expertise was, indeed, unbelievable. Thank you for your time and guidance throughout all this work and for giving me the opportunity to be a member of IBEB.

I would like to extend my sincere thanks to Carlotta Cogoni. Thank you for your invaluable contribution to this project, for your time, experience and unparallel support. Unfortunately, no words can be used to describe your help, but still voglio ringraziarti di cuore per tutto.

This project would not have been possible without the guidance of Ângelo Dias, who taught and mentored me through half of it, where my knowledge was almost “non-significant”.

I’m also very grateful to Vasco Sá for his valuable pieces of advice and for always allowing me one last question.

To my dear friends that I have met throughout these five years at Instituto Superior Técnico, and especially to Çalo, Bico and Nuno, words are not enough to describe your friendship. The coffee breaks, the laughs and all the time we spent together were extraordinary. The second rule of the engineer was, indeed, right.

To Joana, a very special thanks for your support, encouragement, patience, and profound belief in my abilities, especially during the moments where the difficulties seemed unsolvable.

To my family, these five years would not have been possible without your unconditional support. To my mother, thank you for always believing in me, for giving me this opportunity and for even putting me above your own needs. O meu muito obrigado mãe, porque sem ti, nada disto seria possível.

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.

Keywords

Reinforcement learning; reward prediction error; prisoner's dilemma; neuropeptides; striatum; amygdala.

Resumo

Contexto: Oxitocina (OT) e vasopressina (AVP) são neurotransmissores considerados essenciais em interações sociais, interagindo com regiões cerebrais do sistema dopaminérgico, responsável pela aprendizagem por reforço (RL).

Objetivo: Investigar o papel de OT e AVP na RL social e as correlações neuronais deste, em diferentes contextos sociais, avaliando o efeito do sexo do participante.

Métodos: Participantes (148 homens; 144 mulheres) receberam aleatoriamente OT, AVP ou placebo (PB) intranasais, e executaram o jogo do dilema do prisioneiro. Os dados comportamentais foram modelados usando RL, os parâmetros do modelo analisados, e os erros de previsão de recompensa (RPE) foram correlacionados com ativações no cérebro inteiro e em regiões de interesse.

Resultados: OT aumentou o parâmetro α_C (taxa de aprendizagem) em mulheres em jogos com parceiro humano, relativamente a parceiros computacionais. Globalmente, OT promoveu um aumento do parâmetro $Q0$ (propensão inicial) em jogos com humanos comparativamente ao computador. AVP aumentou o $Q0$ em homens quando jogaram com humanos comparando com OT, e com o computador comparando com PB. Tanto a amígdala como o núcleo caudado revelaram maior correlação RPE-ativação cerebral feminina sob OT, em jogos com computador relativamente a humanos. Apenas na amígdala, AVP aumentou esta correlação em homens em jogos com humanos, comparando com o computador.

Conclusões: OT pode promover uma propensão antissocial no início de interações de dilema social, aumentando comportamentos impulsivos femininos após cooperação. Contrariamente, AVP pode promover uma propensão social masculina. Adicionalmente, OT pode aumentar a aprendizagem social feminina em contextos não sociais. AVP pode aumentar a aprendizagem social masculina em contextos sociais.

Palavras Chave

Aprendizagem por reforço; erros de previsão de recompensa; dilema do prisoneiro; neurotransmissores; corpo estriado; amígdala.

Contents

List of Figures

List of Tables

Abbreviations

A	Actions
ADH	Antidiuretic Hormone
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
ASD	Autism Spectrum Disorder
ATP	Adenosine Triphosphate
AVP	Vasopressin
AVPR1A	Vasopressin Receptor 1A
AVPR1B	Vasopressin Receptor 1B
AVPR2	Vasopressin Receptor 2
BIC	Bayesian Information Criterion
BOLD	Blood Oxygen Level Dependent
C	Cooperate
$CMRO_2$	Cerebral Metabolic Rate of O_2
cRPE	Reward Prediction Error Parametric Regressor for the Game with the Computer Partner
D	Defect
DA	Dopamine
EPI	Echo-planar Imaging
EV	Explanatory Variable
FDR	False Discovery Rate
FID	Free Induction Decay
fMRI	Functional Magnetic Resonance Imaging
FOV	Field of View

FWE	Family Wise Error
FWHM	Full Width at Half Maximum
GLM	General Linear Model
GRF	Gaussian Random Field
hRPE	Reward Prediction Error Parametric Regressor for the Game with the Human Partner
LR	Learning Rate
MRI	Magnetic Resonance Imaging
NAcc	Nucleus Accumbens
NMR	Nuclear Magnetic Resonance
OEF	Oxygen Extraction Fraction
OT	Oxytocin
OXTR	Oxytocin Receptor
PB	Placebo
PD	Prisoner's Dilemma
PE	Parameter Estimate
PPC	Posterior Predictive Check
RF	Radiofrequency
RL	Reinforcement Learning
ROI	Region of Interest
RPE	Reward Prediction Error
RW	Rescorla-Wagner
SNc	Substantia Nigra Pars Compacta
STG	Superior Temporal Gyrus
SwE	Sandwich Estimator
TE	Echo Time
TR	Repetition Time
TT	Tit-for-Tat
VBA	Variational Bayesian Analysis
VTA	Ventral Tegmental Area
WB	Wild Bootstrap

1

Introduction

Contents

The present chapter initiates the thesis' subject by introducing its motivation and objectives. Lastly, the dissertation outline is provided, describing the structure of its chapters.

1.1 Motivation

Oxytocin (OT) is a neuropeptide produced by the human brain, which plays an essential role during social interactions. OT is produced by the hypothalamus, specifically by the neurons of the paraventricular and supraoptic nuclei [1], being later secreted to other brain regions, namely the hippocampus, the brainstem, the amygdala, the striatum, and many others [2], not only associated with emotional and behavioural functions, but also with reward [3]. These regions also belong to the dopaminergic system, a system that has important roles in motivation, reward, reproductive and maternal behaviours and Reinforcement Learning (RL) [4]. Thus, those two systems influence and interact with each other, with previous studies showing that OT has multiple binding sites within the dopaminergic system [5, 6].

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 neuropeptides systems, is known to signal reward and prediction, influencing future behaviour [7]. Striatum activation is also associated with the processing of Reward Prediction Error (RPE) signals, which are produced through phasic activation of dopamine (DA) neurons [8], with RPE being the difference between a reward and its prediction. Only two studies have started to investigate this role: *Ide et al.* [9], by using computational modelling and trial-by-trial RPE, found that OT attenuated the RPE encoding during a social interaction; *Kruppa et al.* [10], 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 [2], but also presenting similar affinity for the same receptors [11]. Furthermore, AVP is also reported to interact with the OT and Dopamine (DA) systems [12], so this neuropeptide might also have a role in the RL process in social contexts [13], 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 participant's sex) may have an important role in this process, which has not been studied yet. In order to study those effects, the Prisoner's Dilemma (PD) game can be employed.

To comprehend the human neural mechanisms behind RL, functional neuroimaging emerges as an essential tool to increase previous knowledge, having been extensively used to understand how DA influences decision-making and learning behaviour [14]. Additionally, computational modelling increases

the power of the previous tool, allowing to perceive trial-by-trial variations through the acquisition of variables and parameters, while assuming the occurrence of underlying cognitive mechanisms that influence human behaviour [15].

By understanding these psychological mechanisms, significant insights about the relationship between human social behaviour and mental health can be extracted, relevant for disorders as autism [10], schizophrenia [16], anxiety and depression [17], which are defined by considerable social deficits, most commonly left unadressed by the existing available pharmacological solutions.

1.2 Objectives

The present thesis aims to investigate the roles of OT and AVP on social RL and its neural correlates. To study the social context of RL, a specific task eliciting relationships must be used. A solution is the PD game, a task where two players play with each other, deciding to either cooperate or defect, with several emotions being evoked according to the four possible outcomes [18]. Nonetheless, this task also allows establishing relationships in non-social conditions (*i.e.*, playing with a computer partner instead of a human). Additionally, this thesis aims to understand the participant's sex and partner type (human or computer) effects in the previous roles.

With this purpose, previous behavioural data and pre-processed Functional Magnetic Resonance Imaging (fMRI) data, acquired while participants played a PD game from prior studies [18–21], will be analyzed in three different phases. In a first phase, the behavioural data will be modelled using computational RL models and different model parameters and trial-by-trial RPEs will be extracted. Afterwards, a behavioural analysis will be performed, using the model parameters as dependent variables in a mixed design repeated-measures Analysis of Variance (ANOVA), with the aim of evaluating behavioural differences across different sexes (male, female), partner types (computer, human) and drugs (Placebo (PB), OT and AVP). Finally, an fMRI analysis will be performed to understand which brain regions' activity might be correlated with the RPE signals.

1.3 Thesis Outline

This thesis is structured in 6 chapters. The current chapter 1 comprises the motivation behind this study, as well as its main aims. Afterwards, chapter 2 presents a review of the main theoretical concepts supporting this study, with section 2.1 exploring the mechanisms of social RL. Following, section 2.2 has an overview of computational RL models and their mathematical paradigms. Section 2.3 explains the principles of functional resonance imaging, the Blood Oxygen Level Dependent (BOLD) signal and their general analysis. Section 2.4 provides a summarized state of the art, mentioning the most related

neuroimaging studies to the present research. The last section of this chapter (2.5) consists of the hypotheses of both behavioural and fMRI analyses, based on previous evidence.

Then, chapter 3 starts by characterizing the participants enrolled in this study, the study design, and the fMRI data acquisition and pre-processing process (section 3.1). Afterwards, it focus on the methodology employed, starting by describing the computational RL models created (section 3.2), followed by a description of the behavioural parameter analysis (section 3.3) and by the subject-level and group-level fMRI analysis, including a Region of Interest (ROI) analysis (section 3.4).

In chapter 4, the results acquired are summarized in three sections. The first section (4.1) displays the results from the RL models, including their estimation and evaluation. The second section (4.2) reports the results from the behavioural parameter analysis, together with results from *post hoc* analyses. The last section (4.3) reports the whole-brain and ROI fMRI results. Chapter 5 presents the discussion of the results and is also organized in three sections, one for each section of the previous chapter (sections 5.1, 5.2 and 5.3).

Finally, chapter 6 summarizes the main conclusions drawn from this study, including its limitations, and provides suggestions for future work.

2

Background

Contents

This chapter describes the main underlying theoretical concepts of this study. Specifically, the mechanisms of social RL are introduced, as well as the computational RL models and their mathematical paradigms. In addition, the physical principles of fMRI are also provided alongside a summarized state of the art. Finally, the hypotheses of both behavioural and fMRI analyses are presented.

2.1 Social Reinforcement Learning

The RL expression derived from the concept of reinforcement from the psychology field is highly associated with the theory of operant conditioning developed by the psychologist B. F. Skinner. Skinner [22] defines reinforcement as an operation performed upon an organism that increases the repetition of a behaviour. Specifically, social RL is related to social interactions that increase the frequency of a specific behaviour (*e.g.*, cooperation may act as a reward and reinforce trust [23]).

The beginning of the scientific study of RL dates to the 20th century with the work of Ivan Pavlov. Pavlov [24] made various conditioning studies, and noticed that if a bell was rung multiple times before presenting food to a dog, it started to be conditioned by the sound, and began salivating after the bell was rung, instead of salivating after the food delivery. This experiment led Pavlov to hypothesize that the experience created a connection between the neurons responsible for the bell detection and the preexisting anatomical connection between the salivary glands and the food perception [25].

By the end of the century, Wolfram Schultz performed experiments on monkeys while recording the activity of the midbrain's dopaminergic neurons [26]. By performing a choice task, monkeys had to press one of two levers signalized by a lighting cue, where the correct one always led to a reward (a drop of apple juice) [26]. Initially, monkeys pressed the wrong lever multiple times, while various dopaminergic neurons responded to the delivery of the reward, but remained silent during the cue exposure [26]. After the task performance was established, the accuracy of the correct pressed lever increased, while fewer dopaminergic neurons responded to the reward, and started to respond to the cue illumination [26]. A hypothesis arose, suggesting that the activity fluctuations of DA neurons are behind cognitive and learning behaviour [26].

Furthermore, multiple studies by Pendleton Montague [27–29] on honeybees revealed that fluctuations of DA activity are associated with RPE, regarding future reward predictions made by the cerebral cortex. In other words, DA bursts occur when the rewarding events are unpredicted. Thus, this allows one to understand the reason underlying the increase of DA during the initial training phase, when the reward is received unexpectedly, and also during the later training phase, when a stimulus is already associated with a future reward, and an unexpected stimulus leads to a prediction error.

On the other hand, a negative prediction error occurs when a reward is expected but is not acquired. *Schultz et al.* [30] analyzed this phenomenon on monkeys, by initially training them to predict a reward

after a visual cue and, in some trials, not providing the reward. Physiologically, this study allowed to perceive that the firing rate of DA neurons decreases under the baseline when a negative prediction error occurs [30].

To extract conclusions about the RL system in humans, a non-invasive technique must be used. A solution is the fMRI technique, which will be explained in more detail in Chapter 2.3. At the beginning of the 21st century, new studies using this technique were performed, showing significant activation of the the striatum and the medial orbitofrontal cortex when unpredicted stimuli were presented [31]. Moreover, combining this technique with computational RL models allows to search for hidden variables that might influence human decision and learning (*e.g.*, RPEs) [32].

2.1.1 Dopamine and Striatum

In 1910, DA was first synthesized by James Ewens and George Barger [33]. DA is a neurotransmitter, also known as decarboxylated amine 3,4-dihydroxyphenylethylamine, or 3-hydroxytyramine [34], and, is synthesized in both the central nervous system and periphery [4]. Specifically, its major sources in the central nervous system are the midbrain dopaminergic neurons, located in the diencephalon, mesencephalon and olfactory bulb [35].

As previously described, DA is known to be a promoter of RL, and this idea dates to the 1950s [36] [37], where animals, due to electrical stimulation of the medial forebrain bundle, would execute specific actions, while dopaminergic antagonists would decrease this behaviour [38].

Based on these findings, studies began to focus on the major DA pathways, since they might have an important role in rewarding. Specifically, around 90% of the midbrain dopaminergic neurons are located within the ventral part of the mesencephalon [35]. From this area, the dopaminergic neurons project to several pathways: the nigrostriatal pathway, which starts in the Substantia Nigra Pars Compacta (SNc) and connects to the caudate-putamen (also called dorsal striatum), the mesolimbic pathway, which is originated in the Ventral Tegmental Area (VTA) and reaches the Nucleus Accumbens (NAcc) and olfactory tubercle (also known as ventral striatum), the septum, amygdala, hippocampus, and the mesocortical pathway, which also starts in the VTA and extends to the prefrontal, cingulate and perirhinal cortex [35]. Regarding their functions, while the first pathway is focused on planning skeletomuscular movements, generating behavioural outputs, the last two are associated with behaviour related to emotions, such as motivation and reward [35].

As one might see, the striatum (divided into ventral and dorsal striatum) is highly related to the mesolimbic and the nigrostriatal reward pathways, with the latter influencing the direct and indirect output pathways. These two pathways play an essential role in regulating the basal ganglia output, leading to behaviour regulation. In fact, the activation of the direct (or Go) pathway, that expresses the D1 DA receptor, promotes the inhibition of the basal ganglia's main output, and could lead to the disinhibition of

the thalamus, which has connections with the motor cortex, promoting movement and responding [39, 40]. On the other hand, the indirect (or NoGo) pathway, that expresses the D2 DA receptors, promotes the disinhibition of the basal ganglia's main output, and could lead to the inhibition of the thalamus, suppressing movement and responding [39, 40]. These two pathways are also known as the "Go/NoGo" model [39].

In order to increase the knowledge of how DA and the striatum directly influence decision-making and learning behaviour, a variety of studies focuses on exploiting the connection between reward and go/no-go tasks. These tasks consist of allowing the participant to choose one of multiple stimuli (go), or none (no-go) to win an outcome or not be punished [41]. Other studies focus on go tasks, where participants only need to choose between stimuli, in order to win an outcome [42]. In fact, go choices are associated with increases in DA (positive RPEs [43]), allowing the reinforcement of the direct pathway [40], while the no-go choices are associated with dips in DA (negative RPEs [43]), reinforcing the indirect pathway [40]. By using this type of task, genetic studies revealed that polymorphisms related to the function of D1 and D2 striatal DA receptors are predictors of Go and NoGo learning [42]. Furthermore, the reward expectation also influences the striatum activity, since a larger expectation of reward leads to a decreased striatum activity for a no-go action, and an increased activity for a go action [41]. Also, increased striatum activity occurred when outcomes were better than the ones predicted [41].

2.1.2 Oxytocin and Vasopressin

In 1953 and 1954, Vincent du Vigneaud [44] synthesized and sequenced OT and AVP, also denominated as Arginine–Vasopressin (AVP) or Antidiuretic Hormone (ADH), two neuropeptides of the human brain. These neuropeptides have multiple similarities, both being nine-amino-acid peptides, synthesized in the hypothalamus by neurons of the paraventricular and supraoptic nuclei, and later stored in the posterior lobe of the pituitary gland [1]. Afterwards, they are released to the peripheral circulation and to other brain regions, behaving as neurotransmitters and neuromodulating through their interaction with brain receptors. Specifically, these molecules are currently known for having four receptors, OT has one, the Oxytocin Receptor (OXTR), and AVP has three, the Vasopressin Receptor 1A (AVPR1A), the Vasopressin Receptor 1B (AVPR1B) and the Vasopressin Receptor 2 (AVPR2) [2]. These two neurotransmitters move from the hypothalamus to reach multiple brain regions, namely the hippocampus, the suprachiasmatic nucleus, the brainstem, the bed nucleus of stria terminalis, the amygdala and the striatum [2], which are brain areas associated with reward, emotion and social behaviour [3].

As one might see, these regions are shared by the dopaminergic system, with multiple studies showing that OT and AVP influence and interact with this system [18], not only in social affiliative behaviours [45], but also in maternal provision of care, attachment and motivation [18]. Especially, OT and DA provide reward social behaviours [46, 47], reinforce adult men bonds [18] and modulate cooperation,

trust and revenge behaviours behind decision-making [23].

Due to being present in brain regions associated with social cognition, a considerable number of studies have been focused on the social cognition effects of OT and AVP.

Specifically to OT, early studies have revealed prosocial behavioural effects. In contrast, recent studies revealed that OT's effect might differ, being context-dependent and, due to this reason, multiple hypotheses have been developed in order to understand the role of OT in social contexts. The prosocial hypothesis states that administration of OT leads to prosocial behaviour [48], not only increasing trust [49], even after betrayal [50], but also generosity [51, 52], emotion recognition [53] and the effect, attention and memory of positive social stimuli [54–56]. Furthermore, it promotes cooperation [57] and, in conflict situations, stress reduction and communication [58]. However, OT's antisocial or pro-self behaviours were shown in recent studies, increasing aggression [59], envy, gloating [60], and ethnocentrism, leading to in-group favouritism [61]. In order to accommodate both pro and antisocial behaviours, the social salience hypothesis was theorized, stating that OT increases awareness and concentration to social stimuli, regardless of their valence (positive or negative) [48].

Regarding AVP, it promotes mutual cooperation [62], increases awareness and memory for social behaviours [63], conciliatory gestures in women [64] and the likelihood of reciprocating cooperation among men [18]. As OT, AVP's effects are context-dependent, offsetting male-aggression in affiliative contexts [65] 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 [64]. It is also important to state that, even though these two molecules are very similar, while OT is anxiolytic, AVP is anxiogenic [11].

As previously mentioned, it is worth noting that the social effects of these two peptides are conditioned to multiple environmental and interindividual factors [66]. External factors as the person or group acting with the individual [18], or interindividual factors, namely the sex, attachment style or psychiatric disorders might influence the behavioural response of each individual [67].

2.1.3 A Trust Game - The Prisoner's Dilemma

In order to study the social behaviours of humans, multiple trust games were created, namely the Dictator Game, the Trust Game and the PD [68]. While the first two games focus on allowing the player to divide one resource between themselves and another player, the last one allows the participant to choose to cooperate or defect with a partner.

Specifically, the iterated PD is a model where two players play with each other, in order to elicit relationships established on reciprocal altruism [18]. 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 [18].

Table 2.1: Sequential-choice Prisoner's Dilemma game payoff.

		Player 2	
		Cooperates	Defects
Player 1	Cooperates	\$2 / \$2	\$3 / \$0
	Defects	\$0 / \$3	\$1 / \$1

A specific version of this game is used in multiple studies, the sequential-choice PD game, where Player 1 chooses, and Player 2 is then allowed to see Player 1's decision before choosing [18]. 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 [20]. 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 [18]. It is important to state that the participant may play as Player 1 or as Player 2, and the partner might be a computer or a human (which is also a pre-made algorithm). Table 2.1 summarizes the payoff system.

Furthermore, each possible outcome will elicit different emotions from the participants. *Neto et al.* [23] referred that mutual cooperation is related to trust, love, friendship and obligation, mutual defection is related to anger and rejection, and cooperation followed by defection may elicit indignation or hatred. Moreover, guilt, anxiety or joy after taking advantage of the partner with success are emotions triggered on the defector [23].

This game is commonly used while participants are under the effect of exogenous neuropeptides, in order to perceive behavioural and neural changes. Previous studies [18, 20, 21, 69, 70] using a PD task while participants were under the effect of intranasal OT and AVP reported behavioural and neural changes depending on the participants' sex.

Rilling et al. [18] found that, in men, OT increases the response of the caudate nucleus and amygdala due to reciprocated cooperation (CC), enhancing the reward and promoting learning that the partner can be trusted, and increases the probability of cooperating, after unreciprocated cooperation. OT also promotes male neural responses analogous to females under the effect of PB, and vice versa, which can be explained with an inverted U-shaped curve of OT response [20]. *Feng et al.* [69] reported that OT, in females, decreases the response of the caudate/putamen and may decrease the importance and reward of social interactions. This neuropeptide also decreases the amygdala and anterior insula activation

after unreciprocated cooperation (CD) in men playing with a human partner, and in women playing with a computer partner, reducing the stress caused by negative social interactions [21]. Additionally, OT decreases VTA activity due to reciprocated cooperation in women, reducing the importance of positive social interactions [70], and, in men, avoids the habituation of negative social interactions.

In men, AVP increases the probability of mutual cooperation and activates the AVP circuitry, after partner cooperation, which may be related to the fact that AVP is anxiogenic, avoiding conflict [18], while in women, it increased the probability of cooperating, after partner defection, a conciliatory decision to resume cooperation [20]. AVP also decreases the amygdala and anterior insula activation, due to unreciprocated cooperation among men [21].

Moreover, both neuropeptides lead women to anthropomorphize computers [20].

2.2 Computational Reinforcement Learning to Model Brain behaviour

Over the last decade, the importance of computational modelling has increased in the cognitive neuroscience field. In fact, these models allow to perceive trial-by-trial variations through the acquisition of variables and parameters, while assuming the occurrence of underlying cognitive mechanisms that influence human behaviour [15]. By joining this new information with neuroimaging, it is possible to study the implementation of those decision variables within the brain [15].

In fact, RL models are a type of computational models that are extensively used in the cognitive and social neuroscience field, whose main objective is to describe the way an agent (for example, a human) establishes relations with an environment (for example, a task) by acquiring feedback from it (for example, an outcome or reward), in order to create and process internal decision values (for example, reward expectations), depending on the possible decisions or actions to be performed. In other words, an action that results in positive feedback will increase the probability of that action being repeated, while negative feedback will decrease the chance of that action being performed again [15].

Although RL includes a variety of models, the Rescorla-Wagner (RW) model was one of the first mathematical models to explain learning of conditioning tasks. Moreover, this model has been applied to Pavlovian instrumental learning with punishment [41] or reward [41, 71, 72] conditions, fear conditions [73], but also to social learning from and with other ”humans” [9, 14].

The RW model is based on the difference between the predicted and real outcomes after making a decision, leading to a learning process driven by the error of reward expectations [15]. This difference is denominated as the reward prediction error (RPE or δ) and the following equation calculates it:

$$\delta_t = R - V_{t-1} \tag{2.1}$$

Specifically, R represents the real outcome, which can be positive or negative, and V_{t-1} represents

the predicted one (on the previous trial), also called action value [15]. By using the δ, it is possible to update the predicted outcome for the following trial (V_t):

$$V_t = V_{t-1} + \alpha \times \delta_t \tag{2.2}$$

Where α represents the Learning Rate (LR), a parameter that will be estimated using the model fitting and is constrained between 0 and 1 ($0 < \alpha < 1$) [15]. This parameter will adjust the impact of the δ on the next prediction [15].

Together, Equations 2.1 and 2.2 can be merged into one (2.3), that summarizes the whole process. Note that, if more than one possible action exists, only the action value of that selected action (A_i) will be updated, while the others will assume the values of the previous trial.

$$V_t(A_i) = V_{t-1}(A_i) + \alpha \times [R - V_{t-1}(A_i)] \tag{2.3}$$

It is also worth mentioning that 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. Thus, the property of speed of learning is often attributed to α (also called as the learning step size) [15]. In other words, according to this notion, a higher α leads to faster learning. However, it is important to mention that a higher α cause considerable increases of action values after positive outcomes, and considerable decreases after one negative outcome, leading to oversensitivity [15]. Biologically, a higher α represents impulsive or unstable behaviours mostly based on the previous interaction, while a lower α represents more stable behaviours, based on a continuous-time valuation of multiple interactions.

Furthermore, after updating the action values (using Equation 2.3), the following phase aims to perform a new decision on the subsequent trial, by using those values. This is performed by using the Softmax choice rule [74], which calculates the likelihood of performing each action (A_i) on trial t, $p_t(A_i)$, according to the following equation:

$$p_t(A_i) = \frac{e^{\beta \times V_t(A_i)}}{\sum_i e^{\beta \times V_t(A_i)}} \tag{2.4}$$

When there are only two options (B and D), Equation 2.4 can be simplified into:

$$p_t(B) = \frac{e^{\beta \times V_t(B)}}{e^{\beta \times V_t(B)} + e^{\beta \times V_t(D)}} = \frac{e^{\beta \times [V_t(B) - V_t(D)]}}{e^{\beta \times [V_t(B) - V_t(D)]} + 1} \tag{2.5}$$

While the probability of selecting the action D on the trial t is defined by:

$$p_t(D) = 1 - p_t(B) \tag{2.6}$$

Where $p_t(B)$ is the probability of option B being selected, $p_t(D)$ is the probability of option D being

selected, and β is the inverse temperature parameter ($\beta > 0$), which quantifies the consistency of choices [15]. In other words, the lower the β, the higher the randomness of the choices, while a higher β represents a higher consistency of choices, in detriment of higher reward choices [15]. Biologically, a lower β is associated with exploratory behaviours, while a higher β represents more similar, consistent behaviours. As one might see, the higher the value of $V(B)$, the higher the probability of B being selected on trial t.

Model parameters are estimated by minimizing the difference between the model and behavioural data. This is accomplished by computing the posterior probability distribution over model parameters, given the data ($p(\theta|y,m)$), *i.e.*, the likelihood of a set of parameters, given the observed data y, and a model m [15]. The posterior distribution can be computed based on the Bayes rule:

$$p(\theta|y,m) = \frac{p(y|\theta,m)\,p(\theta|m)}{p(y)} \tag{2.7}$$

Since $p(y)$ will work as a normalizing constant [15], the posterior distribution can be approximated to:

$$p(\theta|y,m) \propto p(y|\theta,m)\,p(\theta|m) \tag{2.8}$$

Where $p(y|\theta,m)$ is the likelihood function which describes how likely a given set of individual choices y is, given the model m and the set of parameters θ, and $p(\theta|m)$ represents the prior distribution of model parameters. The likelihood function corresponds to the choice probabilities calculated which, in RL models, is the softmax function [15].

Regarding model selection, to select the best model from a set of candidates, the posterior probability of a model m, given data y, $p(m|y)$, must be determined [15] [75]. Through the Bayes rule:

$$p(m|y) = \frac{p(y|m)\,p(m)}{p(y)} \tag{2.9}$$

Following the previous procedure, the posterior probability can be approximated to:

$$p(m|y) \propto p(y|m)\,p(m) \tag{2.10}$$

Assuming that all models are *a priori* equally likely ($p(m)$ being equal for all models), the key quantity is $p(y|m)$, known as model evidence, *i.e.*, the probability of the data y under the model m [76]. Thus, the model that best explains data is the one with highest model evidence [76], which can be computed through the following equation:

$$p(y|m) = \int p(y|\theta,m)\,p(\theta|m)\;d\theta \tag{2.11}$$

According to this equation, model evidence is an integration over model parameters and, therefore,

accounts for model complexity, which avoids overfitting [77]. This integration is often computationally intractable, and, therefore, an analytic approach to compute it is needed [76].

Several computational libraries have been designed to help performing parameter estimation and model selection. One of the most widely used libraries is the Variational Bayesian Analysis (VBA) toolbox of MATLAB, which uses variational Bayesian approaches to compute both parameter posteriors and model evidence [78]. Variational Bayesian approaches rest on optimizing a free-energy lower bond $F(q)$ to the model evidence, with respect to an approximate conditional density $q(\theta)$:

$$F(q) =< ln\, p(\theta|m) + ln\, p(y|\theta, m) - ln\, q(\theta) >_q = ln\, p(y|m) - D_{KL}(q(\theta); p(\theta|y, m)) \tag{2.12}$$

with D_{KL} being the Kullback-Liebler divergence. The maximization of $F(q)$ is performed by using the Laplace approximation of the posterior distribution $p(\theta|y, m)$, minimizing the D_{KL}, which is always positive or zero (when the terms $q(\theta)$ and $p(\theta|y, m)$ are equal) [78]. Hence, through iterative optimization, D_{KL} will tend to 0 and $F(q) \approx ln\, p(y|m)$, meaning that the $F(q)$ can be used as an approximation to the log model evidence and $q(\theta)$ as an approximation of the posterior distribution over parameters. Therefore, the higher the free-energy value, the better the model fits to data [78].

It is also worth noting that approximations to model evidence other than free-energy are frequently used. Two of the most widely used analytic approaches of model evidence are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), known as limit approximations [76]. In their different ways, both attempts to resolve the overfitting problem by using penalty terms for the number of parameters (while the BIC scales this number by the log of the number of observations, n, AIC only subtracts the number of parameters p) [76] shown by the following equations:

$$AIC =< log\, p(y|\theta, m) >_q - p \tag{2.13}$$

$$BIC =< log\, p(y|\theta, m) >_q - \frac{p}{2} log(n) \tag{2.14}$$

However, this type of approximation tends to be wrong in multiple cases, namely, when a parameter is added to the model, while having a similar effect as another one [76]. Thus, while the AIC and BIC would increase the complexity of the model, the free-energy would not change its true complexity, meaning that the AIC and BIC are prone to overestimate the impact of increasing or decreasing the number of parameters on the complexity of the model [76].

Although fitting various models and performing model comparison are required processes, so that the models' ability to fit the data and its generalizability are assessed and compared, model validation must also be performed. This process is required since model comparison is performed on a relatively small scale and does not assure that the chosen model predicts and explains the behavioural events of

interest [15]. The solution is to use the Posterior Predictive Check (PPC) procedure, which evaluates the divergences between the model predictions and the observed data [15]. In other words, after fitting the model to the data, the PPC uses the parameters' joint posterior distributions to calculate new predictions [15], which are then compared to the observed data.

Herein, the PPC process can be applied by simulating artificial choices for each subject using the winning model's parameters [15]. Then, simulated and real data are compared, and the same analyses performed on real data are also performed on the simulated data [15].

Moreover, RL models are prone to poor identifiability, which might be exacerbated when the quantity of information is deficient [79]. In order to evaluate these problems, a parameter recovery analysis must be performed using the same process to acquire artificial data. Then, the individual parameters for each artificial participant must be estimated. Finally, the parameters obtained from the real data and the ones from the artificial data are compared, and model changes must be performed if there are considerable divergences between both data [80].

2.3 Functional Magnetic Resonance Imaging

fMRI is a group of imaging techniques based on the BOLD effect [81, 82]. These techniques aim to display changes in brain activity and metabolism within different regions and at different time points in the form of activation maps [82].

Furthermore, fMRI techniques have been increasingly applied to a higher number of studies due to being non-invasive (the injection of a radioisotope or pharmacologic molecule is not required), having a good spatial resolution, availability and relatively low cost (since they can be conducted on clinical scanners with 1.5T) [82].

2.3.1 Principles of Magnetic Resonance Imaging

The Magnetic Resonance Imaging (MRI) techniques intrinsically depend on the properties of the Nuclear Magnetic Resonance (NMR) signal [81]. The NMR research was initiated in 1946 based on the work of two physicists and their groups: Felix Bloch and Edward Purcell [83, 84].

The phenomenon of NMR is based on the spin (intrinsic angular momentum) of protons and neutrons, components of the atomic nucleus [81]. However, not all nuclei possess a net spin (since only some quantized values are allowed). In fact, net spin is a characteristic of nuclei that have an odd number of particles or an odd number of protons and neutrons [81]. Regarding the human body, the most abundant nucleus with a net spin is the hydrogen ^{1}H (due to being a component of H_2O), being the focus of the MRI [81].

When a proton is under the effect of a magnetic field, it is exposed to the torque force, causing a rotation of the dipole until it is aligned with the field [81]. However, if the proton owns an angular momentum, as the ^1H, the proton spin axis will precess around the axis of the magnetic field, instead of aligning with it [81]. The precession frequency of the proton is denominated as Larmor frequency, ω_L, and is directly proportional to the magnetic field [81], described by the following equation:

$$\omega_L = \gamma \times B_0 \tag{2.15}$$

Where γ represents the gyromagnetic ratio, and B_0 the strength of the magnetic field [81].

Furthermore, the spins are separated into two quantized energy states, one with higher energy with the nuclear moment aligned antiparallelly to the magnetic field, and another with lower energy with the nuclear moment being aligned parallelly [85].

In order to measure an NMR signal, a Radiofrequency (RF) pulse is used to create a magnetic field $B1$, substantially smaller and perpendicular to B_0, in order to derange spins into an excited state [85]. This pulse will tip the equilibrium magnetization, M_0 (which results from the difference between the dipoles oriented with and against the magnetic field $B1$) to the transverse plane, in order to oblige that all dipoles precess at the same rate [81]. Furthermore, since the RF frequency is equal to the precession frequency, resonance takes place [81]. A few moments later, the RF field is deactivated, allowing the spins to go back to equilibrium, through a process called relaxation [85]. Relaxation comprises two events: the restoration of the spins' parallel orientation with the B_0 field, denominated as longitudinal relaxation, and their transverse dephasing, denominated as transverse relaxation. While the first requires the T_1 time constant to occur, the latter decay is characterized by the T_2 time constant [81]. Since the different human body tissues possess different T_1, T_2 and densities, ρ, these three constants might be used as the contrast basis of MRI [81].

Due to magnetic field changes, a current is induced in a coil (through the Faraday's law of induction), allowing to measure the NMR signal (denominated as the Free Induction Decay (FID)) which is proportional to the precessing magnetization magnitude [81]. The decay of this signal is proportional to the transverse magnetization decay, which occurs due to variations on the spins' Larmor frequencies of the target [81]. This phenomenon occurs not only due to spin-spin interactions, which are based on interactions between proximal nuclei, causing random fluctuations of low frequency in the local magnetic field, characterized by the T_2 time constant, but also due to magnetic field inhomogeneities that arise from magnetic susceptibilities of proximal tissues, characterized by the T_2^* time constant [81]. Specifically, on tissues close to blood vessels, the deoxyhemoglobin influences the T_2^*, since susceptibility variations are prone to occur [81]. In fact, this phenomenon is responsible for the BOLD contrast that can be used in fMRI [81].

Furthermore, since variations caused by magnetic susceptibilities are non-random and caused by

offsets of constant fields, it is possible to amend the effect of field inhomogeneities, and have a higher signal, which no longer decays with the T_2^* time constant [81].

2.3.2 Blood Oxygen Level Dependent Signal

The BOLD signal is based on the level of oxygen, as the name states. In the blood flow, oxygen is mostly transported by erythrocytes from the lungs to other tissues, through the hemoglobin protein [81]. In a healthy adult, this protein comprises two α and two β subunits, each bound to a heme group containing an iron atom. Each iron atom allows carrying one O_2 molecule, leading to a maximum of four O_2 molecules per hemoglobin [86]. Due to this reason, there are two forms of hemoglobin in the human body: the oxyhemoglobin, which occurs when oxygen binds to a heme group, or the deoxyhemoglobin , which does not have any bounded oxygen [87]. Furthermore, oxygen attachment to hemoglobin changes its magnetic properties, since deoxyhemoglobin is paramagnetic (having high magnetic susceptibility) and oxyhemoglobin is diamagnetic (having low magnetic susceptibility) [87]. Due to being paramagnetic, deoxyhemoglobin changes the susceptibility property of the blood and induces field distortions near vessels when blood becomes more deoxygenated, which not only decreases the transverse relaxation (due to the effect of the T_2^*), but also the signal [88]. On the opposite, blood with full oxygenation has the same susceptibility as other tissues from the brain, leading to a higher signal [88].

As previously stated, the blood's oxygen level is proportional to neuronal activity. In fact, there are four main basic processes to enable neuronal signalling: the preservation of the cell during rest, the production and propagation of action potentials, the pre-synaptic recovery and the post-synaptic recovery of action potentials, with the latter being the one with higher energy costs [81]. This occurs mainly due to the activity of the Na^+/K^+ pump, which requires a great amount of energy to move molecules against their gradient, with Adenosine Triphosphate (ATP) as the main source of energy [81]. In order to produce this molecule, the brain mainly uses glucose and O_2 as precursors via the glucose metabolism [81]. In other words, it was expected that a higher neuronal activity would require an increased O_2 consumption, leading to a smaller signal; however, the opposite occurs [81].

At a physiological level, the activation of a brain area, increasing its activity and its oxygen demand, leads to an increase in the Cerebral Metabolic Rate of O_2 ($CMRO_2$), which also promotes a larger increase of the cerebral blood flow (rCBF) [81]. However, the amount of oxygen extracted from the blood cells, *i.e.*, the Oxygen Extraction Fraction (OEF) decreases, due to a higher local blood velocity in venules, capillaries and arterioles, decreasing the blood transit time, although other hypotheses exist to justify this fact [81]. The OEF decrease causes an increase in venous blood oxygenation (with a decrease of deoxyhemoglobin levels), creating a local increase of the MRI signal [81].

2.3.3 Subject-level (first level) fMRI Data Analysis

After acquiring the BOLD signal (and pre-processing it, which will not be covered), it is necessary to proceed with the statistical analysis of each voxel time-series from each subject's data, in order to detect local brain activations to the task previously executed by each subject [81].

2.3.3.A General Linear Model

General Linear Model (GLM) is, currently, the most used statistical approach to perform fMRI data analysis [89], modelling the data as a linear combination of model responses, regressors, or Explanatory Variables (EVs), $x(t)$ [81]. In fact, this data is obtained from the stimulation period when the subject is inside the MRI scanner, usually performing a task-based paradigm. Thus, a good model fit will plausibly justify that stimuli originated the data while the subject was being examined [89].

The first step of this approach requires the acquisition of a stimulus pattern, $u(t)$, which consists of a function that assumes the value of 1 while the stimulus is on, and the value of 0 while it is off [81]. However, the brain's hemodynamic response is not a basic stimulus function, since it might include short-term variations, such as delays from the stimulus start or undershoots [81]. To solve this issue, a hemodynamic response function, $h(t)$, is required in order to modulate the stimulus pattern, having a gamma function shape [81]. Thus, by assuming linearity, a closer approximation of the brain response to the stimulus, $x(t)$, can be acquired by using the convolution operation [81], through the following equation:

$$x(t) = u(t) \circledast h(t) \tag{2.16}$$

The next step requires not only to account for amplitude variations of the $x(t)$ signal, but also for noise in the acquired data [81] [89], which can be summarized by the following equation:

$$y(t) = \beta \times x(t) + \epsilon(t) \tag{2.17}$$

Where the $y(t)$ represents the time-series of one voxel, the β represents the weight of the explanatory variable, and the $\epsilon(t)$ represents the random noise [81]. This process performs a univariate analysis, assuming that each voxel's timecourse is independent [89]. Nevertheless, this equation could also be extrapolated to the time series of all voxels, where Y is a vector representing the time-series of all voxels, X is the design matrix, with all regressors, B is a vector representing the weights (one per regressor per voxel) and E is a vector representing the residual errors (one per voxel) [81]. Equation 2.18 summarizes the whole process.

$$Y = B \times X + E \tag{2.18}$$

By using the least-squares method, it is possible to calculate the estimates of each β value (also denominated as Parameter Estimate (PE)) that allow the best model fit to the data [81]. Specifically, the goal is to minimize the sum of squared residuals, obtained by subtracting the model from the data [81]. Since PEs will be calculated per voxel, they can be later used as statistical parameters for mapping [81]. Furthermore, the higher the PE, the better the model's fit, and its regressor allows a better explanation of the original data [89].

As stated before, it is also possible to use multiple EVs, each one being specific to one stimulus. With multiple EVs, the interest in understanding which stimulus originated the voxel main activation increased, which can be clarified by subtracting PE, also called "contrasts". In other words, if two EVs were used and must be compared, contrast weights (c, associated with a specific EV and stimuli) can be created [89], leading to the following inequality equation:

$$c_1 * \beta_1 + c_2 * \beta_2 > 0 \tag{2.19}$$

Through the combination of different c values, it is possible to obtain different comparisons and results. In fact, if $c_1 = 1$ and $c_2 = 0$ (also formulated as a contrast vector $c = [1, 0]$) equation 2.19 would become $\beta_1 > 0$, evaluating if the response to the stimuli 1 is greater than 0 [89]. If $c_1 = 1$ and $c_2 = -1$ (contrast $c = [1, -1]$), equation 2.19 would become $\beta_1 > \beta_2$, evaluating if the response to stimuli 1 is greater than the response to stimuli 2 [89]. Again, since new PEs will be created (through the combination of different contrast weights and subtractions), they can be later used for mapping.

Moreover, it is still necessary to perform statistical tests in order to verify if the acquired PEs are statistically relevant [89]. Thus, this can be done through the calculation of a T-test per voxel, which compares the PE with the uncertainty of its estimation, evaluating if the PE is significantly different from zero [89]. In other words, to test for the evidence of one effect (*i.e.*, comparison between two β parameters, or just single β magnitude), T-test can be applied [89]. Thus, the contrasts that originate these effects are also nominated as T-contrasts. However, if the objective is to test several contrasts simultaneously, in order to detect multiple activations generated by different contrasts (any linear combination of them), a F statistic, or F-test can be applied [89]. These contrasts are also nominated as F-contrasts and can be represented by a contrast matrix. For example, $c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ detects any responses to the stimuli 1 or to the stimuli 2 greater than 0 [89].

The acquired T and F values can also be converted into a p-value or a Z statistic [89]. Afterwards, a probabilistic threshold must be applied to the statistical map acquired, in order to select the voxels (*i.e.*, brain areas) that are activated, at a specific significance level [89].

2.3.4 Group-Level (second level) fMRI Data Analysis

In order to augment the sensitivity of an experiment, generalize specific conclusions to the whole population and compare brain activation areas between different groups of subjects or conditions, experiments can be run multiple times on the same subject, in multiple subjects, or in both [89]. Thus, a group-level analysis can be performed, using the action maps of the first-level analysis, after being registered to a standard space [89].

Similar to what has been described before, the GLM can be used to perform the group-level analysis. In this process, first-level effect size statistics are used as the dependent variable to integrate results for a specific group of subjects, or perform comparisons between distinct groups of subjects. To accomplish this, fixed-effects and random-effects analyses are performed [89]. While the first treats the interaction between the session and the effect as a fixed value, comparing the effect with the within-session error [90], the latter also accounts for between-session errors, reducing the number of data assumptions and allowing to extrapolate results to the whole population [89]. Note that the GLM regressors are created to subdivide fMRI data into minor groups or conditions. Afterwards, different statistical designs (*e.g.*, paired samples t-tests, full factorial ANOVAs) can be applied to test different hypotheses [91].

Once again, test statistics must be performed and thresholds applied.

2.3.5 The Multiple Comparison Problem

As previously stated, the outcome of each fMRI data analysis corresponds to a statistical brain map. Later, a probabilistic threshold (*e.g.*, $p < 0.05$) must be applied to each voxel within the map, in order to select the ones that were activated at a specific significance level [89].

In the scientific field, the most common threshold is the 0.05 p-value threshold. However, for this threshold, only by chance, 5% of the total number of voxels would be selected to be activated, leading to a considerable number of false-positive activations [81]. This issue is termed as the multiple comparison problem and, in order to solve it, corrections must be applied [81]. An initial solution was the Bonferroni correction, which decreases the significance level by dividing it by the total number of voxels [89], removing both false and true positive activations [92]. Another approach is the Family Wise Error (FWE) correction, which can be based on the Gaussian Random Field (GRF) theory. This theory considers the spatial smoothness of the map through the estimation of the number of statistically independent voxels, allowing a lower reduction of activations in comparison to the previous correction by a factor of 2–20 [89]. The False Discovery Rate (FDR) correction is another method that targets the expected fraction of false positives on the activations that survived, being more sensitive and less prone to false negatives than the FWE correction techniques [92]. Furthermore, each of these techniques can be used on different levels of inference, namely the voxelwise and the clusterwise inference [92]. Specifically, the voxelwise

inference treats each voxel in a univariate way, removing the ones that did not survive the threshold after applying one of the previous methods [92]. The clusterwise inference does not account for the possibility of each individual voxel being a false positive, but rather accounts for the probability of each region being a false positive as a whole, selecting statistically significant clusters, according to the number of contiguous voxels [92].

2.4 State of the Art

Functional neuroimaging has been considerably used to increase the knowledge of how social RL occurs within the human brain. In the last two decades, the number of studies regarding this topic increased substantially, with most of them focusing on how DA influences decision-making and learning behaviour. However, recent studies [9, 18, 20, 21, 23, 69, 70] have shown that OT and AVP might have a role in this process.

With the possibility of perceiving trial-by-trial variations, computational RL models have been used together with fMRI, allowing a higher comprehension of the human reward system.

Over the literature, the most common method of analyzing human learning behaviour is with go/no-go tasks. In fact, this type of task is applied in DA studies, which may also use computational learning models. In 2019, *McDougle et al.* [93] studied the brain's response to rewards, while performing a physical movement to choose one stimulus. Later, six different RL models were used, in order to perceive the one that best fitted the data. Finally, the best model was used to acquire the RPEs, being then applied to the first level fMRI analysis as a regressor. A similar approach was used by *Katthagen et al.*. [72], while investigating the striatal synthesis capacity and RPEs of unmedicated schizophrenia patients through a task where participants had to decide between two cards that might lead to a loss or win, based on probabilities. Again, six different RL models fitted the data. The best model was used to calculate RPEs, later used on the fMRI analysis as a regressor.

Instead of DA, in 2018, *Ide et al.* [9] studied the OT effect on males during RL, while participants performed a trust game task after inhaling a dose of OT. Afterwards, a computational RL model fitted the behavioural data, and the RPEs were acquired. Again, the RPEs were used on the first level analysis of the fMRI as a regressor. In 2019, *Kruppa et al.* [10] studied the OT effect on adult males with high-functioning ASD during RL, while subjects were required to press a button, in order to choose the category of stimuli, following a similar methodology as the others studies.

Furthermore, the effects of these neurotransmitters might change in different conditions. As previously stated, OT and AVP effects were influenced by the subject's sex, by the entity that established the social interaction, by multiple disorders and many other factors.

At the time of this writing, most of the studies using computational RL models focused on analyzing

the decision-making effect of DA, while using tasks that do not evaluate its social effect. Moreover, only two studies [9, 10] used computational RL models to investigate the OT effect on RL, and none used AVP. None of these studies compared the OT and AVP effects between sexes and between entity types (computer or human) with whom they established the social interaction, and both used a sample with less than 40 participants. Thus, the present study aims to narrow this gap by using a social task (the PD), while applying computational RL models in order to perceive the social effects of OT and AVP, in different groups and conditions, with a considerable sample.

2.5 Hypotheses

Based on previous evidence, hypotheses were created for the possible results from the parameter analysis (Section 2.5.1) and the fMRI analysis (Section 2.5.2).

2.5.1 Parameter hypotheses

RL model parameters (α and β) are important tools for perceiving RL's social effects and understanding its underlying cognitive mechanisms [15]. Thus, one might use those tools to study the effect of OT on the RL process. However, since there are no previous reports in the literature on how OT influences those parameters, the present thesis aims to clarify it. Nevertheless, some hypotheses per parameter can be advanced.

2.5.1.A Non-directional hypothesis that OT has a main effect on α and/or β

OT may lead to lower α: OT has been shown to increase social bonding [94–96] and secure relationship attachment [96]. Thus, one might hypothesize that a main effect of the drug is expected, with OT promoting a continuous-time valuation of all the interactions between two individuals and reducing the impact of the last interaction on the next one (lower α), in comparison to PB.

OT may lead to lower β: Following the same rationale, one might hypothesize that a main effect of the drug is expected, with OT reducing the disparity among choices (higher β), in comparison to PB.

OT may lead to higher α: Since OT is associated with defensive aggression focused on protecting and negating threat induced by out-groups [97], another expected main effect of the drug might be that OT promotes impulsive behaviours in response to interactions with a stranger or acquaintance, leading to a stronger influence of the latter interaction on the next one (higher α), in comparison to PB.

OT may lead to lower β: L-DOPA, a DA precursor, has been shown to decrease the β parameter [98], facilitating exploratory behaviour. Thus, since the OT pathway is closely linked and could mediate the release of DA on the mesolimbic dopaminergic system [99], while affecting the DA role in social salience

attribution [100], OT and L-DOPA might provide a similar effect on the RL circuit. Hence, another main effect of the drug is possible to hypothesize, with OT lowering the β parameter in comparison to PB.

2.5.1.B Non-directional hypothesis that partner type (human, computer) influences the effect of OT on α and/or β

Since the social context plays an important role in the RL process and social information facilitates learning [14], a two-way interaction effect between drug and partner might be hypothesized, with a higher effect of OT on the α and/or β, when playing with a human partner than with a computer partner, in comparison to PB.

2.5.1.C Exploratory analysis of the impact of sex of the participant on the effect of OT on α and/or β

As an exploratory analysis (there being no reports in the literature), a two-way drug by sex interaction on α and/or β will also be tested. The motivation behind this analysis is based on previous evidence reporting sex differences in the baseline levels of OT [101, 102].

2.5.1.D Complementary analyses of the impact of sex and partner on the effect of OT on α and/or β

As complementary analyses, for completeness and help of the interpretation of the latter hypotheses, the main effect of the partner, the main effect of the sex, and the two and three-way interactions (*i.e.*, sex vs partner, drug vs sex vs partner) will also be performed.

2.5.1.E Second exploratory analysis to test the main effect of AVP and all the above effects (*i.e.*, the modulation of sex and partner on its effect) on α and/or β

Since no evidence in the literature was found analyzing the effect of AVP on α and β parameters, an exploratory analysis will be performed. Nevertheless, the AVP pathway is closely linked with the OT pathway and the mesolimbic dopaminergic system [3], both associated with the social reinforcement system, while previous studies [18, 20, 21, 69] reported behavioural AVP effects, modulated by both sex and partner, while using a PD task.

2.5.1.F Further parameter analyses

As later will be explained, in order to achieve a better model fit, and based on the initial results acquisition, new parameters will be created. The alpha parameter will be subdivided into α_C and α_D, and a $Q0$

parameter will be added to the model. Although no hypotheses were performed on these parameters, they will be included on further analysis.

2.5.2 fMRI hypotheses

Throughout the literature, multiple studies have shown that RPE, an essential mechanism in RL, are highly correlated with the activity in the striatum [93,103,104], a brain region that belongs to the dopaminergic pathway.

2.5.2.A Directional hypothesis that OT has a main effect on RPE-striatal activation correlation

L-DOPA, a DA precursor, has been shown to increase the positive correlation between RPE and the NAcc, a striatum region [105, 106]. On the other hand, amisulpride, a DA receptor antagonist at a high dose, and memantine, an antagonist of NMDA receptors (which also regulate DA neuronal activity [107]), have been shown to decrease the positive correlation between RPE and the striatum [108]. In line with the rationale behind a previous hypothesis (parameter hypothesis 2.5.1.A, OT may lead to lower β), one might assume that L-DOPA, and predict that OT, provide a similar effect on the RL circuit, while amisulpride and memantine would have opposite effects in comparison to OT. Thus, in our analyses, a main effect of the drug is expected, with OT increasing the positive correlation between the RPEs and the striatum activity.

2.5.2.B Directional hypothesis that partner type (human, computer) influences the effect of OT on RPE-striatal activation correlation

Moreover, *Kruppa et al.* [10] 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 showed that the striatum region has major importance on the RL process [109], and that social information facilitates learning [14]. Hence, a two-way interaction effect between drug and partner might be hypothesized, with OT increasing the positive correlation between RPEs and the striatum activity, more when the participant is playing with a human partner than with a computer partner, compared to PB.

2.5.2.C Exploratory analysis of the impact of sex of the participant on the effect of OT on RPE-striatal activation correlation

As an exploratory analysis, a two-way drug by sex interaction on RPE-striatal activation correlation will be analyzed. Although there is no previous evidence studying the sex impact on the effect of OT while

performing an RPE fMRI analysis, multiple studies [20, 21, 69] found sex differences on the OT effect on the striatum, using a PD task.

2.5.2.D Complementary analyses of the impact of sex and partner on the effect of OT on RPE-striatal activation correlation

As a complementary analysis, for completeness and help of the interpretation of the latter hypotheses, the main effect of the partner, the main effect of the sex, and the two and three-way interactions (*i.e.*, sex vs partner, drug vs sex vs partner) will also be performed.

2.5.2.E Second exploratory analysis to test the main effect of AVP and all the above effects (*i.e.*, the modulation of sex and partner on its effect) on the RPE-striatal activation correlation

Similar to the first exploratory analysis, there is no previous evidence studying the main effect of AVP and the modulation of the sex and partner on its effect while performing an RPE fMRI analysis. However, following the same rationale used on the fifth parameter hypothesis, AVP might have a main effect on the RPE-striatal activation correlation. Furthermore, previous studies [18, 20, 21, 69] found AVP effects on the striatum, being also modulated by sex and partner, using a PD task.

2.5.2.F Third exploratory analysis to test all the above effects (main effect of OT, AVP and the modulation of sex and partner on their effect) on the RPE-amygdala activation correlation

A third exploratory analysis will also be performed, analyzing the correlation between RPEs and the amygdala activation, while also testing the above effects. The present analysis is not only motivated by the fact that the amygdala has multiple OT and AVP receptors [110–112], but also that the amygdala may play a role in the RL process [8, 113, 114].

3

Methods

Contents

The present chapter introduces the data, its acquisition and pre-processing. On top of this, an explanation of the methodology is also provided.

3.1 Participants and Data Characterization

The present study follows a PD task with fMRI data collected and pre-processed by the studies [18–21]. Thus, sections 3.1.1, 3.1.2, 3.1.3 and 3.1.4 are focused on describing the previous steps executed by those studies, in order to produce the data that is going to be used on this study. Furthermore, only the behavioural and fMRI data of the games with the participant playing as Player 1 is going to be used and analyzed.

3.1.1 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. 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 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.

The exclusion criteria of the initial sample are described in the study [21]. In minor detail, the main reasons to exclude participants were neurological disorders, substance abuses, cardiovascular diseases, endocrine disorders (such as diabetes), head trauma antecedents, psychiatric illnesses, recent persistent and disabling symptoms of asthma or migraine headaches, requiring medication adaptations, usage of medications with psychoactive effects over the previous year and claustrophobia.

3.1.2 Study Design

This study follows a sequential-choice PD game presented in section 2.1.3. Participants started by performing a PD tutorial and two practice trials. Although the beginning of the task and fMRI scan was planned to occur 40 minutes after drug administration, the average time across subjects was 42 minutes.

Before the beginning of each game, a picture of the participant's partner was shown inside the scanner. Each game was composed of 30 trials of a sequential-choice PD game while the participant was being scanned with fMRI, with each participant performing a total of 4 separate game runs. Specifically, participants were informed they were playing with a human partner for two runs, while for the other two runs, participants were informed they were playing with a computer partner. However, although they thought they were playing with the same-sex human introduced before the experiment for two runs, they were always playing with a pre-programmed computer algorithm (later explained in detail). Regarding the order of play, in one of the two runs for both computer and human partners, participants played as first mover (Player 1), while their partner was second mover (Player 2). For the other run, the roles were reversed. Furthermore, the sequence of human and computer sessions was balanced out across participants, with half of the participants being scanned in the sequence: participant as Player 1 playing with a human partner (H1), participant as Player 2 playing with a human partner (H2), participant as Player 1 playing with a computer partner (C1) and participant as Player 2 playing with a computer partner (C2); and the other half being scanned with the sequence: C1, C2, H1, H2. Following the study, participants were asked multiple questions regarding their experience throughout the task. Participants were rewarded with 2/3 of the total amount earned across the games.

While being scanned, participants view the stimuli (presented by the E-prime software, Psychology Software Tools, Pittsburgh, PA, USA) through projection into a mirror installed on the head coil. The response recording was performed by using a response box. Furthermore, each trial followed a timeline summarized in Figure 3.1. At the start of each trial, the trial number and a partner's photo were shown for 2 seconds. Subsequently, Player 1 had 4 seconds to select either to Cooperate (C) or Defect (D) and, if no option were selected within the 4 seconds period, the decision would be defection by default. Afterwards, Player 1's choice was shown to Player 2 for 1 second. A fixation epoch of either 2, 4 or 6 seconds then occurred, represented in Figure 3.1 by a cross symbol. As before, Player 2 had 4 seconds to decide between C or D. The trial outcome was then displayed for 4 seconds. A final fixation epoch of either 2, 4 or 6 seconds ended the trial. Each trial had an approximated duration of 20 seconds, while a session lasted approximately 12 minutes (since five null trials of 14 seconds of fixation occurred through the 30 trials of each session), and all four sessions required around 48 minutes.

The computer algorithm was created to simulate human strategies, differing if the subject was the first mover (Player 1) or the second mover (Player 2).

In the first situation, the computer algorithm always returned the participant's decision in the first trial, while in the other trials it reciprocated cooperation 67% of the time and reciprocated defection 90% of the time, as shown in Table 3.1. Both percentages are realistic, since the response to cooperation has a higher chance of being reciprocated, but also includes a possible sporadic defection, in order to increase the earning gains. In contrast, the response to defection leads to general defection (most

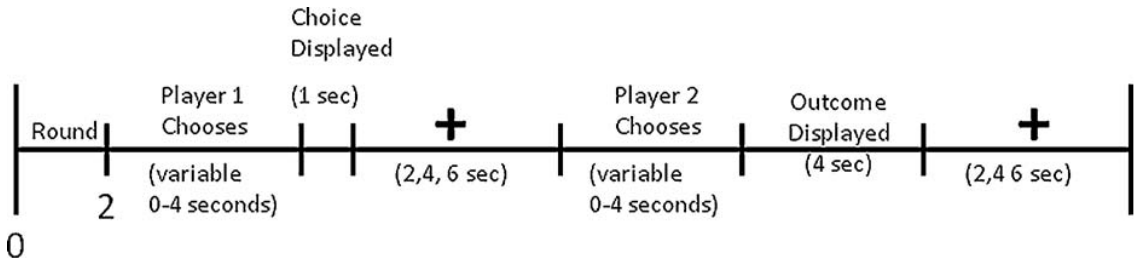


Figure 3.1: Timeline of the sequential-choice Prisoner's Dilemma trials. Figure extracted from [18].

related to feelings of anger and rejection). Furthermore, according to Table 3.1, the expected value of cooperating is $\$1.3$ and of defecting is $\$1.2$. Since those values are almost equal, the participants gain, on average, the same amount whether they defect or cooperate.

Table 3.1: Sequential-choice Prisoner's Dilemma game - partner's decision probabilities as player 2.

		Player 2	
		Cooperates	Defects
Player 1	Cooperates	67%	33%
	Defects	10%	90%

3.1.3 fMRI Data Acquisition

All images were acquired by performing functional scans using an Echo-planar Imaging (EPI) sequence, with a Repetition Time (TR) of 2000 ms, an Echo Time (TE) of 28 ms, a matrix size of 64 $\times$ 64, a Field of View (FOV) of 224 mm, slice thickness of 2.5 mm and with 34 axial slices. Intending to reduce magnetic susceptibility artifacts in the orbitofrontal region, TE was sightly reduced from the common value (32 ms). Each EPI duration was around 12 minutes, the same as the session duration. At end of each session, participants evaluated their emotional response to the four possible outcomes (DD, DC, CD and CC). Thus, sevenpoint Likert scales were utilized to rate the feelings of anger, happiness, fear, guilt, disappointment and relief.

3.1.4 fMRI Data Pre-Processing

The fMRI data was pre-processed using the Brain Voyager QX (version 2.0.8) software (Brain Innovation, Maatricht, The Netherlands). First, a six parameter 3-D motion correction was performed, followed by slice scan time correction applying cubic spline interpolation. Then, spatial smoothing was carried out by using a 8-mm Full Width at Half Maximum (FWHM) Gaussian kernel. Subsequently, a temporal

high-pass filter was applied, conducting linear and nonlinear trend removal, while also applying a cutoff frequency with a length of three cycles per run. Finally, a Talairach brain normalization was done.

3.2 Computational Reinforcement Learning Model Analysis

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 RW models, previously described in Section 2.2, to the present task. In other words, only two Actions (A) are possible to be chosen: C and D for both the participant and partner, and based on their combination, four possible outcomes can be achieved: win 3$ (DC), win 2$ (CC), win 1$ (DD) and do no (CD).

Model estimation was performed using the VBA toolbox [78] 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), which are later going to be detailed in Sections 3.2.1 and 3.2.2.

3.2.1 Simple Family

As previously described, all models calculated the likelihood of the participant performing each of the two actions in each trial, which can be achieved by using the Softmax choice rule. Thus, the probability of cooperating on the trial t, $p_t(C)$, is calculated through the following equation:

$$p_t(C) = \frac{e^{\beta \times V_t(C)}}{e^{\beta \times V_t(C)} + e^{\beta \times V_t(D)}} \tag{3.1}$$

The β parameter is estimated by being restricted to $\beta > 0$. The probability of defecting on the trial t, $p_t(D)$, is achieved by:

$$p_t(D) = 1 - p_t(C) \tag{3.2}$$

Nevertheless, the following five models differ in calculating the predicted outcome for the following trial, which allowed to test different behavioural hypotheses.

3.2.1.A Simple Model

The simple model is the basic RL model, represented by equation 3.3. As each outcome prediction for the present trial 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 for both decisions, represented by equation 3.4, assuming no initial tendency.

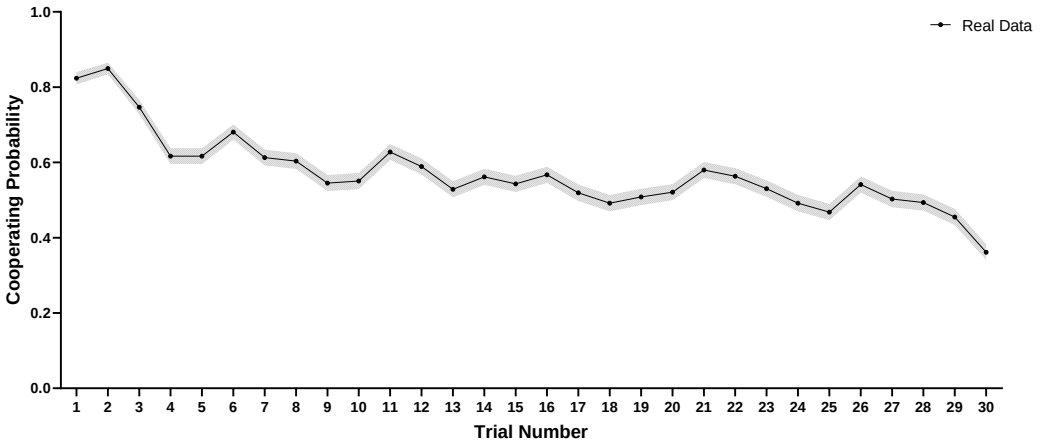


Figure 3.2: Trial-by-trial averaged cooperating probability from all participants obtained from the real data. Solid line indicates the mean, while the shaded error bar represents the standard error.

$$V_t(A) = V_{t-1}(A) + \alpha \times [R - V_{t-1}(A)], \quad t > 0 \tag{3.3}$$

$$V_t(C) = 0 \;\; and \;\; V_t(D) = 0, \quad t = 0 \tag{3.4}$$

Where A is either C or D based on the participant decision and the α parameter is estimated by being restricted to $0 < \alpha < 1$.

3.2.1.B Q0 Model

By analyzing the averaged cooperating probability from all participants in the initial trials, it is possible to conclude that, for most participants, there is a higher tendency to cooperate, as shown in Figure 3.2. Thus, a new parameter $Q0$ was added to the previous model in order to modulate this tendency (equation 3.5), similar to what has been done in the studies [41] and [73]. Biologically, this parameter represents the tendency that the participant has to cooperate or defect, at the beginning of the game and during the initial trials. Furthermore, the predicted outcome was still calculated by equation 3.3.

$$\begin{cases} V_t(C) = Q_0 \quad and \quad V_t(D) = 0, \quad t = 0 \quad and \quad Q_0 > 0 \\ V_t(D) = |Q_0| \;\; and \;\; V_t(C) = 0, \quad t = 0 \quad and \quad Q_0 \leqslant 0 \end{cases} \tag{3.5}$$

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 $-2 < Q0 < 2$.

3.2.1.C 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 3.6, similar to what has been done in the studies [72] and [115]. Thus, as before, the initial tendency modelled by the $Q0$, shown in equation 3.5, was still applied.

$$\begin{cases} V_t(C) = V_{t-1}(C) + \alpha_C \times [R - V_{t-1}(C)], & t > 0 \\ V_t(D) = V_{t-1}(D) + \alpha_D \times [R - V_{t-1}(D)], & t > 0 \end{cases} \tag{3.6}$$

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.

3.2.1.D 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 3.7, similar to what has been done in the studies [73] and [93]. The initial tendency modelled by the $Q0$, shown in equation 3.5, was still applied.

$$\begin{cases} V_t(A) = V_{t-1}(A) + \alpha_{pn(C)} \times [R - V_{t-1}(A)], & t > 0 \quad and \quad A(pn)_t = C \\ V_t(A) = V_{t-1}(A) + \alpha_{pn(D)} \times [R - V_{t-1}(A)], & t > 0 \quad and \quad A(pn)_t = D \end{cases} \tag{3.7}$$

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.

3.2.1.E 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 different importance and influence on the decision of the subsequent trial, and, consequently, induce 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 3.8. Once again, the initial tendency modelled by the $Q0$, shown in equation 3.5, was still applied.

$$\begin{cases} V_t(C) = V_{t-1}(C) + \alpha_{CC} \times [R - V_{t-1}(C)], & t > 0 \quad and \quad A(pn)_t = C \\ V_t(C) = V_{t-1}(C) + \alpha_{CD} \times [R - V_{t-1}(C)], & t > 0 \quad and \quad A(pn)_t = D \\ V_t(D) = V_{t-1}(D) + \alpha_{DC} \times [R - V_{t-1}(D)], & t > 0 \quad and \quad A(pn)_t = C \\ V_t(D) = V_{t-1}(D) + \alpha_{DD} \times [R - V_{t-1}(D)], & t > 0 \quad and \quad A(pn)_t = D \end{cases} \tag{3.8}$$

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.

3.2.2 Tit-for-Tat 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 [23]. This can be translated to the model by adding the parameter TT to the Softmax choice rule (represented in equation 3.1), depending on the partner's decision, originating the equation 3.9.

$$\begin{cases} p_t(C) = \dfrac{e^{\beta \times V_t(C) + TT}}{e^{\beta \times V_t(C) + TT} + e^{\beta \times V_t(D)}}, & A(pn)_{t-1} = C \\ p_t(C) = \dfrac{e^{\beta \times V_t(C)}}{e^{\beta \times V_t(C)} + e^{\beta \times V_t(D) + TT}}, & A(pn)_{t-1} = D \end{cases} \tag{3.9}$$

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 subsequent trial. Similarly, when the partner defected in the previous trial, the TT is added to the defection exponential ($e^{\beta \times V_t(D)}$). 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:

$$p_t(D) = 1 - p_t(C) \tag{3.10}$$

Afterwards, five new models were created (similar to the ones from the Simple family), being fitted using the changed Softmax rule. Table 3.2 summarizes the models information.

3.2.3 Model Selection and Validation

Model comparison and selection was performed using the VBA toolbox. First, the two model families were compared, and one was selected according to their estimated frequencies, which evaluates the likelihood of selecting one model for any subject randomly chosen [76]. Afterwards, the models of the winning family were compared, and, again, one was selected according to their estimated frequencies and exceedance probabilities, which evaluate the likelihood of a model being more suitable than any other [76]. The selected model was then validated through a PPC procedure, and a parameter recovery analysis was also performed.

Table 3.2: Models used to fit the data, their number of parameters and those parameters.

Model name	No. of Parameters	Parameters
Simple	2	α, β
Q0	3	α, β, $Q0$
2LR	4	α_C, α_D, β, $Q0$
2LR Partner	4	$\alpha_{pn(C)}$, $\alpha_{pn(D)}$, β, $Q0$
4LR	6	α_{CC}, α_{CD}, α_{DC}, α_{DD} β, $Q0$
Simple + TT	3	α, β, TT
Q0 + TT	4	α, β, $Q0$, TT
2LR + TT	5	α_C, α_D, β, $Q0$, TT
2LR Partner + TT	5	$\alpha_{pn(C)}$, $\alpha_{pn(D)}$, β, $Q0$, TT
4LR + TT	7	α_{CC}, α_{CD}, α_{DC}, α_{DD} β, $Q0$, TT

3.3 Behavioural Analysis

The behavioural analyses were performed using R (R Foundation for Statistical Computing, Vienna, Austria). Using the model parameters (α_C, α_D β and $Q0$) as dependent variables, a mixed design repeated-measures 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 effects and interactions 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.

3.4 fMRI Data Analysis

After fMRI scans' pre-processing (performed and described by [21]), 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) [116], including custom code written in MATLAB.

Of the 292 participants, only 253 (121 men and 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.

3.4.1 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 Reward Prediction Error 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 Reward Prediction Error 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 $c2 = [0, 0, 0, 1]$).

3.4.2 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 [117], 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 [117]. Furthermore, since our main goal is to analyze the differences in brain activations to RPEs between the different factors, the following twelve 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. A simplified design matrix containing the multiple EVs (as columns) is shown in Figure 3.3. Note that, depending on the group or conditions of the scan input (subject-level contrast input), different weights are assigned in the design matrix, with white representing the weight $+1$, grey the weight 0, and black the weight -1. For example, in order to separate the sexes in the EV7, a weight of $+1$ was assigned to males that were administered with PB, a weight of -1 to females that were administered with PB, and a weight of 0 to other participants that were not administered with PB.

In order to obtain FWE corrected results in SwE, a non-parametric Wild Bootstrap (WB) 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 [117] 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.

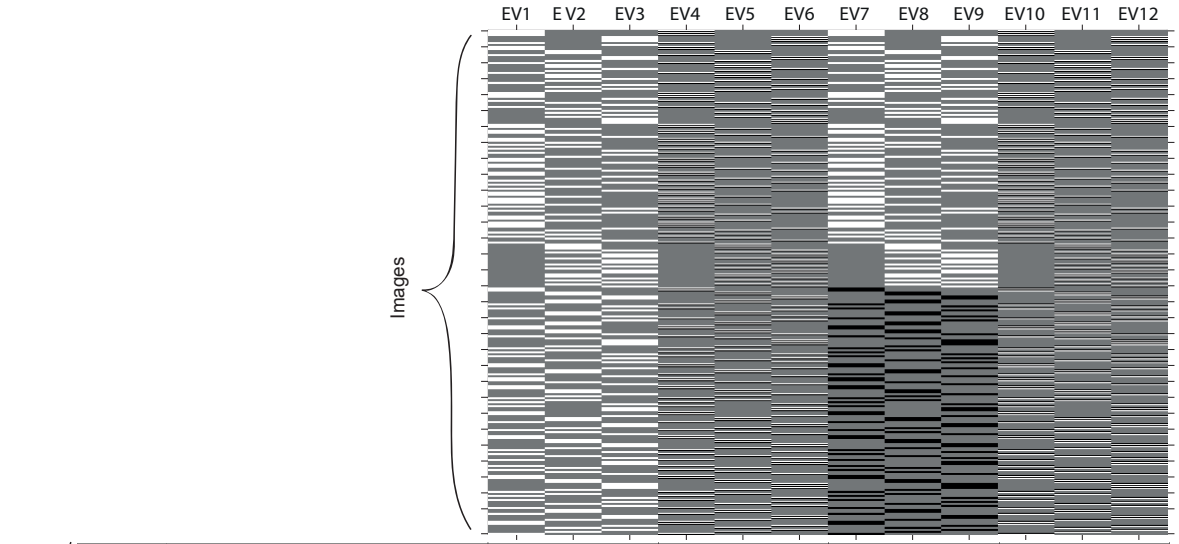
Finally, multiple T-contrasts were defined in order to compare activations between different conditions and groups, which are shown in detail in Figure 3.3. In fact, each contrast was composed of multiple weights, one for each column (EV) of the design matrix. Since the same combination of weights could lead to multiple contrasts, all possible combinations of each interaction are listed in the table.

3.4.2.A Region of Interest Analysis

An 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 [8], a striatum mask (comprising the NAcc, the caudate and the putamen, Figure 3.4(a)) and an amygdala mask (Figure 3.4(b)) 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%) [118], using the MRIcron software.

Two other separated masks, the right and the left caudate (Figures 3.4(c) and 3.4(d), respectively), were also used. Those masks were created by specific studies ([69] and [70]) using the same sample as the one applied on this thesis, 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.

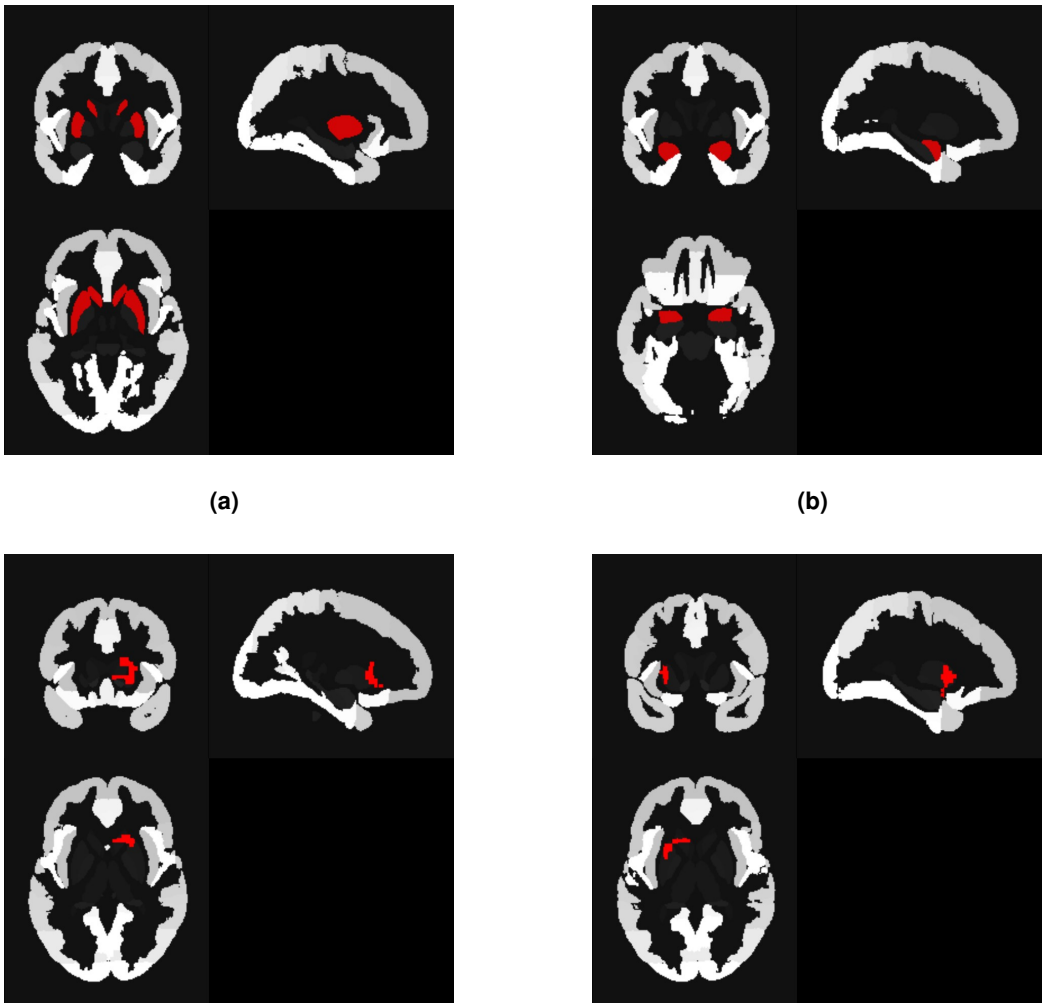
As in the whole-brain analysis, all main effects and interactions were considered statistically significant if the result was FWE corrected and the p-value was below 0.05.



Constrasts

		EV1	EV2	EV3	EV4	EV5	EV6	EV7	EV8	EV9	EV10	EV11	EV12
Player Main Effect	HUM>CPU	0	0	0	1	1	1	0	0	0	0	0	0
	CPU>HUM	0	0	0	-1	-1	-1	0	0	0	0	0	0
Sex Main Effect	Male>Female	0	0	0	0	0	0	1	1	1	0	0	0
	Female>Male	0	0	0	0	0	0	-1	-1	-1	0	0	0
Drug Main Effect	PB>OT	1	-1	0	0	0	0	0	0	0	0	0	0
	PB>AVP	1	0	-1	0	0	0	0	0	0	0	0	0
	OT>AVP	0	1	-1	0	0	0	0	0	0	0	0	0
	OT>PBO	-1	1	0	0	0	0	0	0	0	0	0	0
	AVP>OT	0	-1	1	0	0	0	0	0	0	0	0	0
	AVP>PBO	-1	0	1	0	0	0	0	0	0	0	0	0
2-way Interaction - Drug x Player	PB>OT in HUM>CPU = OT>PB in CPU>HUM	0	0	0	1	-1	0	0	0	0	0	0	0
	OT>PB in HUM>CPU = PB>OT in CPU>HUM	0	0	0	-1	1	0	0	0	0	0	0	0
	PB>AVP in HUM>CPU = AVP>PB in CPU>HUM	0	0	0	1	0	-1	0	0	0	0	0	0
	AVP>PB in HUM>CPU = PB>AVP in CPU>HUM	0	0	0	-1	0	1	0	0	0	0	0	0
	OT>AVP in HUM>CPU = AVP>OT in CPU>HUM	0	0	0	0	1	-1	0	0	0	0	0	0
	AVP>OT in HUM>CPU = OT>AVP in CPU>HUM	0	0	0	0	-1	1	0	0	0	0	0	0
2-way Interaction - Player x Sex	HUM>CPU in Male>Female = CPU>HUM in Female>Male	0	0	0	0	0	0	0	0	0	1	1	1
	HUM>CPU in Female>Male = CPU>HUM in Male>Female	0	0	0	0	0	0	0	0	0	-1	-1	-1
2-way Interaction - Drug x Sex	PB>OT in Male>Female = OT>PB in Female>Male	0	0	0	0	0	0	1	-1	0	0	0	0
	PB>AVP in Male>Female = AVP>PB in Female>Male	0	0	0	0	0	0	1	0	-1	0	0	0
	OT>AVP in Male>Female = AVP>OT in Female>Male	0	0	0	0	0	0	0	1	-1	0	0	0
	OT>PB in Male>Female = PB>OT in Female>Male	0	0	0	0	0	0	-1	1	0	0	0	0
	AVP>OT in Male>Female = OT>AVP in Female>Male	0	0	0	0	0	0	0	-1	1	0	0	0
	AVP>PB in Male>Female = PB>AVP in Female>Male	0	0	0	0	0	0	-1	0	1	0	0	0
3-way Interaction - Drug x Player x Sex	PB>OT in HUM>CPU in Male>Female = PB>OT in CPU>HUM in Female>Male = OT>PB in CPU>HUM in Male>Female = OT>PB in HUM>CPU in Female>Male	0	0	0	0	0	0	0	0	0	1	-1	0
	PB>AVP in HUM>CPU in Male>Female = PB>AVP in CPU>HUM in Female>Male = AVP>PB in CPU>HUM in Male>Female = AVP>PB in HUM>CPU in Female>Male	0	0	0	0	0	0	0	0	0	1	0	-1
	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	0	0	0	0	0	0	0	0	0	0	1	-1
	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	0	0	0	0	0	0	0	0	0	-1	1	0
	AVP>OT in HUM>CPU in Male>Female = AVP>OT in CPU>HUM in Female>Male = OT>AVP in CPU>HUM in Male>Female = OT>AVP in HUM>CPU in Female>Male	0	0	0	0	0	0	0	0	0	0	-1	1
	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	0	0	0	0	0	0	0	0	0	-1	0	1

Figure 3.3: Design matrix of the fMRI task with each column corresponding to one EV and each row corresponding to a subject-level contrast image. The white colour represents a $+1$ weight, while the grey represents 0, and the black -1. The table shows the T-contrasts used (excluding the ones from *post hoc* analysis), with a weight for each column of the design matrix. All possible combinations of each interaction are detailed, since they are produced by the same mathematical contrast. EV: explanatory variable; OT: oxytocin; AVP: vasopressin; PB: placebo; HUM: human partner; CPU: computer partner.



(a) **(b)**

(c) **(d)**

Figure 3.4: The four ROIs masks created using the probabilistic "Harvard Oxford cortical and subcortical structural atlases": (a) Striatum ROI ($x = 30, y = -1, z = -4$), (b) Amygdala ROI ($x = -24, y = -1, z = -19$), (c) Right Caudate ROI ($x = 22, y = 17, z = 0$) and (d) Left Caudate ROI ($x = -26, y = 8, z = 2$). In each figure, 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.

4

Results

Contents

This chapter presents the relevant results obtained from the computational RL models, behavioural and fMRI analyses.

4.1 Models Estimation and Evaluation

With the aim of understanding and quantifying latent processes that influence social learning procedures in each trial, ten computational RL models (summarized in table 3.2) were created and grouped into two major families. Thus, the results of each family were detailed in subsections 4.1.1 and 4.1.2. Subsection 4.1.3 provides the results that allowed the selection of one model.

4.1.1 Simple family

As previously stated, the Simple family is composed of five models (the Simple, the Q0, the 2LR, the 2LR Partner and the 4LR model) and, in order to compare those models with the real data, the probability of cooperating in each trial for all participants can be calculated, being displayed in Figure 4.1.

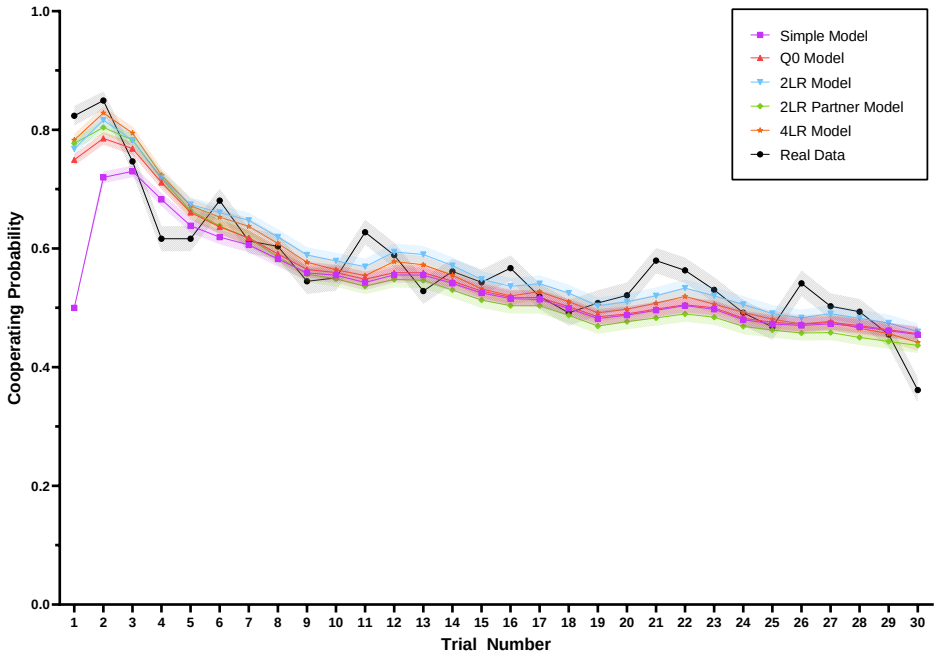


Figure 4.1: Trial-by-trial averaged cooperating probability from all participants, obtained from the real data and from the fit of the Simple model family (including the Simple, the Q0, the 2LR, the 2LR Partner and the 4LR models). Solid lines indicate the mean, while the shaded error bars represent the standard error.

4.1.2 TT family

Similarly, the TT family is also composed of five models (the Simple + TT, the Q0 + TT, the 2LR + TT, the 2LR Partner + TT and the 4LR + TT model), and the comparison of those models with the real data is displayed in Figure 4.2.

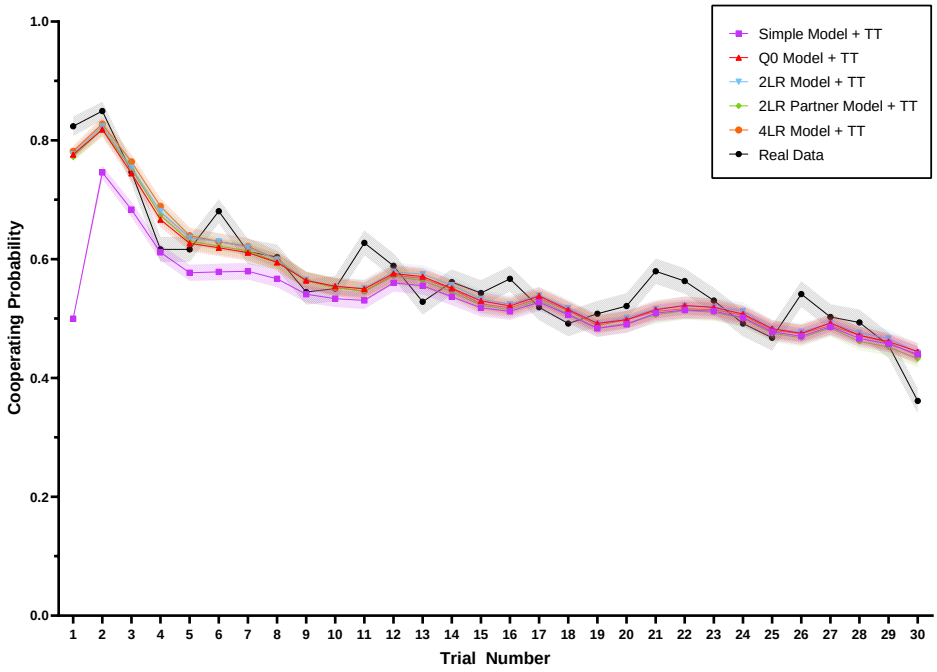


Figure 4.2: Trial-by-trial averaged cooperating probability from all participants, obtained from the real data and from the fit of the TT model family (including the simple + TT, the Q0 + TT, the 2LR + TT, the 2LR Partner + TT and 4LR + TT models). Solid lines indicate the mean, while the shaded error bars represent the standard error.

4.1.3 Model Selection and Validation

After fitting the model to each game data, one model must be selected to be used in the following analysis. Therefore, 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. Figure 4.3 represents the estimated frequencies for the two family models. Note that the higher the frequency, the better the model family's performance. Therefore, the Simple family was selected.

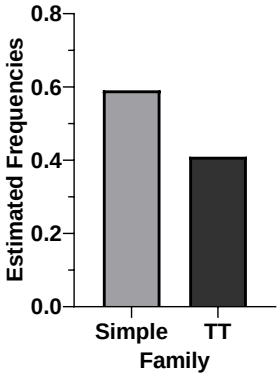


Figure 4.3: Model comparison using the estimated frequencies of the model families based on free-energy as approximation to model evidence.

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. Figure 4.4 represents the estimated frequencies of the five models that belong to the Simple family.

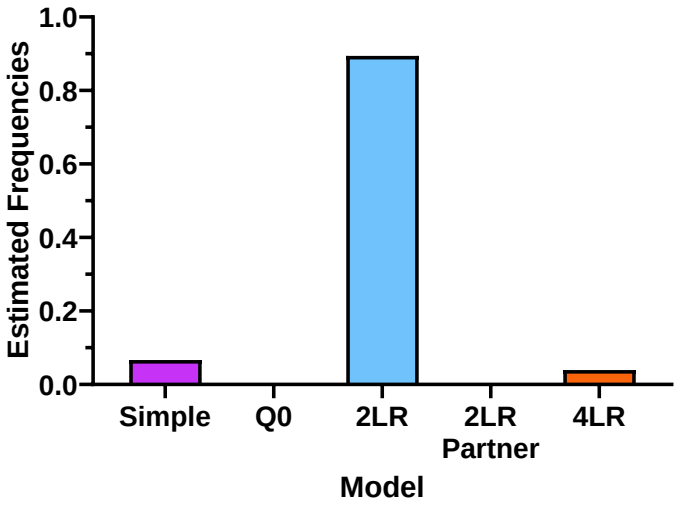


Figure 4.4: Model comparison using the estimated frequencies of the five models that belong to the Simple family based on free-energy as approximation to model evidence.

Furthermore, the exceedance probabilities were also determined to increase the veracity of the previous method. Again, the higher the probability of one model, the more likely a model is to fit the data better. Figure 4.5 represents the exceedance probabilities of the five models that belong to the Simple family.

Therefore, according to the two previous methods, the 2LR model was selected. Figure 4.6 compares the 2LR with the real data, using the trial-by-trial averaged cooperating probability from all participants.

To validate the chosen model, the PPC procedure was performed. Hence, using the 2LR model's

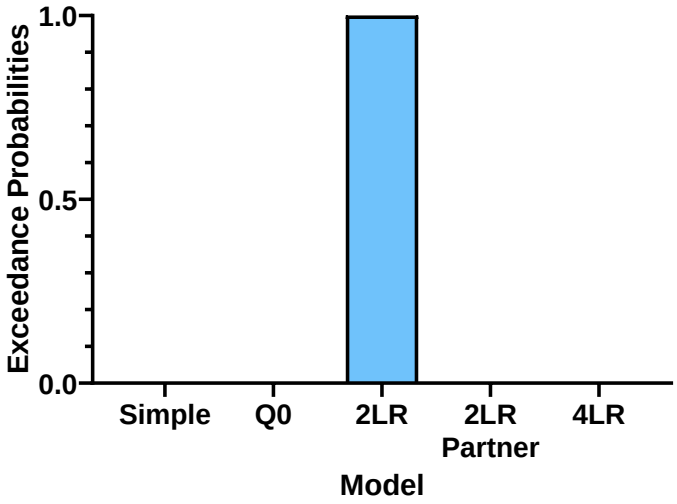


Figure 4.5: Model comparison using the exceedance probabilities of the five models that belong to the Simple family based on free-energy as approximation to model evidence.

parameters, new artificial participant choices were simulated, and a similar comparison using the trial-by-trial averaged cooperating probability was performed, shown in Figure 4.7. In order to assess the individual variation of the 2LR model, the cooperation probability of artificial data was averaged across trials (Figure 4.8). Thus, the closer the mean values are to the identity line, the better the model.

Furthermore, since RL models might not be able to 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 artificially simulated data). Figure 4.9 shows the correlation matrix.

4.2 Behavioural Analysis

To analyze if the drug (OT, AVP, PB), partner type (human, computer) and participant's sex (male, female) had any effect on the model parameters (α_C, α_D, β, $Q0$) acquired for each participant, a mixed design repeated measures ANOVA was performed.

4.2.1 α_C Analysis

No significant main effects or two-way interactions were found for mean α_C ($p > 0.179$), as shown in Table 4.1. However, a significant three-way interaction (drug $\times$ sex $\times$ partner) was found for the mean α_C $[F(2, 247) = 4.25, p = 0.015, \eta^2 = 0.014]$ (Table 4.1). In order to analyze this interaction, *post hoc* analyses were performed, where all combinations of the two-way interactions for each level of the third factor were analyzed, and, afterwards, all simple effects for each level of the second factor were

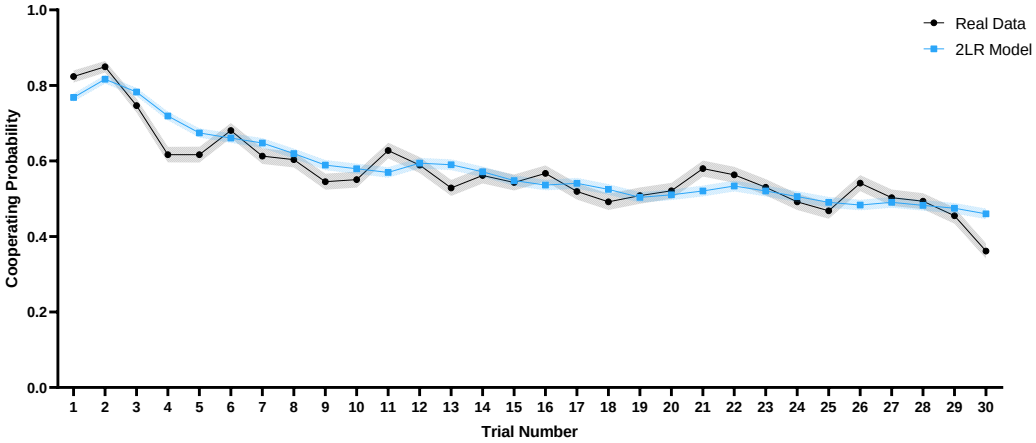


Figure 4.6: Trial-by-trial averaged cooperating probability from all participants from the real data and from the 2LR model. Solid lines indicate the mean, while the shaded error bars represent the standard error.

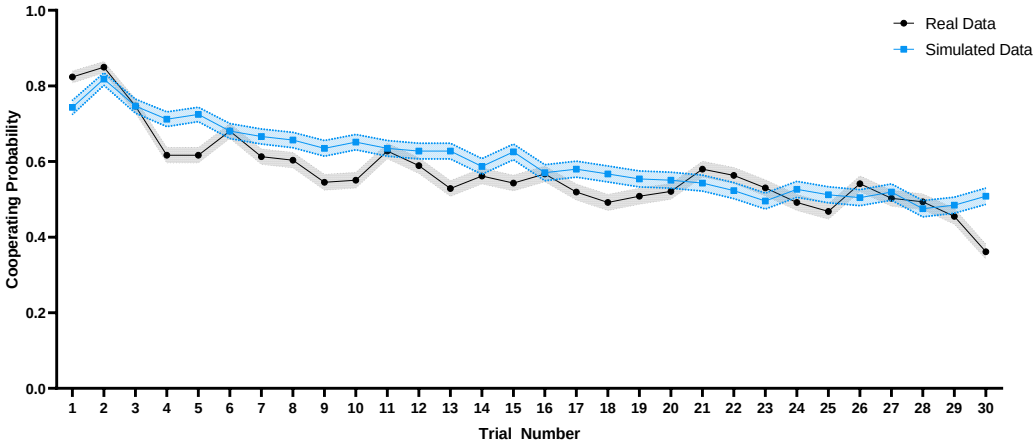


Figure 4.7: Trial-by-trial averaged cooperating probability from all participants from real data and from artificial data acquired using the 2LR model. Solid lines indicate the mean, while the shaded error bars represent the standard error.

explored. Hence, to provide a more comprehensive description of the three-way interaction, two different interpreting approaches will be described in the following two subsections (4.2.1.A and 4.2.1.B).

4.2.1.A Three-way Interaction Analysis - Sex Factor Being Fixed

A significant two-way interaction drug by partner was found for male players ($p = 0.040$) (Table 4.2). For female players, this interaction was not significant ($p = 0.247$) (Table 4.2).

This significant two-way interaction was characterized by male players under PB having a higher α_C when they played with human rather than with computer partners ($p = 0.040$) (Table 4.2, Figure 4.10), with no effect being found in other drug groups ($p > 0.132$) (Table 4.2, Figure 4.10).

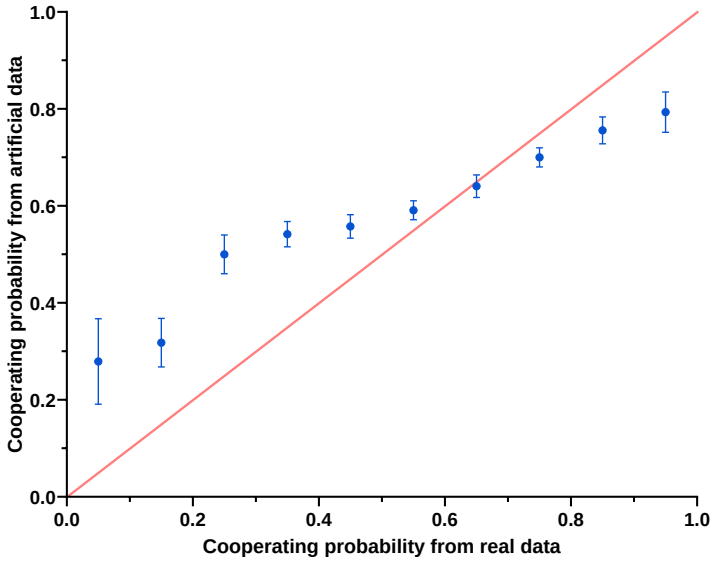


Figure 4.8: Individual's artificial cooperating probability acquired using the 2LR model (in blue, mean ± standard deviation) compared with real data, in relation to the identity line (in red).

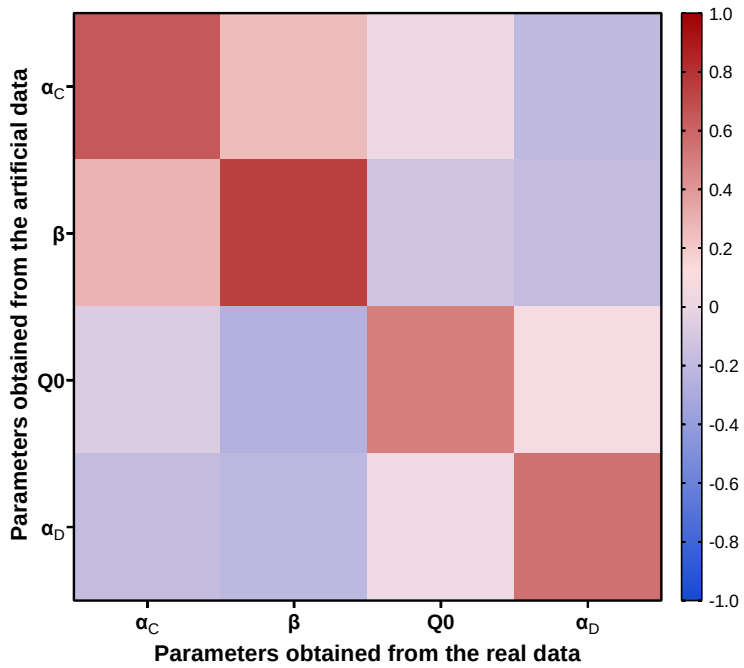


Figure 4.9: Correlation matrix of the 4 parameters of the 2LR model.

Moreover, regarding the same two-way interaction, in males, playing with a computer or a human partner did not affect the α_C difference between drug groups (*i.e.*, the difference of α_C in males playing with a computer partner between different drug groups was non-significant; the same was found for males playing with a human partner; $p > 0.131$) (Table 4.2, Figure 4.10).

4.2.1.B Three-way Interaction Analysis - Drug Factor Being Fixed

A significant two-way interaction sex by partner was found in the OT group ($p = 0.007$) (Table 4.2), while in the other drug groups this interaction was not significant ($p > 0.212$) (Table 4.2).

This significant two-way interaction was characterized by female players under OT having a higher α_C, when they played with a human partner rather than with a computer partner ($p = 0.020$) (Table 4.2, Figure 4.10), with no effect being seen for male players ($p = 0.132$) (Table 4.2, Figure 4.10).

Additionally, the latter significant two-way interaction was also defined by a higher α_C for female over male players, in the OT group, when they played with a human partner ($p = 0.037$) (Table 4.2, Figure 4.10). No effects were not found when they played with computer partners ($p = 0.127$) (Table 4.2, Figure 4.10).

4.2.2 α_D Analysis

No significant main effects or interactions were found for mean α_D ($p > 0.058$), as shown in Table 4.1.

4.2.3 β Analysis

No significant main effects or interactions were found for mean β ($p > 0.180$), as shown in Table 4.1.

4.2.4 $Q0$ Analysis

A significant main effect of partner was found for mean $Q0$ $[F(1, 247) = 3.95, p = 0.048, \eta^2 = 0.007]$, with participants playing with human partners having a higher $Q0$ $(M = 0.84, SD = 0.58)$ than with computer partners $(M = 0.73, SD = 0.74)$, as shown in Tables 4.1 and 4.2. A significant drug $\times$ sex $\times$ partner interaction was also found for mean $Q0$ $[F(2, 247) = 5.63, p = 0.004, \eta^2 = 0.02]$ (Table 4.1) and no other main effects or two-way interactions were acquired ($p > 0.06$) (Tables 4.1 and 4.2).

As before, *post hoc* analyses were performed, in order to analyze the significant three-way interaction. Herein, to provide a more comprehensive description of the three-way interaction, allowing to report different results depending on the fixed factor, three different interpreting approaches will be described in the following three subsections (4.2.4.A, 4.2.4.B and 4.2.4.C).

4.2.4.A Three-way Interaction Analysis - Sex Factor Being Fixed

A significant two-way interaction drug by partner was found for male players ($p = 0.001$) (Table 4.2). For female players, this interaction was not significant ($p = 0.462$) (Table 4.2). Additionally, a simple effect was found, with female participants having higher $Q0$ when playing with human partners than computer partners ($p = 0.034$), while no simple effects were seen in males ($p = 0.571$) (Table 4.2).

The latter significant two-way interaction was characterized by a higher $Q0$ for male participants playing with computer partners, when under AVP than PB ($p = 0.039$) (Table 4.2, Figure 4.10). In the OT group versus the PB and in the OT group versus the AVP, no significant differences were found for male players playing with computer partners ($p > 0.647$) (Table 4.2, Figure 4.10). For male players playing with human partners, not only a significantly higher $Q0$ was found under PB than OT ($p = 0.002$) (Table 4.2, Figure 4.10), but also under AVP than OT ($p = 0.035$) (Table 4.2, Figure 4.10). Under the same conditions, there were no differences between AVP and PB ($p = 1$, n.s.) (Table 4.2, Figure 4.10).

Moreover, the last-mentioned two-way interaction was also defined by male participants under PB having a higher $Q0$ when playing with human partners than computer partners ($p = 0.001$) (Table 4.2, Figure 4.10), with no effects being found in the other drug groups ($p > 0.192$) (Table 4.2, Figure 4.10).

4.2.4.B Three-way Interaction Analysis - Drug Factor Being Fixed

A significant two-way interaction sex by partner was found in the OT group ($p = 0.017$) (Table 4.2), while in the other drug groups this interaction was not significant ($p > 0.051$) (Table 4.2). Additionally, a simple effect was found, with participants under PB having a higher $Q0$ when playing with human than computer partners ($p = 0.009$), which was not significant in other drug groups ($p = 0.555$) (Table 4.2).

The latter significant two-way interaction was characterized by female players under OT having a higher $Q0$ when playing with human than with computer partners ($p = 0.034$) (Table 4.2, Figure 4.10), with no such effect being seen for male players ($p = 0.192$) (Table 4.2, Figure 4.10).

Furthermore, the last-mentioned two-way interaction was also defined by a higher $Q0$ for female over male players, in the OT group, when they played with a human partner ($p = 0.025$) (Table 4.2, Figure 4.10), with no effects being found when playing with a computer partner ($p = 0.269$) (Table 4.2, Figure 4.10).

4.2.4.C Three-way Interaction Analysis - Partner Factor Being Fixed

A significant two-way interaction drug by sex was found when participants played with a human partner ($p = 0.037$). For computer partners, this interaction was not significant ($p > 0.076$) (Table 4.2). Additionally, a simple effect was found, with participants playing with human partners having a higher $Q0$ when under PB than OT. As for taking OT versus AVP and AVP versus PB, no significant differences were found for participants playing with human partners ($p > 0.135$) (Table 4.2). No additional simple effects were found.

Table 4.1: Statistical results from the parameter analysis.

Parameter	Effects	F-test (df)	p-value	η^2
	Drug	0.004 (247)	0.996	$<$0.001
	Sex	0.04 (247)	0.841	$<$0.001
	Partner	1.82 (247)	0.179	$<$0.001
α_C	Drug x Sex	0.094 (247)	0.910	0.003
	Drug x Partner	0.49 (247)	0.617	0.002
	Sex x Partner	1.59 (247)	0.209	0.03
	Drug x Sex x Partner	**4.25 (247)**	**0.015** $\star$	**0.014**
	Drug	0.15 (247)	0.863	$<$0.001
	Sex	0.24 (247)	0.627	$<$0.001
	Partner	2.64 (247)	0.105	0.005
α_D	Drug x Sex	1.39 (247)	0.250	0.006
	Drug x Partner	0.59 (247)	0.553	0.002
	Sex x Partner	3.62 (247)	0.058	0.007
	Drug x Sex x Partner	2.14 (247)	0.119	0.008
	Drug	1.32 (247)	0.269	0.006
	Sex	0.32 (247)	0.574	$<$0.001
	Partner	0.01 (247)	0.917	$<$0.001
β	Drug x Sex	0.44 (247)	0.642	0.002
	Drug x Partner	0.64 (247)	0.527	0.002
	Sex x Partner	1.81 (247)	0.180	0.003
	Drug x Sex x Partner	0.71 (247)	0.491	0.003
	Drug	2.85 (247)	0.060	0.013
	Sex	0.17 (247)	0.685	$<$0.001
	Partner	**3.95 (247)**	**0.048** $\star$	**0.007**
Q0	Drug x Sex	0.73 (247)	0.482	0.003
	Drug x Partner	1.26 (247)	0.286	0.004
	Sex x Partner	1.52 (247)	0.219	0.003
	Drug x Sex x Partner	**5.63 (247)**	**0.004** $\star$	**0.02**

Table 4.2: Mean parameters per experimental condition. OT: oxytocin; AVP: vasopressin; PB: placebo.

Drug	Opponent	Sex	α_C (Mean $\pm$ SD)	α_D (Mean $\pm$ SD)	β (Mean $\pm$ SD)	$Q0$ (Mean $\pm$ SD)
OT	Computer	Male	0,702 $\pm$ 0,239	0,244 $\pm$ 0,253	2,293 $\pm$ 1,250	0,750 $\pm$ 0,636
	Computer	Female	0,617 $\pm$ 0,258	0,325 $\pm$ 0,274	2,352 $\pm$ 1,872	0,571 $\pm$ 0,808
	Human	Male	0,619 $\pm$ 0,288	0,408 $\pm$ 0,281	2,453 $\pm$ 1,577	0,534 $\pm$ 0,772
	Human	Female	0,735 $\pm$ 0,202	0,261 $\pm$ 0,260	2,494 $\pm$ 1,444	0,877 $\pm$ 0,573
AVP	Computer	Male	0,670 $\pm$ 0,242	0,266 $\pm$ 0,247	2,097 $\pm$ 0,987	0,935 $\pm$ 0,515
	Computer	Female	0,661 $\pm$ 0,252	0,257 $\pm$ 0,284	2,177 $\pm$ 1,249	0,734 $\pm$ 0,829
	Human	Male	0,664 $\pm$ 0,220	0,321 $\pm$ 0,287	1,934 $\pm$ 1,058	0,869 $\pm$ 0,575
	Human	Female	0,687 $\pm$ 0,243	0,334 $\pm$ 0,304	2,331 $\pm$ 1,825	0,906 $\pm$ 0,559
PB	Computer	Male	0,627 $\pm$ 0,284	0,255 $\pm$ 0,278	2,603 $\pm$ 1,626	0,571 $\pm$ 0,812
	Computer	Female	0,659 $\pm$ 0,250	0,358 $\pm$ 0,280	2,194 $\pm$ 1,566	0,824 $\pm$ 0,751
	Human	Male	0,722 $\pm$ 0,210	0,295 $\pm$ 0,266	2,073 $\pm$ 1,228	0,979 $\pm$ 0,388
	Human	Female	0,672 $\pm$ 0,247	0,325 $\pm$ 0,304	2,354 $\pm$ 1,146	0,878 $\pm$ 0,527

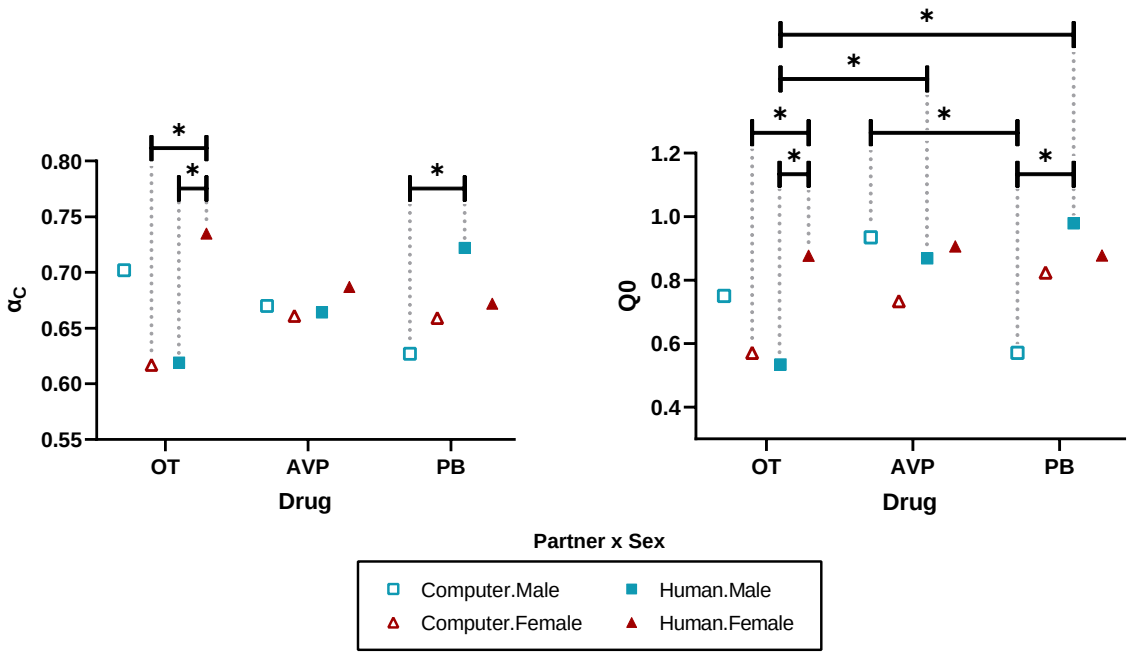


Figure 4.10: Mean values of the α_C (left plot) and $Q0$ (right plot) parameters for the different combination of three factors (drug, partner and sex). OT: oxytocin; AVP: vasopressin; PB: placebo.
$\star$ Statistically significant at $p < 0.05$

4.3 fMRI Data - Whole-Brain Analysis

The model-based fMRI analyses of the present study were focused on perceiving how the RPEs were represented in the human brain, and how the drug (OT, AVP, PB), partner type (human, computer), and participant's sex (male, female) influenced that representation.

First, a whole-brain analysis was performed to identify the different regions that were recruited to generate and integrate the RPEs and compare them across the three different factors (*i.e.*, the drug, the partner and the sex). A significant simple effect was found, with the PB group expressing a higher correlation of the RPE signal with brain activation in the Superior Temporal Gyrus (STG) when playing with a human partner, compared to a computer partner. Results are shown in Table 4.3.

Table 4.3: 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

4.4 fMRI Data - ROI Analyses

To analyze if the drug (OT, AVP, PB), partner type (human, computer) and participant's sex (male, female) had any effect on the neural modulation of RPEs in the amygdala, striatum, right and left caudate (regions that also belong to the striatum), ROI analyses were performed. Note that, the present study's design matrix did not allow to compare the three drug groups simultaneously during a three-way interaction, so, all three-way interactions presented subsequently focused on the differences between two drug groups separately.

4.4.1 Striatum ROI Analysis

To test the hypothesis that different effects were expected in the striatum, a ROI analysis was performed in this region, but no significant effects were obtained ($p > 0.05$).

4.4.2 Amygdala ROI Analysis

An exploratory analysis was also performed using an amygdala ROI, and two significant three-way interactions were found. Results are shown in Table 4.4 and Figure 4.11.

Similar to what was done in section 4.2, to explore the three-way interactions, all combinations of the two-way interactions for each level of the third factor were analyzed, and, afterwards, all simple effects

for each level of the second factor were explored.

The following three subsections (4.4.2.A, 4.4.2.B and 4.4.2.C) will provide a detailed description of each three-way interaction. Note that the first three-way interaction will be described using two different interpreting approaches, since different results were found depending on the fixed factor.

4.4.2.A Three-way Interaction (Drug (OT vs PB) $\times$ Sex $\times$ Partner) - Sex Factor Being Fixed

Regarding the first three-way interaction (drug (OT vs PB) $\times$ sex $\times$ partner), a significant two-way interaction drug by partner was found for female players ($p < 0.020$), while this interaction was not significant in males.

This significant interaction was not only characterized by female participants having a higher RPE-amygdala activation correlation under PB than OT when playing with a human partner ($p = 0.040$) (Figure 4.12(a)), but also under OT than PB when they played with a computer partner ($p = 0.001$) (Figure 4.12(a)).

Additionally, the latter significant two-way interaction can also be defined by female participants under PB having a higher RPE-amygdala activation correlation when they played with human than computer partners ($p = 0.043$) (Figure 4.12(a)), and by female players under OT when they played with the computer than human partner ($p < 0.045$) (Figure 4.12(a)).

4.4.2.B Three-way Interaction (Drug (OT vs PB) $\times$ Sex $\times$ Partner) - Partner Factor Being Fixed

Regarding the same three-way interaction (drug (OT vs PB) $\times$ sex $\times$ partner), a significant two-way interaction drug by sex was found when participants played with human partners ($p = 0.027$), while not being significant when playing with computer partners (Figure 4.12(a)).

This significant two-way interaction was characterized by male players under OT having a higher RPE-amygdala activation correlation when playing with human partners, in comparison to female participants ($p = 0.036$) (Figure 4.12(a)), with no effects being found in the PB group (Figure 4.12(a)).

No additional simple effects were found.

4.4.2.C Three-way Interaction (Drug (AVP vs PB) $\times$ Sex $\times$ Partner) - Sex Factor Being Fixed

Regarding the second three-way interaction (drug (AVP vs PB) $\times$ sex $\times$ partner), a significant two-way interaction drug by partner was found for males ($p < 0.028$), with no interactions found in females.

This significant two-way interaction was characterized by male players under AVP having a higher RPE-amygdala activation correlation when playing with human than computer partners ($p = 0.007$) (Figure 4.12(b)), with no effects being found in the PB group (Figure 4.12(b)).

No additional simple effects were found.

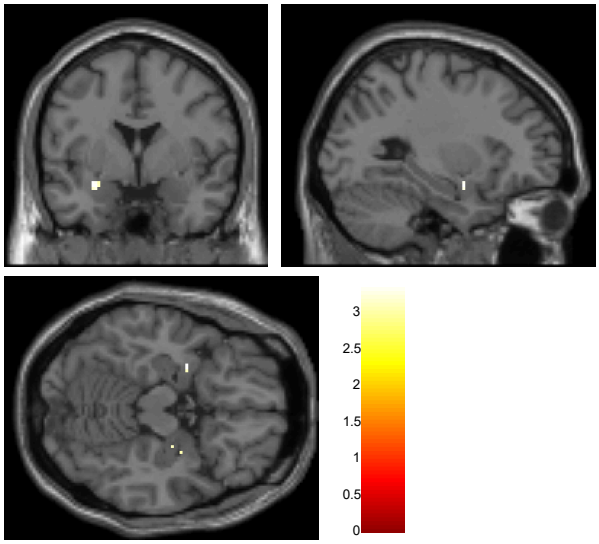


Figure 4.11: Neural representation of RPE-amygdala activation correlation ($x = -28, y = 0, z = -16$). The most representative coronal (on the upper left), sagittal (on the upper right), and transverse (on the lower left) slices are shown, all obtained with the MRIcron software. Display thresholded at $p < 0.05$, family-wise error (FWE) corrected.

Table 4.4: 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

4.4.3 Right Caudate ROI Analysis

The right caudate ROI analysis found a significant three-way interaction (Table 4.5). However, no significant two-way interactions or simple effects were obtained while analyzing this significant interaction

(drug (OT vs AVP) $\times$ sex $\times$ partner) (Figure 4.12(c)).

4.4.4 Left Caudate ROI Analysis

Regarding the left caudate ROI analysis, two significant three-way interactions were found (Table 4.6).

The following two subsections (4.4.4.A and 4.4.4.B) will provide a detailed description of each three-way interaction.

4.4.4.A Three-way Interaction (Drug (OT vs PB) $\times$ Sex $\times$ Partner) - Sex Factor Being Fixed

Regarding the first three-way interaction (drug (OT vs PB) $\times$ sex $\times$ partner), a significant two-way interaction drug by partner was found for female players ($p = 0.001$), while this interaction was not significant in males.

This significant interaction was not only characterized by female participants having a higher RPE-left caudate activation correlation under PB than OT when they played with a human partner ($p = 0.009$) (Figure 4.12(d)), but also under OT than PB when they played with a computer partner ($p = 0.016$) (Figure 4.12(d)).

Additionally, the latter significant two-way interaction can also be defined by female participants under PB having a higher RPE-left caudate activation correlation when playing with human than computer partners ($p = 0.008$) (Figure 4.12(d)), and by female participants under OT having a higher RPE-left caudate activation correlation when playing with computer than human partners ($p < 0.013$) (Figure 4.12(d)).

4.4.4.B Three-way Interaction (Drug (OT vs AVP) $\times$ Sex $\times$ Partner) - Drug Factor Being Fixed

Regarding the second three-way interaction (drug (OT vs AVP) $\times$ sex $\times$ partner), a significant two-way interaction partner by sex was found in the OT group ($p = 0.014$), while in the AVP group this interaction was not significant.

This significant two-way interaction was defined by female participants under OT having a higher RPE-left caudate activation correlation when playing with computer than human partners ($p < 0.013$) (Figure 4.12(d)), similar to what was found in the previous section, since the same brain area was being analyzed.

No additional simple effects were found.

Table 4.5: 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

Table 4.6: 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

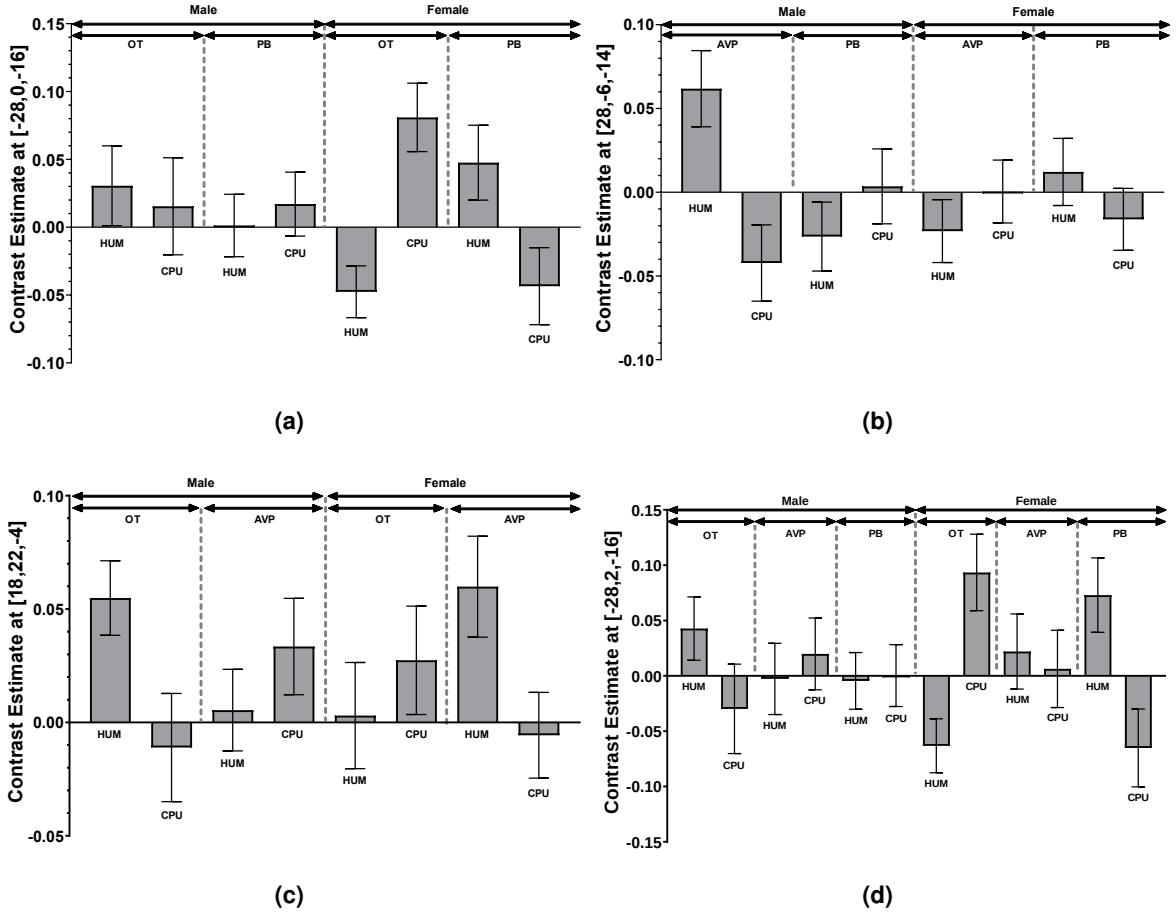


Figure 4.12: Neural correlates at the corresponding peak voxel for the three-way interactions drug $\times$ sex $\times$ partner: (a) contrast estimate for the amygdala ROI from the three-way interaction ($x = -28, y = 0, z = -16$); (b) contrast estimate for the amygdala ROI ($x = 28, y = -6, z = -14$); (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.

5

Discussion

Contents

This chapter discusses the results provided in the previous chapter.

As stated above, the present thesis aimed to study 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. To accomplish it, cumulative data from several fMRI studies was used [18–21]. In fact, each study increased the data set of the previous one, implementing the same sequential-choice PD task. Thus, the present study used the data of 292 participants (148 males and 144 females) for the behavioral analysis, and the data of 253 participants (121 men and 132 women) for the fMRI analysis, that were randomly administered with either intranasal OT, AVP or PB. Afterwards, computational RL modelling of behavioural data was performed, and trial-by-trial RPE signals were calculated and then correlated with the brain activation in an fMRI analysis.

Despite not being the first study to apply computational RL models to investigate intranasal OT effects on RL, this is the first study, at the time of this writing, to also investigate AVP's effect, while comparing those effects between partner types and participants' sex, with a considerable sample size. It is also important to note that, in comparison to other tasks applied in the literature (*e.g.*, the go/no-go tasks or the trust game), the PD task is mainly focused on eliciting socially-relevant behaviour, while allowing them to explicitly cooperate or defect the other's choices, which is essential to mimic real-life social interactions. Again, at the time of this writing, no other study performed computational RL modelling of PD's behavioural data, neither analyzed the intranasal OT and/or AVP effects on the previous modulation in both behavioural and fMRI data.

5.1 RL Model Analysis

RL models are key tools to perceive trial-by-trial variations of cognitive mechanisms that guide human behaviour. Thus, ten different models were created and used to fit the behavioural data.

To analyze the fitting performance of these models, the cooperation probability of each model was calculated. As one might see, the simple models fit (in both Simple and TT family, Figures 4.1 and 4.2, respectively) was considerably different from the real data in the first five trials, which exhibited the importance of the $Q0$ parameter in the model fitting procedure, similar to previous studies [73]. However, between each family, the cooperating probability of the other models was very similar. Herein, in order to select the best-fitting model, model comparisons were performed using the exceedance probabilities and the estimated frequencies based on the free-energy values, leading to the selection of the 2LR model from the Simple family.

In the context of the present task, the best-fitting model induced the following inferences. First, participants had an initial tendency to cooperate or defect at the beginning of the task ($Q0$). Additionally, participants had asymmetries in learning from trials when they cooperated and from trials when they

defected (α_C and α_D), meaning that the participant's choice had a stronger influence in their learning process, rather than their partner's choice (2LR Partner model) or trial outcome (4LR model). Lastly, the fact that the TT family did not have a better fit goes against previous evidence, since *Neto et al.* analyzed the same data used in this study and found a general preference for the tit-for-tat strategy. However, this might be related to the fact that an additional parameter (*i.e.*, the TT) could promote parameter overfitting, a penalizing factor taken into account during model comparison.

A PPC analysis was performed in order to evaluate the differences between model predictions and observed data. Thus, by comparing the trial-by-trial cooperating probability from real data and from artificial data acquired using the 2LR model, a similar behavioural pattern was found with some reduced differences between trials 10 and 20 (Figure 4.7). Furthermore, another analysis was performed at the participant level, in order to compare the individual's artificial cooperating probability with the real data. As one might see (Figure 4.8), the 2LR model produced artificial data with reduced cooperating probabilities, when participants cooperated less, while the artificial cooperating probability increased when participants cooperated more, leading to a behavioural pattern similar to the identity line.

Finally, a parameter recovery analysis was performed to analyze the accuracy and selectivity ability to conduct parameter identification. Similar to previous evidence [115], the diagonal of the correlation matrix (Figure 4.9) revealed significantly high Spearman coefficients (highest $p - value < 6.34 \times 10^{-34}$), while the rest of the matrix exhibited low coefficients. Herein, it is possible to conclude that the 2LR model was able to recover its parameters accurately.

It is also worth noting that previous studies using RL models used tasks with a considerable number of trials (*e.g.*, *Mcdougle et al.* [93] performed model fitting with 300 trials, *Zhang et al.* [14] with 100 trials and *Murray et al.* [119] with 120 trials), which is an important factor to take into account in order to acquire stable parameter estimations [15]. Nevertheless, with a reduced number of trials (30), the present study was able to achieve a good model fitting. However, an increase in the number of trials would allow an even better fit and parameter estimation.

5.2 Behavioural Analysis

The behavioural analysis performed in this thesis aimed to study the impact of intranasal OT, AVP, partner type and participant sex on the RL model parameters. In the following subsections, the statistically significant results on each parameter will be discussed.

5.2.1 α_C Analysis

As previously described, the α_C parameter represents the learning rate (or learning speed) from trials when the participant cooperated and it will adjust the impact of the RPE on the next prediction. When

the participant cooperated, a higher α_C value will, indeed, represent a higher influence of the latter PD interaction on the next one.

Contrary to the hypotheses of this study, no main effects of OT in comparison to PB or a two-way interaction between drug and partner were found for α_C. Nevertheless, a three-way interaction, two-way interactions and main effects were found.

Thus, intranasal OT increased the α_C among women, when they played with human partners than with computer partners, but not under PB. As previously stated, OT is associated with defensive aggression focused on protecting and negating threats induced by out-groups [97], causing impulsive actions, since higher importance of the previous interaction is weighted on the next decision. Thus, considering the present results and previous evidence, the social context (*i.e.*, the partner being human) might augment that effect, leading to increased defensive aggression and impulsive behaviours in women. The increase of aggression after intranasal OT intake during social tasks has indeed been reported for a Social Orientation Paradigm, which measured real-time aggressive behaviour with participants playing with a same-sex player [59]. Furthermore, an increase of aggression after intranasal OT intake during a social task was also found in women by a previous study [120].

Different from females, a higher α_C was found for male participants under PB, when they played with human partners than with computer partners, but not under OT. Extraneous OT might produce different responses according to the participant sex, since there are sex variations in the baseline levels of OT, with women having higher OT in the cerebrospinal fluid than men [101]. Herein, *Rilling et al.* [20, 69] proposed that neural activity might follow an inverted-U shaped response as a function of the OT levels dose, with an increase in OT levels in males shifting the neural activity closer to the maximum, while the same effect would decrease the neural activity of females, moving it to the right of the maximum. Following this idea, lower levels of OT in males would promote the same effect as extraneously increased levels of OT in females.

Under OT, a higher α_C was also found in female than male players, when playing with a human partner, but not in PB. A study [121] using prairie voles revealed that untreated males have higher intrasexual aggression than females, while OT administration would promote the opposite effect. Hence, considering the present results and previous evidence, during social contexts, OT might promote intrasexual defensive aggression and impulsive behaviours in females, while reducing them in males.

The fact that there were significant results for α_C reveals that the cooperation decision had high importance on the next trial's decision. Since cooperation implies trusting the other person, every time the 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 next 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 contexts and men.

5.2.2 $Q0$ Analysis

As previously referred, the $Q0$ parameter represents the participant's tendency, at the beginning of the task, to cooperate or defect. The higher the $Q0$ value, the higher the tendency to cooperate.

Multiple significant simple effects and a main effect 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 [122] 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 [20, 123] report that OT may promote anthropomorphism of computer partners in women, the present results showed a significant 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, OT might increase the number of cooperating choices when playing with a computer partner throughout the game, leading to anthropomorphism.

Significant drug effects were also found, with participants having a higher $Q0$ under the effect of PB than 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 [61], and may also reveal that a pro-self bias and threat identification might occur at the beginning of a social 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 imply that AVP might induce a pro-social bias in males, at the beginning of an interaction in a non-social context. Previous evidence [62] showed that AVP increased the willingness to cooperate in males using the Stag Hunt task, by increasing the desire to take risks, compared to PB. Herein, by enhancing the willingness to take risks, the present results suggest that AVP might promote a pro-social bias in males, inducing cooperation, at the beginning of an interaction in a non-social context.

5.3 fMRI Analysis

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

5.3.1 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 correlation between RPEs and the striatal activity 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 partner than with a computer partner. Although the STG is traditionally associated with language and auditory processing [124, 125], studies [126, 127] have reported that it has an important role in processing social stimuli. Specifically, a previous study [128] reported that STG has an essential role in behavioural monitoring and reappraisal and another [129] 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.

5.3.2 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 DA neuronal projection that codes RPEs [8]. 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 [69] and [70]) 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.* [10], 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

[130]), 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.* [23], a study using the same behavioural data as the present thesis, 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 [14], also using RL models and trial-by-trial RPEs.

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 emotional learning [131] and processing of emotional information [132], while recognizing the stimulus for the needs and goals of the organism [132]. Furthermore, previous studies also reported that the amygdala applies social attention, information and emotions in decision-making [133, 134], and also has an important role in the RL process [8, 114].

In fact, similar significant results to the ones found in the caudate were found in the amygdala. A previous study [9] using computational RL models and trial-by-trial RPEs reported identical 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 [8, 135], with both structures receiving multiple DA projections [136, 137]. Physiologically, studies have shown that the stimulation of the basolateral amygdala may increase DA release in the ventral striatum due to glutamatergic input signals [138] 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 [139]. 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.

Comparing both sexes, the present amygdala results suggest that OT enhances social learning more in males than females when they play with a human partner. As reported in subsection 5.2.1, and by *Rilling et al.* [20], the hypothesis of neural activity following an inverted-U shaped response as a function of the OT levels, with men's baseline OT levels being left to the maximum, while women levels being closer to the maximum, might be the reason behind the difference between sexes. This hypothesis also

states that enhancing the OT levels in males would shift their neural activity closer to the maximum, while the opposite would occur in females. Thus, OT might increase the RPE-amygdala correlation activation in males, while decreasing it in females, especially in social contexts, where the OT effect might be enhanced.

Additionally, an AVP simple effect was also 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 [140], this hypothesis is relatively unexplored. Herein, further research should be conducted to study the role of AVP in the RL process.

Table 5.1 summarizes the most important results obtained throughout this thesis, and their interpretations.

Table 5.1: Summary of the most important results obtained throughout this thesis, and their interpretations.

Results	Interpretation
Under OT, female players had a higher α_C when playing with human than with computer partners.	OT might increase defensive aggression and impulsive behaviours in females, after cooperating, an effect that was augmented during social contexts.
Under PB, male players had a higher α_C when playing with human than with computer partners.	Lower OT levels in males might promote the same effect as extraneously increased levels of OT in females.
Under OT, female players had a higher α_C when playing with a human partner, than male players.	During social contexts, OT might promote intrasexual defensive aggression and impulsive behaviours in females, while reducing them in males.
When playing with human partners, participants had a higher $Q0$ than when playing with computer partners.	When playing with human partners, an enhanced tendency to cooperate might occur.
Participants had a higher $Q0$ under the effect of PB than OT, when playing with human partners.	OT might induce a pro-self bias and threat identification at the beginning of a social interaction.
A higher $Q0$ was found for male players 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.	AVP might promote a pro-social bias in males, inducing cooperation, at the beginning of an interaction in a non-social context.
Under PB, participants had a higher RPE-STG activation correlation when playing with human than with computer partners.	STG might have an additional role in the social RL process.
Under OT, female players had a higher RPE-amygdala/left caudate activation correlation when playing with computer than with human partners.	OT might enhance social learning in females when playing with a computer partner, compared to a human partner, by compensating and reinforcing learning from non-social contexts or by increasing learning from social partners to a point where it also increases the learning from non-social partners.
Under PB, female players had a higher RPE-amygdala/left caudate activation correlation when playing with human than with computer partners.	Under PB, playing with a human partner might enhance social learning in females, compared to computer partners.
Under OT, male players had a higher RPE-amygdala activation correlation when playing with human partners, than female players.	OT might enhance social learning, more in males than females, when they play with a human partner.
Under AVP, male players had a higher RPE-amygdala activation correlation when playing with a human than with computer partners.	AVP, in social contexts, might increase social learning in males.

6

Conclusion

Contents

The present thesis suggests new specific roles for OT and AVP in the social RL process, 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, the present study results revealed that OT may promote a pro-self (non-social) bias and threat identification prior to 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 non-social contexts and men. On the other hand, AVP might promote a pro-social bias prior to 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 RPEs. 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 OT may enhance social learning more in males than females in social contexts, but also 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.

6.1 Limitations and Future Perspectives

Throughout this thesis, some limitations exist. First, the PD task used in this study was composed of thirty trials per game. Although it might have been enough for previous analyses using this data, for RL models, it might not be sufficient to provide a superb fit. In fact, previous studies used a considerably larger number of trials to perform similar analyses (*e.g.*, *Mcdougle et al.* [93] performed model fitting with 300 trials, *Zhang et al.* [14] with 100 trials and *Murray et al.* [119] with 120 trials). Thus, future studies using a similar paradigm, but with a higher number of trials, would be desirable. Secondly, the present data was acquired from young participants (ages between 18 and 22 years), which is not representative of the whole population. Due to this fact, the present conclusions may not be generalizable to other age groups. It is also worth noting that all interactions with human partners occurred with same-sex participants (*i.e.*, participants were told they were playing with a same-sex human), which also limits their generalizability for the sake of reducing heterogeneity and noise in the current study. Therefore, future research is required in order to further test the generalization of the current findings.

Finally, this study's scientific approach (model-based behavioural and neuroimaging) is very recent,

representing the first steps on the journey to understand the roles of OT and AVP in the social RL process. Thus, further studies are required to increase confidence in the field. Additionally, future research should also focus on exploring an amygdala role in the social RL process, which is supported by the present preliminary results with little precedent in the literature so far.

In conclusion, the value of understanding those psychological mechanisms is, indeed, incalculable, which might not only provide insights about the neurological pathways that occur in the human brain and influence social behaviour, but also in multiple social disorders, being essential to understand if and how those two neuropeptides can be used to enhance treatment procedures.

Bibliography

[1] L. Peled-Avron, A. Abu-Akel, and S. Shamay-Tsoory, “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,” *Neuroscience & Biobehavioral Reviews*, vol. 114, pp. 70–95, 2020.

[2] A. Meyer-Lindenberg, G. Domes, P. Kirsch, and M. Heinrichs, “Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine,” *Nature Reviews Neuroscience*, vol. 12, no. 9, pp. 524–538, 2011.

[3] K. MacDonald and T. MacDonald, “The peptide that binds: a systematic review of oxytocin and its prosocial effects in humans,” *Harvard Review of Psychiatry*, vol. 18, no. 1, pp. 1–21, 2010.

[4] M. Klein, D. Battagello, A. Cardoso, D. Hauser, J. Bittencourt, and R. Correa, “Dopamine: functions, signaling, and association with neurological diseases,” *Cellular and Molecular Neurobiology*, vol. 39, no. 1, pp. 31–59, 2019.

[5] G. Gimpl and F. Fahrenholz, “The oxytocin receptor system: structure, function, and regulation,” *Physiological Reviews*, vol. 81, no. 2, pp. 629–683, 2001.

[6] R. Stoop, “Neuromodulation by oxytocin and vasopressin,” *Neuron*, vol. 76, no. 1, pp. 142–159, 2012.

[7] S. Hyman, R. Malenka, and E. Nestler, “Neural mechanisms of addiction: the role of reward-related learning and memory,” *Annual Review of Neuroscience*, vol. 29, pp. 565–598, 2006.

[8] B. Averbeck and V. Costa, “Motivational neural circuits underlying reinforcement learning,” *Nature Neuroscience*, vol. 20, no. 4, pp. 505–512, 2017.

[9] J. Ide, S. Nedic, K. Wong, S. Strey, E. Lawson, B. Dickerson *et al.*, “Oxytocin attenuates trust as a subset of more general reinforcement learning, with altered reward circuit functional connectivity in males,” *Neuroimage*, vol. 174, pp. 35–43, 2018.

[10] J. Kruppa, A. Gossen, E. Weiß, G. Kohls, N. Großheinrich, H. Cholemkery *et al.*, "Neural modulation of social reinforcement learning by intranasal oxytocin in male adults with high-functioning autism spectrum disorder: a randomized trial," *Neuropsychopharmacology*, vol. 44, no. 4, pp. 749–756, 2019.

[11] M. Heinrichs, B. von Dawans, and G. Domes, "Oxytocin, vasopressin, and human social behavior," *Frontiers in Neuroendocrinology*, vol. 30, no. 4, pp. 548–557, 2009.

[12] E. Keverne and J. Curley, "Vasopressin, oxytocin and social behaviour," *Current Opinion in Neurobiology*, vol. 14, no. 6, pp. 777–783, 2004.

[13] M. Lim, Z. Wang, D. Olazábal, X. Ren, E. Terwilliger, and L. Young, "Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene," *Nature*, vol. 429, no. 6993, pp. 754–757, 2004.

[14] L. Zhang and J. Gläscher, "A brain network supporting social influences in human decision-making," *Science Advances*, vol. 6, no. 34, p. eabb4159, 2020.

[15] L. Zhang, L. Lengersdorff, N. Mikus, J. Gläscher, and C. Lamm, "Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices," *Social Cognitive and Affective Neuroscience*, vol. 15, no. 6, pp. 695–707, 2020.

[16] E. Abohamza, T. Weickert, M. Ali, and A. Moustafa, "Reward and punishment learning in schizophrenia and bipolar disorder," *Behavioural Brain Research*, vol. 381, p. 112298, 2020.

[17] E. Reilly, A. Whitton, D. Pizzagalli, A. Rutherford, M. Stein, M. Paulus *et al.*, "Diagnostic and dimensional evaluation of implicit reward learning in social anxiety disorder and major depression," *Depression and Anxiety*, vol. 37, no. 12, pp. 1221–1230, 2020.

[18] J. Rilling, A. DeMarco, P. Hackett, R. Thompson, B. Ditzen, R. Patel *et al.*, "Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men," *Psychoneuroendocrinology*, vol. 37, no. 4, pp. 447–461, 2012.

[19] C. Feng, A. DeMarco, E. Haroon, and J. Rilling, "Neuroticism modulates the effects of intranasal vasopressin treatment on the neural response to positive and negative social interactions," *Neuropsychologia*, vol. 73, pp. 108–115, 2015.

[20] J. Rilling, A. DeMarco, P. Hackett, X. Chen, P. Gautam, S. Stair *et al.*, "Sex differences in the neural and behavioral response to intranasal oxytocin and vasopressin during human social interaction," *Psychoneuroendocrinology*, vol. 39, pp. 237–248, 2014.

[21] X. Chen, P. Hackett, A. DeMarco, C. Feng, S. Stair, E. Haroon *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," *Brain Imaging and Behavior*, vol. 10, no. 2, pp. 581–593, 2016.

[22] B. Skinner, *The Behavior of Organisms*, 1st ed. The Century Psychology Series, 1938, pp. 3–8.

[23] M. Neto, M. Antunes, M. Lopes, D. Ferreira, J. Rilling, and D. Prata, "Oxytocin and vasopressin modulation of prisoner's dilemma strategies," *Journal of Psychopharmacology*, vol. 34, no. 8, pp. 891–900, 2020.

[24] I. Pavlov, *Conditioned Reflexes: An investigation of the physiological activity of the cerebral cortex*. Dover Publications, 1960, pp. 1–15.

[25] P. Glimcher, "Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis," *Proceedings of the National Academy of Sciences*, vol. 108, no. Supplement 3, pp. 15 647–15 654, 2011.

[26] W. Schultz, P. Apicella, and T. Ljungberg, "Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task," *Journal of Neuroscience*, vol. 13, no. 3, pp. 900–913, 1993.

[27] P. Montague, P. Dayan, and T. Sejnowski, "Foraging in an uncertain environment using predictive hebbian learning," in *Advances in neural information processing systems*, 1994, pp. 598–605.

[28] P. Montague, P. Dayan, C. Person, and T. Sejnowski, "Bee foraging in uncertain environments using predictive hebbian learning," *Nature*, vol. 377, no. 6551, pp. 725–728, 1995.

[29] P. Montague, P. Dayan, and T. Sejnowski, "A framework for mesencephalic dopamine systems based on predictive Hebbian learning," *Journal of Neuroscience*, vol. 16, no. 5, pp. 1936–1947, 1996.

[30] W. Schultz, P. Dayan, and P. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, no. 5306, pp. 1593–1599, 1997.

[31] G. Berns, S. McClure, G. Pagnoni, and P. Montague, "Predictability modulates human brain response to reward," *Journal of Neuroscience*, vol. 21, no. 8, pp. 2793–2798, 2001.

[32] Y. Niv, "Reinforcement learning in the brain," *Journal of Mathematical Psychology*, vol. 53, no. 3, pp. 139–154, 2009.

[33] O. Hornykiewicz, “Dopamine miracle: from brain homogenate to dopamine replacement,” *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 17, no. 3, pp. 501–508, 2002.

[34] S. Fahn, “The history of dopamine and levodopa in the treatment of Parkinson’s disease,” *Movement Disorders: Official Journal of the Movement Disorder Society*, vol. 23, no. S3, pp. S497–S508, 2008.

[35] S. Chinta and J. Andersen, “Dopaminergic neurons,” *The International Journal of Biochemistry & Cell Biology*, vol. 37, no. 5, pp. 942–946, 2005.

[36] J. Olds and P. Milner, “Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain,” *Journal of Comparative and Physiological Psychology*, vol. 47, no. 6, p. 419, 1954.

[37] J. Olds, “Self-stimulation of the brain: Its use to study local effects of hunger, sex, and drugs,” *Science*, vol. 127, no. 3294, pp. 315–324, 1958.

[38] R. Wise, “Addictive drugs and brain stimulation reward,” *Annual Review of Neuroscience*, vol. 19, no. 1, pp. 319–340, 1996.

[39] J. Cox and I. Witten, “Striatal circuits for reward learning and decision-making,” *Nature Reviews Neuroscience*, vol. 20, no. 8, pp. 482–494, 2019.

[40] M. Frank, L. Seeberger, and R. O’reilly, “By carrot or by stick: cognitive reinforcement learning in parkinsonism,” *Science*, vol. 306, no. 5703, pp. 1940–1943, 2004.

[41] M. Guitart-Masip, Q. Huys, L. Fuentemilla, P. Dayan, E. Duzel, and R. Dolan, “Go and no-go learning in reward and punishment: interactions between affect and effect,” *Neuroimage*, vol. 62, no. 1, pp. 154–166, 2012.

[42] M. Frank, A. Moustafa, H. Haughey, T. Curran, and K. Hutchison, “Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning,” *Proceedings of the National Academy of Sciences*, vol. 104, no. 41, pp. 16 311–16 316, 2007.

[43] H. Bayer, B. Lau, and P. Glimcher, “Statistics of midbrain dopamine neuron spike trains in the awake primate,” *Journal of Neurophysiology*, vol. 98, no. 3, pp. 1428–1439, 2007.

[44] M. Ottenhausen, I. Bodhinayake, M. Banu, P. Stieg, and T. Schwartz, “Vincent du Vigneaud: following the sulfur trail to the discovery of the hormones of the posterior pituitary gland at Cornell Medical College,” *Journal of Neurosurgery*, vol. 124, no. 5, pp. 1538–1542, 2016.

[45] D. Skuse and L. Gallagher, “Dopaminergic-neuropeptide interactions in the social brain,” *Trends in Cognitive Sciences*, vol. 13, no. 1, pp. 27–35, 2009.

[46] J. Swain, J. Lorberbaum, S. Kose, and L. Strathearn, “Brain basis of early parent–infant interactions: psychology, physiology, and in vivo functional neuroimaging studies,” *Journal of Child Psychology and Psychiatry*, vol. 48, no. 3-4, pp. 262–287, 2007.

[47] T. Insel, “The challenge of translation in social neuroscience: a review of oxytocin, vasopressin, and affiliative behavior,” *Neuron*, vol. 65, no. 6, pp. 768–779, 2010.

[48] M. Piva and S. Chang, “An integrated framework for the role of oxytocin in multistage social decision-making,” *American Journal of Primatology*, vol. 80, no. 10, p. e22735, 2018.

[49] M. Kosfeld, M. Heinrichs, P. Zak, U. Fischbacher, and E. Fehr, “Oxytocin increases trust in humans,” *Nature*, vol. 435, no. 7042, pp. 673–676, 2005.

[50] T. Baumgartner, M. Heinrichs, A. Vonlanthen, U. Fischbacher, and E. Fehr, “Oxytocin shapes the neural circuitry of trust and trust adaptation in humans,” *Neuron*, vol. 58, no. 4, pp. 639–650, 2008.

[51] P. Zak, A. Stanton, and S. Ahmadi, “Oxytocin increases generosity in humans,” *PloS One*, vol. 2, no. 11, p. e1128, 2007.

[52] J. Barraza, M. McCullough, S. Ahmadi, and P. Zak, “Oxytocin infusion increases charitable donations regardless of monetary resources,” *Hormones and Behavior*, vol. 60, no. 2, pp. 148–151, 2011.

[53] G. Domes, M. Heinrichs, A. Michel, C. Berger, and S. Herpertz, “Oxytocin improves “mind-reading” in humans,” *Biological Psychiatry*, vol. 61, no. 6, pp. 731–733, 2007.

[54] A. Guastella, P. Mitchell, and F. Mathews, “Oxytocin enhances the encoding of positive social memories in humans,” *Biological Psychiatry*, vol. 64, no. 3, pp. 256–258, 2008.

[55] G. Domes, M. Sibold, L. Schulze, A. Lischke, S. Herpertz, M. Heinrichs *et al.*, “Intranasal oxytocin increases covert attention to positive social cues,” *Psychological Medicine*, vol. 43, no. 8, pp. 1747–1753, 2013.

[56] M. Gamer, B. Zurowski, and C. Büchel, “Different amygdala subregions mediate valence-related and attentional effects of oxytocin in humans,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 20, pp. 9400–9405, 2010.

[57] C. Declerck, C. Boone, and T. Kiyonari, “The effect of oxytocin on cooperation in a prisoner’s dilemma depends on the social context and a person’s social value orientation,” *Social Cognitive and Affective Neuroscience*, vol. 9, no. 6, pp. 802–809, 2014.

[58] B. Ditzen, M. Schaer, B. Gabriel, G. Bodenmann, U. Ehlert, and M. Heinrichs, “Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict,” *Biological Psychiatry*, vol. 65, no. 9, pp. 728–731, 2009.

[59] R. Ne’eman, N. Perach-Barzilay, M. Fischer-Shofty, A. Atias, and S. Shamay-Tsoory, “Intranasal administration of oxytocin increases human aggressive behavior,” *Hormones and Behavior*, vol. 80, pp. 125–131, 2016.

[60] S. Shamay-Tsoory, M. Fischer, J. Dvash, H. Harari, N. Perach-Bloom, and Y. Levkovitz, “Intranasal administration of oxytocin increases envy and schadenfreude (gloating),” *Biological Psychiatry*, vol. 66, no. 9, pp. 864–870, 2009.

[61] C. De Dreu, L. Greer, G. Van Kleef, S. Shalvi, and M. Handgraaf, “Oxytocin promotes human ethnocentrism,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 4, pp. 1262–1266, 2011.

[62] C. Brunnlieb, G. Nave, C. Camerer, S. Schosser, B. Vogt, T. Münte *et al.*, “Vasopressin increases human risky cooperative behavior,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 8, pp. 2051–2056, 2016.

[63] A. Guastella, A. Kenyon, G. Alvares, D. Carson, and I. Hickie, “Intranasal arginine vasopressin enhances the encoding of happy and angry faces in humans,” *Biological Psychiatry*, vol. 67, no. 12, pp. 1220–1222, 2010.

[64] R. Thompson, K. George, J. Walton, S. Orr, and J. Benson, “Sex-specific influences of vasopressin on human social communication,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 20, pp. 7889–7894, 2006.

[65] J. Goodson, “Deconstructing sociality, social evolution and relevant nonapeptide functions,” *Psychoneuroendocrinology*, vol. 38, no. 4, pp. 465–478, 2013.

[66] J. Bartz, J. Zaki, N. Bolger, and K. Ochsner, “Social effects of oxytocin in humans: context and person matter,” *Trends in Cognitive Sciences*, vol. 15, no. 7, pp. 301–309, 2011.

[67] M. Olff, J. Frijling, L. Kubzansky, B. Bradley, M. Ellenbogen, C. Cardoso *et al.*, “The role of oxytocin in social bonding, stress regulation and mental health: an update on the moderating effects of context and interindividual differences,” *Psychoneuroendocrinology*, vol. 38, no. 9, pp. 1883–1894, 2013.

[68] I. Thielmann, G. Spadaro, and D. Balliet, “Personality and prosocial behavior: A theoretical framework and meta-analysis,” *Psychological Bulletin*, vol. 146, no. 1, p. 30, 2020.

[69] C. Feng, P. Hackett, A. DeMarco, X. Chen, S. Stair, E. Haroon *et al.*, "Oxytocin and vasopressin effects on the neural response to social cooperation are modulated by sex in humans," *Brain Imaging and Behavior*, vol. 9, no. 4, pp. 754–764, 2015.

[70] X. Chen, P. Gautam, E. Haroon, and J. Rilling, "Within vs. between-subject effects of intranasal oxytocin on the neural response to cooperative and non-cooperative social interactions," *Psychoneuroendocrinology*, vol. 78, pp. 22–30, 2017.

[71] N. Daw, S. Gershman, B. Seymour, P. Dayan, and R. Dolan, "Model-based influences on humans' choices and striatal prediction errors," *Neuron*, vol. 69, no. 6, pp. 1204–1215, 2011.

[72] T. Katthagen, J. Kaminski, A. Heinz, R. Buchert, and F. Schlagenhauf, "Striatal Dopamine and Reward Prediction Error Signaling in Unmedicated Schizophrenia Patients," *Schizophrenia Bulletin*, vol. 46, no. 6, pp. 1535–1546, 2020.

[73] B. Lindström, A. Golkar, S. Jangard, P. Tobler, and A. Olsson, "Social threat learning transfers to decision making in humans," *Proceedings of the National Academy of Sciences*, vol. 116, no. 10, pp. 4732–4737, 2019.

[74] R. Sutton and A. Barto, *Reinforcement learning: An introduction*, 2nd ed. MIT press, 2018, pp. 25–45.

[75] V. Vieland, "Bayesian linkage analysis, or: how i learned to stop worrying and love the posterior probability of linkage," *American journal of human genetics*, vol. 63, no. 4, p. 947, 1998.

[76] K. Stephan, W. Penny, J. Daunizeau, R. Moran, and K. Friston, "Bayesian model selection for group studies," *Neuroimage*, vol. 46, no. 4, pp. 1004–1017, 2009.

[77] T. Lodewyckx, W. Kim, M. Lee, F. Tuerlinckx, P. Kuppens, and E. Wagenmakers, "A tutorial on Bayes factor estimation with the product space method," *Journal of Mathematical Psychology*, vol. 55, no. 5, pp. 331–347, 2011.

[78] J. Daunizeau, V. Adam, and L. Rigoux, "VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data," *PLoS Computational Biology*, vol. 10, no. 1, p. e1003441, 2014.

[79] L. Fontanesi, S. Gluth, M. Spektor, and J. Rieskamp, "A reinforcement learning diffusion decision model for value-based decisions," *Psychonomic Bulletin & Review*, vol. 26, no. 4, pp. 1099–1121, 2019.

[80] J. Kruschke, "Posterior predictive checks can and should be Bayesian: Comment on Gelman and Shalizi,'Philosophy and the practice of Bayesian statistics'," *British Journal of Mathematical and Statistical Psychology*, vol. 66, no. 1, pp. 45–56, 2013.

[81] R. Buxton, *Introduction to functional magnetic resonance imaging: principles and techniques*, 2nd ed. Cambridge University Press, 2009, pp. 368–389.

[82] G. Glover, "Overview of functional magnetic resonance imaging," *Neurosurgery Clinics*, vol. 22, no. 2, pp. 133–139, 2011.

[83] F. Bloch, "Nuclear induction," *Physical Review*, vol. 70, no. 7-8, p. 460, 1946.

[84] E. Purcell, H. Torrey, and R. Pound, "Resonance absorption by nuclear magnetic moments in a solid," *Physical Review*, vol. 69, no. 1-2, p. 37, 1946.

[85] M. Levitt, *Spin dynamics: basics of nuclear magnetic resonance*, 2nd ed. John Wiley & Sons, 2013, pp. 23–37.

[86] Y. Farid and P. Lecat, "Biochemistry, hemoglobin synthesis," *StatPearls*, 2019.

[87] C. Eldeniz, M. Binkley, M. Fields, K. Guilliams, D. Ragan, Y. Chen *et al.*, "Bulk volume susceptibility difference between deoxyhemoglobin and oxyhemoglobin for HbA and HbS: A comparative study," *Magnetic Resonance in Medicine*, vol. 85, no. 6, pp. 3383–3393, 2021.

[88] S. Ogawa, D. Tank, R. Menon, J. Ellermann, S. Kim, H. Merkle *et al.*, "Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging," *Proceedings of the National Academy of Sciences*, vol. 89, no. 13, pp. 5951–5955, 1992.

[89] P. Jezzard, P. Matthews, and S. Smith, *Functional MRI*, 1st ed. Oxford University Press, 2001, pp. 6–12.

[90] K. Friston, K. Stephan, T. Lund, A. Morcom, and S. Kiebel, "Mixed-effects and fMRI studies," *Neuroimage*, vol. 24, no. 1, pp. 244–252, 2005.

[91] J. Ashburner, G. Barnes, C. Chen, J. Daunizeau, G. Flandin, K. Friston *et al.*, *SPM12 manual*, 2464th ed. Wellcome Trust Centre for Neuroimaging, 2014, pp. 61–92.

[92] H. Han and A. Glenn, "Evaluating methods of correcting for multiple comparisons implemented in SPM12 in social neuroscience fMRI studies: an example from moral psychology," *Social Neuroscience*, vol. 13, no. 3, pp. 257–267, 2018.

[93] S. McDougle, P. Butcher, D. Parvin, F. Mushtaq, Y. Niv, R. Ivry *et al.*, "Neural signatures of prediction errors in a decision-making task are modulated by action execution failures," *Current Biology*, vol. 29, no. 10, pp. 1606–1613, 2019.

[94] H. Lee, A. Macbeth, J. Pagani, and W. Young 3rd, "Oxytocin: the great facilitator of life," *Progress in Neurobiology*, vol. 88, no. 2, pp. 127–151, 2009.

[95] A. Campbell, “Attachment, aggression and affiliation: the role of oxytocin in female social behavior,” *Biological Psychology*, vol. 77, no. 1, pp. 1–10, 2008.

[96] T. Baskerville and A. Douglas, “Interactions between dopamine and oxytocin in the control of sexual behaviour,” *Progress in Brain Research*, vol. 170, pp. 277–290, 2008.

[97] C. De Dreu and M. Kret, “Oxytocin conditions intergroup relations through upregulated in-group empathy, cooperation, conformity, and defense,” *Biological Psychiatry*, vol. 79, no. 3, pp. 165–173, 2016.

[98] N. Kroemer, Y. Lee, S. Pooseh, B. Eppinger, T. Goschke, and M. Smolka, “L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action,” *Neuroimage*, vol. 186, pp. 113–125, 2019.

[99] S. Loiodice, P. Winlow, S. Dremier, E. Hanon, D. Dardou, O. Ouachikh *et al.*, “Pramipexole induced place preference after L-dopa therapy and nigral dopaminergic loss: linking behavior to transcriptional modifications,” *Psychopharmacology*, vol. 234, no. 1, pp. 15–27, 2017.

[100] T. Love, “Oxytocin, motivation and the role of dopamine,” *Pharmacology Biochemistry and Behavior*, vol. 119, pp. 49–60, 2014.

[101] M. Altemus, K. Jacobson, M. Debellis, M. Kling, T. Pigott, D. Murphy *et al.*, “Normal CSF oxytocin and NPY levels in OCD,” *Biological Psychiatry*, vol. 45, no. 7, pp. 931–933, 1999.

[102] K. Dumais and A. Veenema, “Vasopressin and oxytocin receptor systems in the brain: sex differences and sex-specific regulation of social behavior,” *Frontiers in Neuroendocrinology*, vol. 40, pp. 1–23, 2016.

[103] L. Krugel, G. Biele, P. Mohr, S. Li, and H. Heekeren, “Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 42, pp. 17 951–17 956, 2009.

[104] K. Aberg, K. Doell, and S. Schwartz, “The left hemisphere learns what is right: hemispatial reward learning depends on reinforcement learning processes in the contralateral hemisphere,” *Neuropsychologia*, vol. 89, pp. 1–13, 2016.

[105] R. Chowdhury, M. Guitart-Masip, C. Lambert, P. Dayan, Q. Huys, E. Düzel *et al.*, “Dopamine restores reward prediction errors in old age,” *Nature Neuroscience*, vol. 16, no. 5, pp. 648–653, 2013.

[106] M. Pessiglione, B. Seymour, G. Flandin, R. Dolan, and C. Frith, “Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans,” *Nature*, vol. 442, no. 7106, pp. 1042–1045, 2006.

[107] L. Wang, F. Li, D. Wang, K. Xie, D. Wang, X. Shen *et al.*, "NMDA receptors in dopaminergic neurons are crucial for habit learning," *Neuron*, vol. 72, no. 6, pp. 1055–1066, 2011.

[108] G. Jocham, T. Klein, and M. Ullsperger, "Differential modulation of reinforcement learning by D2 dopamine and NMDA glutamate receptor antagonism," *Journal of Neuroscience*, vol. 34, no. 39, pp. 13 151–13 162, 2014.

[109] M. Liljeholm and J. O'Doherty, "Contributions of the striatum to learning, motivation, and performance: an associative account," *Trends in Cognitive Sciences*, vol. 16, no. 9, pp. 467–475, 2012.

[110] D. Furman, M. Chen, and I. Gotlib, "Variant in oxytocin receptor gene is associated with amygdala volume," *Psychoneuroendocrinology*, vol. 36, no. 6, pp. 891–897, 2011.

[111] P. Veinante and M. Freund-Mercier, "Distribution of oxytocin-and vasopressin-binding sites in the rat extended amygdala: a histoautoradiographic study," *Journal of Comparative Neurology*, vol. 383, no. 3, pp. 305–325, 1997.

[112] K. Ebner, O. Bosch, S. Krömer, N. Singewald, and I. Neumann, "Release of oxytocin in the rat central amygdala modulates stress-coping behavior and the release of excitatory amino acids," *Neuropsychopharmacology*, vol. 30, no. 2, pp. 223–230, 2005.

[113] J. Paton, M. Belova, S. Morrison, and C. Salzman, "The primate amygdala represents the positive and negative value of visual stimuli during learning," *Nature*, vol. 439, no. 7078, pp. 865–870, 2006.

[114] V. Costa, O. Dal Monte, D. Lucas, E. Murray, and B. Averbeck, "Amygdala and ventral striatum make distinct contributions to reinforcement learning," *Neuron*, vol. 92, no. 2, pp. 505–517, 2016.

[115] J. Wang, L. Zhu, V. Brown, R. De La Garza II, T. Newton, B. King-Casas *et al.*, "In cocaine dependence, neural prediction errors during loss avoidance are increased with cocaine deprivation and predict drug use," *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 4, no. 3, pp. 291–299, 2019.

[116] B. Guillaume, X. Hua, P. Thompson, L. Waldorp, T. Nichols, Alzheimer's Disease Neuroimaging Initiative *et al.*, "Fast and accurate modelling of longitudinal and repeated measures neuroimaging data," *NeuroImage*, vol. 94, pp. 287–302, 2014.

[117] B. Guillaume, "Accurate non-iterative modelling and inference of longitudinal neuroimaging data," Ph.D. dissertation, Université de Liège, Liège, Belgique, 2015.

[118] D. Kennedy, C. Haselgrove, B. Fischl, J. Breeze, J. Frazier, L. Seidman *et al.*, "Harvard Oxford cortical and subcortical structural atlases," 2016, last accessed 30 September 2021.

[Online]. Available: https://ftp.nmr.mgh.harvard.edu/pub/dist/freesurfer/tutorial_packages/centos6/fsl_507/doc/wiki/Atlases.html

[119] G. Murray, F. Knolle, K. Ersche, K. Craig, S. Abbott, S. Shabbir *et al.*, "Dopaminergic drug treatment remediates exaggerated cingulate prediction error responses in obsessive-compulsive disorder," *Psychopharmacology*, vol. 236, no. 8, pp. 2325–2336, 2019.

[120] L. Kubzansky, W. Mendes, A. Appleton, J. Block, and G. Adler, "A heartfelt response: oxytocin effects on response to social stress in men and women," *Biological Psychology*, vol. 90, no. 1, pp. 1–9, 2012.

[121] K. Bales and C. Carter, "Sex differences and developmental effects of oxytocin on aggression and social behavior in prairie voles (Microtus ochrogaster)," *Hormones and Behavior*, vol. 44, no. 3, pp. 178–184, 2003.

[122] K. Miwa and H. Terai, "Impact of two types of partner, perceived or actual, in human–human and human–agent interaction," *Computers in Human Behavior*, vol. 28, no. 4, pp. 1286–1297, 2012.

[123] D. Scheele, C. Schwering, J. Elison, R. Spunt, W. Maier, and R. Hurlemann, "A human tendency to anthropomorphize is enhanced by oxytocin," *European Neuropsychopharmacology*, vol. 25, no. 10, pp. 1817–1823, 2015.

[124] N. Mesgarani, C. Cheung, K. Johnson, and E. Chang, "Phonetic feature encoding in human superior temporal gyrus," *Science*, vol. 343, no. 6174, pp. 1006–1010, 2014.

[125] A. Aeby, X. De Tiège, M. Creuzil, P. David, D. Balériaux, B. Van Overmeire *et al.*, "Language development at 2 years is correlated to brain microstructure in the left superior temporal gyrus at term equivalent age: a diffusion tensor imaging study," *NeuroImage*, vol. 78, pp. 145–151, 2013.

[126] S. Baron-Cohen, H. Ring, S. Wheelwright, E. Bullmore, M. Brammer, A. Simmons *et al.*, "Social intelligence in the normal and autistic brain: an fMRI study," *European Journal of Neuroscience*, vol. 11, no. 6, pp. 1891–1898, 1999.

[127] E. Bigler, S. Mortensen, E. Neeley, S. Ozonoff, L. Krasny, M. Johnson *et al.*, "Superior temporal gyrus, language function, and autism," *Developmental Neuropsychology*, vol. 31, no. 2, pp. 217–238, 2007.

[128] R. Adolphs, "Cognitive neuroscience of human social behaviour," *Nature Reviews Neuroscience*, vol. 4, no. 3, pp. 165–178, 2003.

[129] S. Gregory, R. Blair, A. Simmons, V. Kumari, S. Hodgins, N. Blackwood *et al.*, "Punishment and psychopathy: a case-control functional MRI investigation of reinforcement learning in violent antisocial personality disordered men," *The Lancet Psychiatry*, vol. 2, no. 2, pp. 153–160, 2015.

[130] H. Feldman, “How young children learn language and speech: Implications of theory and evidence for clinical pediatric practice,” *Pediatrics in Review*, vol. 40, no. 8, p. 398, 2019.

[131] M. Gallagher and A. Chiba, “The amygdala and emotion,” *Current Opinion in Neurobiology*, vol. 6, no. 2, pp. 221–227, 1996.

[132] T. Brosch, K. Scherer, D. Grandjean, and D. Sander, “The impact of emotion on perception, attention, memory, and decision-making,” *Swiss Medical Weekly*, vol. 143, p. w13786, 2013.

[133] K. Bickart, B. Dickerson, and L. Barrett, “The amygdala as a hub in brain networks that support social life,” *Neuropsychologia*, vol. 63, pp. 235–248, 2014.

[134] C. Mosher, P. Zimmerman, and K. Gothard, “Neurons in the monkey amygdala detect eye contact during naturalistic social interactions,” *Current Biology*, vol. 24, no. 20, pp. 2459–2464, 2014.

[135] P. Namburi, A. Beyeler, S. Yorozu, G. Calhoon, S. Halbert, R. Wichmann *et al.*, “A circuit mechanism for differentiating positive and negative associations,” *Nature*, vol. 520, no. 7549, pp. 675–678, 2015.

[136] P. Garris and R. Wightman, “Distinct pharmacological regulation of evoked dopamine efflux in the amygdala and striatum of the rat in vivo,” *Synapse*, vol. 20, no. 3, pp. 269–279, 1995.

[137] S. Haber, J. Fudge, and N. McFarland, “Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum,” *Journal of Neuroscience*, vol. 20, no. 6, pp. 2369–2382, 2000.

[138] S. Floresco, C. Yang, A. Phillips, and C. Blaha, “Association basolateral amygdala stimulation evokes glutamate receptor-dependent dopamine efflux in the nucleus accumbens of the anaesthetized rat,” *European Journal of Neuroscience*, vol. 10, no. 4, pp. 1241–1251, 1998.

[139] J. Jones, J. Day, B. Aragona, R. Wheeler, R. Wightman, and R. Carelli, “Basolateral amygdala modulates terminal dopamine release in the nucleus accumbens and conditioned responding,” *Biological Psychiatry*, vol. 67, no. 8, pp. 737–744, 2010.

[140] N. Cilz, A. Cymerblit-Sabba, and W. Young, “Oxytocin and vasopressin in the rodent hippocampus,” *Genes, Brain and Behavior*, vol. 18, no. 1, p. e12535, 2019.