FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Sara Dias dos Santos

The role of the core and extended face networks in visual perception and high level social cognition - Trust game, an fMRI study

Supervisor: Miguel de Sá e Sousa de Castelo Branco (M.D., PhD)

Co-Supervisor: Inês Alexandra Teixeira de Almeida (PhD)

Dissertation presented to the University of Coimbra to obtain
a Master's degree in Biomedical Engineering.

Dissertação de Mestrado apresentada à Universidade de Coimbra para
prestação de provas de Mestrado em Engenharia Biomédica

Coimbra, 2014

This project has been developed in colaboration with:

FMUC FACULDADE DE MEDICINA
UNIVERSIDADE DE COIMBRA

iBiLi Institute for Biomedical Imaging and Life Sciences
University of Coimbra

ICNAS INSTITUTO DE CIÊNCIAS NUCLEARES APLICADAS À SAÚDE
UNIVERSIDADE DE COIMBRA

Aos meus pais e irmãs.

Acknowledgements

Começo por agradecer, em primeiro lugar, ao Professor Miguel Castelo-Branco. Por toda a orientação e apoio, pelas sugestões e por prover todas as oportunidades e recursos para que o projecto se desenrolasse da melhor forma. Por, no meio de tantos compromissos, ter sempre o tempo necessário para esclarecer todas as dúvidas e ter sempre novas ideias que enriqueceram o trabalho.

À Inês, por todo o acompanhamento ao longo do projecto. Pelas inúmeras "discussões", sugestões e correcções. Porque me ajudou a compreender que é normal e não devo recear o facto de ainda não saber e conhecer tantas coisas que, neste meio e para muitos, são verdadeiramente simples. Porque sempre foi tão paciente com todas as minhas dúvidas e porquês, e me ajudou a não desanimar quando a tarefa e as análises pareciam nunca mais ficar concluídas.

Ao José Rebola, pela sua eterna paciência e simpatia. Pelas suas explicações tão terra-a-terra que tornaram tantas coisas tão mais fáceis de compreender.

À Catarina Duarte e João Duarte, cuja ajuda foi tão preciosa quando o código em Matlab teimava em não funcionar!

Ao Professor Miguel Morgado, que como nosso coordenador de curso, sempre defendou com tanta garra os Engenheiros Biomédicos, bem como o seu importante papel no presente e no futuro.

À Miriam Santos, Marta Pinto, Joana Paiva, Sofia Prazeres, Daniela Martins, Patrícia Santos, Diana Capela e Carolina Queijo pelos bons momentos passados ao longo destes 5 anos.

Ao Enoque Moisés, pela sua tremenda paciência, ajuda e dedicação, principalmente neste último ano. Por ser um exemplo vivo de integridade no meio de tantas coisas menos boas.

Às minhas irmãs, Raquel e Olívia. Uma com a sua sensatez e pés tão bem assentes na terra e a outra com os seus devaneios e desvarios de artista que é. Sem vocês jamais seria quem sou!

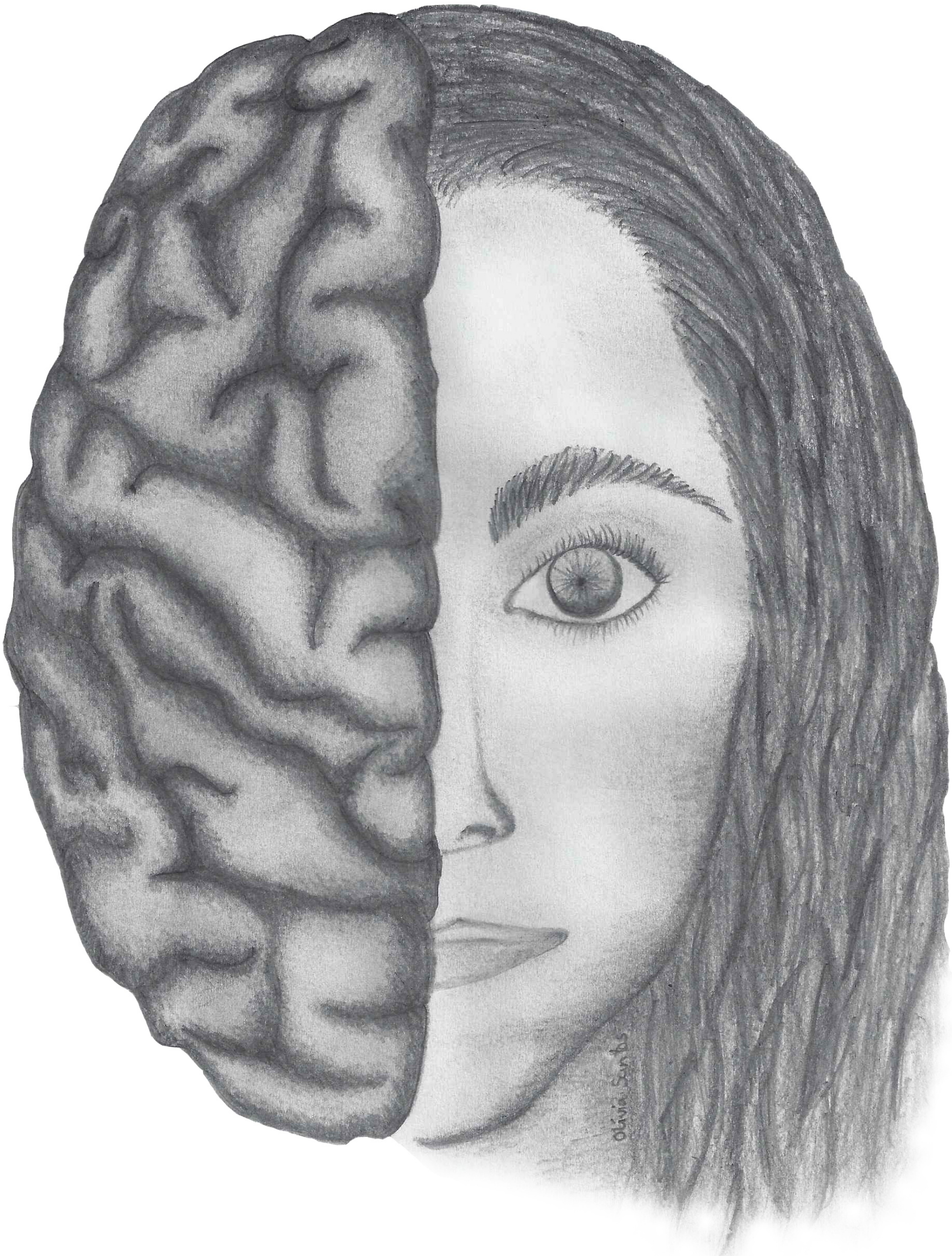
Aos meus pais, que sempre me mostraram pelo seu extraordinário exemplo, esforço e dedicação que o único sítio em que o sucesso vem antes do trabalho é no dicionário. Porque acima de tudo, sempre me ensinaram que "tudo posso nAquele que me fortalece" (Filipenses 4:13).

E por fim, mas de forma alguma menos importante, agradeço a Deus. Porque "o temor do Senhor é o princípio da sabedoria" (Prov. 9:10), e porque "todas as coisas foram feitas por intermédio dEle, e sem Ele, nada do que foi feito se fez." (João 1:3). Incluindo esta tese.

Financing

This project was supported by BIAL Foundation, PT/FB/BL-2012-133 - The role of the core and extended face networks in visual perception and high level social cognition.

Bial



"Your eyes are windows into your body.

Mattews 6:22

Abstract

The present project aims to have a better understanding of the visual processing during the perception of 3D faces and the cognitive processing of considered important social stimuli. The fundamental question of this study aims to disentangle how is the perception of faces and facial cues integrated in the social cognition network to extract information regarding trustworthiness, a dimension which of most importance during social interactions.

Other question is related to the role of the amygdala in dynamic face processing, (namely, facial characteristics such as the eye gaze) and reward expectations, since the amygdala has been implicated not only in processes regarding high-level decision-making responses to complex visual stimuli, but also in processes involved in reward expectations. Furthermore, this project seeks to understand the neural correlates of face perception when facial signals, more precisely gaze direction, are used as social attention cues to other human agents.

To achieve the goals above presented, and in order to understand the neural bases of face processing (e.g. amygdala response), functional Magnetic Resonance Imaging (fMRI) was used. Participants interacted in a real-time social interaction known as the Trust Game.

Results here presented reveal amygdala involvement for faces presenting Directed rather than Averted gaze. Also, middle frontal frontal gyrus, precentral gyrus and left superior temporal gyrus were shown to be involved in reward expectation processes.

If eye gaze is able to influence trustworthiness judgements, this can be of a major impact, especially if the amygdala is involved in the extraction of this type of information from social interactions.

Keywords: face, eye gaze, amygdala, social cognition, trust, reward expectations, fMRI.

Resumo

Este projecto tem como objectivo obter uma melhor compreensão entre o processamento visual durante a percepção de faces 3D e o processamento cognitivo destes estímulos sociais. A questão fundamental aqui apresentada tem como objectivo compreender como é que a percepção de faces e características faciais se encontra integrada na rede da cognição social, extraindo desta rede informação relacionada com processos de confiança, uma dimensão deveras importante no âmbito das interações sociais.

Outra questão relacionada com este assunto diz respeito ao papel da amígdala no processamento de faces dinâmicas (nomeadamente de características como o *eye gaze*) e de expectativas de recompensas, uma vez que se tem revelado o envolvimento da amígdala não apenas em processos relativos à tomada de decisão após a visualização de estímulos complexos, mas também em processos relacionados com expectativas de recompensa. Para além do já referido, este estudo também procura proporcionar uma melhor compreensão dos processos neurais associados à percepção de sinais faciais, mais precisamente a direcção do olhar, quando usados ao interagir com outras pessoas.

De forma a atingir os objectivos acima propostos, de forma a compreender as bases neurais do processamento de faces (como a resposta da amígdala), foi utilizada ressonância magnética functional, enquanto os participantes interagiam num jogo conhecido como o Trust Game.

Os resultados obtidos revelam o envolvimento da amígdala para faces que apresentavam Directed gaze quando comparando com faces que apresentavam Averted gaze. Este estudo também revelou o envolvimento de estruturas como o middle frontal gyrus, precentral gyrus e left superior temporal gyrus em processos de expectativas de recompensa.

Se o *eye gaze* de alguma forma influenciar julgamentos de confiança, então isto poderá ter um grande impacto, principalmente se a amígdala estiver envolvida na extracção deste tipo de informação a partir de interacções sociais.

Palavras-chave: face, *eye gaze*, amígdala, cognição social, confiança, expectativas de recompensa, fMRI.

Abbreviations

AC-PC	Anterior commissure - posterior commissure
ANOVA	Analysis of variance
Av.	Averted
Av. Dir. Av.	Averted directed averted
BA	Brodmann area
BOLD	Blood-oxygen-level-dependent
Dir. Av.	Directed Averted
Dir.	Directed
EEG	Electroencephalography
FFA	Fusiform face area
FFX	Fixed Effects
fMRI	Functional magnetic resonance imaging
GLM	General linear model
Hb	Hemoglobin
IPS	Intraparietal Sulcus
ISI	Inter-stimulus interval
ITI	Intertrial interval
MEG	Magnetoencephalography
MPF	Medial prefrontal cortex
MPRAGE	Magnetization prepared rapid gradient echo
MR	Magnetic resonance
MRI	Magnetic resonance imaging
NMR	Nuclear magnetic resonance
PET	Positron emission tomography
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
ROI	Region-of-interest

RT	Response time
SPECT	Single photon emission computerized tomography
SD	Standard deviation
SE	Standard error
sec	seconds
SPSS	Statistical Package for the Social Sciences
STS	Superior temporal sulcus
TAL	Talairach
TG	Trust Game
TR	Repetition time
TTrust	Trustworthy Trustee
TUntrust	Untrustworthy Trustee
UG	Ultimatum Game

List of Figures

List of Tables

Page

Table of Contents

Motivation

During social interactions we are confronted with the need to make decisions regarding others intentions. The study of decision-making allows understanding how can we choose a certain option taking into account all the alternatives available [1]. Considering that all our choices and decisions impact our lives, it is important to comprehend how can some mechanisms, namely facial and in particular gaze signals, influence our perception of trust, leading us to choose what we consider to be the best option.

One of the key components that influence our daily decisions relates to social interactions and perception of affective signals in others [2, 3]. Little information is often sufficient for us to make rapid inferences about others, such as body posture or even facial impressions [4-6]. In fact, the face is one of the most relevant sources of information from which judgements of competence, aggressiveness and trustworthiness can be made. Such judgements are known to impact real world behaviour, predetermining the outcome of social interactions [2].

Several neuroeconomic games are currently used in order to study decision-making. The Trust Game is one of the strategies in which mathematical decision models are applied to investigate psychological and neuronal correlates concerning social decisions [1].

The myth that people generally behave in a rational and self-interested manner is now known to not hold. In fact, people do not always behave rationally, and base their decision on strong affective and emotional signals. For this reason sometimes do not take decisions according to what could be the predicted just based in rational decision [1, 7]. What are then the factors that influence behavioural outcomes? What are the facial characteristics which influence others' trustworthiness judgements? Some studies point the direction of eye gaze as one of the most relevant cues by which people understand others minds and determine their intentions. Also, gaze direction is known to influence our social perceptions and evaluations of others [8]. However, not many studies have been performed in order to have a better understanding of the influence of eye gaze in these processes.

When evaluating traits like the trustworthiness of faces, some neuroimaging studies highlight the amygdala as one of the most frequently recruited brain regions [9, 10]. Nevertheless, as far as we know, the role of the amygdala has never been studied when developing impressions of others over a social interaction, while direction and duration of eye gaze is modulated. Consequently, the way amygdala and other brain structures process this information and the way this is reflected in decision-making is still unknown.

Since perception of the trustworthiness of others has a considerable impact in decision-making, it is important to have a better understanding of the influence of face perception in social cognition. Thus, in order to comprehend the way social behaviour unfolds, direction and duration of eye gaze will be experimentally manipulated.

Also, clinically it can be relevant, since it may help to understand pathologies of social cognition and behaviour such as psychopathy.

Different methodologies were then applied to help to unveil this problem, so that appropriate answers could be found to the questions previously addressed.

Objectives

This report, regarding the project *The role of the core and extended face networks in visual perception and high level social cognition*, is presented to conclude the Master Degree in Biomedical Engineering.

This project was performed in the area of neurosciences, and it was developed with the aims to learn how to identify and define a relevant problem in the complex fields of imaging and social neuroscience, and in addition to develop and test a new experimental design and paradigm suitable for the raised neuroscientific questions. To achieve these goals, formal methods were used such as systematic reviews of the literature for definition of the problem and quantitative meta-analyses, to test the defined hypothesis.

This study is focused on the definition of hierarchical relations between low level visual and intermediate level visual processing of 3D faces as well as high level cognitive processing of these important social stimuli and their contextual dependence.

Therefore, in order to understand the neural bases of face processing, and in particular the amygdala response, participants will interact in a real-time social interaction known as the Trust Game, using functional Magnetic Resonance Imaging (fMRI).

The fundamental questions this project attempts to respond are: how is perception of faces, which are pivotal social stimuli, integrated with cognitive processing with high level regions involved in social cognition? How do bottom-up and top-down mechanisms integrate in such processing?

The dissertation is structured in 4 chapters. The first chapter, the introduction, consists in a systematic review and meta-analysis that was conducted to have a better understanding of what is known about the role of the amygdala regarding social cognition and the evaluation of trustworthiness. Also, other concepts considered relevant for the study that were not taken into account in the systematic review are presented, such as the importance of eye gaze in social interactions. The second chapter, Methods, is focused in the design of the experiment, namely in the statistical methods and investigated experimental contrasts tested, the experiment timeline, the experiment itself and the analysis model. In the third chapter, results will be presented and, in the forth, these will be discussed within the context of the literature. Chapters 5 presents the concluding remarks and chapter 6 suggestions of future work.

The project was developed under the scientific orientation of the Professor Miguel

Castelo-Branco (M.D., Ph.D.) and Inês Almeida (Ph.D) at Institute for Biomedical Imaging and Life Sciences (IBILI) and Institute of Nuclear Sciences Applied to Health (ICNAS).

1 Introduction

1.1 Systematic Review

In our daily life, we are constantly confronted with decisions and choices, with inherent repercussions in future outcomes. In turn, these decisions are influenced by social interactions and perception of others' intentions [2, 3].

The mechanisms that allow us to make such inferences are related to social cognition. Social cognition is defined as a mental process that underlies social interactions, in which our ability to understand the intentions and dispositions of others is included [11]. In fact, interacting with others concedes us the opportunity to collect information about their behaviour. These social inferences can be processed in very different ways. Each person will process the social information considered more important and build a model of the social world. Some of the social cues here included are gaze direction, which can reveal the attentional focus of others; facial expressions, like fear or disgust, which can signal a potentially dangerous outcome; or even body posture, from which it can be inferred someone's intention of doing or planning a particular action [4, 12].

Previous studies showed that first personal impressions can be built based on brief exposures in the order of milliseconds of another person's face [13, 14]. Although much evidence comes from the use of emotional expressions, trait judgements such as trustworthiness, competence, and aggressiveness can be performed even upon exposure to neutral faces [15]. These types of inferences are essential since they can predetermine the outcome of social interactions. For example, approach or avoidance behaviours towards someone can result from previous trustworthiness judgements [2].

It has been argued that detection of trustworthiness signals is crucial for human survival [16]. A set of brain regions is specifically involved in the perception of trustworthiness, with previous functional neuroimaging (fMRI) studies showing that facial trustworthiness is related with the activation of areas such as the amygdala, the insula and the fusiform gyrus [16-19].

The amygdala in particular has been associated with social judgement and perception, more specifically with social, emotional and reward processing [20]. Some studies have found that the human amygdala is highly implicated when evaluating other people's intentions and affective state, responding to social cues like fearful faces [21] and variations in eye gaze [22]. Studies also revealed that his structure plays a very important role in the perceived trustworthiness of faces [2, 23, 16, 24]. Adolphs et al. showed that patients with amygdala lesions or dysfunction were not capable of judging others' trustworthiness. Patients with bilateral amygdala damage judged untrustworthy-looking faces as if they were more approachable and trustworthy comparing to neurologically normal subjects [10, 28].

Additionally, some fMRI studies indicate that the activity evoked in the amygdala by untrustworthy-looking faces is higher than for trustworthy-looking ones [14]. In other words, the amygdala response to faces increases with the decrease of their perceived trustworthiness, even when subjects are performing tasks that do not require explicit evaluation of faces [2, 16, 25, 26]. This stronger response of the amygdala for untrustworthy faces is sometimes described as following an ordinal quasilinear trend, while other studies have showed U-shaped, quadratic responses in this structure with higher responses at the extremes of the trustworthiness dimension [2, 26-28].

In order to study decision-making related to social cognition and inherent trustworthiness judgments, various strategies are used. Several hypotheses have been given when trying to explain the role of the amygdala regarding the trustworthiness of faces, but some remain unsolved. In the current study we planned to answer to the following questions: does the amygdala respond more to trust or untrustworthy faces; is this a quasilinear or quadratic response pattern? An additional important question is whether the amygdala is considered a core structure in the social cognition network, i.e., does it modulate, or is it a modulated region in the social information processing network?

Considering the above mentioned, a systematic review was conducted with the purpose to answer to the questions previously addressed. PRISMA statements guidelines were followed [30]and the articles were retrieved from three databases, according to a predefined strategy. The keywords used were "face AND trust* AND amygdala" in order to evaluate the role of the amygdala in social cognition, namely in the context of trustworthiness judgements of faces.

Data sources and literature search

A systematic review was performed adhering to the principles of the PRISMA statement [29]. Two preliminary searches were conducted in order to properly define the search terms. The final search reported herein was undertaken in July 2014. Publications were searched on three databases, notably on MEDLINE, via PubMed (http://www.ncbi.nlm.nih.gov/pubmed), on Science Direct (Elsevier, http://www.sciencedirect.com/) and Web of Science (https://webofknowledge.com/), using the search string "face AND trust* AND amygdala", with the filter "title+abstract+keywords" being imposed. Abstracts were screened for English language publications and fMRI studies only. References included in the articles deemed appropriate for full-text revision were hand-searched for retrieving other relevant publications.

Eligibility criteria

For a study to be considered as eligible, it had to meet the following criteria:

1. Be written in English language;
2. Involve clinical or healthy participants (animal studies were excluded);
3. Involve original research articles (review studies were excluded);
4. Use of brain imaging techniques, namely functional neuroimaging (fMRI),
5. Goal of the study: to assess normal performance without introducing sources of disturbance such as transcranial magnetic stimulation (TMS),
6. Include direct measurements made in the amygdala, with activation reported;;
7. Activity in the amygdala had to be specifically and separately reported (e.g. without being included in a general "medial temporal lobe" label).

Study selection and data extraction

Eligible studies selection was performed by two authors (I.A. and S.S.), according to the following phases:

Identification phase - Data collected and duplicates eliminated.

Screening phase - Titles and abstracts of the remaining records independently screened and assessed for eligibility by the two authors (see Figure 1).

Eligibility phase - Records considered potentially eligible for criteria 1 to 7 by at least one of the reviewers included for further full paper assessment.

The selected studies (behavioural, neuroimaging) presented face stimuli in a trustworthiness task under an fMRI procedure with measurements of amygdala's activity. Relevant features of these articles were extracted, being the following factors of interest were registered:

1. Type of task (implicit or explicit; and trustworthiness judgements or gender categorization),
2. Stimulus duration,
3. Stimulus type (faces: real or avatars),
4. Direction of gaze (directed, averted),
5. Nature of stimuli presentation (static pictures or dynamic videos),
6. Task design (event-related, block design),
7. Neuroimaging statistical analysis (whole brain; regions of interest, ROI),
8. Type of neuroimaging analyses: contrasts of interest, correlation or connectivity analysis,
9. Amygdala lateralization (if mentioned).

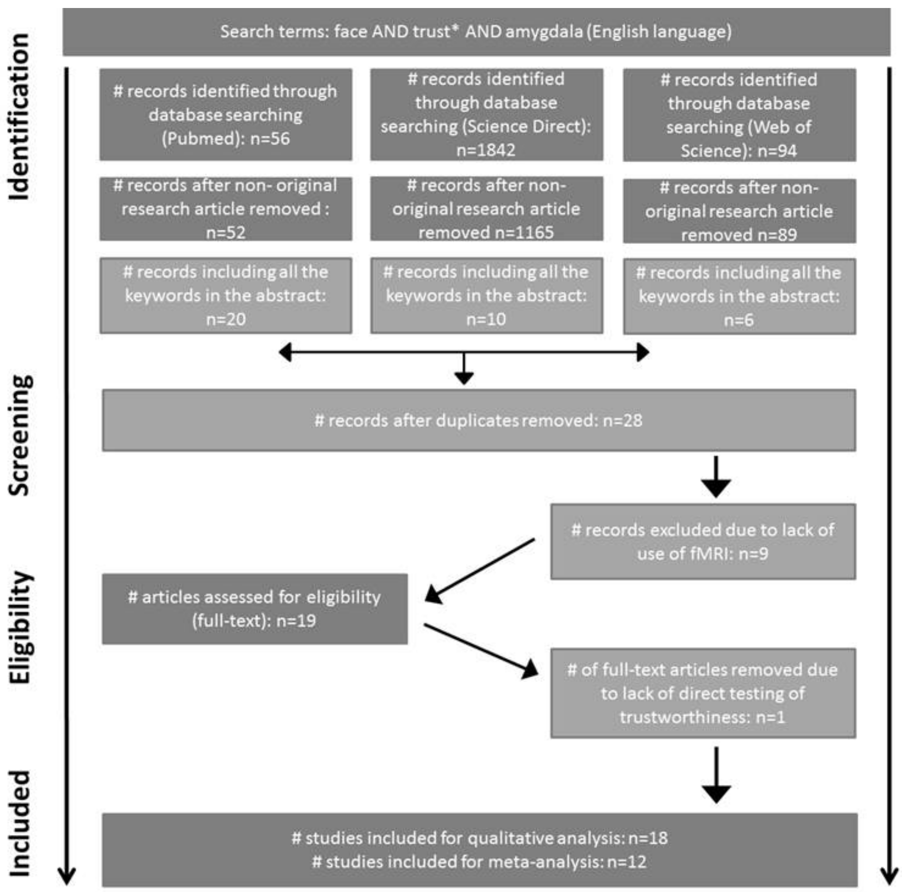


Figure 1: Flow of information describing the different phases of the systematic review.

Data analysis

The review provides a quantitative and qualitative summary of neuroimaging (fMRI) findings. Meta-analysis of effects was undertaken either for the specific contrast "untrustworthy > trustworthy faces" or for the linear correlation "untrustworthy - trustworthy" outcome using correlation coefficients to assess the effect of the different factors in the outcome.

In order to obtain the summary measures, we performed both fixed and random-effects analyses. Heterogeneity was assessed both with the inconsistency (I^2) statistic and the Q coefficient. I^2 is a standard test that measures the degree of inconsistency across studies. This test results in a range from 0% to 100%, which describe the proportion of variation in treatment effect estimates due to inter-study variation [30]. Values higher than 50% indicate a substantial level of heterogeneity. Q coefficient was also used to calculate the homogeneity of effect sizes [31]. If the studies only differ by the sampling error ($I^2 < 50\%$, homogeneous case), a fixed-effects model is applied in order to obtain an average effect size. If the study results differ by more than the sampling error ($I^2 > 50\%$, heterogeneous case) a random-effects model is preferred instead [31].

Both t and Z statistical values taken from the original research articles were considered to estimate the effect sizes. t-values were taken from 8 studies, whereas 3 of the studies reported Z-scores, with the remaining one reporting r values. Z-scores were used to estimate t-scores. The meta-analysis was performed after r determination from the t-values according to the following equation: $\mathrm{r} = \frac{t}{(n-2+t^2)}$.

All the estimates included were recomputed from original articles descriptions, potentially resulting in slightly different values.

The meta-analysis was performed with the software package MedCalc (R) (version 12.7.2.0 – 64 bit, Copyright 1993-2013, MedCalc Software bvba, Mariakerke, Belgium, www.medcalc.org) which was used for calculation of the power, inconsistency and the 95% CI of the study.

Results

Considering the search items above discriminated, and as presented in Figure 1, this review of the literature identified 1992 potential target studies, from which 10 were duplicates. According to the information available in the abstracts and taking into account criteria (1) to (7), 1973 records were excluded.

From the 19 articles selected to this point, a full text assessment was carried out, resulting in the exclusion of 1, since there was no direct assessment of trustworthiness in the reported study.

The remaining 18 studies were submitted to a qualitative data extraction (e.g. amygdala response to trust/untrustworthy faces; connectivity analysis of the amygdala in the brain social network). Since 12 of the 18 studies satisfied the criteria for inclusion in a meta-analysis, the same was conducted, and we present hereafter its results.

To note that all of the studies used functional magnetic resonance imaging (fMRI), besides reporting behavioural data. Only one study reported effect sizes [24]. All were published in the last 6 years, except one dating from 2002 [16].

Analysis

In the following section, amygdala responses to untrustworthy versus trustworthy faces found in the selected studies will be reported and analysed, as well as results evaluating the role of the amygdala as a core or modulated region.

Amygdala responses to untrustworthy versus trustworthy faces

Table 1 presents right amygdala activation for faces responding more to untrustworthy > trustworthy faces regarding all studies included in the systematic review.

Table 1: Meta-analysis: amygdala activation for faces untrustworthy > trustworthy and response type

Number	Author	Year	R Amygdala
1	Baas et al.	2008	n.a.
2	Baron et al.	2011	Linear
3	Bos et al.	2012	Linear
4	Bzdok et al.	2012	n.a.
5	Doallo et al.	2012	Linear
6	Engell et al.	2007	Linear
7	Killgore et al.	2013	n.a.
8	Kim et al.	2012	Linear
9	Kragel et al.	2014	Linear
10	Mattavelli et al.	2012	Quadratic*
11	Pinkham et al.	2008	n.a.
12	Pinkham et al.	2008	n.a.
13	Rule et al.	2013	Quadratic
14	Said et al.	2009	Linear
15	Todorov et al.	2008	Linear
16	Todorov et al.	2008	Linear
17	vanRijn et al.	2012	Linear
18	Winston et al.	2002	Linear

(*) For this study, the R-squared values for the quadratic polynomial were significantly higher than the R-squared for the linear regression. However, linear results were also presented, being these used for the meta-analysis.

Results shown in Table 1 only present right amygdala responses for untrustworthy > trustworthy faces. Studies number 6, 16 and 18 also reported left amygdala activations, however mainly for trustworthy > untrustworthy faces. Thus, it is worth to notice that the amygdala behaviour seems to present distinct response patterns for different levels of trustworthiness. While the right amygdala presented higher responses with the decrease of trustworthiness, left amygdala revealed higher activations for trustworthy stimuli [2] [16] [25]. Since there was no sufficient data for computing a meta-analysis, a separate analysis for left amygdala was not conducted. For the meta-analysis computation, only linear response models were included.

Twelve studies were included in the final meta-analysis to measure the amplitude of amygdala responses for the contrast "untrustworthy > trustworthy" of judged faces. Results are presented in Table 2 and Figure 2.

Table 2: Meta-analysis: sample size, correlation coefficient and 95% CI for the contrast "untrustworthy > trustworthy" faces for the amygdala response of the selected studies.

Number	Author	Year	Sample size	Correlation coefficient	95% CI
2	Baron et al.	2011	24	0.654	0.340 to 0.837
3	Bos et al.	2012	16	0.655	0.236 to 0.869
5	Doallo et al.	2012	12	0.750	0.309 to 0.926
6	Engell et al.	2007	15	0.884	0.679 to 0,961
8	Kim et al.	2012	12	0.638	0.101 to 0.887
9	Kragel et al.	2014	43	0.377	0.086 to 0.608
10	Mattavelli et al.	2012	20	0.761	0.480 to 0.900
14	Said et al.	2009	32	0.507	0.192 to 0.727
15	Todorov et al.	2008	15	0.832	0.557 to 0.943
16	Todorov et al.	2008	21	0.506	0.095 to 0.770
17	vanRijn et al.	2012	18	0.724	0.388 to 0.890
18	Winston et al.	2002	16	0.754	0.412 to 0.910

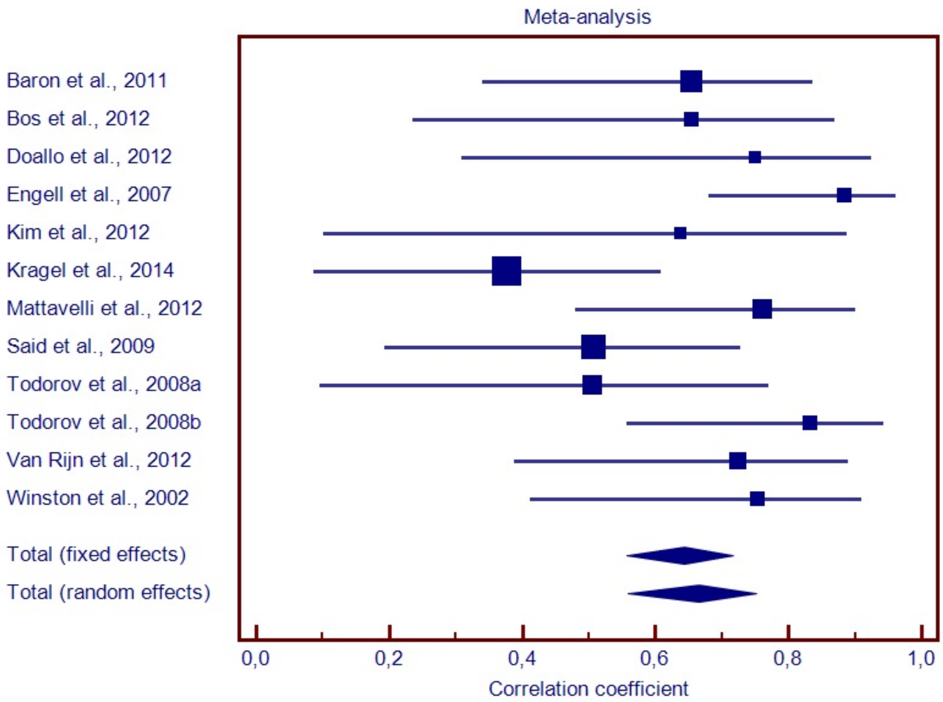


Figure 2: Forest plot resulting from the meta-analysis for the contrast "Untrustworthy > Trustworthy" faces presenting central values of correlation coefficients (square markers) and their confidence intervals (horizontal lines). The size of the square markers varies with the sample size. Diamond markers represent pooled effects. The location of the diamond represents the estimated effect size and the width of the diamond reflects the precision of the estimate.

Since there was no evidence of statistical heterogeneity among the studies ($I^2 = 34.06$ %; Q = 16.6830; DF = 11; p = 0.1176; 95% CI: 0.00 - 75.38), the fixed-effect model approach was chosen. This model showed that amygdala responses in healthy controls are higher to untrustworthy compared to trustworthy faces (ffx: 0.637; N = 244; IC 95%: 0.556 a 0.716) (see Figure 2 and Table 2).

Amygdala: core or modulated region during trustworthiness judgements?

Another purpose of this study was to assess if amygdala was reported as a core or modulated region in the social information processing network.

The following table (Table 3) presents the results obtained from the 18 selected studies. Only 3 of these 18 studies have performed connectivity analysis in order to directly evaluate the role of the amygdala as a core or modulated region.

Table 3: Papers from which amygdala was considered a core or modulated region

Number	Author	Year	Core region	Modulated region
1	Baas et al.	2008	n.a.	n.a.
2	Baron et al.	2011	x	
3	Bos et al.	2012	x	
4	Bzdok et al.	2012	n.a.	n.a.
5	Doallo et al.	2012	n.a.	n.a.
6	Engell et al.	2007	n.a.	n.a.
7	Killgore et al.	2013	n.a.	n.a.
8	Kim et al.	2012	n.a.	n.a.
9	Kragel et al.	2014	n.a.	n.a.
10	Mattavelli et al.	2012	n.a.	n.a.
11	Pinkham et al.	2008	n.a.	n.a.
12	Pinkham et al.	2008	n.a.	n.a.
13	Rule et al.	2013	n.a.	n.a.
14	Said et al.	2009	n.a.	n.a.
15	Todorov et al.	2008	n.a.	n.a.
16	Todorov et al.	2008	x	
17	vanRijn et al.	2012	n.a.	n.a.
18	Winston et al.	2002	n.a.	n.a.

According to the results presented in Table 3, 3 studies identified the amygdala as a seed region, modulating other areas (Dorsal Medial Prefrontal Cortex, Orbital Frontal Cortex, Occipital and Temporal Regions in face network).

Thus, this systematic review concerning amygdala responses to trustworthiness traits in faces returned two major conclusions. First, by means of a meta-analysis, it was concluded that amygdala responses are higher for untrustworthy faces. Second, all studies which directly assessed the role of the amygdala in the network related with trustworthiness judgements reported a core role of this region, modulating other regions such as the Dorsal Medial Prefrontal Cortex, Orbital Frontal Cortex, Occipital and Temporal Regions.

1.2 Trustworthiness, amygdala and eye gaze

As above mentioned, the amygdala is one of the brain regions with a relevant role when evaluating novel stimuli, especially when developing first affective impressions [26, 35, 81]. Several studies revealed the recruitment of this structure (see Figure 3), particularly when evaluating traits like the trustworthiness of faces [2, 9, 10].

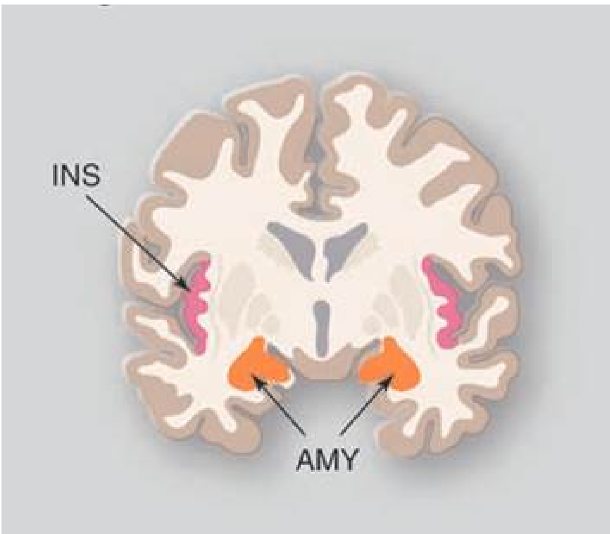


Figure 3: Localization of amygdala and insula, some of the brain areas commonly activated in social decision-making studies. (Adapted from Sanfey, A. G. (2007) [1]).

The face is one of the pivotal sources of the information collected during social interactions, allowing the formation of judgements and subsequent decisions [17]. These judgements and decisions have also been found to predict real world reward outcomes [81].

Accordingly, decision-making consists in choosing a certain option while taking into account all the alternatives available. Another important aspect of decision-making within the neuroeconomic context is related with the existence of rewards expectations [33]. In this field, areas such as the ventral striatum, nucleus accumbens, insula and amygdala are usually reported as having a major role [34-36]. Also, it is known that decision-making is highly influenced by others behaviour in our social interactions [1,2]. Thus, mathematical decision models are frequently applied in order to investigate the psychological and neuronal correlates concerning our social decisions [1]. One of the tasks currently used to evaluate particularly the amygdala role in decision-making is the Trust Game (TG), while participants are developing and expressing trust [17, 20]. The TG is a multi-round economic game played with two anonymous individuals that will interact with each other. One, the investor or trustor, has a certain amount of money and will choose how much to keep for himself and how much to transfer to the partner, also referred to as trustee [1, 17]. After deciding the value - that can be the entire endowment, only a part, or nothing at all - the money transferred by the investor will be multiplied by a previous determined factor (usually 3 or 4), and then given to the trustee [7, 17]. The trustee has then two options: to return some of the multiplied amount to the investor or to keep the whole amount to himself. By transferring back part of the value, the trustee is honouring the trust placed by the investor, and so, both players will finish the game with a higher monetary pay-off comparing to the initial value. Nevertheless, if the trustee decides not to share the money with the investor, abusing thus his trust, the investor will end up the game with less money than initially, being the trustee the only one with a higher profit [1, 17].

Importantly, people do not always play in a rational manner and for this reason sometimes do not take decisions in accordance to the predicted [6, 69].

According to the Game Theory, a rational and selfish trustee will betray the trust given by the investor [37]. Becoming aware of this, the investor will not trust his partner, and because of this fact he/she will not invest in the beginning of the game. However, regardless of these predictions, a great number of investors end up sharing some of the initial value with the trustee, being this trust honoured in most of the cases [6].

Which are then the motives that influence decisions and their outcomes? How do gaze direction, facial expression or other facial characteristic signal influence the perception of others trustworthiness?

One of the most relevant cues by which people understand others' minds and determine their intentions is the direction of eye gaze (see Figure 4), since it is known to influence feelings, intentions and is a signature of personality traits [38]. In general, it does shape our social perception and evaluation of others [39].

Studies show that directing gaze to others, instead of looking away from them, makes people to e perceived as more persuasive [40]. Also, subjects who make direct eye contact are in general considered more trustworthy and attractive than those who avoid eye contact [41, 42]. However, the direction of eye gaze is not the only thing that matters. In order to interpret gaze behaviour, there are other parameters that need to be taken in consideration, like gaze duration [39, 43, 44]. Some studies demonstrated that the longer a person looked into the observer's eyes, more favourable would their opinion of that person be [39]. Nevertheless, other researches also concluded that when direct gaze is prolonged, it can be interpreted as a threat [42, 44].

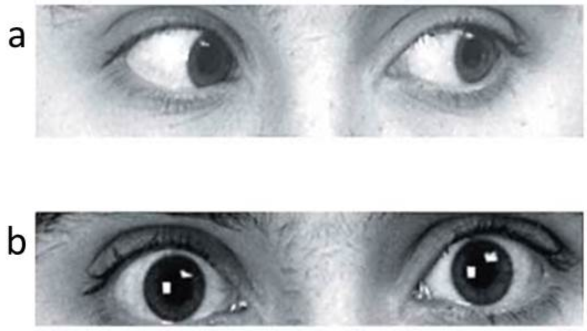


Figure 4: Example of types of gaze: averted (a) and direct gaze(b). (Adapted from Adolphs, R. (2003) [18]).

However, when the opposite happens, i.e., when a person avoids the direct gaze of someone else, it can be seen as a sign of embarrassment, sorrow or disgust. In fact, most of the times, averted gaze will be associated with negative qualities like deception or even as a demonstration of dominance [40, 44].

Several neuroimaging studies demonstrated that the processing of direct and averted gaze recruit distinct neural regions [45]. While direct gaze has been shown to increase activation in neural regions such as the amygdala and the fusiform gyrus [45-48], averted gaze caused a greater activation of areas such as the superior temporal sulcus (STS) and intraparietal sulcus (IPS) [45]. Other researchers point the medial prefrontal cortex (MPF) as processing both direct and averted gaze, although the latter is primarily involved when observing averted gaze [45, 49].

Therefore, we hypothesized if eye gaze could be one of the mechanisms through which the amygdala is modulated, by interacting with others trustworthiness appearance. Some of the questions this study attempts to answer are therefore:

(a) How does the amygdala process direction and duration of eye gaze?

(b) How do these eye gaze features modify decision-making within the TG context?

However, before the task in the scanner, another game, called the Ultimatum Game (UG), was performed, in order to establish a "social reputation" for the trustee. UG is a popular economic experiment, commonly used to evaluate the effect of perceived unfairness in experimental settings that involve economic decisions. In this game, participants react emotionally to proposals, especially if they are unfair, leading them to reject or accept those offers [5]. Thus, before playing the TG, participants will have created a sense of fairness for each trustee.

In sum, in order to have a better understanding of the neural mechanisms underlying the relation between different types and duration of eye gaze and the amygdala within the context of a social interaction, this study was conducted using functional Magnetic Resonance Imaging (fMRI) while participants were interacting in a TG with a trustee with a preestablished reputation.

2 Methods

2.1 Theoretical considerations regarding neuroimaging methods

The brain is considered by many as the most complex organ in the human body. Despite many years of studying and developing techniques which allow to have a better understanding about how it works, its functioning remains in many aspects poorly understood.

Early on, the brain was considered like a black box: somehow the information was transferred from external to the internal environment, but the type of processing involved and its underpinnings were unknown. Also, there were not many technical possibilities available so that scientists could understand the structure and function of the human brain. One of the existing possibilities was by means of leasion studies and by performing *post-mortem* dissections of parts of the brain of healthy individuals and of individuals with specific functional deficits. By comparing both, they were able to make inferences about structure-function correlations. However, and despite all the information that was obtained through these type of methods, they were quite restricted in scope.

Nevertheless, in the last decades, there have been great advances in science, namely through imaging techniques, especially in which concerns to brain research. These technologies have allowed scientists to observe and define some of the brain's structural and functional systems, and even at a neurochemical level in a non invasive manner [50].

2.1.1 Functional Magnetic Resonance Imaging

Over the past few decades several human brain mapping methods have been developed. Among the mapping techniques that have emerged since then, there are two basic classes: those allowing for mapping or localizing brain's underlying electrical activity and those that map local structural, physiological or metabolic changes in brain's electrical activity. Non-invasive neural electromagnetic techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) are included in the first category. These procedures allow a temporal resolution in the order of 10-100 milliseconds, however, with a low spatial resolution [51]. Structural, spectroscopic and functional Magnetic Resonance Imaging (fMRI) are methods included in the second category. Th latter neuroimaging procedure extends the use of MRI technology, providing not only anatomical information, but also information about biological function [55].

Functional Magnetic Resonance Imaging is an imaging technique that allows the study of biological tissue *in vivo* using strong magnetic fields and gradients [53]. This non-invasive imaging technique is able not only to measure but also to localize specific functions of the human brain without the need of injecting radioactive isotopes into the object of study for signal generation, such as in Positron Emission Tomography (PET) and Single-Photon Computed Tomography (SPECT) [54]. FMRI methods are able to detect neuronal activity through changes in regional blood perfusion, volume or even blood oxygenation [55]. Blood oxygen level dependent (BOLD) technique, which will be later explained, uses blood as an intrinsic contrast [52] [56] [57], allowing the generation of functional images from the correspondence between BOLD-signals and neuronal activity [58]. Considering that it can be limited by the spatiotemporal range of hemodynamic response, fMRI allows having images with a spatial resolution of the order of a few millimetres and a temporal resolution of a few seconds [55].

BOLD effect

The BOLD signal is one of the most commonly used imaging modality to generate images in functional MRI (fMRI) studies, relying on regional differences in cerebral blood flow. The origin of the BOLD fMRI signal change is related to the state of oxygenation of the hemoglobin [52]. Depending on the concentration of oxygen, this molecule will present different magnetic properties. When oxygenated, hemoglobin (Hb) behaves as diamagnetic substance, while deoxygenated hemoglobin is paramagnetic [53], changing the local magnetic susceptibility, which leads to the creation of magnetic field distortions [58]. Vessels that contain oxygenated arterial blood will, consequently, cause little or no distortion to the magnetic field in the surrounding tissue, while capillaries and veins containing blood partially deoxygenated will cause distortions in the magnetic field of the surrounding tissue [52] [59].

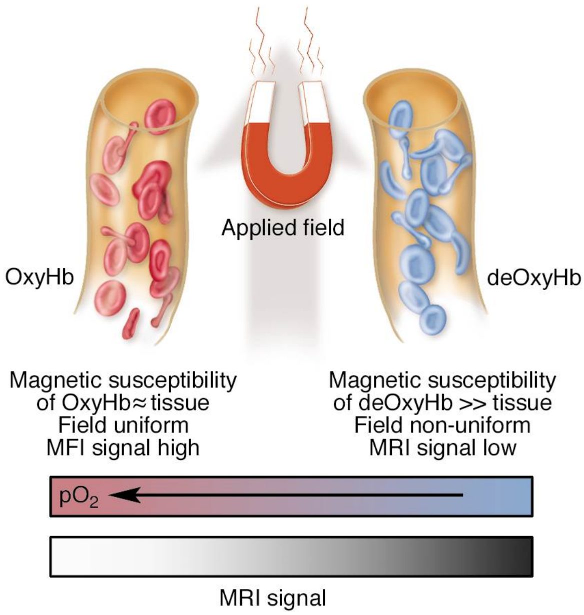


Figure 5: Illustration of the origin of the BOLD effect. Arterial blood has similar magnetic properties to tissue, while deoxygenated blood is paramagnetic, inducing inhomogeneities within the magnetic field in tissue. (Adapted from Gore, J. C. (2003) [60])

The increase of the proportion of oxyhemoglobin relative to deoxyhemoglobin during the hemodynamic response will result in a more homogeneous local magnetic field. This will be reflected in the obtained BOLD image, since areas with higher concentration of oxyhemoglobin will present a brighter image (due to higher signal) than areas with low concentration[51]. Thus, the change in the ratio of local oxyhemoglobin/deoxyhemoglobin and in its magnetic field homogeneity will act as an endogenous marker of neural activity [58].

The hemodynamic response corresponds to the change in the MR signal that results from neuronal activity, considering that its shape can vary according to the properties of the evoking stimulus [53]. This response time course is well-studied, and can be represented by the waveform of Figure 6.

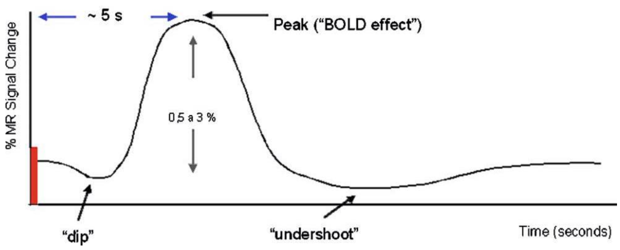


Figure 6: Example of hemodynamic response function (HRF) from a short duration stimulus. (Adapted from Amaro Jr, E., Barker, G.J. (2006)) [51])

The stimulus processing in a certain brain region will be accompanied by a transient increase in deoxyhemoglobin concentration, causing a significant decrease in the BOLD signal, known as the initial dip [61]. After this period, the oxy/deoxyhemglobin ratio will increase, leading to a higher MR signal, which reflects the increase of neural activity [62]. If the stimulus is maintained for a sufficient time, it can even reach a plateau. When the stimulus ceases, the BOLD signal starts to decrease, and eventually underpasses even the original baseline, being this phenomena known as the undershoot effect [56] [63].

Therefore, while a subject is performing a specific task, the increase in neuronal activity can be indirectly detected, by using the BOLD effect.

2.1.2 Experimental designs for fMRI

Considering that fMRI does not measure absolute neural activity [55], neuroimaging studies have to be designed in order to quantify relative changes of activity. There are several presentation schemes that can be applied when designing a fMRI experiment, and the most commonly applied ones are the block and event-related designs. These two types of design will be briefly hereafter described, along with the mixed design.

Block Design

This type of paradigm consists in maintaining cognitive engagement in a task during relatively long periods, alternating with other periods, when different conditions are presented [51] (see Figure 7). By doing so, variations due to scanner sensitivity and others, such as patient movements or changes in their attention, may have a similar impact on the signal responses when comparing both states [55]. Nevertheless, for relatively long periods, it can be difficult to control cognitive states. Also, sometimes this type of design may not be appropriate, for example when trials depend on subject's performance or when trials inherently need to be presented in a non-blocked fashion [55].

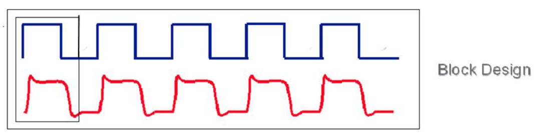


Figure 7: Block design: for the same condition, stimulus are presented sequentially, alternating with periods or blocks when different conditions are presented. (Adapted from Amaro Jr, E., Barker, G.J. (2006)) [51])

Event-related Design

In event-related designs, stimuli are presented as individual events or trials (see Figure 8). There are two main differences when comparing this type of design to the block design: conditions are presented in a randomized order, and the time between stimulus presentation (interstimulus interval - ISI) can also vary, which allows to better maintain the attention level throughout the experience. This happens because the subject's ability to predict when and what will happen will be reduced, avoiding cognitive adaptation strategies from the subjects [51]. Another advantage of event-related designs is the analysis of individual responses to trials, allowing the analysis of neural correlates of individual behavioural responses, such as errors when performing challenging paradigms [64].

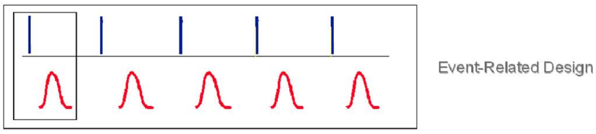


Figure 8: Event-related-design: conditions are presented as individual events or trials, in a randomized order, with possibility of manipulation of the inter-stimulus interval (ISI). (Adapted from Amaro Jr, E., Barker, G.J. (2006)) [51])

Event-related designs with long temporal intervals between individual trials are named slow event-related designs, being 12 seconds the optimal intertrial interval (ITI) for stimuli of 1-2 seconds. In this type of event-related design, after each stimulus, the hemodynamic response decays to baseline, allowing the individualization of the response to each trial. Faster designs, with shorter temporal intervals between trials, are called rapid event-related designs. For this type of design, special analysis procedures are needed, in order to separate (deconvolute) the hemodynamic responses of different events [53], since they are overlapped due to the closely spaced trials. The isolation of condition-specific event-related time courses can be done using deconvolution analysis. To perform a deconvolution analysis, trials are required to have randomized intertrial intervals ("jitter"), which can be accomplished by adding "null" (baseline) trials [58].

Mixed designs

Nearly all fMRI studies are performed using one of the two design categories above presented. However, mixed design is one of the alternatives, combining features from both types, since it uses both repetitive sets of stimuli presented under the form of a block design, and transient responses detected by event-related approaches [51] (see Figure 9).

This design is then able to provide information that allows the concomitant analysis of the sustained and more transient neural activity while performing a paradigm [65].

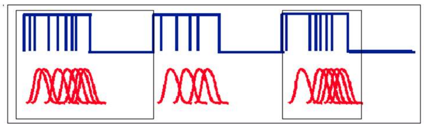


Figure 9: Mixed design: combination of stimuli closely presented, alternating with control condition. (Adapted from Amaro Jr, E., Barker, G.J. (2006)) [51])

2.2 The General Linear Model

The General Linear Model (GLM) is mathematically identical to a multiple regression analysis, suited to implement parametric statistical tests with one dependent variable. It can include factorial ANOVA designs and designs presenting multiple qualitative and quantitative variables. Due to the possibility of incorporating qualitative and quantitative independent variables, the GLM is one of the most commonly applied tools for fMRI data analysis.

When considering a multiple regression analysis, the GLM attempts to predict the variation of a dependent variable in terms of a linear combination of several reference functions. The dependent variable is relative to the observed fMRI time course of a voxel, being the reference functions relative to time courses of the expected fMRI response for the different conditions present in the experimental paradigm. These reference functions can also be designated as predictors, regressors, explanatory variables, covariates or basis functions. The design matrix will then be formed by a set of specified predictors, also known as model. To obtain a predictor time course, a convolution of a condition box-car time course with a standard hemodynamic response function is usually performed. One way of defining a condition box-car time course is by setting to 1 time points at which the modelled condition is defined ("on"), setting to 0 all other time points. After this, each predictor time course will then get an associated coefficient or beta weight b that quantifies its potential contribution in explaining the voxel time course y. This voxel time course y results from the sum of the defined predictors, being each one multiplied by its associated beta weight b. However, due to noise fluctuations, this linear combination does not perfectly explain the data. Because of this, an error value e is added to the GLM equations presenting n data points and p predictors:

$$y_1 = b_0 + b_1X_{11} + ... + b_pX_{1p} + e_1$$
$$y_2 = b_0 + b_1X_{21} + ... + b_pX_{2p} + e_2$$
$$y_3 = b_0 + b_1X_{31} + ... + b_pX_{3p} + e_3$$
$$...$$
$$y_n = b_0 + b_1X_{n1} + ... + b_pX_{np} + e_n$$

The left side variable y corresponds to the data, more specifically to the measured time course of a single voxel. In the left column y1 corresponds to the measured value at time point 1, y2 to the measured value at time point 2, and so on. The first column presented in the right side corresponds to the first beta weight b_0. Its corresponding predictor time course X_0 is constant, with a value of 1 for each time point. It does not explicitly appears in the equation, since multiplication with 1 does not alter the value of b_0. Also, the value of b_0 typically represents the signal level of the baseline condition, and despite its value being not informative, it is important the inclusion of the constant predictor in a design matrix. This inclusion allows other predictors to model small fluctuations condition-related. Predictors on the right side are responsible for the modulation of the expected time courses of different conditions.

Regarding multi-factorial designs, predictors can be defined by combining condition levels so that main and interaction effects can be estimated. The beta weight of each condition predictor, as above mentioned, allows the quantification of the contribution of its time course in explaining the voxel time course. A considerable part of positive/negative beta weights typically indicate the stronger activation/deactivation exhibited by voxels during while an experimental condition is modelled relative to baseline. The last column presents error values, also known as residuals, prediction errors or noise. These values quantify the deviation that occurs between measured voxel time course and the predicted time course [66].

2.3 The current project

2.3.1 Pilot Study

Considering that the purpose of this study was to understand the processing of faces in social interactions, it was clear that the stimuli selected needed to include faces. Since the goal of the study was to evaluate the influence of direction and duration of eye gaze in social interactions, it was decided that dynamic stimuli (e.g., videos) would be used instead of static pictures, in clear contrast with the majority of studies that evaluate trustworthiness from pictures of faces [2] [67] [19].

2.3.1.1 Participants

Fourteen healthy individuals, 6 males and 8 females, right-handed, with a mean age of 25.79 (SD=5.85), having a mean of 15.86 (SD=2.93) years of education participated in the pilot study.

2.3.1.2 Videos recording and Pilot study

For the creation of a set of videos with the desired characteristics, namely neutral faces [26] [19] with variations in duration and direction of eye gaze, the recording of several eye gaze combination videos was conducted.

Five male identities (S01, S02, S03, S04 and S05) participated in the recordings. These were done in the same room, in the same day period (afternoon), using the same artificial illumination, and controlling the distance individual face - camera. Also, all individuals were wearing a black t-shirt, since it is known that clothing style can impact the formation of first impressions [68]. One of the individuals (S03) had to be excluded due to facial characteristics which could result as a distraction factor.

Recordings were done using a Panasonic HDC-SD707 Video camera with a tripod and videos were edited with Adobe Premiere Pro CS5.5 setting the duration of each video to 4 seconds.

After of the videos edition, a pilot study was performed for stimulus selection using Presentation software (Neurobehavioral Systems, USA, www.neurobs.com). A total set of 70 videos was presented in 4 runs with the following number for each fixation-saccade combination: averted N=17, averted-directed (see definition below) N=14, directed N=24, directed-averted N=15. Participants (N=14) rated the videos in terms of Valence (-2 to 2 scale), Arousal (0 to 4 scale), Eye Gaze duration (short, normal or long) and Trustworthiness (untrustworthy, neutral or trustworthy). A trial example of the pilot study can be seen in Figure 10.

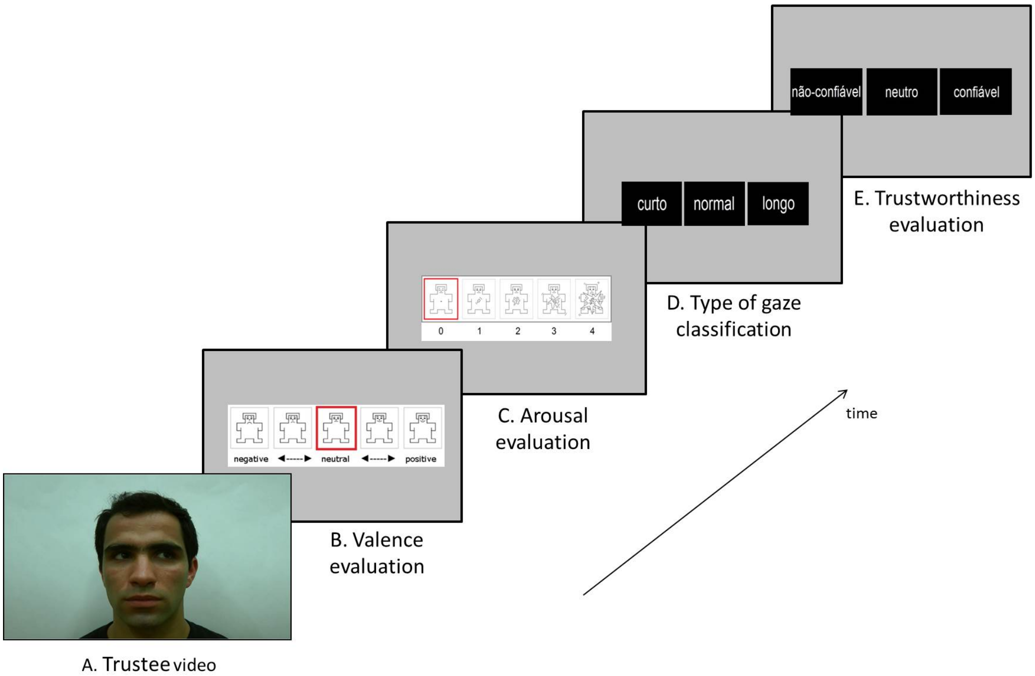


Figure 10: Trial example of the pilot study.

The following tables reflect the results of the pilot study. Ratings concerning the evaluation of Valence, Arousal and Trustworthiness for each individual are presented in Table 4. Given the core role of Trustworthiness in our study, videos were selected based on these ratings alone, and therefore statistics are only presented for this factor.

Table 4: Pilot study: Evaluation of each video (individual face) in terms of Arousal, Valence and Trustworthiness.

Individual	Valence	Arousal	Trustworthiness
S01	-0.43 (0.765), N=242	0.74 (0,850), N=242	-0.22 (0.698), N=242
S02	-0.06 (0.808), N=269	0.52 (0.678), N=269	0.08 (0.726), N=269
S04	-0.08 (0.825), N=212	0.68 (0.826), N=212	0.04 (0.756), N=212
S05	-0.43 (0.794), N=214	0.74 (0.859), N=214	-0.29 (0.717), N=214

Valence: -2 a 2, Arousal: 0 a 4, trust: -1=untrust., 0=neutral, 1=trust. Values display M(SD).

According to Table 4, all individuals presented near zero values regarding Valence evaluation, with individual S01 and S05 presenting the lowest value and S05 the highest. As to Arousal ratings, none of the four individuals presented values superior to 1. Individual S02 presented the lowest value for Arousal evaluation, with individuals S01 and S05 presenting the highest. Concerning the Trustworthiness evaluation, S05 was rated as the most untrustworthy individual and S02 as the most trustworthy individual (main effect: $F(3,39)=8.756$, $p=0.000$; post-hoc contrasts: S02 > S05, $t(13)=4.231$, $p=.001$; S02 > S01,

$t(13)$=3.147, p=0.008; S04 > S05, $t(13)$=3.358, p=0.005; corrected for multiple comparisons).

Different types of Eye Gaze, namely directed (Dir.), averted (Av.) and combinations alternating between directed and averted gaze (Dir. Av or Av. Dir. Av.) were also one of the evaluated features, with results being presented in Table 5. Videos labelled as "directed" referred to videos with total presentation (4sec) of directed gaze, while when presenting averted gaze, the trustee would never direct his gaze to the camera. Also, videos presenting combinations of Av. Dir. Av., had approximately 4 seconds of averted gaze, intercalated with a direct fixation at the 2 seconds point. The Dir. Av. combination presented 2 seconds of averted, followed by 2 seconds of directed gaze.

Table 5: Pilot study: Evaluation of type of eye gaze in terms of Arousal, Valence and Trustworthiness.

Eye gaze type	Valence	Arousal	Trust
Directed	-0.10 (0.890), N=295	0.74 (0.850), N=242	0.14 (0.722), N=295
Averted	-0.44 (0.797), N=238	0.63 (0.789), N=238	-0.30 (0.682), N=238
Av. Dir. Av.	-0.41 (0.757), N=195	0.89 (0.842), N=195	-0.26 (0.764), N=195
Dir. Av.	-0.06 (0.698), N=209	0.60 (0.772), N=209	-0.01 (0.704), N=209

Valence: -2 a 2, Arousal: 0 a 4, trust: -1=untrust., 0=neutral, 1=trust. Values display M(SD).

As seen in Table 5, Dir. Av. combination showed the highest value for valence, with averted gaze showing the lowest value. As to arousal evaluations, Dir. Av. combination resulted in the lowest value, with Av. Dir. Av. combination presenting the highest. Trustworthiness judgements revealed Directed eye gaze as the more trustworthy eye gaze type and Averted gaze as the most untrustworthy (main effect: F(1.714,22.280)=4.357, p=0.005; post-hoc contrasts: Directed > Averted, $t(13)$=3.442, p=0.004; Dir. Av. > Averted, $t(13)$=3.161, p=0.008; corrected for multiple comparisons).

Interestingly,differences for the side where eye gaze was directed at were found as function of Valence, Arousal and Trustworthiness judgements. When subjects presented averted eye gaze, namely when directing their eye gaze to the left side, results regarding Valence and Trustworthiness judgements presented lower values and Arousal higher values, comparing to when directing their gaze to the right side (see Table 6). Trustworthiness judgements reflected these differences (main effect: F(1.319,17.150)=8.168, p=0.007; post-hoc contrasts: Directed > left, $t(13)$=3.217, p=0.007; corrected for multiple comparisons).

Table 6: Pilot study: Evaluation of the eye gaze direction of the Trustee (directed, or averted - to left or to right visual fields) in terms of Arousal, Valence and Trustworthiness.

Eye Gaze direction	Valence	Arousal	Trustworthiness
Directed (frontal)	-0.10 (0.890), N=295	0.59 (0.794), N=295	0.14 (0.722), N=295
Left	-0.37 (0.751), N=322	0.72 (0.810), N=322	-0.25 (0.711), N=322
Right	-0.24 (0.789), N=320	0.68 (0.807), N=320	-0.14 (0.736), N=320

Valence: -2 a 2, Arousal: 0 a 4, trust: -1=untrust., 0=neutral, 1=trust. Values display M(SD).

Also, as shown in Table 7, participants rated the different types of eye gaze regarding its duration. When directed eye gaze was presented, it was rated as of normal or long duration. When only presenting averted eye gaze, it was rated for all participants as having a short duration. When alternating between averted, directed and averted gaze, eye gaze was also rated as having a short duration. However, when alternating only between directed and averted gaze, this combination was rated as of normal duration.

Table 7: Pilot study: Evaluation of the eye gaze duration, depending on its type.

Eye gaze type	Eye gaze duration	Label
Directed	0.360 (0.622), N=295	Normal - Long
Averted	-0.100 (0.000)(*), N=238	Short
Av. Dir. Av.	-0.560 (0.634), N=195	Short
Dir. Av.	-0.180 (0.659), N=209	Normal

Eye gaze duration: -1=short, 0=normal, 1=long; Values display M(SD).
*all videos presenting averted gaze were considered as of short duration.

2.3.1.3 Videos selection

After analysing the results regarding the pilot study, the final selection of the videos for the stimuli was done according to the judgements of Trustworthiness, since this is the trait in which this study is focused on.

Two individuals were selected: one considered as the most trustworthy and the other as the most untrustworthy. According to the results shown in Table 4, participant number 2 (S02) was considered the most trustworthy comparing to other individuals. Therefore, from this moment on, S02 will be labelled as the *Trustworthy* trustee (TTrust). Taking this into account, two videos presenting higher values of trustworthiness as well as other two rated as the most untrustworthy from the same individual were selected (see Table 8).

Table 8: Pilot study: Videos selection of the Trustworthy trustee presenting higher (videos 1, 2) and lower (videos 3, 4) values of trustworthiness.

Trustee	Video nr	Eye gaze direction	Trust(mean)	Trust(SD)
Trustworthy	1	Directed	0.75	0.452
	2	Directed	0.58	0.515
	3	Averted	-0.36	0.745
	4	Averted	-0.43	0.646

Trust: -1=untrust., 0=neutral, 1=trust.

Participant number 5 (S05) was rated as the most untrustworthy individual (see Table 4), and therefore, will from now on be labeled as the *Untrustworthy* trustee (TUntrust). Similarly, two videos presenting higher values of trustworthiness and other two rated as the most untrustworthy were selected (see Table 9).

Table 9: Pilot study: Videos selection of the Untrustworthy trustee presenting higher (videos 1, 2) and lower (videos 3, 4) values of trustworthiness.

Trustee	Video nr	Eye gaze direction	Trust(mean)	Trust(SD)
Untrustworthy	1	Directed	0.25	0.866
	2	Directed	0.08	0.669
	3	Averted	-0.50	0.76
	4	Averted	-0.64	0.497

Trust: -1=untrust., 0=neutral, 1=trust.

2.3.2 Main Study

After stimulus selection, two experimental tasks were designed to test social neuroeconomic interactions with both trustees.

The first task was designed to mimic a social interaction in which first impressions are formed, namely, a reputation associated to each trustee. For this, a Ultimatum Game (UG) was used and performed outside the scanner. The second task referred directly to the Trust Game (TG) and was performed inside the scanner for neuroimaging data collection evaluating decision-making based on trustworthiness judgements.

2.3.2.1 Participants

Six healthy individuals, 2 males and 4 females, right-handed, with a mean age of 25.83 (SD=6.24), and a mean of 15.50 (SD=1.52) years of education participated in the main study. However, results regarding two of the participants (1,5) were not considered for fMRI data as well as for statistical analysis of behavioural data colleted after the scanning session. Participant number 1 results were excluded due to methodological issues. Participant number 5 was unable to complete the task due to excessive tiredness.

Participants were given the informed consent, according to the Declaration of Helsinki.

Before performing the tasks, participants were asked to fill in some scales, namely the Geschwind-Oldfield Questionnaire to evaluate participants handedness and the Rotter Trust scale (1967), with a mean of 70(SD=9.50) points.

The following sections describe the exact procedures and task designs.

2.3.2.2 Ultimatum Game experimental design

2.3.2.2.1 Stimuli and apparatus

Videos of faces (untrustworthy and trustworthy trustees) were used as stimuli. Videos presented in this task were the videos selected from the pilot study.

The task was presented using the Psychotoolbox from MatLab R2014a (Neurobehavioral SystemsMathWorks, USA, http://www.mathworks.com/). Responses were given by pressing the right or left arrow of the keyboard.

2.3.2.2.2 Task design and procedure

In UG, two players must divide a sum of money, being the offer made by the proposer. The responder is then able to accept or reject the proposal. If the responder accepts, the amount of money will be divided as initially proposed. However, if the responder rejects, neither player receives anything. If people were entirely rational, they should accept even the smallest positive offer, since the alternative is getting nothing [69]. Knowing this, the proposer could offer the smallest amount possible. Nevertheless, and despite these predictions, studies show that proposals below 30% or 20% will be rejected most of the time [69][1]. Thus, the UG here designed will present several types of proposals, some considered as fair and others considered as unfair.

The UG task consisted in 48 trials, in which half of them were played with the trustee considered Trustworthy (TTrust), and the other half with the trustee rated as Untrustworthy (TUntrust).

Each trial started with a fixation cross (1.5 s), followed by a video (4 s) of one of the trustees. After this, a screen would appear informing about the amount to be divided (4 s).

Thereafter, a screen with a division proposal was shown. With the proposal on the screen, participants could press a key, accepting or rejecting the offer. The last screen presented the result, according to the chosen option of the participant (see trial example in Figure 11).

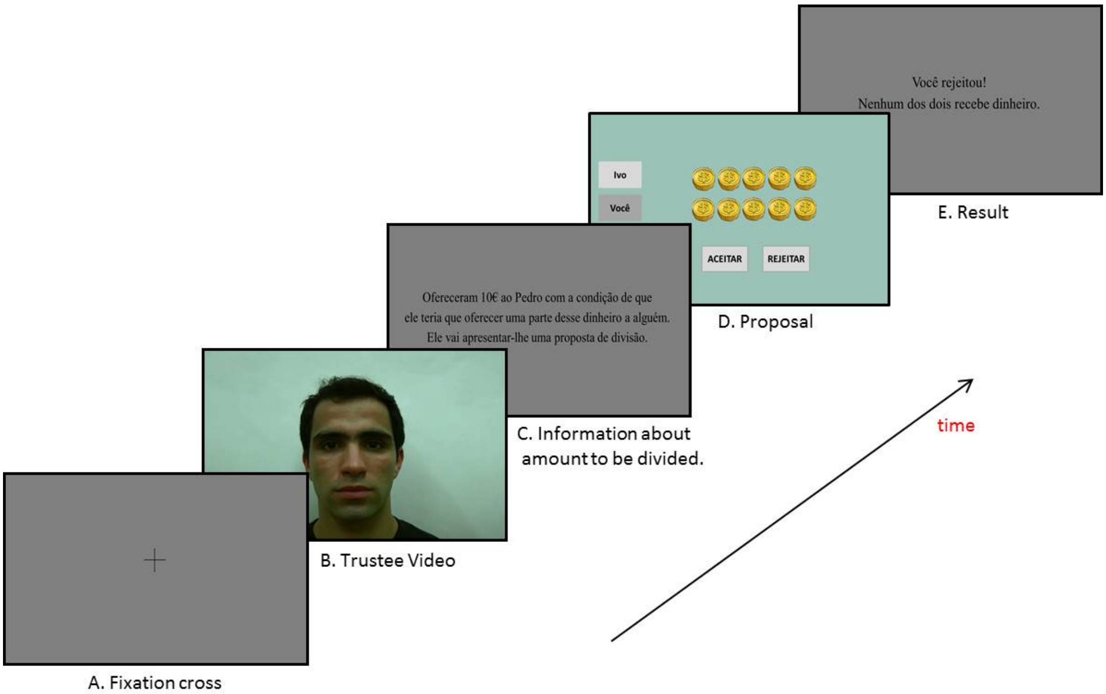


Figure 11: Trial example of the Ultimatum Game.

Trials were randomly presented, and total amounts to be divided were 10€, 7.5€, 2.5€ or 0.2€. Fair proposals consisted in a model offer of 60/40% and 50/50%, while the unfair ones presented division proposals of 80/20% or 90/10%, being the higher value always for the proposer.

For the TTrust, 16 of his 24 proposals were fair, while the other 8 were unfair (see Table 10).

Table 10: UG experimental design: type and number of trials for each proposal regarding the Trustworthy trustee.

Trustee	Total Nr of Trials	Proposal(%)	Nr Trials x Amount
	8	60/40	2 x 10€, 2 x 7.5€, 2 x 2.5€, 2 x 0.2€
Trustworthy	8	50/50	2 x 10€, 2 x 7.5€, 2 x 2.5€, 2 x 0.2€
	4	80/20	1 x 10€, 1 x 7.5€, 1 x 2.5€, 1 x 0.2€
	4	90/10	1 x 10€, 1 x 7.5€, 1 x 2.5€, 1 x 0.2€

Regarding the TUntrust, the opposite would happen: 8 of his 24 proposals were fair, while the remaining 16 were unfair (see Table 11).

Table 11: UG experimental design: type and number of trials for each proposal regarding the Untrustworthy trustee.

Trustee	Total Nr of Trials	Proposal(%)	Nr Trials x Amount
	4	60/40	2 x 10€, 2 x 7.5€, 2 x 2.5€, 2 x 0.2€
Untrustworthy	4	50/50	2 x 10€, 2 x 7.5€, 2 x 2.5€, 2 x 0.2€
	8	80/20	1 x 10€, 1 x 7.5€, 1 x 2.5€, 1 x 0.2€
	8	90/10	1 x 10€, 1 x 7.5€, 1 x 2.5€, 1 x 0.2€

To notice that not only the number of trials for each proposal and amount of money were variable. Another goal of this study, like previously reported, was to evaluate the role of direction and duration of eye gaze while performing these neuroeconomic games. In order to do so, and according to the results previously shown, direct eye gaze was associated with a trustworthy behaviour, and averted eye gaze with an untrustworthy behaviour. So, for videos displaying a trustee with direct eye gaze, a fair proposal (50/50% or 60/40%) would appear on the screen. The exact opposite would happen when averted gaze was shown in the videos, i.e., the presented offer would be unfair (80/20% or 90/10%).

The duration of the task was about 15 minutes, and after performing it, participants were asked to rate each of the 8 videos in a scale of -3 to 3, doing fairness judgements of the trustee. They were also asked to rate each global trustee behaviour concerning its fairness in a scale of -3 to 3. This was done so that when each participant started the Trust Game, he had already created a first impression of the fairness or unfairness of each trustee. It was also expected that while performing the UG task, participants had learnt the influence of directed gaze (associated with fairness) and averted gaze (related to unfairness).

2.3.2.3 Trust Game experimental design

2.3.2.3.1 Stimuli and apparatus

Videos of faces (untrustworthy and trustworthy trustees) and one picture of a computer (PC) were used as stimuli. Videos presented were the same used in the the UG task. The picture of the computer was selected from the internet, and transformed in order to remove any reference to companies or trademarks.

Inside the scanner, the stimuli were back projected using an AVOTEC (www.avotec.org) projector on a 20(w) x 15(h) (1024 x 768 pixels) screen pad that was placed at a viewing distance of 50.5 cm by means of a head coil mounted mirror. The task was presented using the Psychtoolbox from MatLab R2014a (Neurobehavioral SystemsMathWorks, USA, http://www.mathworks.com/), and originally displayed on a monitor with a 60Hz refresh rate. Responses were given in a response box (Cedrus Lumina LP-400 response pad for fMRI, www.cedrus.com). In the end, a fMRI localizer task presenting videos of faces, bodies and objects was performed.

2.3.2.3.2 Task design and procedure

The Trust Game was the chosen game to evaluate parameters that influence trust in economic transactions. In this game, a certain amount of money was given to the participant. The participant had to decide how much he wanted to invest (0.02€, 0.10€ or 0.20€) and send to the trustee, knowing that by doing so, the value would be tripled. Then, the trustee decided how much he wanted to return to the participant.

The TG task had two types of trials. The most frequent one started with a fixation cross (4s) followed by a video of one of the trustees or the picture of the computer. After the video presentation, an inter-stimulus interval (ISI) was shown (2s). Next, a screen asking how much the participant wanted to invest appeared, showing the three investment options (4s). Other ISI was shown (4 or 6s), followed by a screen informing the participant of how much he had won in that trial. Finally another ISI (2 or 4 s) was presented. The other trial type was very similar, however, an extra screen regarding an evaluation was presented between the video and the first ISI (see Figure 12). Here, participants were asked to evaluate the Trustee or Computer behaviour, asking how much did they thought the Trustee/PC would return in the end of that trial. Two options were presented, namely the minimum or the maximum value possible. The duration of this evaluation period corresponded to participants response time (RT). These evaluation trials only appeared 6 time per run: in the middle and at the last trial of each Trustee/PC performance, i.e., twice for each.

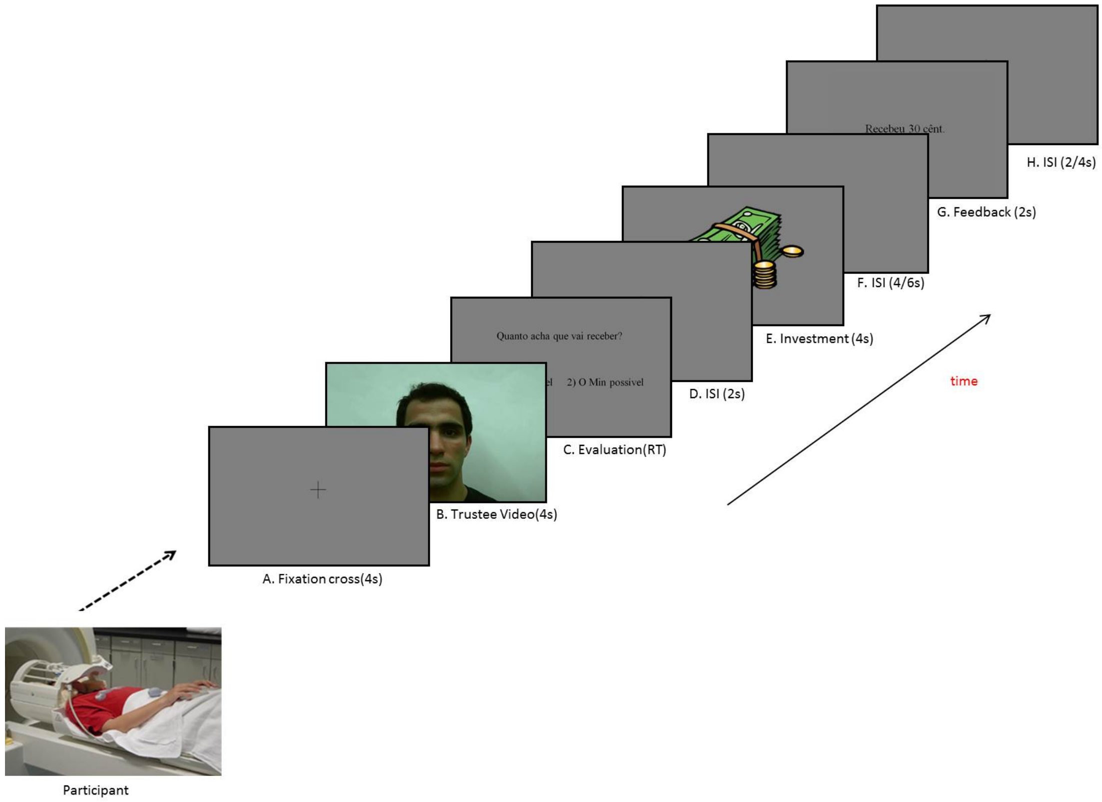


Figure 12: Example of a Trust Game trial with evaluation.

The TG task consisted in 8 runs. Runs 1 to 4 were congruent runs, since when Trustee videos presented direct eye gaze, the returned value was the maximum possible, i.e., 3 x 50% participants investment. However, when Trustee videos presented averted gaze, the returned value was the minimum possible, in this case, 25% of participants investment (see Table 12).

Table 12: TG experimental design for congruent run: Number of trials for each trustee, eye gaze type and associated returned value.

Eye Gaze type	Nr of TTrust trials	Nr of TUntrust trials	Nr of PC trials	Returned Value(*)
Directed	4	2	-	3 x 50%
Averted	2	4	-	25%
-	-	-	2	3 x 50%
-	-	-	2	25%

(*) Value returned considering participants investment at the beginning of each trial

For incongruent runs, 5 to 8, the opposite happened. If the Trustee presented direct eye gaze, participants received the minimum value, while when Trustee was not looking directly to participants, they received the maximum value (see Table 13). For both type of runs, the Computer behaviour would be the same: half of the times returned the maximum amount, while the other half returned the minimum amount possible.

Trials were randomly presented, with the following distribution regarding the number of trials for each Trustee/PC:

Table 13: TG experimental design for incongruent run: Number of trials for each trustee, eye gaze type and associated returned value.

Eye Gaze type	Nr of TTrust trials	Nr of TUntrust trials	Nr of PC trials	Returned Value(*)
Directed	2	4	2	25%
Averted	4	2	2	3 x 50%
-	-	-	2	3 x 50%
-	-	-	2	25%

(*) Value returned considering participants investment at the beginning of each trial

Participants were asked to remain as still as possible during the testing session. It was emphasized that this would be important to minimize data artefacts. Participants were also told that each response should be given using the 3 response buttons. Furthermore, participants were informed that they would receive the amount earned during the Trust Game task.

2.3.2.3.3 Imaging data acquisition and preprocessing

Functional images were acquired in a 3T Siemens TimTrio scanner using BOLD contrast echo planar imaging (EPI, TR=2 sec, TE=30 msec, 34x3 mm-thick-slices, in-plane matrix 92 x 92 voxels, 223 volumes) covering the entire brain. The scanning session also included a high resolution T1 weighted anatomical scan (MPRAGE sequence, 1x1x1 mm3 voxel size, TR=2.3 sec, TE=2.98 msec, 160 slices) to help the transformation of the functional images into standard space. Data were preprocessed and analysed using BrainVoyager QX v2.6 (Brain Innovation, www.brainvoyager.com). Preprocessing included slice scan time corrections, temporal filtering and motion correction. Before group analysis, images were spatially smoothed using a 6-mm full-width-half-maximum Gaussian kernel and then transformed into Talairach space.

Eye tracker EyeLink 1000 Plus (SR Research, www.sr-research.com) was also used to measure eye movement.

2.3.2.4 Statistical analyses

All statistical analyses were performed using IBM SPSS Statistics 21 (IBM, USA, http://www.ibm.com/software/analytics/spss/) and Brain Voyager QX v2.6 software.

Pilot Study

Data concerning Trustworthiness ratings were collected and analysed performing ANOVA Repeated Measures tests and paired-samples t-Tests in order to see if there were statistical differences for each Individual and Eye Gaze type and direction. Greenhouse-Geisser corrections were used whenever sphericity assumptions were violated.

Main study: Behavioural data

Behavioural data concerning fairness and trustworthiness judgements were collected and analysed performing ANOVA Repeated Measures tests in order to see if there were statistical differences between different levels of Trustee and Eye gaze.

Data concerning reward expectations were separated in two distinct analysis, one regarding the match between video and reward expectations and the other for the match between reward expectations and investment decisions. For the match between video and reward expectations, a Chi-Squared test was performed to evaluate if there were global differences in match and non-match trial frequencies. Also, another Chi-Squared test was performed to assess if there were differences in match and non-match trial frequencies between the two type of tasks (congruent and incongruent). All runs were considered for these two analyses, however, for the last one, trials regarding the PC performance were excluded. This was made because it was not possible for participants to predict PC behaviour through video/image presentation, no existing, thus, correct expectations. As to data regarding the match between reward expectations and investment decisions, all runs and trials were considered for the analysis, since its purpose was to see if participants' investment was coherent with their previous expectations.

Main study: Functional data

For the analysis regarding main effects of Trustee and Eye gaze, a FFX ANOVA 2x2 was performed for congruent and incongruent task separately. Also, contrasts analysis between TTrust an TUntrust were carried using a FFX general linear model (GLM). The first run of the incongruent task was excluded from this analysis since it was considered a run were participants were learning to deal with the inversion between eye gaze type and its associated returned values.

The analysis concerning reward expectations and decision-making was performed with both tasks, congruent and incongruent, together. For this analysis, FFX-GLM contrasts were carried considering 4 main predictors - Video, Evaluation, Investment and Feedback.

All contrasts were performed using the Cluster Threshold plugin (BrainVoyager). Corrections for multiple comparisons were made using the Cluster Threshold plugin (BrainVoyager) using 1000 Monte Carlo simulations, with minimum cluster sizes corresponding to significance at a threshold of $p < 0.05$ for each contrast.

3 Results

3.1 Behavioural Results

3.1.1 Outside the Scanner: Fairness and Trustworthiness

In this section, results regarding fairness and trustworthiness ratings performed by the participants outside the scanner will be presented.

Fairness judgements

After performing the UG task, participants (N=6) were asked to rate all videos regarding each Trustee fairness in a scale from -3 to 3 (see Figure 13). For detailed informations regarding evaluations at the participant level, please see Table 25 in Appendix A.

According to Figure 13, participants seem to consider TTrust videos more fair than TUntrust videos, however these results do not reach statistical significance (but see global analysis below). Also, for each Trustee, videos presenting direct gaze were rated as being more fair than those presenting averted gaze ($F(1,5)=8.197$, $p=0.035$). These results are similar to the results obtained in the pilot study.

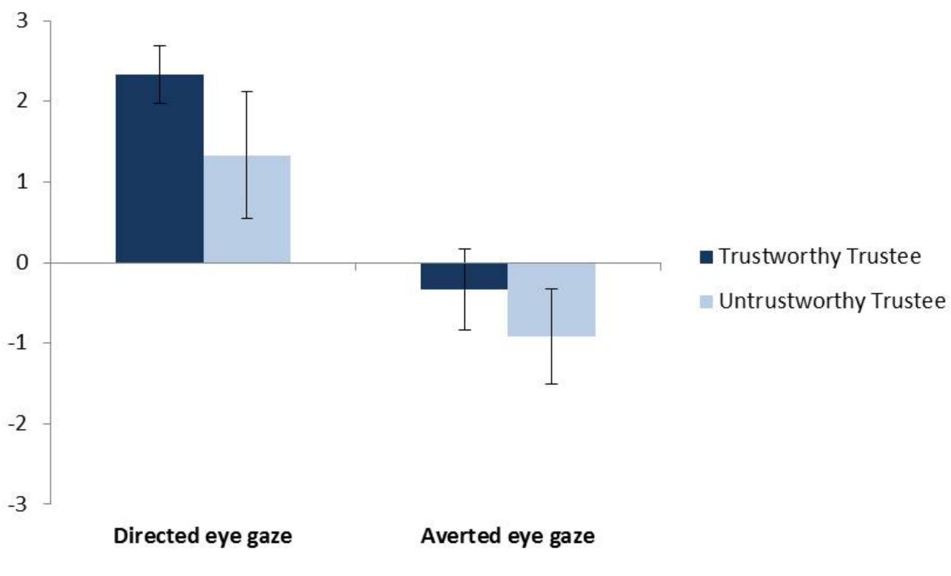


Figure 13: Trustee's fairness judgements considering eye gaze.

Fairness: -3=unfair, 0=neutral, 3=fair.

Participants were also asked to make overall judgements regarding each Trustee fairness. TTrust fairness in average was rated as fair (mean=1.80, SE=0.37) in a scale of -3 to 3, while TUntrust fairness was classified as moderately unfair (mean=-0.20, SE=0.66) (see Figure 14). These results support our predictions and manipulations.

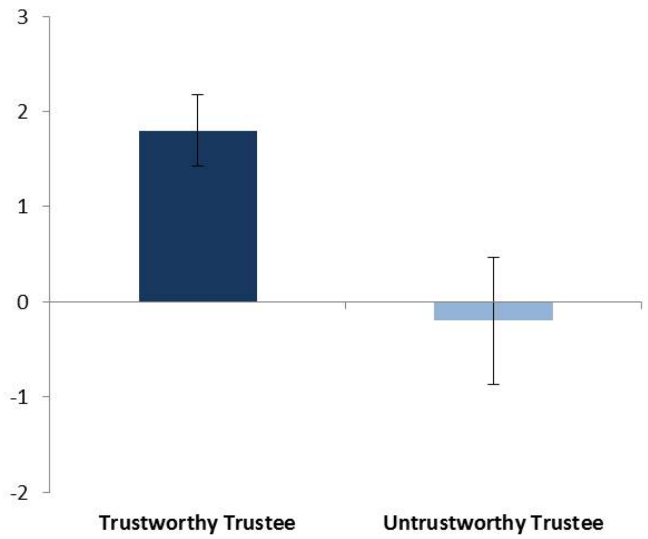


Figure 14: Trustee's overall fairness evaluation.

Fairness: -3=unfair, 0=neutral, 3=fair.

Trustworthiness judgements

After performing the TG task inside the scanner and just after the incongruent task, participants were asked to rate each Video, concerning this time not their fairness, but the Trustee's trustworthiness. Results regarding these evaluations are presented in Figure 15.

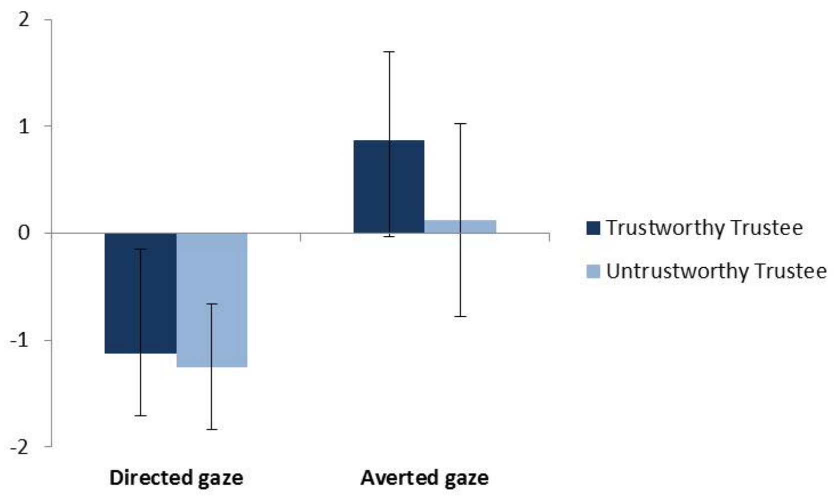


Figure 15: Trustee's trustworthiness judgements considering eye gaze.

Trustworthiness: -3=untrust, 0=neutral, 3=trust.

In trustworthiness judgements, participants have not considered all TTrust videos as trustworthy, neither TUntrust videos as untrustworthy, a clear disagreement with the previous fairness judgements, which might seem surprising at first sight, but may be explained by the fact that it followed the incongruent task (for more detail see Table 26 in Appendix A). Another difference from the judgements previously performed concerns eye gaze direction. Figure 15 shows that participants were able to extinguish the initial association between eye gaze direction and Trustee's behaviour, e.g., directed eye gaze associated to a trustworthy behaviour, after the transition between the congruent and incongruent run. Participants not only extinguished their initial association, but were also able to invert it, rating videos presenting direct gaze as more untrustworthy than videos presenting averted gaze (Figure 15). However, these results have not reached statistical significance.

Despite the inversion between eye gaze type and returned values, and according to the experience design, when participants were asked to rate once again each Trustee in a scale of -3 to 3, TTrust was still rated in average as trustworthy (mean=1.0, SE=0.50), while TUntrust was once again classified in average as untrustworthy (mean=-0.67, SE=0.58) (see Figure 16).

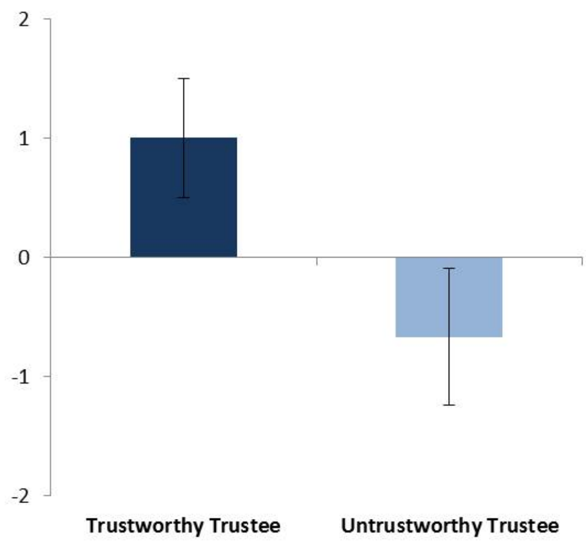


Figure 16: Trustee's overall trustworthiness evaluation.

Trustworthiness: -3=untrust, 0=neutral, 3=trust.

After the trustworthiness evaluation, a final debriefing was made in order to understand the strategy used by each participant when deciding how much to invest. Participants reported that, when watching a new video for the first time, they tended to choose the second option of investment (medium), and according to the returned value, they would adjust their investment value for the next trials. Also, some participants referred that their primary investment decisions were based in the UG task, and only later based on values returned by each trustee. Regarding the direction of eye gaze, participants associated directed gaze with higher returned values, and averted gaze with smaller amounts returned. Also, all participants noticed that the relation between eye gaze and the amount returned (trustee's behaviour) was inverted in the middle of the task. In general, participants were not able to predict the PC performance, as expected. For this reason participants' were unable to classify its behaviour as trustworthy or untrustworthy.

3.1.2 Inside the Scanner: Reward expectations

As previously described, both in the middle and at the end of the TG task, participants had to press a button concerning what they thought Trustee would return at the end of that trial, namely the minimum or maximum possible returned value. The following results aim to show if participants had matching expectations in the middle and at the end of the run.

Two types of analysis were performed with these results. The first analysis aimed to see if there was a match between the video presented and participants' reward expectations, i.e., what participants thought they would receive after watching a particular video. The second analysis seeked to understand if a match existed between reward expectations, namely through participants' evaluation, and their effective investment.

Matching video – reward expectations

Within this first analysis, the following scoring method was used to represent the match between Evaluation and Video: 0 was given to perfect matches between Evaluation and Video, and 1 was given to non-matches, e.g., when Trustee's eye gaze (displayed in the video) was directed, but participants were expecting the minimum return and the opposite happened. Since for the Trustee "PC" the probability of receiving the highest amount was 50%, these results were not included in the statistical analysis, given that it was impossible to predict the PC behaviour since it performed at chance level and there were no matching expectations.

Afterwards, frequency differences in the match and non-match categories concerning each task (congruent, incongruent) were tested. This analysis was performed in order to see if there were any global differences between match ("0") and non-match ("1") frequencies, independent of task type. A Chi-Square test was performed, showing that there were statistical differences between frequencies in the match and in the non-match categories, with match trials presenting higher frequencies than non-match trials (X^2=32.929, p=0.000).

Another analysis was performed to see if there were differences in match ("0") and non-match ("1") frequencies between the two types of task (congruent and incongruent). Tables 14 and 15 present match and non-match percentages for both tasks.

Table 14: Behavioural results: Match between reward expectations and presented video for congruent task.

Congruent Task			
Participant nr	Trials nr*	non-match(%)	match(%)
2	2	0	100
3	16	31.25	68.75
4	16	6.25	93.75
6	16	6.25	93.75

*valid trials of Evaluation.

Table 15: Behavioural results: Match between reward expectations and presented video for incongruent task.

Incongruent Task			
Participant nr	Trials nr*	non-match(%)	match(%)
2	16	37.5	62.5
3	16	62.55	37.5
4	16	25	75
6	16	12.5	87.5

*valid trials of Evaluation.

Statistical differences between tasks in match and non-match frequencies were found. Non-match trials presented higher frequencies in the incongruent task, comparing to the congruent task (X^2=5.659, p=0.024).

Thus, for the congruent task, participants were more aware of the relation between eye gaze type and its associated outcome, than for the incongruent task.

Matching reward expectations – investment decisions

Regarding the second analysis here performed, it was expected that participants investment would be related to the evaluation they made in the beginning of the trial, e.g., participants expecting a higher reward would choose the higher investment option. Results concerning this analysis are presented in Table 16.

Table 16: Behavioural results: Match between reward expectations and Investment option.

Participant nr	Trials nr*	non-match(%)	match(%)
1	N=46 trials	0.00	100.00
2	N=26 trials	26.92	73.08
3	N=48 trials	43.75	56.25
4	N=48 trials	18.75	81.25
5	N=22 trials	54.55	45.45
6	N=46 trials	13.04	86.96

*valid trials of Investment.

In general, participants presented a high percentage match between the evaluation moment and the investment option chosen, i.e., between what they thought they would receive and what they have effectively decided to invest. However, results also show that not all participants decided to invest according to their own evaluations regarding each Trustee. Once again, the same participants which had lower matches between Evaluation and Video presented lower values for the match between trustee's evaluation and their investment option.

3.2 Functional results

The purpose of this study was to have a better understanding of the influence of facial characteristics, namely eye gaze direction, in the modulation of neural mechanisms involved in social cognition. It was hypothesized that eye gaze is one of the mechanisms through which the amygdala is modulated, besides trustworthiness cues. First, areas responding to different levels of Trustee and Eye gaze direction regarding the Video period were analysed. A second analysis was performed with the purpose to understand which were the areas involved in reward expectations regarding evaluation, investment decision and feedback periods.

This section will be divided in two subsections, the first describing the results for main effects of Trustee and Eye gaze (analysis 1) and the other one presenting the results for Reward Expectation and Decision-Making (analysis 2).

3.2.1 Analysis 1: Main effects of Trustee and Eye gaze

An FFX ANOVA 2x2 was performed for each task (Congruent, Incongruent) in separate, having as factors Trustee (trustworthy, untrustworthy) and Eye gaze (directed, averted). The list of regions, peak voxel coordinates and statistics are described in Tables 17, 18, 19 and 20.

Main effects

For the main effect of Trustee, the following list of activated regions found is described in Table 17 for the congruent and incongruent task.

Table 17: Results of fixed-effects (FFX)-ANOVA 2x2 main effect, outputs and statistics of Trustee, depending on task.

Region	**Peak X (TAL)**	**Peak Y (TAL)**	**Peak Z (TAL)**	**Nr of voxels**	**t**	**p**
Congruent task*						
Trustee: Trustworthy > Untrustworthy						
R Inferior Parietal Lobule	62	-23	24	18479	-3.936	0.000085
R Middle Temporal Gyrus	56	-20	-12	3619	4.491	0.000007
R Supramarginal Gyrus	56	-59	33	2823	3.173	0.001523
R Precentral Gyrus	44	16	36	4988	3.608	0.000313
R Middle Frontal Gyrus	35	61	9	4218	3.686	0.000232
R Cuneus	14	-89	9	3199	-3.742	0.000186
L Superior Frontal Gyrus BA 8	-29	31	51	2621	4.188	0.000029
L Precentral Gyrus BA 44	-61	7	12	16190	-4.1380	0.000036
Incongruent task**						
Trustee: Trustworthy> Untrustworthy						
R Pons (Brainstem)	2	-14	-30	581	4.269	0.000020

*Contrasts were performed at $p<0.05$ using cluster threshold correction.
**Contrasts were performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

Right middle temporal gyrus and precentral gyrus responded more to TTrust than to TUntrust, which was not expected, since these regions were shown to present higher activations for untrustworthy faces [2] [70] [26]. Right cuneus presented higher activations for TUntrust, which is in accordance with studies showing higher responses for untrustworthy face stimuli [2] [71].

Regions which showed a main effect of Eye gaze are listed in Table 18 for the congruent and incongruent task.

Table 18: Results of fixed-effects (FFX)-ANOVA 2x2 main effect, outputs and statistics for Eye gaze, depending on task.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Congruent task*						
Eye gaze: Directed > Averted						
R Declive (Cerebellum)	26	-59	-12	2984	3.708	0.000213
R Superior Frontal Gyrus	23	49	9	2812	-4.442	0.000009
R Cuneus	20	-80	18	12121	4.054	0.000052
L Medial Frontal Gyrus	-1	61	9	10703	3.659	0.000257
L Middle Occipital Gyrus BA 19	-31	-93	15	7207	3.748	0.000181
Incongruent task**						
Eye gaze: Directed > Averted						
R Medial Frontal Gyrus BA 11	5	34	-15	1331	3.739	0.000190
L Uvula	-16	-83	-24	1584	4.038	0.000056

*Contrasts were performed at $p<0.05$ using cluster threshold correction.
**Contrasts were performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

Table 18 results for the congruent task reveal higher activations of right cuneus and left occipital gyrus as responding more to Directed than to Averted Eye Gaze. Regarding the incongruent task, results once again reveal the activation of right medial frontal gyrus, but also of the left uvula for Directed > Averted Eye gaze. These brain regions were pointed by some studies as being involved in perceiving the trustworthiness of faces [27] [2] and will later be discussed.

Specific contrasts for the interaction Trustee x Eye gaze

Contrasts between Eye Gaze type within each Trustee were performed. Table 19 summarizes the list of activated areas for the contrast "Trustworthy Trustee : Directed > Averted Eye gaze".

Table 19: Results of fixed-effects (FFX)-ANOVA 2x2 interaction effect, outputs and statistics regarding Trustworthy Trustee Eye gaze, depending on task.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Incongruent task*						
Trustworthy Trustee: Directed > Averted Eye gaze						
L Inferior Parietal Lobule	-52	-50	39	1064	3.384	0.000726
L Middle Frontal Gyrus	-46	19	33	1298	3.878	0.000108

*Contrast performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

It was found that, for the TTrust, left middle frontal gyrus, involved in trustworthiness processes, responded more to Directed than to Averted Eye gaze for the incongruent task, while results of the congruent task have not shown differences between Eye gaze direction, being thus not presented.

For the contrast "Untrustworthy Trustee: Directed > Averted eye gaze", the results are displayed in Table 20 for the congruent and incongruent tasks.

Table 20: Results of fixed-effects (FFX)-ANOVA 2x2 interaction effect, outputs and statistics regarding Untrustworthy Trustee Eye gaze, depending on task.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Congruent task*						
Untrustworthy Trustee: Directed > Averted Eye gaze						
L (amygdala)	-10	7	-6	1927	3.514	0.000448
Incongruent task**						
Untrustworthy Trustee: Directed > Averted Eye gaze						
L Middle Frontal Gyrus	-46	19	33	1168	-3.713	0.000210

*Contrasts were performed at $p<0.05$ using cluster threshold correction.
**Contrasts were performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

Interestingly, for the TUntrust (Table 20), the left middle frontal gyrus showed higher activation to Averted compared to Directed Eye gaze, in opposition to what was previously found for the TTrust (Table 19). Most importantly, it was found that left amygdala responded more for Directed than to Averted Eye Gaze (x=-16 y=3 z=-17), but only for the TUntrust (Figure 17).

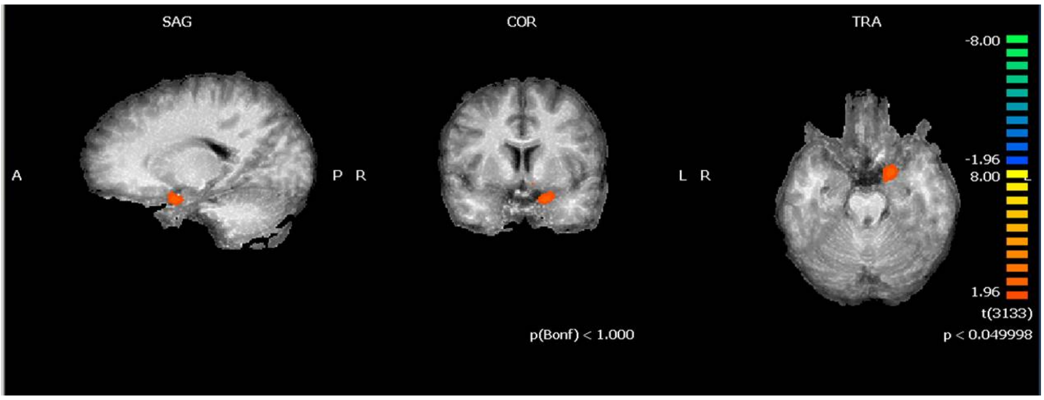


Figure 17: Activated region yielded by the FFX analysis (within the congruent task) for the contrast Untrustworthy Trustee: Directed > Averted Eye gaze. Left amygdala showed larger responses to directed compared to averted gaze for the Untrustworthy Trustee. Cluster threshold correction was set at $p<0.01$ with a minimum cluster size of 64 voxels.

3.2.2 Analysis 2: Reward expectations and Decision-making

In this FFX analysis, 4 main predictors were considered - Video, Evaluation, Investment and Feedback.

Specific contrasts were performed within Video and Evaluation periods, considering High and Low reward expectations (Table 21).

Table 21: Results of fixed-effects (FFX)-GLM contrasts, outputs and statistics for High and Low reward expectations regarding Video and Evaluation periods.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Video: high reward expectations > low reward expectations						
R Middle Frontal Gyrus	24.0	64.0	18.0	740	3.721	0.0002002

All contrasts were performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

During Video presentation and for high reward expectations > low reward expectations, middle temporal gyrus was the only reported region, which is very interesting, since this is a region shown to be involved in reward anticipation [72, 73].

For the Investment predictor, contrasts between High and Low Investment values were performed and presented in Table 22.

Table 22: Results of fixed-effects (FFX)-GLM contrasts, outputs and statistics for High and Low investment values regarding the Investment period.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Investment: high investment value > low investment value						
R Postcentral Gyrus	41	-26	48	5218	-4.667	0.000003
L Postcentral Gyrus BA 40	-40	-26	45	5401	4.277	0.000019

All contrasts were performed at p<0.01 using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

According to the results presented in Table 22, left postcentral gyrus presented higher responses for higher investments, which is very curious, since this brain structure was shown by previous studies to be involved in choice selection and reward anticipation [72, 76].

Finally, for the Trustee Feedback, both High > Low Feedback and High > Low Reward Expectations were compared. Tables 23 and 24 summarize the results.

Table 23: Results of fixed-effects (FFX)-GLM contrasts, outputs and statistics between High and Low Feedback regarding reward expectations.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
High > Low Feedback for high and low expectations						
R Postcentral Gyrus	44	-29	51	598	3.506	0.000458
High > Low Feedback for high expectations						
L Superior Temporal Gyrus	-58	-8	9	1107	4.259	0.000021
High > Low Feedback for low expectations						
R Middle Temporal Gyrus	32	-74	21	958	-3.476	0.000513
R Angular Gyrus	41	-59	33	654	-3.664	0.000251
R Posterior Cingulate	2	-56	18	619	-3.326	0.000884
R Precuneus BA 7	11	-56	45	1556	-3.695	0.000222

All contrasts were performed at p<0.01 using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

For the contrast "High > Low Feedback for low expectations", the angular gyrus, posterior cingulate and precuneus were brain regions that deactivated, i.e., presented higher activation when the Feedback was coherent with participants' expectations.

Table 24: Results of fixed-effects (FFX)-GLM contrasts, outputs and statistics between High and Low reward expectations regarding the Feedback.

Region	Peak X (TAL)	Peak Y (TAL)	Peak Z (TAL)	Nr of voxels	t	p
Low Feedback for high expectations > low expectations						
L Precentral Gyrus BA 6	-61	-8	30	1382	-3.598	0.000324
L Middle Temporal Gyrus	-53	-74	12	737	-3.760	0.000171
L Superior Temporal Gyrus	-58	-8	6	699	-4.016	0.000060
High and Low Feedback for high expectations > low expectations						
R Postcentral Gyrus BA 1	53	-17	49	925	-3.758	0.000173

All contrasts were performed at $p<0.01$ using cluster threshold correction.
X, Y and Z represent Talairach coordinates. R, right; L, left.

When participants reported having high expectations and the Feedback was lower than expected, precentral gyrus, middle temporal gyrus and superior temporal gyrus deactivated. Some of these regions were shown to be involved in reward processes, namely when expectations are different from the outcomes [36].

4 Discussion

The major aims of this project were to understand how can facial characteristics influence others' trustworthiness judgements and how can eye gaze in particular influence the process of decision-making. Since the amygdala has been pointed as a core structure regarding the perceived trustworthiness of others, this and other brain regions were studied when developing impressions of others during a social interaction, the Trust Game. While performing the TG, eye gaze direction was modulated in order to evaluate its effect in participants behaviour. Also, considering that decision-making takes into account the possibility of associated rewards, neural regions involved in reward processing were simultaneously studied.

Main effect of Trustee and Eye gaze

Regarding the neuroimaging results obtained, the first analysis focused on the identification of brain areas responding to different types of Trustee, Eye Gaze, and in the interaction between these two factors.

This part major findings were the amygdala and occipital gyrus activation for faces presenting Directed rather than Averted gaze. Right cuneus presented higher activations for the TUntrust, with left middle frontal gyrus presenting higher responses both for Directed and Averted gaze, for TTrust and TUntrust, respectively.

For the main effect of Trustee, the congruent task revealed, among others, the dactivation of right middle temporal gyrus, precentral gyrus and deactivation of R cuneus for the TTrust comparing to the TUntrust,in opposition o fthe two other activated regions. Studies reporting implicit and explicit trustworthiness judgements of faces revealed the increase of right cuneus response with the decrease of facial trustworthiness [2, 71], which is in accordance with these results, as behavioural results regarding Trustworthiness judgements after the Trust game task showed that the TUntrust was rated as the most untrustworthy trustee. However, the activation of the other structures is interesting, since middle temporal gyrus and precentral gyrus were thought to be involved in perceiving the trustworthiness of faces, and were described as responding more to untrustworthy than to trustworthy faces [2, 26, 70]. It may be that somehow Eye gaze interacts with the Trustee identity, since the precentral gyrus was shown to be active to meaningful gaze cues. Also, it has been identified as one of the structures of the called 'mirror system', which is known to be active while observing and executing an action, facilitating, thus, the understanding of others behaviour [74]. For the contrast Trustworthy > Untrustworthy Trustee of the incongruent task, our results do not point any core area involved in trustworthiness processes. One possible explanation may be that in the second (incongruent) task the participants had to relearn a new contingency between Eye gaze type and Trustee outcomes, which may have introduced noise to accurately find differences. Accordingly, behavioural results obtained during and after the TG revealed that participants were more aware of the association between eye gaze direction and Trustee behaviour during the congruent task than during the incongruent one (section 3.1.2). The lack of significant findings for the amygdala concerning this main effect was a surprise, since amygdala is known to be highly implicated in perception of trustworthiness in faces [16, 23]. It may be that participants, after performing the first trials of the task, were more focused on the Trustee's eye gaze direction than in the Trustee itself, since behavioural results demonstrated that they were aware of the existing relation between eye gaze direction and the following outcome.

When testing Eye gaze main effects for the congruent task, regions such as the right cuneus, left medial frontal gyrus and left occipital gyrus revealed increased neural responses to Directed compared to Averted Eye gaze. All these regions were pointed by some studies as being involved in the perception of trustworthiness in faces [2, 27, 70]. However, both right cuneus and left middle occipital gyrus displayed higher responses to untrustworthy stimuli (e.g. faces), and our results reveal higher activations for Directed Eye gaze, rated by participants as trustworthy stimuli. It is possible that different stimuli types (eye gaze, faces) may modulate this relationship. Regarding the incongruent task, an interesting result was the activation of the left uvula, a structure shown to present quadratic responses when evaluating the trustworthiness of faces [27], meaning it responds higher both to trustworthy and untrustworthy stimuli, compared to more neutral stimuli.

Interaction effects between Trustee and Eye gaze revealed the left middle frontal gyrus as one of the activated regions for the incongruent task, both for the TTrust and the TUntrust. However, it revealed higher activations for the Directed gaze of the TTrust, and lower activations for Directed gaze of the TUntrust both compared to Averted gaze. There seems to be no consensus between the role of this structure concerning trustworthy stimuli, since some studies demonstrated higher neural responses for trustworthy, but also for untrustworthy stimuli [70] [27]. Thus, it seems that the left middle frontal gyrus may be involved in processes related to the analysis of both trust and untrustworthy stimuli.

For the congruent task, the interaction between the Untrustworthy Trustee and Eye gaze led to increased activation of the left amygdala for the contrast "Directed > Averted Eye gaze", which is an accordance with studies revealing higher involvement of the amygdala in directed gaze processing [45, 46]. According to several studies, and as pointed in the systematic review here presented (see section 1.1), the amygdala is a core structure concerning the evaluation of others' trustworthiness, namely through facial characteristics. However, these studies also reveal that the left and the right amygdala present distinct responses to different levels of trustworthiness. Higher activations of left amygdala seem to be a consequence of trustworthiness increase, while right amygdala responses are known to increase with the decrease of perceived trustworthiness [2, 16, 25, 75]. For this contrast, responses to directed versus averted eye gaze were compared. According to the pilot study and behavioural results of the congruent task, directed eye gaze was rated as being more trustworthy and fair than averted gaze, respectively. Thus, the activation of the left amygdala for Directed > Averted gaze also validate our previous ratings considering direct eye gaze more trustworthy, since left amygdala responses are known to be higher for trustworthy comparing to untrustworthy stimuli. However, the same pattern did not occur for the TTrust. It may be that differences between directed and averted eye gaze for the TTrust were not so notorious comparing to TUntrust, not reaching thus statistical significance.

Reward Expectations

For the second analysis regarding neuroimaging data, only trials with evaluation were considered. This analysis focused in brain regions involved in reward expectations, in this particular case, related to decision-making processes. Major findings of this second part were the involvement of middle frontal gyrus in periods anticipating rewards, and of the precentral gyrus and left superior temporal gyrus for the Feedback periods, since precentral gyrus revealed higher activations when participants expectations and outcome were not coincident, and left superior temporal gyrus responses were higher for outcomes matching with participants expectations, both for high and low values.

The process of decision-making can be divided in three major steps. First, individuals evaluate and form preferences regarding available options; second, an option needs to be selected and executed; and third, individuals evaluate the outcome or the consequences of their chosen action. During the first step, individuals attribute values to the available options, selecting one of them. During the second step, an action is initiated, performed and completed according to the preferences shown in the first stage. In the third step, the individual will compare the expected to the experienced outcome [76].

In this study, the first step above described can be associated with video presentation and trustee evaluation. The second step can be considered the choice of an investment option, with the last step being the observation of the returned value, comparing it to the initially expected reward.

When a potential reward is perceived in the environment, brain's system of reward motivation, also known as the reward system, is activated. This reward system starts in the midbrain and passes through the limbic system, ending in the neocortex. Neuroimaging studies have pointed ventral striatum, nucleus accumbens, amygdala and insula as some of the most commonly activated brain structures when a task involving reward stimuli is performed [34] [77] [35].

Regarding the Expectations of reward, when watching the video, for the contrast high expectations > low expectations, results revealed the middle frontal gyrus as the only significant area. According to studies evaluating choice selection and reward anticipation, activation in brain regions such as the dorsal frontal gyrus, middle frontal gyrus and inferior parietal lobe, among others, was found [72] [36]. This may suggest that even though participants were not yet in the evaluation and decision period of the task, they could be making their own decisions during video presentation.

As to the evaluation period, there were no statistically significant activated brain regions. One possible reason for this to happen could be that participants had already made their decisions in the previous period. However, another possible explanation may be related to the fact that these results were calculated only for trials in which evaluation of the trustee was performed (6 for each 16 trials of each run), having thus not enough statistical power.

Regarding the Investment period, for the contrast high investment value > low investment value, left and right postcentral gyrus were some of the reported brain regions. Studies evaluating structures involved in reward anticipation also reported the activation of left postcentral gyrus as one of the involved regions after choosing an option and anticipating a certain reward [72] [76] [78]. Thus, it seems that the period given to choose an investment option can be considered a phase in which anticipation of reward already occurs.

In the beginning of the trial, when participants reported having low expectations regarding the result and the feedback was higher than expected (Contrast High > Low feedback for low expectations), four brain regions deactivated, most importantly the angular gyrus, the precuneus and the posterior cingulate. These areas have been directly implicated in reward processing, and in studies where participants had to perform tasks entailing monetary rewards when giving correct answers. The occipital lobe, the angular gyrus and the precuneus/cingulate cortex were pointed as responding with higher intensity to higher payments [79] [80]. These results seem to point for the opposite direction, since the reported regions deactivated for higher Feedback values. Nevertheless, there is another component not evaluated by the previously referred studies, the existence of expectations which were formed earlier. Thus, these regions, known to be implicated in reward processing, may be involved not only in processes of higher payments, but also of lower payments, depending on the previously formed expectations.

Furthermore, when Trustee feedback was low and participants were expecting a higher value (Contrast Low Feedback for high expectations > low expectations), areas such as the precentral gyrus and middle temporal gyrus deactivated. A meta-analysis concerning reward-related tasks in healthy adults identified the precentral gyrus as one of the activated regions when reward expectations were higher than the outcome values [36]. Thus, it seems that this region may be involved in reward processes when outcomes are not coincident with individuals expectations.

Also interesting was the response of the left superior temporal gyrus. First, its activation was not expected for an outcome phase, since studies show that this brain region is usually more related with assessment and action selection during decision-making [76]. Thus, one can suppose that participants, while observing the returned value during the Feedback period, could have been already updating their evaluation of the Trustee and consequently selecting the next investment value. It is also worth to notice that this brain region presented higher activations for higher Feedback values when participants reported having high Expectations. However, for low Feedback values it deactivated when participants had higher expectations, i.e., for low Feedback values, it presented higher activation for low expectations. It seems, thus, that this region can also be involved in reward processes, namely when participants expectations are coincident with Feedback values.

The Reward expectations analysis allowed to establish a connection between the different stages of decision-making. Given our results and what has been referred in the literature, we can define the phase of the Video presentation and the phase of Investment of our task as an Anticipation phase. This statement is supported by the fact that some of the activated brain regions during these periods, such as the middle frontal gyrus and the postcentral gyrus, were described to be involved in reward anticipation [72] [76] [36]. Regarding the third decision-making step (individuals evaluation of the outcome), when individuals compare predicted to experienced outcomes, areas such as the precentral gyrus, angular gyrus, precuneus, posterior cingulate and middle temporal revealed higher activations for outcomes that were not coincident with previously formed expectations. For outcomes coinciding with participants expectations, left superior temporal gyrus was shown to be activated not only for low Expectations and low Feedback values, but also for high Expectations and high Feedback values.

However, and despite some similar areas reported in previous studies, these results do not reveal activation in areas considered core regions when evaluating reward expectations, such as the ventral striatum, nucleus accumbens, insula and amygdala, among others. This could be due to the small number of participants, and to the fact that these analysis were performed only for the trials in which Trustee evaluation occurred, minimizing thus its statistical power.

Importantly, one should point that through the debriefing performed with participants in the end of the tasks, participants referred that in the beginning of the TG task, they would choose their next investment option according to the returned amount of the last trial played with that same Trustee. These results are in agreement with studies showing that our actions and behaviours are based in our previous social interactions with others [2] [3]. However, after this initial moment, participants' investment choices were based in the evaluation that each participant performed of each Trustee behaviour during the Trust Game. In particular, eye gaze direction of the Trustee for each trial was taken as an important cue related with the next Trustee Feedback value. This seems to reveal that facial characteristics, in this particular case eye gaze direction, can influence people choices, in accordance with studies that demonstrated the impact of facial characteristics when performing social interaction judgements and decisions able to predict real world outcomes [17] [81].

Overall, for the main effect of Trustee and Eye gaze, the most important finding was the left amygdala activation for faces presenting Directed rather than Averted gaze, suggesting Directed gaze as a trustworthy stimuli, since left amygdala was shown to present higher responses for these type of stimuli. Regarding the Reward Expectation analysis, the middle frontal gyrus activation for the Video presentation suggests an anticipation of the evaluation and decision period. Furthermore, areas such as the precentral gyrus and the left superior temporal gyrus seem to be involved in reward expectations, since the precentral gyrus revealed higher activations when Feedback values were not coincident with participants' expectations, while the left superior temporal gyrus activations were superior when the returned amount matched the expectations of the participants, both for high and low values.

5 Conclusions

This project was developed with the purpose to learn how to identify a relevant problem in the field of social neuroscience (systematic review) and to develop (task development in Matlab) and test a new experimental design and paradigm suitable to respond to the raised neuroscentific questions (task performance with acquisition of behavioural and functional data).

After identification of the question to study by performing a systematic review of the literature, the paradigm was designed and tested. According to the results, this was validated as appropriate for responding to the question regarding the role of the amygdala when performing trustworthiness judgements while modulating the direction of eye gaze.

The experimental paradigm here designed allowed to evaluate which neural regions were involved in the main and interaction effects between Trustee and Eye gaze and in Reward Expectations. For the interaction between Trustee and Eye gaze, our results have shown higher left amygdala activation when observing Trustee's Directed rather than Averted Eye gaze. Some studies reported left amygdala involvement when watching trustworthy stimuli (e.g., faces), which is in accordance with our results, since participants rated Directed gaze as more trustworthy than Averted gaze. As to Reward Expectation analyses, the middle frontal gyrus activated during Video presentation. Since this region is known to be involved in choice selection and reward anticipation, it seems that during this period, participants were making their own decisions for the following phases, anticipating the expected reward. As to the Feedback period, results seem to point the involvement of areas such as the precentral gyrus and the left superior temporal gyrus in reward expectations. Precentral gyrus higher activations were found when Feedback values were not matching with individuals expectations, while the left superior temporal gyrus revealed superior activations when the returned amount matched the expectations of the participants, not only for high but also for low values.

Summarizing, results here presented validate the experimental paradigm designed in this study, since it is suitable to provide a better understanding of the role of the amygdala in our perceptions of trustworthiness, taking into account the influence of some mechanisms, in particular eye gaze signals. There was no activation of areas considered core regions in the field of reward expectations, such as the ventral striatum, nucleus accumbens, insula or amygdala. This can somehow be related to the low number of participants and the reduced number of trials of evaluation from which analyses were performed.

6 Future work

The purpose of this project was to master systematic and metaanalytic techniques and develop and test an experimental paradigm suitable to provide a better understanding of the amygdala role in social interactions, and this was accomplished. However, results here presented were obtained from a small sample, since the FMRI task was only applied to 4 participants successfully. Thus, in order to obtain more robust data, the task should be applied to a higher number of participants.

Thereafter, with a higher number of participants, it would be interesting to separate them in different groups, after analysing the results of behavioural data. Participants revealing high risk investments could be analysed separately from participants whose investment choices where of low risk, to assess if there are any differences in their neural responses.

Another future and pertinent analysis would be to perform ROI analysis, particularly in the amygdala, since it was the focus of this study. However, ROI analysis could also be performed in other regions part of the face network (e.g. occipital and fusiform face area (FFA), superior temporal sulcus (STS)).

Eye tracker data will be considered in the future. These data analyses, in particular of eye tracker data, could provide relevant information, namely concerning participants' eye gaze while observing Trustee's faces.

Finally, given the role of the amygdala in social cognition, and in particular in the perception of trustworthiness cues, it is also planned to perform connectivity analyses. In fact, it is planned to see in this particular task (trust game) and considering trustworthiness judgements of faces and eye gaze, which regions are directly influenced by the amygdala, and by which regions is this structure modulated.

7 Appendix

Appendix A

Table 25: Behavioural results: fairness judgements regarding Trustworthy and Untrustworthy Trustees.

		Trustworthy	Trustee			Untrustworthy	Trustee	
Participant*	Video 1	Video 2	Video 3	Video 4	Video 1	Video 2	Video 3	Video 4
1	3	2	0	-1	2	3	-1	-1
2	2	2	0	-1	1	1	-1	-2
3	2	0	0	1	-1	1	2	0
4	3	3	-2	-3	2	2	-3	-3
5	2	3	-2	-2	3	2	-2	0
6	3	3	3	3	0	0	0	0
Fairness mean(SD)	2.50(0.55)	2.17(1.17)	-0,17(1.83)	-0.50(2.17)	1.17(1.47)	1.50(1.05)	-0.83(1.72)	-1.00(1.26)

Fairness: -3=unfair, 0=neutral, 3=fair.
*All participants were considered.

Table 26: Behavioural results: Trustworthiness judgements regarding Trustworthy and Untrustworthy Trustees.

		Trustworthy	Trustee			Untrustworthy	Trustee	
Participant*	Video 1	Video 2	Video 3	Video 4	Video 1	Video 2	Video 3	Video 4
1	2	2	-1	-2	1	1	0	1
2	-2	-2	0	2	-2	-1	-1	1
3	2	2	-1	-2	0	-2	-2	-2
4	-2	-3	2	2	1	2	-1	2
5	3	2	1	-2	1	2	-2	1
6	-3	-2	2	2	-2	0	2	2
Trustworthiness mean(SD)	0.17(0.98)	-0.17(0.98)	0.50(0.56)	0.00(0.89)	-0.17(0.60)	-1.33(0.49)	-0.67(0.61)	0.83(0.60)

Trustworthiness: -3=untrust, 0=neutral, 3=trust.
*Only participants 2, 3, 4 and 6 were considered for data representation in Figures 15 and 16.

8 References

[1] Sanfey, A. G. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, 318(5850), 598-602.

[2] Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, 19(9), 1508-1519.

[3] Frith, C. D., & Singer, T. (2008). The role of social cognition in decision making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1511), 3875-3886.

[4] Frith, C. D., & Frith, U. (2007). Social cognition in humans. *Current Biology*,17(16), R724-R732.

[5] Kim, H., Choi, M. J., & Jang, I. J. (2012). Lateral OFC activity predicts decision bias due to first impressions during ultimatum games. *Journal of cognitive neuroscience*, 24(2), 428-439.

[6] Kragel, P. A., Zucker, N. L., Covington, V. E., & LaBar, K. S. (2014). Developmental trajectories of cortical-subcortical interactions underlying the evaluation of trust in adolescence. *Social cognitive and affective neuroscience*, nsu050.

[7] Costa-Gomes, M. A. (2010). Behavioral Economics And Game Theory. *Wiley Encyclopedia of Operations Research and Management Science.*

[8] Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., & Vogeley, K. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, 46(4), 1154-1163.

[9] Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393(6684), 470-474.

[10] Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature neuroscience*, 5(3), 277-283.

[11] Brothers, L. (2002). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Foundations in social neuroscience*, 367-385.

[12] Frith, C. D. (2008). Social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1499), 2033-2039.

[13] Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. Emotion, 6(2), 269.

[14] Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27(6), 813-833.

[15] Willis, J., & Todorov, A. (2006). First impressions making up your mind after a 100-ms exposure to a face. Psychological science, 17(7), 592-598.

[16] Cosmides, L., & Toobey, J. (2000). The cognitive neuroscience of social reasoning. In M. S. Gazzaniga (Ed.), *The new cognitive neuroscience* (pp. 1259–1276). Cambridge; Massachusetts: MIT Press.

[17] Van't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. Cognition, 108(3), 796-803.

[18] Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, 4(3), 165-178.

[19] Baas, D., Aleman, A., Vink, M., Ramsey, N. F., de Haan, E. H., & Kahn, R. S. (2008). Evidence of altered cortical and amygdala activation during social decision-making in schizophrenia. *Neuroimage*, 40(2), 719-727.

[20] Koscik, T. R., & Tranel, D. (2011). The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia*,49(4), 602-611.

[21] Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of neuroscience*, 18(1), 411-418.

[22] Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Advances in experimental social psychology*, 32, 201-271.

[23] Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Cognitive Neuroscience*, Journal of, 14(8), 1264-1274.

[24] Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambady, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of personality and social psychology*, 104(3), 409.

[25] Baron, S. G., Gobbini, M. I., Engell, A. D., & Todorov, A. (2011). Amygdala and dorsomedial prefrontal cortex responses to appearance-based and behavior-based person impressions. *Social cognitive and affective neuroscience,* 6(5), 572-581.

[26] Todorov, A., & Engell, A. D. (2008). The role of the amygdala in implicit evaluation

of emotionally neutral faces. *Social Cognitive and Affective Neuroscience*, nsn033.

[27] Said, C. P., Baron, S. G., & Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, 21(3), 519-528.

[28] Mattavelli, G., Andrews, T. J., Asghar, A. U., Towler, J. R., & Young, A. W. (2012). Response of face-selective brain regions to trustworthiness and gender of faces. *Neuropsychologia*, 50(9), 2205-2211.

[29] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, et al. (2009) The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med* 6(7): e1000100.

[30] Huedo-Medina, T.B., Sánchez-Meca, J., Marín-Martínez, F., & Botella J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, 11(2), 193-206.

[31] Field, A. P., & Gillett, R. (2010). How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology*, 63, 665–694.

[32] Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48(2), 175-187.

[33] Marsh, A. A., Blair, K. S., Vythilingam, M., Busis, S., & Blair, R. J. R. (2007). Response options and expectations of reward in decision-making: the differential roles of dorsal and rostral anterior cingulate cortex. *Neuroimage*,35(2), 979-988.

[34] McClure, S. M., York, M. K., & Montague, P. R. (2004). The neural substrates of reward processing in humans: the modern role of FMRI. *The Neuroscientist*, 10(3), 260-268.

[35] Peterson, R. L. (2005). The neuroscience of investing: fMRI of the reward system. *Brain research bulletin*, 67(5), 391-397.

[36] Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 35(5), 1219-1236.

[37] Tzieropoulos, H. (2013). The Trust Game in neuroscience: A short review. *Social neuroscience*, 8(5), 407-416.

[38] Wyland, C. L., & Forgas, J. P. (2010). Here's looking at you kid: Mood effects on

processing eye gaze as a heuristic cue. *Social Cognition*, 28(1), 133-144.

[39] Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., & Vogeley, K. (2009). Duration matters: dissociating neural correlates of detection and evaluation of social gaze. *Neuroimage*, 46(4), 1154-1163.

[40] Bailenson, J. N., Beall, A. C., Loomis, J., Blascovich, J., & Turk, M. (2005). Transformed social interaction, augmented gaze, and social influence in immersive virtual environments. *Human Communication Research*, 31(4), 511-537.

[41] Normoyle, A., Badler, J. B., Fan, T., Badler, N. I., Cassol, V. J., & Musse, S. R. (2013, November). Evaluating perceived trust from procedurally animated gaze. *In Proceedings of the Motion on Games* (pp. 119-126). ACM.

[42] Bayliss, A. P., & Tipper, S. P. (2006). Predictive Gaze Cues and Personality Judgments Should Eye Trust You?. *Psychological Science*, 17(6), 514-520.

[43] Droney, J. M., & Brooks, C. I. (1993). Attributions of self-esteem as a function of duration of eye contact. *The Journal of social psychology*, 133(5), 715-722.

[44] Argyle, M., & Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, 289-304.

[45] Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contactERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, 45(13), 3024-3037.

[46] George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *Neuroimage*, 13(6), 1102-1112.

[47] Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., & Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring A PET study. *Brain*, 122(4), 779-783.

[48] Adams, R. B., Franklin, R. G., Kveraga, K., Ambady, N., Kleck, R. E., Whalen, P. J., & Nelson, A. J. (2011). Amygdala responses to averted vs direct gaze fear vary as a function of presentation speed. *Social cognitive and affective neuroscience*, nsr038.

[49] Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christoffels, I., & Young, A. W. (2002). Reading the mind from eye gaze. *Neuropsychologia*, 40(8), 1129-1138.

[50] Carter, M., & Shieh, J. (2009). Guide to research techniques in neuroscience.

Academic Press.

[51] Amaro Jr, E., & Barker, G. J. (2006). Study design in fMRI: basic principles. *Brain and cognition*, 60(3), 220-232.

[52] Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, 87(24), 9868-9872.

[53] Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging* (Vol. 1). Sunderland, MA: Sinauer Associates.

[54] Santarelli, M. F., Landini, L., & Positano, V. (Eds.). (2005). *Advanced image processing in magnetic resonance imaging.* Taylor & Francis.

[55] Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(1), 6-12.

[56] Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, 89(13), 5951-5955.

[57] Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., & Ugurbil, K. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophysical journal*, 64(3), 803-812.

[58] Stippich, C. (2007). Clinical functional MRI. S*pringer-Verlag Berlin Heidelberg.*

[59] Turner, R., Bihan, D. L., Moonen, C. T., Despres, D., & Frank, J. (1991). Echo-planar time course MRI of cat brain oxygenation changes. *Magnetic Resonance in Medicine*, 22(1), 159-166.

[60] Gore, J. C. (2003). Principles and practice of functional MRI of the human brain. *The Journal of clinical investigation*, 112(1), 4-9.

[61] Yacoub, E., Shmuel, A., Pfeuffer, J., De Moortele, V., Adriany, G., Ugurbil, K., & Hu, X. (2001). Investigation of the initial dip in fMRI at 7 Tesla. *NMR in Biomedicine*, 14(7-8), 408-412.

[62] Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magnetic resonance imaging*, 22(10), 1517-1531.

[63] Buxton, R. B. (2009). Introduction to functional magnetic resonance imaging:

principles and techniques. *Cambridge university press.*

[64] Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex*, 11(9), 825-836.

[65] Donaldson, D. I., Petersen, S. E., Ollinger, J. M., & Buckner, R. L. (2001). Dissociating state and item components of recognition memory using fMRI. *Neuroimage*, 13(1), 129-142.

[66] Smith, S. M., Matthews, P. M., & Jezzard, P. (Eds.). (2001). Functional MRI: an introduction to methods. Oxford University Press, chapter 11.

[67] Scharlemann, J. P., Eckel, C. C., Kacelnik, A., & Wilson, R. K. (2001). The value of a smile: Game theory with a human face. *Journal of Economic Psychology*, 22(5), 617-640.

[68] Howlett, N., Pine, K., Orakçioglu, I., & Fletcher, B. (2013). The influence of clothing on first impressions: Rapid and positive responses to minor changes in male attire. *Journal of Fashion Marketing and Management*, 17(1), 38-48.

[69] Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, 289(5485), 1773-1775.

[70] Doallo, S., Raymond, J. E., Shapiro, K. L., Kiss, M., Eimer, M., & Nobre, A. C. (2012). Response inhibition results in the emotional devaluation of faces: neural correlates as revealed by fMRI. *Social cognitive and affective neuroscience*, 7(6), 649-659.

[71] van Rijn, S., Swaab, H., Baas, D., de Haan, E., Kahn, R. S., & Aleman, A. (2012). Neural systems for social cognition in Klinefelter syndrome (47, XXY): evidence from fMRI. *Social cognitive and affective neuroscience*, 7(6), 689-697.

[72] Ernst, M., Nelson, E. E., McClure, E. B., Monk, C. S., Munson, S., Eshel, N., & Pine, D. S. (2004). Choice selection and reward anticipation: an fMRI study. *Neuropsychologia*, 42(12), 1585-1597.

[73] Berns, G. S., McClure, S. M., Pagnoni, G., & Montague, P. R. (2001). Predictability modulates human brain response to reward. The Journal of Neuroscience, 21(8), 2793-2798.

[74] Hooker, C. I., Paller, K. A., Gitelman, D. R., Parrish, T. B., Mesulam, M., & Reber, P. J. (2003). Brain networks for analyzing eye gaze. Cognitive Brain Research, 17(2), 406-418.

[75] Killgore, W. D., Schwab, Z. J., Tkachenko, O., Webb, C. A., DelDonno, S. R., Kipman, M., & Weber, M. (2013). Emotional intelligence correlates with functional responses to dynamic changes in facial trustworthiness. *Social neuroscience*, 8(4), 334-346.

[76] Paulus, M. P., Feinstein, J. S., Leland, D., & Simmons, A. N. (2005). Superior temporal gyrus and insula provide response and outcome-dependent information during assessment and action selection in a decision-making situation. *Neuroimage*, 25(2), 607-615.

[77] O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Current opinion in neurobiology*, 14(6), 769-776.

[78] Bjork, J. M., & Hommer, D. W. (2007). Anticipating instrumentally obtained and passively-received rewards: a factorial fMRI investigation. *Behavioural brain research*, 177(1), 165-170.

[79] Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, 318(5854), 1305-1308.

[80] McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., & Cohen, J. D. (2007). Time discounting for primary rewards. *The Journal of Neuroscience*, 27(21), 5796-5804.

[81] Rule, N. O., Moran, J. M., Freeman, J. B., Whitfield-Gabrieli, S., Gabrieli, J. D., & Ambady, N. (2011). Face value: Amygdala response reflects the validity of first impressions. *NeuroImage*, 54(1), 734-741.