N°d'ordre NNT : xxx

**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**
opérée au sein de
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale ED 476 – NSCo**
**L'École Doctorale Neurosciences et Cognition**

Soutenue publiquement le 14-Septembre-2016, par :
**Manuela Costa**

---

# Social traits and facial information

## Behavioral and neuronal evidence within the framework of phylogenetic and clinical studies

---

Sous la direction de: Angela Sirigu (DRCE, CNRS)

Devant le jury composé de:

Dr. Deruelle, Christine, rapporteur
Pr. Ferrari, Pier Francesco, rapporteur
Dr. Strange, Bryan, examinateur
Dr. Pascalis, Olivier, examinateur
Pr. Gervais, Remi, examinateur
Dr. Angela Sirigu, directrice

# UNIVERSITE CLAUDE BERNARD - LYON 1

| | |
|---|---|
| **Président de l'Université** | **M. le Professeur Frédéric FLEURY** |
| Président du Conseil Académique | M. le Professeur Hamda BEN HADID |
| Vice-président du Conseil d'Administration | M. le Professeur Didier REVEL |
| Vice-président du Conseil Formation et Vie Universitaire | M. le Professeur Philippe CHEVALIER |
| Vice-président de la Commission Recherche | M. Fabrice VALLÉE |
| Directeur Général des Services | M. Alain HELLEU |

## *COMPOSANTES SANTE*

| | |
|---|---|
| Faculté de Médecine Lyon Est – Claude Bernard | Directeur : M. le Professeur J. ETIENNE |
| Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux | Directeur : Mme la Professeure C. BURILLON |
| Faculté d'Odontologie | Directeur : M. le Professeur D. BOURGEOIS |
| Institut des Sciences Pharmaceutiques et Biologiques | Directeur : Mme la Professeure C. VINCIGUERRA |
| Institut des Sciences et Techniques de la Réadaptation | Directeur : M. le Professeur Y. MATILLON |
| Département de formation et Centre de Recherche en Biologie Humaine | Directeur : Mme la Professeure A-M. SCHOTT |

## *COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE*

| | |
|---|---|
| Faculté des Sciences et Technologies | Directeur : M. F. DE MARCHI |
| Département Biologie | Directeur : M. le Professeur F. THEVENARD |
| Département Chimie Biochimie | Directeur : Mme C. FELIX |
| Département GEP | Directeur : M. Hassan HAMMOURI |
| Département Informatique | Directeur : M. le Professeur S. AKKOUCHE |
| Département Mathématiques | Directeur : M. le Professeur G. TOMANOV |
| Département Mécanique | Directeur : M. le Professeur H. BEN HADID |
| Département Physique | Directeur : M. le Professeur J-C PLENET |
| UFR Sciences et Techniques des Activités Physiques et Sportives | Directeur : M. Y.VANPOULLE |
| Observatoire des Sciences de l'Univers de Lyon | Directeur : M. B. GUIDERDONI |
| Polytech Lyon | Directeur : M. le Professeur E.PERRIN |
| Ecole Supérieure de Chimie Physique Electronique | Directeur : M. G. PIGNAULT |
| Institut Universitaire de Technologie de Lyon 1 | Directeur : M. le Professeur C. VITON |
| Ecole Supérieure du Professorat et de l'Education | Directeur : M. le Professeur A. MOUGNIOTTE |
| Institut de Science Financière et d'Assurances | Directeur : M. N. LEBOISNE |

## Thanks

*After three years of living in Lyon, with a great time at the Institute des Sciences Cognitives (ISC), I need to thanks many people that have been fundamental to arrive to this first objective, hopefully, of my research life. First of all, I want to thank my supervisor Angela Sirigu that gave me the opportunity to start this interesting experience of a PhD in Cognitive Neuroscience. She believed in me and supports me during the entire path. Alice and Guillaume, cat and dog, the angel and the devil. This thesis is also yours. Alice, thank you for the support and for being so close to me in the difficult moments of my thesis. Thank you for your generosity and for transmitting me your knowledge. Thanks Guillaume, who teach me the rigor and the criticism in science. Thanks also to Andres that dedicated his time to my questions and doubts.*

*Martina and Mina: what to say, we started together this experience, sharing more than the simply PhD life, we lived together, as a family and shared difficult and happy moments along these three years. Your presence was impressively important to arrive to the end. This is only one of the important moments that we will live together. Thanks to all the colleagues of the Angela's team, Arthur, Song, Flavia, Gloria, Nathalie, Nadia, Michel, that shared their knowledge with me and with whom I had formative and interesting conversations. How to forget the people that supported me with their energy and their curiosity when I was tired and full of work: Mirjana, Noemi, Silvia, Antoine, Joanne. The moments spent with you help me to understand how lucky I was and make my motivation stronger. Thanks to Jean Rene Duhamel that gave me the opportunity to approach the monkeys kingdom; to Caroline Demily for letting me discover this impressive pathology of Williams's syndrome that also allowed me to meet many extraordinary people. Thanks to all the persons that nicely accepted to participate in my studies, impossible to remember all of you but each of you shaped this work. Thanks to all the patients and families and also the monkeys work.*

*Not only colleagues but also friends from the ISC, Alex, Miguel, David, Laure, Elodie, Manu, Gustavo, Thomas. I'm glad to have meet you and I hope our path will cross again soon. Thanks Audrey for the nice time spent together. Thanks to Fabio, Michela, Roberta, Giada and all my friends that support me during these years of living abroad, thanks for your visits, chat and love. Far but extremely close, my family is always fundamental in all my experiences; they make me feel so special. My mother that always attended my questions and doubts; with her patience and determination she teaches me how to deal with life. My brothers Antonio and Gianmarco tireless and hard workers. Thanks to Dad I'm sure you will clap me as well. And now you, the best gift of my PhD, I can't realize that is done, passion, support, braveness, patience, restoration, my scape, love...this and much more. In the life paths, side by side. Now, I can breathe, we made it!*

# Table of contents

**Resumé**

Le visage est probablement l'une des parties les plus saillante et pertinente de notre corps. En effet, les visages fournissent à l'observateur un ensemble dynamique et complexe d'informations physiques, émotionnelles et sociales qui déterminent notamment la manière dont les gens vont interagir entre eux. Grâce à cette information faciale, l'humain peut se faire très rapidement une première impression de la personne, comme par exemple former un jugement de confiance ou d'attractivité.

La capacité de former des jugements de nature sociale à partir d'un visage est au centre de ce travail de thèse. Une attention toute particulière est donnée à la perception de la fiabilité d'autrui. La confiance implique de prendre le risque de mettre son sort dans les mains de quelqu'un d'autre. Au quotidien, les individus prennent constamment des décisions basées sur la confiance dans différents contextes sociaux (i.e. où acheter de la nourriture, où s'asseoir dans le métro, à qui demander une information quand il est perdu). Ce type de jugements intervient également dans des contextes beaucoup plus engageants, comme dans le choix d'un partenaire sexuel, les collaborations professionnelles et beaucoup d'autres. L'évaluation de la fiabilité d'autrui est donc une capacité essentielle tant à un niveau individuel que dans la société.

Les recherches présentées dans cette thèse se focalisent sur la manière dont une information sociale, et particulièrement la confiance peut être détectée spontanément à partir d'un visage. Dans ce travail, j'ai employé des techniques d'oculométrie (eye-tracking), d'électrophysiologie (EEG) et des mesures comportementales. Avant d'expliquer en détail les objectifs de chaque étude, le premier chapitre offrira un résumé succinct de l'organisation faciale d'un point de vue anatomique. Puis, une section présentera la littérature actuelle sur les bases neuronales du traitement des visages. Je décrirai différentes théories qui proposent l'existence d'un réseau neuronal distinct et de régions spécifiquement dédiées au traitement des visages chez l'homme et le singe. Puis, je me focaliserais sur la littérature lié à la notion de confiance, en décrivant les principales études portant sur les corrélats neuronaux de la perception de traits sociaux véhiculés par un visage.

Dans une première étude (Chapitre 2), j'ai tenté de déterminer si le sentiment immédiat de confiance est universel. J'ai étudié la notion de confiance d'un point de vue évolutif. Le but est de tester si les singes (Macaca Mulatta, Macaca Fascicularis) peuvent montrer une préférence spontanée envers des visages humains classifiés au préalable comme confiants, suggérant ainsi une capacité à détecter les indices faciaux impliqués dans la

construction du sentiment de confiance déjà observés dans l'espèce humaine. Ceci nous a permis de déterminer si l'utilisation de certaines caractéristiques morphologiques du visage est commune aux deux espèces primates. Pour ce faire, les mouvements des yeux ont été enregistrés chez des sujets volontaires sains et chez un groupe des singes durant un paradigme de préférence visuelle de visages appariés. Nous avons observé chez les deux espèces un temps de regard supérieur pour les visages inspirant confiance par rapport à ceux n'inspirant pas confiance. Nous avons également trouvé une corrélation significative entre ce temps de regard préférentiel et certains traits des visages, à savoir la joie, la féminité et le ratio largeur/hauteur du visage (RLHv), connus pour être des prédicteurs du jugement de confiance.

Dans une deuxième étude, afin de mieux comprendre la contribution du RLHv, une mesure du visage prédictive du jugement humain (Chapitre 3), j'ai réalisé une seconde étude qui avait pour objectif de distinguer les effets de chacune des composantes du RLHv sur la perception des traits sociaux. Cette mesure de la morphologie faciale est calculée à partir de la hauteur de la partie supérieure du visage (composante verticale) et de la largeur des bi-zygomatiques (composante horizontale). Pour cela, j'ai manipulé de manière orthogonale ces deux composantes afin de tester leur effet spécifique et combiné dans la formation de différentes impressions. J'ai observé que l'effet de la composante verticale dans les différentes expériences et conditions était plus important que l'effet de la composante horizontale.

Dans une troisième étude (chapitre 4), je présente trois expériences conduites chez des patients atteints d'un syndrome de Williams-Beuren (WS). Cette pathologie a été utilisée comme modèle neurobiologique humain dans lequel il existe un comportement d'appétence sociale. Les expériences visaient à répondre aux questions suivantes: les patients Williams sont-ils capable de détecter la confiance à partir d'un visage? Comment les patients Williams se représentent-il un visage qui inspire confiance? J'ai également étudié comment les corrélats neuronaux du traitement facial, les sources électrophysiologiques localisées dans les régions du sulcus temporal supérieur (STS), sont modulées par les indices sociaux d'un visage dans cette pathologie.

Dans une première éxperience, j'ai cherché à étudier comment la perception de la confiance est altérée chez ces patients. J'ai observé que les patients Williams regardent mois longtemps les visages qui inspirent confiance, suggérant qu'ils ont une tendance à davantage faire confiance à tout le monde. Nous avons également observé une grande variabilité inter patients durant l'exploration spontanée des visages.

Afin de mieux comprendre comment les patients Williams forment une représentation d'un visage digne de confiance, j'ai utilisé un paradigme de corrélation inverse. L'intérêt de cette technique réside dans le fait qu'elle prévient tout biais expérimental a priori sur les indices faciaux (tels que la forme de sourcils, des pommettes tel que modulé dans la base de donnée de Todorov) connu pour être des facteurs importants lors de la détection de confiance. Cette procédure force les sujets à sélectionner à partir du bruit les caractéristiques faciales qu'ils pensent être importants pour le jugement demandé. Nos résultats démontrent qu'en comparaison à un groupe sain apparié chaque patient Williams peut former une représentation différente du hasard. Par contre, au niveau du groupe ils ne présentent pas une image stéréotypique d'un visage qui inspire confiance en raison d'un faible accord inter-sujet. Cette technique n'a jamais été employée avec des patients et je pense qu'elle pourrait être utile comme outil clinique pour évaluer les représentations mentales subliminales dans différentes pathologies.

Dans une dernière étude, en utilisant une méthode de séparation de source à l'aveugle j'ai étudié si les sources neuronales éléctrophysiologiques, en particulier celles localisés dans le sulcus temporal supérieur (STS), pouvait expliquer la variabilité entre patients. Puis je voulais déterminer si cette activité était anormale en comparaison d'un groupe contrôle. Dans cette étude, les sujets devaient regarder des visages. Dans chaque essai, la fovéa du sujet étais dirigé directement dans une région différente du visage tel que les yeux, le nez, les sourcils, les pommettes, la mâchoire, etc.. J'ai observé que l'activité de la source localisée dans le STS survenant à environ 240ms était modulée de manière significative par rapport à la proximité des yeux: lorsque la fovéa des patients était alignée avec cette région l'activité de la source augmentait tandis qu'elle diminuait lorsque le regard s'éloignait des yeux. Ce pattern d'activité cérébrale est similaire à celui observé chez les sujets contrôles, suggérant que le comportement social altéré des patients Williams ne serait pas lié à un dysfonctionnement de cette région.

Dans la discussion générale, j'ai essayé: de montrer les liens et les différences entre les études que j'ai réalisées; d'apporter des interprétations et des conclusions des résultats obtenus; d'ouvrir d'autres horizons vers lesquelles les hypothèses de recherches pourraient s'orienter.

## Summary

The face is probably the most salient part of the body. Faces provide a large, complex and dynamic set of physical, emotional, and social information to the observer that essentially determines how people will interact with others, hence the importance of facial details underlying these abilities. From facial information, human subjects can form rapid, first impression judgments such as trustworthiness or attractiveness.

The ability to create social judgments from faces is the core topic of this work, with a special focus on trust detection from facial appearance. Trust is taking the risk of putting one's own fate in someone else's hands and is highly implicated in social exchanges. In daily life people constantly make rapid trust decision in different social contexts (e.g., where to sit in the metro, where to buy food, to whom to ask information in the street) and trust decision can be considered the basis of people' relationship as it is fundamental in mate choice, business cooperation and many others social exchanges. Thus, trustworthiness is clearly an essential skill useful at the individual and societal level.

The research presented in this thesis will focus on how social information and specifically trust is spontaneously detected from faces. In my studies I used eye tracking procedure, electrophysiology (EEG) and behavioral measures. Before explaining in detail the objective of each study, the first chapter will provide a short summary on facial anatomy: structure, muscles and nerves. Successively, a section will present the more recent literature on the neural bases of face processing. I will describe different theories proposing different neural networks and core regions associated with face processing in human and monkeys. Moreover, I will focus on the trustworthiness literature, by discussing important studies that have investigated face processing and the neural correlates associated with the detection of social traits from faces.

In a first experiment (Chapter 2), I investigated the evolutionary origin of trustworthiness detection in humans and non-humans primates. The objective here was to test whether monkeys (Macaca Mulatta, Macaca Fascicularis) have a spontaneous preference towards trustworthy human faces, thus suggesting a capacity to detect facial cues similar to those used by humans when making fast judgments of trust. Eye movements were recorded from both humans and monkeys during a face paired preference visual paradigm. We observed that both species spent more time looking at trustworthy faces than untrustworthy ones. We also found a significant correlation between facial features that predict judgments of

trustworthiness, i.e. happiness, femininity and facial width to height ratio, and humans' and monkeys' looking time.

To better understand the contribution of facial width to height ratio (fWHR), a facial metrics that predict humans' trustworthiness judgments (Chapter 3), I performed a second study with the goal to disentangle the effect of the two components of fWHR, namely, upper facial height (vertical component) and byzigomatic width (horizontal component), on the perception of different social traits. To this aim, I orthogonally manipulated the upper facial height (vertical component) and the bizygomatic width (horizontal component) to test the selective and the combined effect of fWHR components in the formation of different impression. Interestingly, I observed that the effect of vertical component on trust in all experiments and conditions was higher compared to the horizontal component effect.

In Chapter 4, I present three studies conducted in patients affected by Williams-Beuren syndrome (WS). This pathology can be considered a neurobiological human model for the overexpressed social behavior. I addressed the following question: are Williams-syndrome patients able to detect trustworthiness from faces? How WSP represent a trustworthy face? I have also investigated how neural correlates of face processing, particularly electrophysiological brain sources localized in the superior temporal sulcus, are modulated by social facial cues in this pathological condition.

In the first experiment I aimed to investigate how perception of trustworthiness is disrupted in this genetic and neurodevelopmental disorder. I observed that WS patients looked less the trustworthy faces compared to control group. This implicit behavior supports patients' tendency to trust everybody. I also observed great variability between patients during spontaneous face exploration. To better understand how WS patients form a representation of trustable faces I used a reverse correlation paradigm. The interest of using this technique is to avoid any experimental a priori bias on facial cues (such as shape of the eyebrow or chin as it was the case in Todorov's database) known to be important for the detection of trustworthiness. This procedure, in fact, pushes subjects to select from noise the facial features that they believe are important for a specific judgment. Our findings demonstrate that, although each WS patient could form a representation that was above chance, at group level they did not show a stereotypical image of trustworthy faces compared to healthy controls, probably because of a low inter-subject agreement. This technique has never been used in patients and I believe our result may be useful in designing novel clinical tools to investigate face subliminal representation of social traits in different pathologies. In a final study I investigate whether electrophysiological brain sources, with particular attention to one source localized in the superior temporal sulcus (STS), could explain patients' variability and whether this activity was abnormal in WS patients. In this study subjects looked at faces displaying a direct gaze. In each trial, subject's fovea was directed towards different parts of the face (eyes, nose, chin, eyebrow, cheek). I found that the activity of a source localized in the STS was significantly modulated by eye proximity. In other words, when patients' fovea was aligned with the eye region of the image activity increased in STS while the same activity decreased when patients' visual attention was brought fairway from the eyes. This pattern of brain activity was similar to that found in healthy controls, suggesting that exaggerated social behavior in WS may not be linked to dysfunctional activity in the STS region. Finally, in the last chapter of this thesis I propose a general discussion with the aim of highlighting links and differences between the studies I performed, drawing conclusions and open future avenues of research hypotheses.

# Chapter 1

# Face processing in evolutionary psychology and neuropsychology

## 1.1. Face anatomy: structure, muscles and nerves

The face is defined by bones, which constitute the structure, by muscles, which organize movements and expression and by nerves, which trigger muscle activity in response to brain signals or brain activity in response to face skin stimulation.

The face is a complex body organ presents in numerous species. Although similarity exists across individuals in terms of facial skeleton and muscles, the face remains the most important body part that allows people to discriminate different identities. In humans, the face extends from the forehead to the chin with different levels building up the facial components. Facial shapes are largely influenced by the skull size and principal bones' structure. Accordingly, the face can be divided into two main sections: upper facial height and byzigomatic width (Figure 1). The maxilla length and the nasion bone determine the upper facial height i.e. the length of the face. Zygomatic bones' variation and mandible's size change facial width.

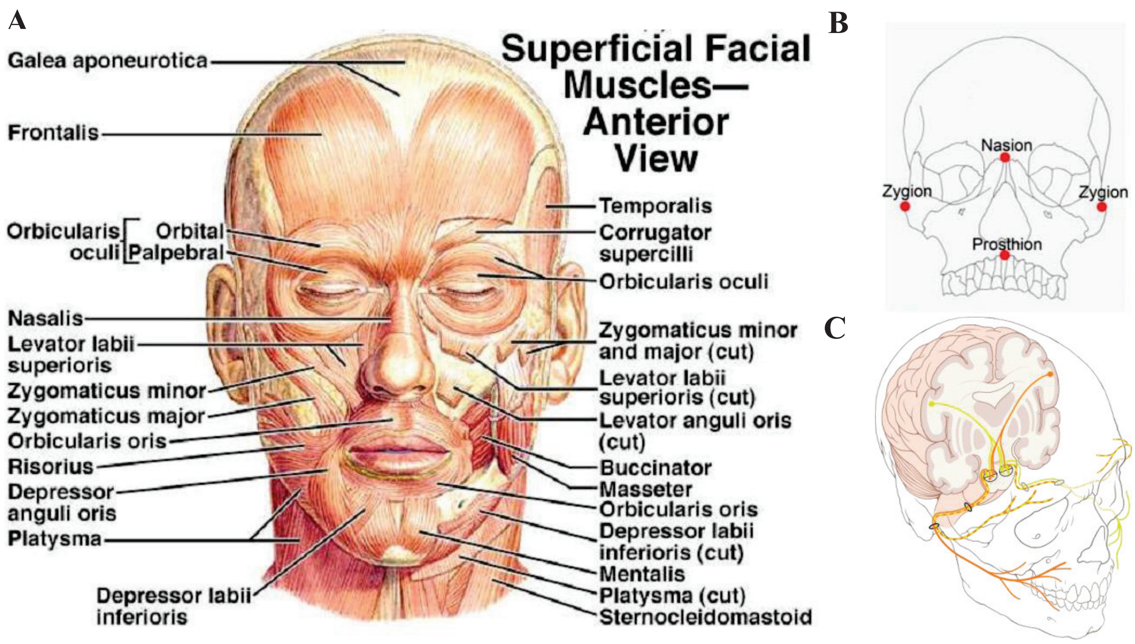



*Figure 1*. **A.** Anterior view of main facial muscles of the human face (Latham, 2012); **B.** Skeletal craniofacial variables relate to facial appearance (Weston, Friday, & Liò, 2007) **C.** Cranial nerve VII connection from the face to the brain.

The facial appearance is also determined by the muscle organization composed by more than one hundred muscles. Human facial muscles have been developed for surviving functions (self-protection) such as blinking, suckling, biting, or chewing. However, the same muscles shift functionality during social interactions as they wink, smile, or while producing other rather complex facial expressions to communicate emotions and intentions (social function).

On the upper face (Figure 1), the Frontalis muscle, is located on the frontal bone, attached to the eyebrows skin, it is mainly used for lifting the eyebrows, allowing movements such as looking up or expression of feelings. Other important muscles are the Temporalis that are involved in chewing and mastication; they are located one on each side of the head over the temporal bone. Moving on the front of the face, two muscles are named Orbicularis and the Orbicularis Oculi and the Orbicularis Oris. The former is posed around the eyes, it arises from the nasal part of the frontal bone, and it is a muscle involved in closing and opening the eyelids. Moving on the mid-section of the face the Orbicularis Oris is located around the mouth and contributes to mouth's movements. Surrounding the Orbicularis Oris, the Zygomaticus major and minor muscles are strongly involved in facial expressions. While, the major is also called the "smiling muscles" because it draws the mouth angle allowing a smile, the minor draws the upper lip backward, allowing expressions of sadness. Hence these muscles play antagonistic roles in facial expression. Another important muscle is the Buccinator, also called fish or kiss muscles, a large muscle along the Maxilla and the Mandible, with a function to protrude the lips for kissing or blowing, but also useful in chewing and keeping food. The Masseter is close to the Buccinator and represents the bigger muscle used for chewing and mastication. Still in the mid-part of the face the Levator labii superioris, between the nose and the zygomatic bone, useful for elevating the upper lip, also involved in sadness expressions. Straight down from the Orbicularis Oculi, the Depressor labii inferioris allows the bottom lip to go down. Close to this one, the Depressor anguli oris muscle is used for frowing while the antagonist of the Byzigomatic major muscles is used for smiling. Finally, the Sternocleidomastoid and splenius capitus are fundamental for head motion.

In sum, facial muscles can be classified for their main functions: the Masseter, Temporalis, Medial pterygoid and lateral pterygoid are used for mastication. Importantly, the muscles primarily used to produce facial expression are Frontalis and Occipitalis, which are differently coordinated to generate facial expressions.

The face is marked by an extremely sensitive function with 4 out of the main 5 senses located in the face. Human perception in part is produced thanks to the presence of five principal senses: touch, smell, taste, hear and vision. Cranial nerves I, II, VIII are respectively called olfactory, optic and vestibulocochlear nerves for their function linked with humans' sense (smell, vision, and hearing). All other facial nerves have motor and sensory functions. Amongst them, the most important for facial movements is the Cranial nerve VII, also named as facial nerve. It emerges from the brainstem, between the pons and the medulla, and controls muscles involved in facial expressions. The facial nerves have three nuclei:

1. Facial motor nucleus receives inputs from the primary motor cortex, innervating the muscles of facial expression.

2. Superior salivatory nucleus receives afferent inputs from the thalamus, while the lacrimal nucleus receives afferent signals from the hypothalamus and regulates emotional responses.

3. Solitary nucleus receives as input taste and sensorial information and transports information to the cortex.

For the present work, I will consider the importance of the facial structure in social judgments but also the link between social traits and emotion. The underlying interests of the thesis it is to further understand how social traits are expressed in the face by the combination of structural and emotional information and how this information is processed at the brain level.

## 1.2. Face processing in non-human and human primates

The importance of facial information in social interactions and survival has been revealed in human and monkey's kingdom suggesting that similar mechanisms co-exist in both species. Non-human primates are strongly attracted by faces (Deaner, Khera, & Platt, 2005; Ferrari, Paukner, Ionica, & Suomi, 2009; Sugita, 2008). This faculty is innate as newborn rhesus macaques deprived from seeing their mother's or caregivers faces showed an innate preference for faces compared to objects (Sugita, 2008). Other animal research has shown that newborns monkeys imitate facial movement such as tongue protrusion or movements with emotional content like lip-smacking (Ferrari et al., 2006).

Several studies showed that Macaque monkeys and humans are sensitive to some common facial features. Rhesus macaques are sensitive to identity (Katalin M Gothard, Brooks, & Peterson, 2009; Parr, 2011), they show preferences for conspecific compare to other species (Dufour, Pascalis, & Petit, 2006; Méary, Li, Li, Guo, & Pascalis, 2014; Sigala, Logothetis, & Rainer, 2011), although also interested in human faces. Rhesus macaques are sensitive to familiarity (Leonard, Blumenthal, Gothard, & Hoffman, 2012), and detect facial expression in conspecific (Mosher, Zimmerman, & Gothard, 2014). They are able to discriminate gazes (Allison, Puce, & McCarthy, 2000; Ferrari, Kohler, Fogassi, & Gallese, 2000), showing a preference for direct gaze compared to adverted gaze (Gibboni, Zimmerman, & Gothard, 2009; Mosher, Zimmerman, & Gothard, 2011).

In humans, face processing is enriched with complex social details and norms. The ability to readily access information about a person is essential to determine how people will interact with other individuals. With this aim, human faces are processed in a fast and accurate way, as they are the most informative source of social information.

Facial preference is an innate skill evolved early in development (Pascalis, de Haan, & Nelson, 2002; Pascalis & Kelly, 2009; Paul C Quinn et al., 2011). After few days of life, newborns already prefer to spend time looking at faces compared to objects (Johnson, Dziurawiec, Ellis, & Morton, 1991; Johnson, 2005). Newborns also imitate others' facial movement (Meltzoff & Moore, 1977) such as tongue protrusion, show an early interest (2 months) for others´ eyes (Warren Jones & Klin, 2013; Maurer & Salapatek, 1976) and develop the ability to follow others' gaze from 3 to 12 months (Corkum & Moore, 1998; Infants, Yu, Smith, Yu, & Smith, 2016). This innate preference for faces in humans is then shaped by social interactions. Babies for instance preferred to look at happy faces compared to neutral or fearful ones (Farroni, Menon, Rigato, & Johnson, 2007) and chose to spend longer times staring feminine faces (P C Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002; Paul C. Quinn et al., 2010). These results may be explained by the fact that babies are more exposed to females faces from the first months of their life as they represent rewarding stimuli satisfy, important to satisfy the infants' needs.

The ability to learn new faces improves with aging (until 30 years old)(Germine, Duchaine, & Nakayama). The information detected in faces largely influences human social behavior and relies on different cognitive processes. First, to establish whether a face is familiar or unfamiliar is an essential skill. Compared to children, adults are exposed to an incredible number of faces; famous characters from television, politicians, colleagues from different work's context. When encountering a familiar face different processes are immediately activated: remembering the name, the emotional feelings associated to the person, memories and information about the person's life (semantic knowledge) and so on

(Bruce & Young, 1986; Todorov, Gobbini, Evans, & Haxby, 2007). Adults develop a vocabulary of facial gesture that they can use to understand, predict and interact with people of their own social circle.

Importantly, faces are informative even when we see someone for the first time in our life. Humans have evolved mechanisms to process facial features that are dissociated from the recognition of a specific identity, such as detection of emotions, intentions or personality traits. Individuals do not need to know a person to understand if the person is happy or unhappy, if something scared her/him or if she/he is angry for unknown reasons. Interestingly, by looking at the eyes or following the gaze, individuals rapidly understand others' will and can predict others' actions in the imminent future(Saxe & Kanwisher, 2003; Schilbach et al., 2006). Lastly, but central for the current work, individuals quickly create first impression judgments based on facial inferences. People infer social traits such as trustworthiness, attractiveness or dominance after short time of exposure to faces; this information driven by faces' structure does not necessarily gain consciousness, though it strongly bias people' perception. (Olivola & Todorov, 2010b; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006).

Cognitive psychology and neuroscience are still puzzling on how face processing is acquired, used for social interactions and modified in non-human and human primates. I will start describing the first cognitive model on face processing that has been proposed by Bruce and Young (1986). Thanks to the increasing number of neuroimaging and electrophysiological experiments, different models have been designed with the aim of explaining which brain regions are selectively or differently implicated in each of the steps involved in facial processing with social motives.

Thus, I will present the brain networks that are specialized in face processing in non-human and human primates with a detailed summary of the more recent literature in the domain. I will then focus on the ability to make trustworthiness judgments from faces. Finally, I will provide some evidence about face processing in patients with brain lesions and in a neuropsychiatric disorder as Williams-Beuren syndrome.

### 1.2.1. Cognitive models of face perception and recognition

Over the past three decades, scientists have shown interest on how human process faces. Several years ago, Bruce and Young (1986) proposed the first model of face processing, still considered an influential one.

Theoretically, facial processing can be considered as the result of different modules in the brain each exploiting a different function (Bruce & Young, 1986). Among the models proposed, Bruce and Young categorized four principal modules: the Face Recognition Unit (FRU), the Person Identity Nodes (PIN), the Semantic information units (SIUs), the Name Unit.

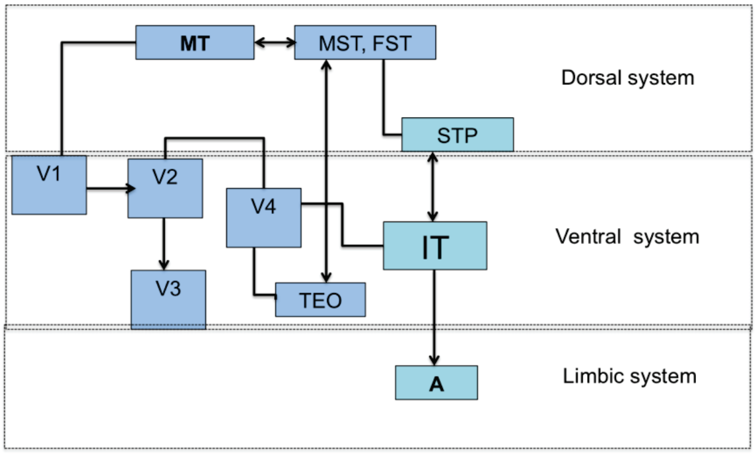
In their behavioral model, Bruce and Young suggest the different facial modules organization is hierarchical. After a *structural encoding of faces*, which depends on vision and facial configuration, facial information is processed by separate systems hierarchically. PIN perceives personal identity; once the identity has been established, the Name Unit modules retrieve name and semantic information about the person (SIU). At the same time, different modules are able to process facial expressions and speech-related mouth movements in parallel. The first modules (FRU, PIN, SIUs) are activated by the presence of a familiar person, the latter are sensible to unfamiliar faces too. Few years later, 1990, Burton proposed an updated version of Bruce and Young model indicating the presence of an independent pathway for the emotional response within facial detection that is parallel to the pathway to access person identity (Bauer, 1984). The contribution of these models is unique since they proposed dissociation between recognition of identity and detection of emotion, being one of the most ancient dissociation in face processing research.

The presence of dissociated functions between these modules was based on behavioral observations and described as a priming effect that facilitates one task and not the other. It has been shown for instance that familiarity priming facilitates performance during an identity recognition task but not during a task involving detection of facial expression (Bruce & Young, 1986) The dissociation between different face processing have been further supported by neuroimaging studies and evidences from patients. These two points will be discussed in the next sections.

## 1.3. Brain circuits of facial processing in non-human and human primates

Neuroimaging and electrophysiological studies lead to support, discuss and extend the cognitive models previously proposed to pinpoint several brain regions that are selectively responsive to faces or implicated in processing facial information in human and non-human primates. In the first section I will propose a comparison of the core region of face processing observed in both species; that is the brain regions that selectively respond to faces or to specific facial features.  In the second one, I will present the brain regions that have been observed as a part of the extended system of face processing (Bernstein & Yovel, 2015; Gobbini & Haxby, 2007; Natu & O'Toole, 2011). In other words, the regions that connected to the core system exploit more complete face processing tasks as the process of social, emotional and semantic content of faces. A synthesis that provides the entire picture of these networks is provided for monkeys and humans in Figure 2.

**A** **Monkeys network**




**B** **Human network**

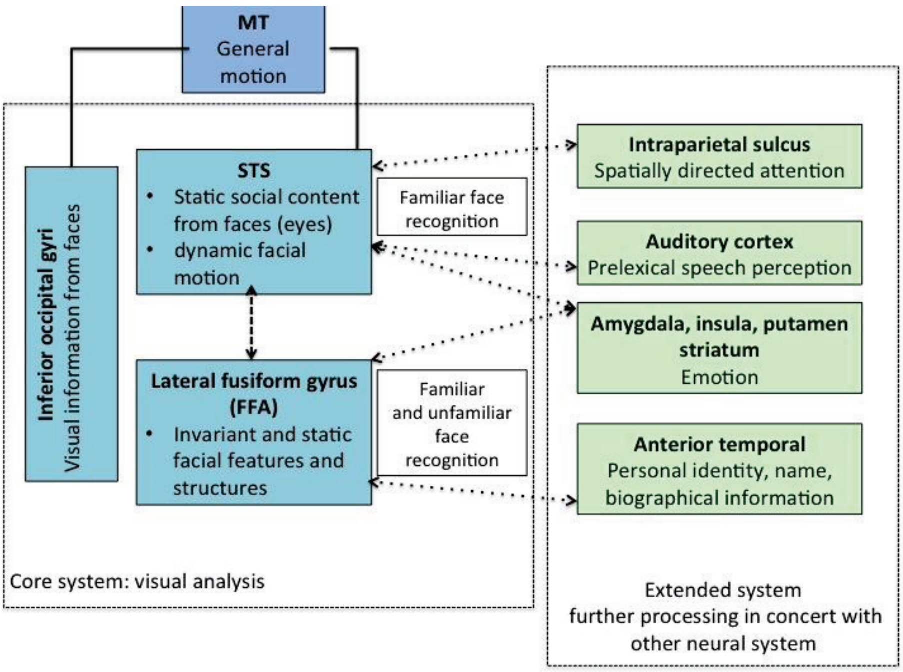



*Figure 2.* **Face processing networks in monkeys and human A.Monkey network.** Some connections between V1 and brain areas in dorsal visual system, ventral visual system and limbic system. In light blue the regions in which face selective neurons have been found and that can be considered the 'core system' of face processing. **B.Human network**. Core system and extended system in human brain and their related functions. Integration of models as reported by (Bernstein & Yovel, 2015; Haxby, Hoffman, & Gobbini, 2000; Natu & O'Toole, 2011)

### 1.3.1. The core system of face processing in monkeys and humans

From the visual system to specific facial processing areas, several similarities have been identified between monkeys and the human brain(Sereno & Tootell, 2005; D. Y. Tsao, Moeller, & Freiwald, 2008; Van Essen, 2004). Different studies have tried to understand which face-selective areas in the monkey's brain may be considered homologues to the ones found in humans. Such information is key for understanding early markers of facial cues in human evolution.

In both species, one of the most used neuroimaging paradigm to detect regions that are selective to face presentation is to display sequences of faces compared to sequences of other objects (in monkeys, (Doris Y Tsao, Freiwald, Knutsen, Mandeville, & Tootell, 2003) (in humans, (Kanwisher, McDermott, & Chun, 1997; Sergent, Ohta, & Macdonald, 1992). In both species, functional magnetic resonance revealed the cortical regions that display an increased blood flow when images of faces were compared to other stimuli identifying face-selective patches. Then, in non–human primates electrophysiological recording have been used as main technique to characterize face cells properties (Figure 3).

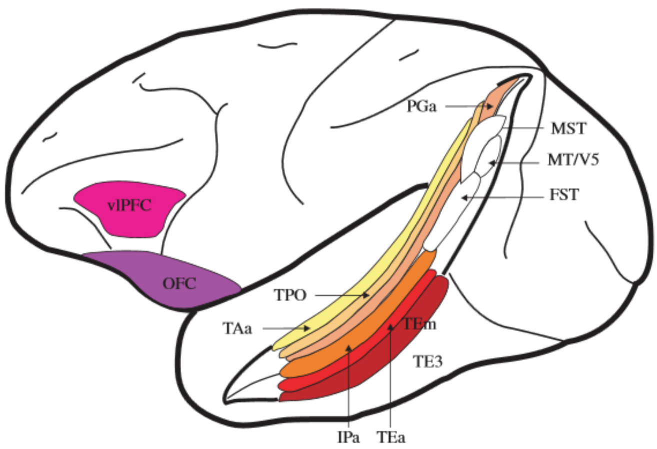



*Figure 3*. Cortical regions in the monkey's brain that contain face-sensitive cells (Nick E Barraclough & Perrett, 2011).

*The core system in monkeys*

The superior temporal sulcus (STS) and the inferior temporal (IT) cortex are considered the primary regions in the monkey brain, both regions implicated in the dorsal and ventral visual streams respectively. Face-selective patches in monkeys suggest dissociation between face cells in the superior temporal sulcus (STS) and the inferior temporal (IT) cortex. Face patches in monkeys located nearby or within the middle STS patch were reported as the most active areas sensitive for facial information (Doris Y Tsao et al., 2003).These neurons are involved in facial movement perception and changeable aspects of static faces, such as expression or different orientations of eyes and head showing a similar function of the dorsal pathway in humans (Hasselmo, Rolls, & Baylis, 1989; Mormann et al., 2008; Perrett, Smith, & Potter, 1985)

Differently, neurons in the monkey´s IT region are more likely to respond to facial identity (Hasselmo et al., 1989; Ku, Tolias, Logothetis, & Goense, 2011; Sliwa, Plante, Duhamel, & Wirth, 2016). However less clear evidence is available on which brain regions may correspond to the ventral stream homologous in humans. A 7T fMRI study reported several regions along the ventral temporal cortex that may correspond to the ventral stream in humans (Ku et al., 2011), including the TGa, AMTS, perirhinal and enthorhinal cortices. Some of them were already found in previous electrophysiology or fMRI studies (Nakamura and Kubota 1996, Logothesis 1999, Tsao 2003). Moreover, they found face selective areas around the AMTS at AP 17-21, paraH, vV4.

*The core system in humans*

In the human brain, three bilateral regions in the occipito-temporal extra striate cortex are consistently reported in face processing experiments: *inferior occipital gyri* (IOG) (OFA, Halgren et al 1999, Haxby 1999, Kanwisher 1997, Kanvisher & Yovel 2006; McCarthy et al 1997; Puce, Allison, Asgari, Gore, McCarthy, 1996), the *lateral fusiform gyrus* (FFA)(Kanwisher et al., 1997; Sergent et al., 1992) and the *superior temporal sulcus* (STS) (Fox, Moon, Iaria, & Barton, 2009; Haxby et al., 2000) ( Figure 4).

Justine Sergent in 1992 for the first time described the specificity of a region in the lateral fusiform gyrus that responded specifically to faces. Later, this region was named by Nancy Kanwisher, in 1997, fusiform face area (FFA). She proposed that the existence of the FFA was specific for face detection, confirming the presence of domain specificity in the visual system. The STS role in face processing is also largely accepted (Puce et al. 1996; Haxby et al. 2000; Pitcher et al. 2011), mainly supporting a function sensitive to eye and biological motion (Bonda et al. 1996; Allison et al. 2000; Grossman et al. 2000; Grossman and Blake 2002; Pelphrey, Mitchell, et al. 2003; Pelphrey et al. 2005).


*Figure 4.* **Functional correspondence between human and monkey face areas**. Six face patches that have been consistently found by fMRI studies. They are referred as the posterior later (PL), medial lateral (ML, medial fundus MF), anterior lateral (AL), anterior fundus (AF), and anterior medial (AM) face patch. B. Anatomical location of the three core-face selective areas in humans: the fusiform face area (FFA), the occipital face area (OFA) and the face selective area in the posterior temporal sulcus (p-STS-FA). Pictures from (Yovel & Freiwald, 2013).

**Table 1.** *Correspondence between human and monkey face areas based on relative and absolute anatomical location and connectivity from* (Yovel & Freiwald, 2013).

| **Scenario 1** | | | | | | |
|---|---|---|---|---|---|---|
| | | **Posterior-anterior axis** | | | | **Species** |
| **Dorso-Ventral Axis** | dorsal | | **MF** | **AF** | | Macaque |
| | | | **pSTS** | **aSTS** | | Human |
| | ventral | **PL** | **ML** | **AL** | **AM** | Macaque |
| | | **OFA** | **pFFA** | **mFFA** | **AFP** | Human |

| **Scenario 2** | | | | | |
|---|---|---|---|---|---|
| | | **Posterior-anterior axis** | | | **Species** |
| **Dorso-Ventral Axis** | dorsal | | **mSTS** | **aSTS** | Macaque |
| | | | **pSTS** | **aSTS** | Human |
| | ventral | **vV4** | **TF** | **AMTS** | Macaque |
| | | **OFA** | **FFA** | **AFP** | Human |

| **Scenario 3** | | | |
|---|---|---|---|
| | | **Posterior-anterior axis** | **Species** |
| **Dorso-Ventral Axis** | dorsal | **–** | Macaque |
| | | **STS-FA(s)** | Human |
| | ventral | **PL-ML/MF-AL/AF-AM** | Macaque |
| | | **OFA - FFA - AFP** | Human |

Three scenarios indicating putative correspondences between macaque and human face areas along the posterior-anterior and dorso-ventral anatomical axis

Haxby and colleagues (2000) proposed a model based on neuroimaging data to explain the organization of the core brain regions selective for facial processing in humans. In continuity with the earlier cognitive models (Bruce and Young's 1986), they defined a hierarchical view of the visual system with a core system for facial visual analysis and an extended system that encodes the content and meaning of this information. More precisely, facial information is sent from visual areas (OFA) towards the STS and FFA, specialized in different aspects of face processing. The role of the ventral pathway (including OFA-FFA) is in charge of static and *invariant information* from faces, to create a representation of facial identities. Differently, the posterior STS (p-STS) is sensitive to dynamic and *changeable aspects* and is implicated in detection of facial expressions, eye movements and speech (Puce et al. 1996; Haxby et al. 2000; Pitcher et al. 2011; Bonda et al. 1996; Allison et al. 2000; Grossman et al. 2000; Grossman and Blake 2002; Pelphrey, Mitchell, et al. 2003; Pelphrey et al. 2005). According to this model (summarized in Figure 2), key information obtained from familiar faces are processed mainly by the ventral pathway, while the dorsal visual stream processes information from unfamiliar faces.

The current dominant neural models in human studies have accepted the division of these brain structures divided in two neural pathways for facial processing (Haxby et al., 2000). The ventral visual stream, with the FFA as central hub, is dedicated to facial trait encoding and the dorsal visual stream, being the STS the principal region, is in contrast specialized in processing facial states (Haxby et al., 2000). However, new models have been proposed to integrate other areas and different functions of the regions mentioned above during face processing tasks.

In 2002, O'Toole et al proposed an extension of Haxby's model expanding the STS role in facial information (O'Toole, Roark, and Abdi 2002; for a more recent review Natu and O'Toole 2011). They claimed that both FFA and STS have a role in recognition of *familiar faces*. Nevertheless, these two brain regions use different strategies. The STS is sensitive in perceiving dynamic information and can have access to dynamic features of familiar people, already stored in the brain. Differently, the FFA uses facial structural information and stable facial features to recognize identity and store the new ones. The selective ability of FFA in identity recognition is not therefore questioned; in fact, this region is active during processing of familiar and unfamiliar faces.

Interestingly, O'Toole and colleagues also consider a possible role of MT in general motion that may influence face processing. Evidence from patients with bilateral damage to the ventral occipital cortex, including MT and fusiform gyrus, suggests these patients are better in matching identity of moving faces compared to static ones (Lander & Butcher, 2015; Lander, Humphreys, & Bruce, 2004). Their results also support the existence of MT and STS connectivity. The STS location in the brain, close to MT, support the presence of high connectivity between these two regions and the role of STS for processing dynamic facial cues or feature that signal potential social interaction.

### 1.3.2. The extended face-processing system in monkeys and human

*The extended system in monkeys*

Connection between core regions and extended ones may be similar in both species (Schwierdrzik, 2015). In monkeys, the core regions previously presented (STS-TE-TEO-IT-AF-AM) are embedded in a larger network that represents the extended system of face processing (Figure 2). For instance, previous single-unit studies in monkeys showed increased neural amygdala responses selective for faces (Leonard et al. 1985; Rolls 1984; Sanghera et al. 1979, Logothethis 1999, Hoffman, Gothard 2007). Gothard and colleagues (2007) found in the amygdala a large subsection of face-selective neurons (64%) responsive to both expression and identity. They suggested that expression and identity are joint processes rather than separate independent components of facial processing. Overall, they concluded that neurons in the amygdala are able to detect complex visual stimuli that are socially relevant and that this region is sensitive to detect stimuli valence (they also found response for human faces). Finally, face selective response in the amygdala was confirmed in several fMRI studies (Ku et al., 2011). Ku and colleagues found activation in several subcortical regions not only in the amygdala but also in the hippocampus, cingulate and insula. See figure 5 for percent signal change according to different cue presentation (face, fruit, fractal or houses) showing increase activity over the regions mentioned above, while monkeys were passively viewing different stimuli.

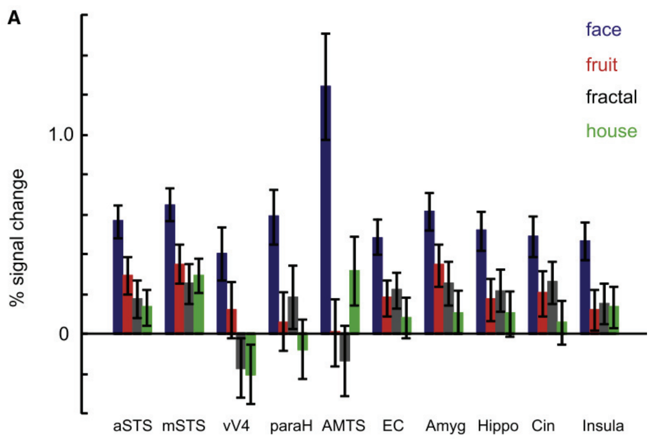



*Figure 5.* Percentage of BOLD signal change found in selective areas in response to different stimulus categories in two monkeys (Ku et al., 2011).

This “valence effect” was confirmed by Hadj-Bouziane (Hadj-Bouziane et al., 2012). In their work, control monkeys showed an increased bilateral activity in IT cortex, TE and TEO, STS and amygdala for facial expressions of emotion (lipsmack, fear and treat) compared with neutral stimuli. Interestingly, amygdala lesions disrupted the activity in the IT cortex while perceiving facial expression. On the contrary, the lesion did not produce a deficit in facial detection as neurons still responded more to normal faces compared to scrambled faces mainly in two region TE and TEO. Overall, these findings confirm the modulatory role of the amygdala over IT cortex during valence perception in faces and suggest that connections between the core system and the extended one may be similar in across species.

*The extended system in humans*

The extended system in humans has been also conceptualized into models (Haxby 2000, 2007, O’Toole 2011) to explain face processing task. Meeting a person we know, in fact, determines different reactions such as emotional responses or intentional attributions that support the activation of social brain regions. To recognize a familiar person doesn’t call only on the activity of visual areas (OFA, FFA) but also on regions implicated in social and cognitive functions. Lateral FG is connected to the anterior temporal regions involved in coding personal identity, name and biographical information. Connections from the STS to the intra-parietal sulcus (IPS) are implied in coding spatial directed attention and gaze. The STS is also connected to the auditory cortex for facial speech, and to the amygdala, insula and limbic system for processing emotional information from faces (see extended system in Figure 2). In a work published in 2007, Gobbini and Haxby, emphasized the importance of semantic, episodic and emotional information that are triggered by familiar faces, placing models on face processing in a wider social and cognitive context. Anterior paracingulate cortex (PAC), posterior superior temporal sulcus (pSTS), temporal parietal junction (TPJ), and the precuneus (Pcun) are some of the main regions. These regions have different functions: PAC is mainly involved in interpreting the mental state of other (Calder 2002, Frith and Frith 1999) and in the representation of personal trait (see what does it means in Mitchell et al 2002). Differently, pSTS and TPJ have been found to have a role in social cognition such as the evaluation of of others’ intentions (Allison, 2000; Hoffman & Haxby 2000, Puce and Perret, 2003, Winston, Strange & Dolan 2002, Saxe 2003).

For example, activities in PAC, pSTS and Pcun were found when contrasting personally familiar face, (family or friends) vs famous familiar faces (Gobbini et at, 2004) or while seeing own child vs unrelated child (Leibenluft, 2004). In both experiments, FFA was also active and more activated by personal familiar faces (Gobbini & Haxby 2007). Moreover higher amygdala activation has been found when mothers saw their child compared with the other categories of faces. These results showed activity in regions that strongly correlated with emotional modulation (as shown by amygdala and insula activity) but also have a role in perceiving familiar faces that are emotionally relevant (Shah et al 2001, Pierce, Haist, Sedaghat and Courchesne, 2004; Gobbini et al 2004; Leibenluft, Gobbini, Harrison and Haxby 2004, Sugiura 2001).

## 1.4. Multiple function of the superior temporal sulcus (STS) in human domain

The role of STS in face processing is largely accepted (Puce et al. 1996; Haxby et al. 2000; Pitcher et al. 2011). The STS is sensitive to eye and speech-mouth movements (in linguistic processing (Binder et al. 1997; Vigneau et al. 2006; Fedorenko et al. 2012) but also to voice (Belin et al. 2000), and biological motion (Bonda et al. 1996; Allison et al. 2000; Grossman et al. 2000; Grossman and Blake 2002; Pelphrey, Mitchell, et al. 2003; Pelphrey et al. 2005). Moreover activations have been found during task implying understanding of others' intentions and mental states (Fletcher et al. 1995; Gallagher et al. 2000; Saxe and Kanwisher 2003; Saxe and Powell 2006; Ciaramidaro et al. 2007). STS have been also pointed out as an integrative region that process social information in multimodalities audiovisual integration (Calvert et al. 2001; Beauchamp et al. 2004; Taylor et al. 2006), and the control of visual attention (Corbetta and Shulman 2002).

Interestingly in a recent work, Ben Deen and colleagues tried to provide a functional organization of the STS testing the same subjects in different tasks already used in previous experiment (Deen, Koldewyn, Kanwisher, & Saxe, 2015a). They compare the activity of STS during face perception, biological motion perception, mental state understanding, and linguistic processing and voice perception (Figure 6).

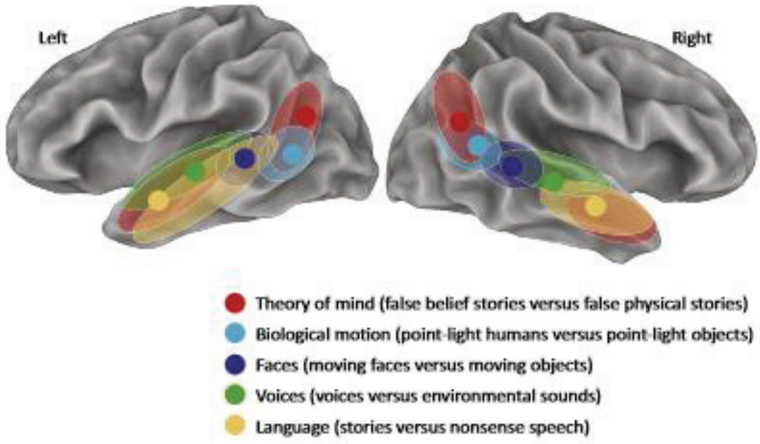



*Figure 6.* **Organization of social perception and cognition within the superior temporal sulcus (STS).** Results of Deen et al., shown on an inflated cortical surface model of the left and right hemisphere of the N27 atlas brain. Filled circles show the location of the peak activation, averaged across subjects, for each contrast. Colored regions show the extent of the activation for each contrast. Figure based on information and data in (Deen, Koldewyn, Kanwisher, & Saxe, 2015b).

Overall this work needs to be considered a step forward in the understanding of the role of STS in social processing confirming that STS is a region strongly implicated in social cognition tasks. They showed that it is possible to discriminate subregions that have a specialized function while others areas of overlapping, i.e. sub regions that are implicated in multiple tasks, or that play an integrative role. In sum, results reveled a bilaterally responses in the angular gyrus (region previously called TPJ, Saxe and Kanwisher 2003) during Tom contrast. Responses to language activated different regions from the angular gyrus to middle to anterior STS that as expected were higher in the left hemisphere. Voice contrast activated the middle STS. Interestingly, face processing and biological motion contrast showed activities in the pSTS (Pelphrey, Mitchell, et al.2003; Shultz et al. 2011). The region for biological motion was centered slightly posteriorly in most subjects. In sum the results showed the following STS functional regions: the TPJ response to ToM, pSTS response to biological motion, pSTS response to faces, middle STS respond to voices and anterior STS respond to language. Different overlapping spots have been found. The most strongly significant contrasts have been found between face and voice responses (mean 36%); for face and biological motion responses (30%). Dynamic face and vocal sound was the higher contrast detected, suggesting that STS is fundamentally an audiovisual area.

## 1.5. Overlapping function between the dorsal and ventral visual streams in human

The dissociation between the ventral and the dorsal stream on facial processing presented before may be an oversimplification on how the brain implements face processing. The main point that has been addressed is whether perception of facial identity and expression reflect a straightforward bifurcation along the visual pathway. Neurophysiological, neuropsychological and imaging studies in non-human primates showed that there is a degree of neuronal separation between these two mechanisms. In a review, Calder and Young (2005) proposed that the separation between identity and expression recognition is relative rather than absolute (Calder & Young, 2005).

Overall, the neuroimaging results agree on the fact that different functions can be attributed to these two regions (Bernstein & Yovel, 2015). However subtle but essential changes can be proposed considering the results of the last decade.  Some claims of Haxby and O'Toole model are still accepted. Recent studies confirmed for instance the selective role of FFA in detect facial identity from static faces. FFA presented higher activity when different identities were presented, differently less activity was recorded during the repetition of the same one (fMR-adaptation paradigm, Baseler 2013). Coherently, STS did not show any difference in the response to different face then same identity face (Mazard et al 2005). Interestingly, aTMS study reported no effect on facial identity recognition task after stimulation was applied over the STS, while impairment was reported for detection of facial expression (Pitcher 2014).

Another accepted claim is that STS has a role in processing dynamic facial features. Importantly, the STS is not sensitive to any kind of motion (as MT do), it plays a selective role in biological motion (Grossman et al 2000) that explain the extension of its function in social and communicative tasks (Beauchamp, 2002). Another important function of the STS is the integration of different cognitive, social and perceptual modalities(N E Barraclough, Xiao, Baker, Oram, & Perrett, 2005; Beauchamp, Argall, Bodurka, Duyn, & Martin, 2004). Calder and Young proposed a main role for STS in coding facial expression but also in those changeable facial cues which might require online monitoring during social exchange, as opposed to facial identity (Simons & Levin, 1998). However, STS do not only show preferential activity during dynamic faces compared to static ones (Lee et al 2010; Fox 2009, Pitcher 2011) but this region is also sensitive to facial cues that are important for social interaction such as eyes and emotional aspects, that are susceptible to move (Puce 2000).

In sum, as suggested by Bernstein et al (2015), it can be considered the existence of a motion sensitive stream, i.e. the dorsal one that includes pSTS-FA, aSTS-FA and IFG-FA and a ventral stream that is not sensitive to motion, including the OFA and FFA. Recently, it has been shown a larger amount of connections between the OFA and FFA as compared to the OFA and STS (Gschwind et al 2012; Pyles 2013), which has been confirmed using connectivity analysis (Avidan 2014). A causality approach using TMS (Pitcher, Duchaine, & Walsh, 2014) showed that disruption of the rOFA reduced the neural response towards static and dynamic faces in the fusiform gyrus. On the contrary response in the rpSTS was still present for dynamic faces only.

## 1.6. Face processing after brain lesions and in psychiatric disorders

Perception of faces can be directly altered after focal brain lesions or following psychiatric states. The most direct evidence is obtained from patient's behavior. Prosopoagnosic patients, for instance, no longer recognize faces but they can still recognize emotional expressions and trust from faces (Sergent et al., 1992; Todorov & Duchaine, 2008). These patients also present skin conduction changes when presented with familiar faces (Bauer, 1984, 1986; De Haan, Bauer & Greve, 1992, Tranel& Damasio, 1985). In contrast, patients with the Capgras's syndrome deny recognizing familiar faces; they are in fact obsessed thinking that all familiar people has been replaced by alien and impostors (Capgras&Reboul-Lachaux, 1923; Ellis and Young, 1990; Hirstein & Ramarchandran, 1997). These patients for instance do not either show skin conductance changes signaling that also the emotional pathway is impaired (Ellis, Young, Quayle & DePauw, 1997).

The emotional content processing is somehow preserved in patients affected by amnesia. For instance, Karsakoff patients cannot hold memories of having seen certain faces but they are able to categorize faces as either pleasant or unpleasant (Johnson, Kim& Risse 1985). In Alzheimer's disease, patients forget sematic information about people or they even forget to have met the person, however they still manage to detect facial emotions or perform judgments of trustworthiness faces (Burnham, H & Hogervorst, E, 2004). Thus, patients' behavior confirms the dissociations already suggested in Bruce and Young model and provide clear evidence for facial detection failures associated with certain altered neural circuits. Most of these aspects of face processing received support by neurological findings ten years later (Haxby et al., 2000). Following studies conducted using neuroimaging techniques collected wider information on which brain regions and networks are responsible in different face processing behaviors and how the brain orchestrates processing of different facial information.

Neurologically damaged patients are strongly informative about the dissociated functions of face processing as they establish brain-behavior interactions.

Patients who present face recognition impairments suffer from lesions over the right inferior occipital cortex, along the territory of the right "occipital face area" (OFA), suggesting that this region is causally implicated in facial identification (Steeves, Dricot, & Rossion, 2004). Patients with lesions on ventral visual cortex proofed that FFA is not crucial for the perception of biological motion (Gilaie-Dotan et al 2015). Damasio and Tranel (1993) as described in a single-case report with damage over medial and lateral prefrontal cortex. In normal conditions, these types of lesions impair patients in learning new identities or semantic information about a new person. Nevertheless, the patient reported behaved coherently with the positive or negative experiences he had with the person he interacted with. On the other hand, patients with bilateral amygdala damage are impaired in an emotional recognition task or while judging social traits from faces, but they do not exhibit problems in recognizing identities (Adolphs, Tranel, & Damasio, 1998). Interestingly, the ability to perform trustworthiness judgments has been considered in different pathologies. For instance, the involvements of the amygdala in the evaluation of trustworthiness faces have been firstly pointing out by Adolph and colleagues (1998), who showed that the ability to detect untrustworthy faces was impaired in patients with bilateral amygdala damage.

The ability to make judgments of trustworthiness is preserved in prosopoagnosic patients, which are severely impaired in the perception of facial identity and in face memory task (Todorov & Duchaine, 2008). This result confirmed the dissociation between mechanisms that are involved in forming first person impression and mechanisms for encoding face identity.

### 1.6.1. Abnormal face processing in Williams-Beuren syndrome

Face processing is also abnormal in neuropsychiatric disorder such as Williams-Beuren syndrome (WS). WS is a rare neurodevelopmental disorder caused by chromosomal abnormality (S. Porter, ten Brinke, & Gustaw, 2010). WS affects 1 in around 20,000 people worldwide. Children affected by this pathology present special facial appearance (Figure 7), complex clinical dysfunctions and distinctive cognitive profile (Barak & Feng, 2016; Bellugi, Bihrle, Jernigan, Trauner, & Doherty, 1990; Meyer-Lindenberg, Mervis, & Berman, 2006).

*Figure 7.* **The facial appearance of children with WS**. Composite picture derived by using photographs of 12 different individual with WS (Tiddeman, Brut Perret 2001).

From the clinical aspects they display cardiovascular and gastrointestinal problems but also neurological problem that impact coordination and walk. They also present hypersensitivity to sounds. The cognitive profiles, defined with peaks and valleys, reveal impaired cognitive functions such as visuo-spatial construction and attentional deficit. On the other hand, language skills are unaffected and sometimes above average (Losh et al. 2000, Reilly 1999). In the field of face processing, recognition and discrimination between faces are preserved in WS (Deruelle 1999, Martens 2008). Generally, they display a preference in processing local elements and features compared to global stimuli (Bellugi 2000) a behavior which is also found for the processing of faces although WS are less impaired in detection of faces during inversion tasks of faces (Karmiloff-Smith et al 2004) compared to a control population.

However, WS ability on processing facial emotion in disrupted in this pathology. This pattern has been observed in recognition of emotional expression tasks (Gagliardi, 2003, Plesa-Skwerer, 2006) also when emotions were expressed by the eyes only.  This impaired ability to detect facial expression has been often linked to patient's social behavior but yet its precise cause and the link with the brain regions differently involved in this pathology is largely unknown.

WS social behavior is characterized by overfriendliness and appetence to create social ties and social interactions.  At least two explanations have been provided for this over social behavior. One is a general absence of inhibition, confirmed using fMRI showing the absence of connectivity between frontal region mPFC and the amygdala (Meyer-Lindenberg et al., 2005). Another explanation is more linked with face processing, suggesting an atypical neural activity while processing faces and social interactions (Meyer-Linderberg 2005, Meyer-Linderberg 2005 Haas 2009). Behaviorally, WS have difficulties in detecting fear (Barak & Feng, 2016) and they show less arousal when looking at angry faces (Skwerer et al., 2011). This deficit is probably linked to an altered amygdala activity. An fMRI study showed that amygdala activity in WS only increased while looking at threatening scenes but not processing threating faces (Meyer-Linderberg 2005; Haas 2009). Similarly, atypical amygdala activation during processing of happy (WS>controls) and fearful faces (WS<controls) was reported (Haas 2009). This atypical process of emotion may biases their ability to perform social judgments (Frigerio 2006). WS tend in fact to rate unfamiliar and negative faces as more approachable than controls. Again, also in this study they found an over positive bias for positive expression, that is ratings were higher than controls for positive faces too.

Yet, the atypical social behavior represents a mystery in WS behavior and its neural correlation is largely unclear. To dissociate cognitive and emotional deficits present in WS, the current work explores behavioral and neural signatures of WS during an implicit trust task (see chapter 4).

## 1.7. Fooled by first impression: trustworthiness judgments from facial appearance

In the previous section I focused on the interest of face in social perception presenting cognitive and neural models about face processing in human and non-human primates. In this section I will shift the attention on the ability to make social judgment from faces to reach a trustworthiness judgment from facial information.

From the ancient Greece to modern physiognomy, the face was identified as important source for humans inferences about emotion and personality traits (Lavater, 1880). From Darwin's studies, the pioneer in this field, a variety of studies have been done to show the structural prototypes for emotional expressions (Ekman, 1993; Russell 1997; Russell, Bachorowsky, & Fernandez-Dols, 2003) and social traits are still a field of interest of the contemporary neuroscience. Physiognomy or morphology never became a scientific field; however scientists are investigating the facial cues that help guide the identifications of social traits. Great contributions in this field have been provided by Todorov and colleagues that in the last years proposed several studies showing which facial features are essential for different social judgments, that impact face perception in social context. Moreover he pinpointed the neural basis of the ability to form first impression judgments.

When faced to unfamiliar faces, human form fast first impression judgments based on facial appearance. First impression effect, can bias decisions, often unconsciously, in a variety of domains i.e. mate choice (Olivola et al., 2009) political election, business context and finance (Gorn, Jiang, & Johar, 2008; Naylor, 2007; Pope & Sydnor, 2008; Ravina, 2008; Rule & Ambady, 2008b), law/forensic-science (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Zarkadi, Wade, & Stewart, 2009; Zebrowitz & McDonald, 1991), and military actions (Mueller & Mazur, 1996). First impression is present continuously in humans' daily life and especially in new contexts where one needs to evaluate others for the first time. Indeed, the main survival function of this kind judgment is to assess quickly whether strangers can be approached or avoided (Nikolaas Oosterhof & Todorov, 2008; Todorov, 2008b).

Todorov and colleagues showed that humans rapidly form first impression from faces (Bar, Neta, and Linz 2006; Willis and Todorov 2006). Upon facial detection, opinion is built after 100ms of exposition to faces and these judgments will not change if subjects are exposed to faces for longer time (i.e. 500ms). This result was replicated and extended considering only trustworthiness judgments (Todorov et al., 2009). In a first study, authors aimed to understand how time exposure of faces judged as more or less trustworthy by an independent group of subjects, may influence subjects' perception. Subjects were exposed to faces either for 50, 100 or 500 ms, then faces were masked and participants had to respond using their gut reaction rating the trustworthiness of the faces. Although confidence with the judgments increased with time exposition, after already 50 ms people were able to perform judgments of trustworthiness that correlated with the prediction. In a second study (Todorov et al., 2009), faces were presented for multiple times and minimal time of exposition before the mask appeared was decreased at 33 ms and 17ms. Differently in this task, participants were informed that the experiment was on first impression of trustworthiness. After 33 ms of facial exposure, participants were able to perform judgments of trustworthiness (minimal threshold).

In sum, they conclude that making trustworthiness judgments, not necessarily accurate, is based on a fast (33ms) and automatic cognitive process (Todorov et al., 2009).

A different set of studies showed that facial evaluation on social dimensions is not restricted to conscious appraisal but also happens at a preconscious level, hence the fast outcome in social judgements (Getov, Kanai, Bahrami, & Rees, 2014; Stewart et al., 2012).

The ability to detect trustworthiness from faces is strongly related to the perception of facial features (Peter Mende-Siedlecki, Said, & Todorov, 2013; Nikolaas Oosterhof & Todorov, 2008). Todorov and colleagues created and validated (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013) a database of computer-generated faces (Caucasian male identities randomly generated using FaceGen software) that displays specific facial features that predict trustworthiness's judgments (Figure 8).

*Figure 8* .**Model of trustworthiness judgments** from Todorov database (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015).

The novelty of this software is the setting of emotional content of each face with respect to a neutral face parameter. Transformations from the neutral facial features predict social judgments. Interestingly they found that the facial cues that better predict participants' judgments were shape of the eyebrows and shape of the mouth. This result was also confirmed using psychophysical reverse correlation methods (Dotsch & Todorov, 2012), were random noise distorts the faces and subjects were still asked to perform a trustworthiness's judgment. Using this technique the authors observed that subjects choose the noise that modulated mouth, eye and eyebrow shape (Figure 9).

*Figure 9.* Classification images of judgments of trustworthiness (left image) and untrustworthiness (right image) using reverse correlation method (Dotsch & Todorov, 2012).

Although trustworthy faces are emotionally neutral, the extreme positive variation of trustworthy faces are perceived as happy, while the extreme negative variation of untrustworthy faces resemble angry faces (Nikolaas Oosterhof & Todorov, 2009; NN Oosterhof & Todorov, 2008; Todorov, Baron, & Oosterhof, 2008; Todorov, 2008b). This phenomenon has been explained using the overgeneralization hypothesis. (Knutson, 1996; Montepare & Dobish, 2003; Todorov, 2008b) (Zebrowitz & Montepare, 2008). According to this hypothesis, resemblance of neutral faces to emotional expressions is perceived as indicating the trait attributes or behavioral tendencies associated with these emotions. Engell and colleagues, in a study where they used a behavioral adaptation paradigm found that adaptation to angry or happy faces biased judgments towards neutral faces, but fearful faces did not generate the same effect (Engell, Todorov, & Haxby, 2010). Resemblance of trustworthy faces to happiness is only one of the facial features that can predict trustworthiness judgments. As reported by Todorov and colleagues, other facial features such as facial femininity, resemblance to yourself or maturity are predictors of trustworthiness judgments (Todorov et al., 2015).

In a third set of experiments, Todorov and colleagues (Todorov et al., 2009) used computer generated faces to test priming effects. Participants were more likely to judge neutral faces as untrustworthy if the neutral face was primed by an untrustworthy face. Interestingly, the positive priming (trustworthy face) did not significantly biased judgments of neutral faces.

What is the neural basis of trust decisions? Neuroimaging studies have tried to signal brain regions underling the formation of first impressions and trust. Trustworthiness inferences have been linked to significant activity in activity over the amygdala supporting that detection of trustworthiness immediately active a region that is also involved in the automatic processes of assigning valence to stimuli (Engell, Haxby, & Todorov, 2007; Todorov et al., 2008; Winston, Strange, O'Doherty, & Dolan, 2002).

Using an implicit trustworthiness task (Engell et al., 2007), greater amygdala activity was reported for both trustworthy and untrustworthy faces, suggesting that the amygdala is sensible to the valence expressed by faces. Coherently, less activity over the amygdala was observed during presentation of neutral faces. Trustworthiness dimension is the social trait that more correlates with the valence (Oostorof and Todorov 2008). The implicit task was a face memory task, where participants were instructed to remember the first 11 faces presented and they had to respond whether the 12$^{th}$ was an already seen face or not. The same subjects rated (from 1 to 9) after the scanner session their level of trustworthiness on all the faces seen during the fMRI experiment.

Authors reported a linear activity of the right amygdala, showing an increasing response proportional to changes in the level of untrustworthiness of faces. The left putamen and the right insula also showed similar patterns. Left amygdala and medial prefrontal cortex and the precuneus showed a negative quadratic response, that is the activity was higher for both very untrustworthy and very trustworthy faces (Todorov et al., 2008).

Trustworthiness judgments are strongly related with valence perception (Oostorof and Todorov 2008), however we can still try to understand whether specificity exists between social traits (see Chapter 2) and which brain areas are more selective while forming first impression judgments while controlling for the valence. Schiller and colleagues (2009) tried to disentangle the brain regions selectively involved in the forming of first impression about others from those areas involved in more general processing of social information. Authors found that the posterior cingulate cortex (PCC) and the amygdala were stronger activated while encoding social information suggesting that these two regions are crucial in forming first impression. The study however did not focus on trustworthiness but participants had to explicit express their impression after listening some sentences about a person and say whether they liked or not the character (Figure 10). Differently from previous results they observed that dmPFC, was involved in the processing of personal-descriptive information. Nevertheless the signal was not different when the information was relevant or not to perform a social judgment. On the contrary the amygdala and the PCC were selectively activated for relevant information and not for the irrelevant one. Moreover they do not find any difference between positive and negative valence judgments.

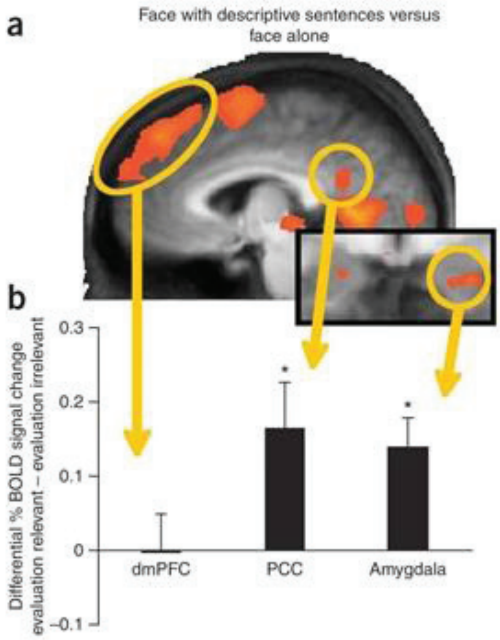



*Figure 10.* Brain regions demonstrating the difference in evaluation effect out of regions engaged in the impression-formation task. As shown by Schiller and colleagues (2009) posterior cingulate cortex (PCC) and the amygdala were crucial in forming first impression.

A similar attempt to separate general social neural processes from trust, a meta-analysis of 29 neuroimaging studies on social facial evaluation for judgments of trustworthiness and attractiveness was done (Mende-Siedlecki et al., 2013) (Table 2). Results were divided into negative and positive judgements of trustworthiness and attractiveness. During negative evaluations, untrustworthiness and unattractiveness were found as the most consistent activity in the right amygdala. However, the left amygdala, right anterior insula, right IFG, right vlPFC and right globus pallidus were also found but less consistent across studies. Considering positive judgments of trustworthiness and attractiveness they found highly consistent activity over the left caudate. The same consistent activity was observed over the right amygdala, insula, IFG and vlPFC. This activity was coherent with what has been found for faces with happiness content. It has also been argued that positive valuation of faces depends in part on the structures that underline reward processing as suggested by the activity of the left caudate and nucleus accumbens. Overall, this review did not find a significant difference of neural activity related to trustworthiness and attractiveness judgments which tends to suggest that a common cognitive source is present while performing both behaviors.

**Table 2.** *Neuroimaging studies on trustworthiness judgments (P. Mende-Siedlecki, Said, & Todorov, 2013).*

| Study | Included contrasts | N | Task nature[a] | | | Study type | Valence | | | ROI?[c] | Stimulus category |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Implicit | Explicit | Collapsed[b] | | Negative | Positive | Non-Linear | | |
| Aharon, et al., 2001 | 1 | 6 | X | | | attractiveness | | X | | | extreme |
| Aron, et al., 2005 | 1 | 17 | X | | | attractiveness | | X | | | average |
| Baas, et al., 2008 | 4 | 21 | X | X | X | trustworthiness | X | X | | | average |
| Blasi, et al., 2009 | 1 | 43 | | X | | trustworthiness | X | | | | average |
| Bray, et al., 2007 | 2 | 25 | X | | | attractiveness | X | X | | | extreme |
| Chatterjee, et al., 2009 | 4 | 13 | X | X | | attractiveness | X | X | | | computer generated |
| Cloutier, et al., 2008 | 2 | 48 | | X | | attractiveness | X | X | | | extreme |
| Engell, et al., 2007 | 1 | 15 | X | | | trustworthiness | X | | | | average |
| Gordon, et al., 2009 | 2 | 6 | X | | | trustworthiness | | X | X | X | average |
| Hampshire, et al., 2011 | 1 | 19 | | X | | attractiveness | | X | | | extreme |
| Iaria, et al., 2008 | 1 | 10 | | X | | attractiveness | | X | | | extreme |
| Kampe, et al., 2001 | 1 | 16 | X | | | attractiveness | | X | | | average |
| Kim, et al., 2007 | 1 | 25 | | X | | attractiveness | | X | | | computer generated |
| Liang, et al., 2010 | 3 | 17 | X | | | attractiveness | X | X | X | X | extreme |
| O'Doherty, et al., 2003 | 2 | 25 | X | | | attractiveness | X | X | | | extreme |
| Pinkham, et al., 2008 | 1 | 12 | | X | | trustworthiness | X | | | X | average |
| Pochon, et al., 2008 | 1 | 17 | | X | | attractiveness | | X | | | extreme |
| Said, et al., 2009 | 3 | 31 | | X | | trustworthiness | X | X | X | | average |
| Said, et al., 2010 | 1 | 24 | X | | | trustworthiness | | | X | | computer generated |
| Smith, et al., 2010 | 1 | 26 | | X | | attractiveness | | X | | | average |
| Todorov, et al., 2008 | 2 | 14 | X | | | trustworthiness | X | | X | X | computer generated |
| Todorov, et al., 2010 | 2 | 22 | | X | | trustworthiness | X | | X | | computer generated |
| Todorov, et al., 2010 | 2 | 22 | X | | | trustworthiness | X | | X | | computer generated |
| Tsukiura, et al., 2010a | 3 | 22 | | X | | attractiveness | X | X | X | | average |
| Turk, et al., 2004 | 1 | 18 | | X | | attractiveness | | X | | | extreme |
| VanRijn, et al., 2011 | 3 | 18 | | X | | trustworthiness | X | X | | X | average |
| Winston, et al., 2002 | 2 | 14 | | | X | trustworthiness | X | X | | | average |
| Winston, et al., 2007 | 2 | 26 | | | X | attractiveness | | X | X | | extreme |
| Zaki, et al., 2011 | 1 | 14 | | X | | attractiveness | | X | | | extreme |

# Chapter 2

# Implicit preference for human trustworthy faces in monkeys

This chapter is a modified version of

Manuela Costa, Elodie Barat, Alice Gomez, Guillaume Lio, Jean-René Duhamel, Angela Sirigu. Tacit preference for trustworthy faces in monkeys. Submitted to *Elife*.

**Abstract**

In numerous species, trust is a basic prerequisite of group living. As it does not come without a risk, the ability to select trustworthy partners is an essential survival skill. Although prior social experience influence perception, research in social psychology has demonstrated that human judgments of trustworthiness are based on subtle processing of specific perceptual features. However, it is not known if this ability has an evolutionary origin. Here we show that macaque monkeys, like humans, have a preferential attention to trustworthiness-associated facial cues portrayed in computer generated human faces. They look significantly longer at faces categorized *a priori* as trustworthy as compared to untrustworthy. We further found significant correlations between facial width-to-height ratio – a morphometric feature that predicts trustworthiness' judgments in humans –and looking time in both species. These findings reveal the importance of facial cues in providing social information in human and non-human primates suggesting the presence of common evolutionary mechanisms for first impression of trustworthiness.

## 2.1. Introduction

Trust is a fundamental psychological dimension, influencing people's decisions in social interactions such as cooperation (Cosmides & Tooby, 1992; Ross & Lacroix, 1996), voting intentions (Olivola & Todorov, 2010a), economic decision-making (Olivola, Funk, & Todorov, 2014; Ross & Lacroix, 1996) . Trusting is taking the risk of putting one's own fate in someone else's hands, hence the importance of trustworthiness assessment to minimize this risk. Surprisingly, research in social psychology has demonstrated that rather than being based solely on rational criteria (reputation, prior social interactions) judgments of trustworthiness are robustly related to specific perceptual features (Dotsch & Todorov, 2012; Todorov et al., 2015; Willis & Todorov, 2006). For instance, different shapes of eyebrows, cheekbones and chin can trigger different level of trustworthiness from faces. These facial features automatically capture observers' attention and lead to trustworthiness judgments in less than 33ms, the so-called first impression effect (Todorov et al., 2009). This effect is based on detection of facial cues and also on holistic processing of a face's appearance such as the resemblance of neutral faces to typical expressions of anger and happiness (Nikolaas Oosterhof & Todorov, 2009), femininity/masculinity (Nikolaas Oosterhof & Todorov, 2008); facial maturity (Montepare & Zebrowitz, 1998) and physical similarity to the self (DeBruine, 2005).Finally, it has been shown that the facial width-to-height ratio (FWHR)(Weston et al., 2007), a morphometric measure that rely on face structure, predicts explicit trustworthiness judgments (Stirrat & Perrett, 2010, 2012).

Mechanisms of face processing are shared between humans and non-human primates (D. Y. Tsao et al., 2008). Because the capacity to perform judgments of trustworthiness strongly involves processing of facial cues, and given the adaptive value of this skill in strongly cooperative societies, one may wonder whether such ability has an evolutionary origin. We therefore asked if non-human primates are responsive to trustworthiness-associated facial cues by recording eye movements in a preferential looking test.

In this study we showed macaque monkeys (n=8) pairs of parameterized human faces drawn from Todorov et al's image database (Todorov et al., 2013), each displaying a most (+3SD from the baseline) and a least (-3SD from the baseline) trustworthy version of the same facial identity. These computer-generated faces only vary on the facial features that predict judgments of trustworthiness. Human faces with high inner eyebrows, pronounced cheekbones, wide chins and shallow nose sellion are perceived as more trustworthy than faces with low inner eyebrows, shallow cheekbones, thin chins and deep nose sellion (Todorov et al., 2008). We presented the two extreme variants of the same facial identity in each trial to ensure that monkey's preference towards one face or another could depend from the only difference between the two stimuli on trustworthiness-associated facial cues. To ensure spontaneous preferences, monkeys were not rewarded to look at specifically at the faces. They freely moved their eyes about and were periodically given juice rewards to maintain gaze within the limits of the computer screen surface where the images were displayed (see Methods).

The rationale for using this approach relies on the known fact that non-human primates are highly sensitive to, and use facial cues during social interaction. Studies on infant rhesus monkeys have shown that these animals exhibit both innate and early experience-dependent preferences for monkey as well as human faces (K M Gothard, Battaglia, Erickson, Spitler, & Amaral, 2007; Sugita, 2008). Furthermore, monkeys raised in captivity develop considerable expertise about humans. For instance, rhesus macaques form spontaneous face-voice associations for familiar humans (Sliwa, Duhamel, Pascalis, & Wirth, 2011), and capuchin monkeys are sensitive to observed human interactions. They show avoidance of humans who are not helpful and do not reciprocate in social exchanges (Anderson, Kuroshima, Takimoto, & Fujita, 2013; Anderson, Takimoto, Kuroshima, & Fujita, 2013) and they approach and look more at humans who are imitating them (Paukner, Suomi, Visalberghi, & Ferrari, 2009). Finally, eye tracking studies in chimpanzees further show that non-human primates look more frequently and longer at positive valence stimuli (approach behavior) and less frequently and shorter at negative stimuli (withdrawal behavior)(Braccini, Lambeth, Schapiro, & Fitch, 2012).

We therefore reasoned that monkeys might be able to use facial features and discriminate between trustworthy and untrustworthy human faces (Todorov et al., 2013). Because faces in each trial only differed for the level of trustworthiness, we hypothesized that attention towards one of the two faces could be a sign of detection of the features that differently characterized the pair of faces. Moreover, because monkeys were not rewarded to look at faces we hypothesized that a greater looking time towards trustworthy faces may be interpreted as an approach behavior towards that faces.

To establish across-species comparisons, and assess if human would also spontaneously and preferentially allocate gaze toward trustworthy faces, we assessed the performance of human subjects (*n=56*) following the same procedure as in monkeys.

## 2.2. Methods

All experimental procedures were in accordance with the local authorities (Direction Départementale des Services Vétérinaires, Lyon, France) and the European Community standards for the care and use of laboratory animals [European Community Council Directive (1986), Ministère de l'Agriculture et de la forêt, Commission Nationale de l'Expérimentation Animale].

**Participants**

*Monkeys*: eight adult monkeys (Macaca Mulatta, one female and four males 4-17 years old, and Macaca fascicularis, three males 6 years old) have been tested. All animals were born in outdoor enclosures and were then socially housed indoor, so they have been exposed to both conspecific and humans.

*Humans participants*: fifty-six healthy subjects, (28 women, M age= 26.9 years, SD = 5.9), with normal vision, took part in the experiment. They were blind about the aim of the study; participants only knew to take part to a first impression study. To perform the analysis we excluded two subjects because they significantly differ from the rest of the population (avarage±2*SD).

**Stimuli**

The stimuli used in the experiment were 48 computer-generated male faces created with the FaceGen software development kit (Singular Inversions, Toronto, Canada) selected from the Todorov's well-controlled quantitatively validated stimulus repertoire of faces. All faces were bald and Caucasian. The Trustworthiness database is composed of facial identities varying on 7 levels of trustworthiness (Todorov et al., 2013). Todorov et al' work showed that human explicit judgments of trustworthiness match with the model's prediction (Todorov et al., 2013) For the current study we selected from the 24 identities the two most extreme versions (-3 SD and +3 SD).

**Task procedure**

*Monkeys:* to assess preference formation, we used a preferential looking paradigm. For monkeys, pairs of faces were presented in a random order. Each face pair was presented twice to counterbalance for side of presentation. During the experiment, monkeys were seated in a primate chair inside a darkened room with their head restrained. Stimuli were presented on a 15-inch color monitor (1024 x 768 pixels) at a viewing distance of 24cm. A trial began with the appearance of a single fixation point in the center of the screen. Once the monkey fixated this point, two face stimuli subtending 13° x 21.2° of visual angle (207 x 340 pixels) were displayed and remained on the screen for up to 2s. The monkey was free to move its eyes over the images and received a juice reward provided its gaze stayed within the boundaries of the video monitor for the entire 2s period, otherwise the stimuli were extinguished and the trial discarded. Monkeys could choose to look outside of the face and still receive reward. The monkey's gaze position was monitored by ISCAN infrared eye tracking system at 200-Hz. Experimental control, stimulus presentation, data sampling and storage was done with

REX/VEX software system (57). Before the experiment, monkeys underwent a 5-point eye position calibration and were trained until they understood the visual exploration task using 8 different pairs of non-face biological and non-biological stimuli.

*Humans participants:* healthy participants (*n=56*) were instructed to look at the same pairs of faces during 5 s. Stimuli were displayed on a 17-inch computer screen at a resolution of 1280 x 1024 pixels using Presentation® software (Version 14.9, www.neurobs.com). The viewing distance from the participant's eyes to the screen on which stimuli subtending 7.8° x 12.5° of visual angle (377 x 604 pixels) were displayed was 73cm. Humans' eye positions were recorded using an infrared video-based tracker (Tobii 1750) at a 60-Hz sampling rate and Clearview 2.7.0 allowed online recording of eye-gaze data. The two systems were synchronized using the Tobii extension for Presentation. In a second session humans were asked to explicit choose.  During all sessions the experimenter monitored on-line the position of the subject's eye gaze that was projected on a second screen in the same room but placed far from the location of the participants. Prior to the experiment, humans underwent a 5 point-calibration task. The final experimental set comprised 48 trials.

**Statistical Analysis**

*Pre-processing and data analysis.* ClearView fixation filter was used to filter the data for humans (with a visual angle of 1° and duration of 100ms). An in-house Matlab script was used to pre-process and filter monkeys' eye-tracking data. First, eye velocity for each location was computed as the angular distance traversed by the eye within a 5ms moving window. Next, for each trial, a velocity threshold was set at three times the median during the 2s window. Data points that exceeded this threshold were considered as saccades. Fixation times were considered as the interval between two saccades with a minimum duration of at least 100ms, and fixation locations were defined as the eye position at the central fixation time point. In order to quantify allocation of attention to faces, regions of interest (ROI) delimiting each face were defined manually. The mean looking time was calculated as the average of the total time spent within each ROI during a trial. Only trials with at least one fixation at one of the two faces were included in the dataset. For the main statistical analysis, mean looking times on each face were calculated for each participant and for each trial. The results from the different trials in each condition were averaged for each participant.

*Temporal dynamics and statistical analysis.* In order to identify the time windows showing significant differences between trustworthy and untrustworthy faces, a large scale multiple testing procedures was designed. First, statistical differences between the number of fixations in trustworthy and untrustworthy ROIs were tested using successive non-parametric Wilcoxon Rank-sum tests at each time point (15ms). Then, multiple comparisons were performed using cluster-based permutation test (Maris & Oostenveld, 2007).

*Spatial distribution of fixations.* To provide information on the spatial distribution of the fixations, the barycenter of fixations and a heat map representation were calculated for each face at the subject level. Heat maps were calculated using Gaussian kernel density mapping of the fixations, weighted by the fixations' duration (Caldara & Miellet, 2011). Then, at the group-level, individual heat-maps were normalized and averaged to visualize the spatial distribution of the fixations of the studied population.

**Facial attributes**

*Facial width-to-height ratio (FWHR).* To obtain a score for the facial width-to- height ratio, two independent raters measured the distance between the lip and brow (upper facial height) and the left and right zygion (bizygomatic width) of each face from the trust database. FWHR was calculated as width divided by height(Carré & McCormick, 2008). Inter-rater reliability was high for all measures (all Rs>.79, all Ps<0.001).

*Facial emotion and gender evaluation.* Perceived happiness and femininity of face stimuli contribute to trustworthiness judgements *(10, 11)*. This was checked through a quantification of these two attributes in two separate experiments, by independent human observers (*n*=7) in the following manner. Pairs of randomly selected faces of all identities were presented on the screen and subjects had to choose either the most 'happy' or the most 'feminine' one by pressing a key on a keyboard. Each face stimulus was presented at least 16 times by subject (1225 trials). Happiness and femininity scores for each face were given by the percentage of instances the stimulus was selected as the happiest and the most feminine, respectively. Mean inter-subject reliability in these evaluations was high for happiness, $r(48)=0.88$ and for femininity, $r(48)=0.60$. We computed correlation between happiness, femininity and FWHR and explicit trustworthiness judgements (value used: -3; +3). As expected, all correlations were significant: happiness, $r(48)=0.91$ ($P<0.000001$); femininity, $r(48)=0.88$ ($P<0.000001$); FWHR, $r(48)=-0.54$ ($P<0.00005$). Correlation between happiness and femininity was significant $r(48)=0.76$ ($P<0.000001$); correlation between happiness and FWHR was significant $r(48)=-0.46$ ($P<0.0005$). At the same way femininity was correlated with FWHR $r(48)=-0.60$ ($P<0.00001$) (see Table1).

## 2.3.Results

### 2.3.1. Monkeys' and humans' visual preferences

In order to quantify gaze allocation, regions of interest (ROIs) encompassing the trustworthy and untrustworthy faces were defined (Figure 10). Ocular fixations within and outside these ROI were recorded during each trial (see Methods). The mean looking time was calculated as the average of the total time spent within trustworthy and untrustworthy faces for all stimulus pairs presented.

The first analysis revealed that monkeys were attracted to human faces, spending more time on these stimuli than predicted by random exploration of the video monitor (chance level=160ms; trustworthy: M=512.89ms, SD=286.75, $T_7=4.45$, $P<0.01$; untrustworthy: M=292.60ms, SD=262.09, $T_7=2.30$; $P=0.054$). Because the presented faces differed for the level of trustworthiness we further tested whether monkeys discriminated between the two stimuli presented by looking longer at one than the other and, indeed, monkeys spent significantly more time looking at trustworthy than untrustworthy faces ($T_7=3.29$; $P<0.05$, Figure 10A).

Humans exhibited the same pattern, spending most of the viewing time looking at the faces (chance level=853ms; trustworthy: M=2311.96ms, SD=208.75, $T_{53}=46.86$, $P<0.001$, untrustworthy: M=2177.11 ms, SD=228.72, $T_{53}=46.59$, $P<0.001$) and showing a significant bias in favor of the trustworthy face category ($T_{53} = -2.96$, $P<0.005$) (Figure 11D).

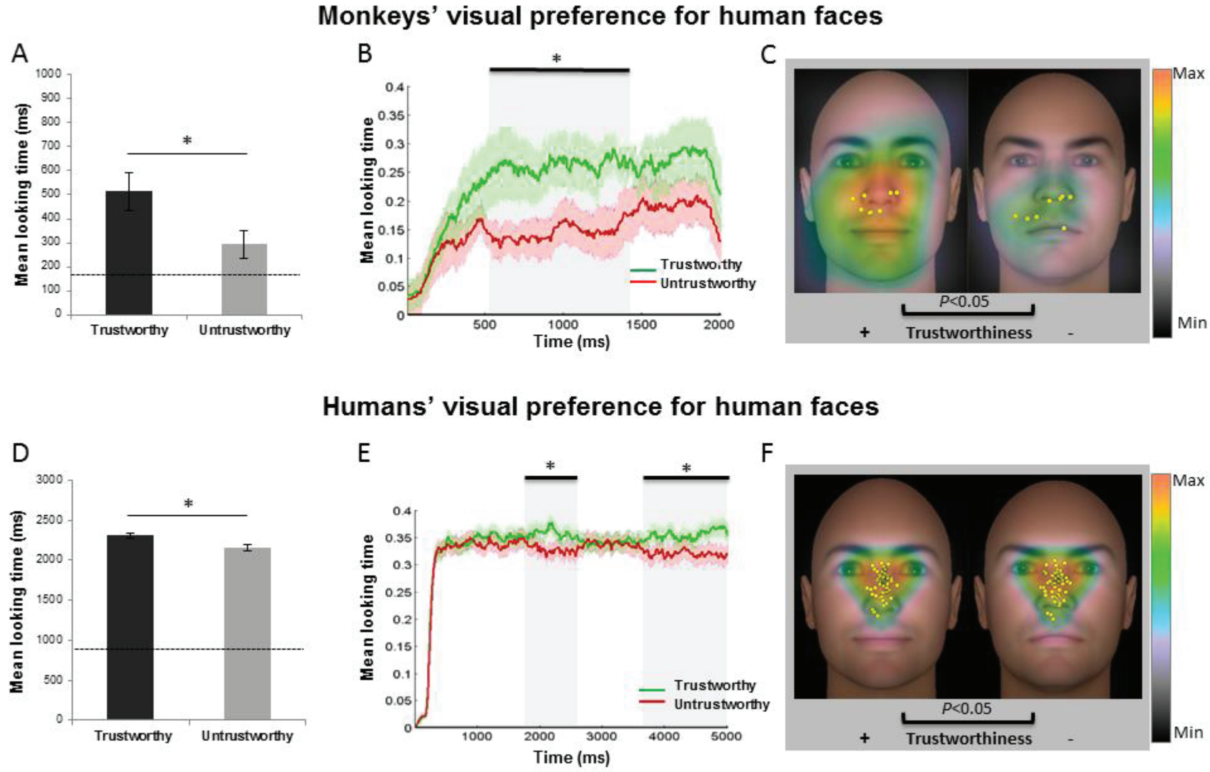



*Figure. 11.* **Looking preference for trustworthy vs. untrustworthy faces by rhesus macaques and human subjects. MONKEYS (n=8): (A)** *Mean looking time in milliseconds* (ms) for the most trustworthy (+3SD of the neutral face) and the least trustworthy (-3SD of the neutral face) versions of the same facial identities. The error bars denote standard error of the mean. **P*<0.05. Monkeys looked significantly longer at the two faces than predicted by chance and looked more at trustworthy than untrustworthy faces (chance level represented with dotted line was 160ms for each face region of interest). (**B**) *Time course of looking preference*. Mean viewing time on each facial prototype plotted each 15ms. A cluster-based permutation test showed that preference for the trustworthy faces (green line) was significant between 510ms and 1485ms (*P*<0.05 corrected for multiple comparison). **C.** *Gaze heat maps for trustworthy and untrustworthy faces* averaged across subjects (trustworthy face on the left by convention, facial prototype spatial location was counterbalanced within and between subjects). Yellow dots show fixation centers of gravity for each subject. **HUMANS (n=56) (D-E-F)** Plots show (D) significantly longer mean looking times at trustworthy than untrustworthy faces and (E) later onset of preference for trustworthy faces (1800ms to 2640ms and again from 3615ms to the end of the trial). Note that the average barycenter of fixation was located in the region surrounding the nose in monkeys whereas it is around the eye region in humans (C, F).

In order to determine whether longer looking times are actually related to perceive trustworthiness, in a separate session, humans were asked to look for the most trustworthy face while we recorded their eye movements. We found a significant correlation between mean looking times between the first (implicit) and the second (explicit) preferential looking tasks ($R$=0.30 $p$=0.027, Figure 12).

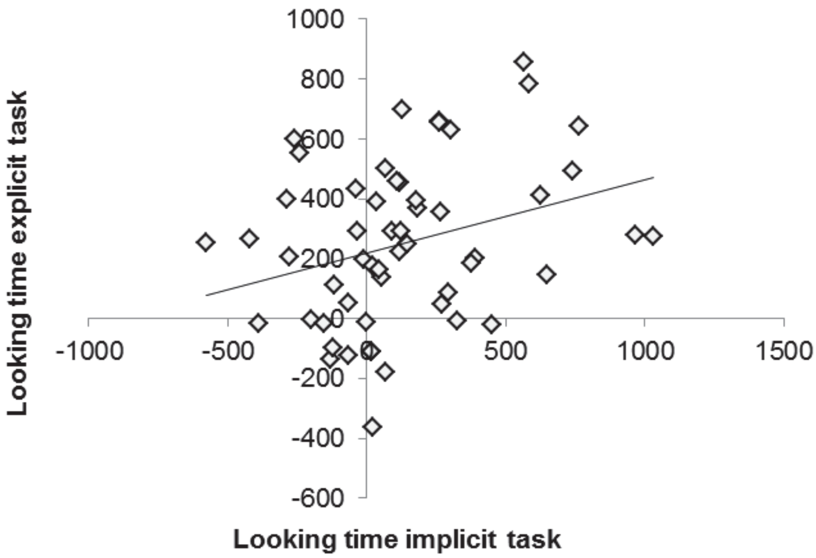



*Figure 12*. **Human correlation between difference of looking time (trust - untrust) during the implicit and explicit task**. We observed a positive correlation (R=0.30, p=0.027). The subjects that looked more the trustworthy faces during the implicit task also looked more the same faces during the explicit task, while they were asked to select the most trustworthy face. This result support that eye movements in adult normal population predict their explicit judgments.

To further corroborate our finding on the spontaneous preference for trustworthy stimuli we performed a T-test using identity instead of subjects as random effect. We compared looking time for trustworthy and untrustworthy faces for each identity. We found a significant effect in both species, monkeys ($T_{23}=4.85$; $P<0.0001$) and humans ($T_{23}=10.03$; $P<0.0001$).

Hence, our results reveal that both macaque monkeys and humans detected and preferred to look at human faces displaying trustworthiness-associated facial cues.

Because of this common preference across species, we explored whether monkeys and humans used similar eye gaze strategies with a focus on temporal dynamics and spatial distribution of fixations. A cluster-based permutation test (see Method) showed that, in monkeys, preference for the trustworthy faces occurred from 510ms to 1485ms after image onset ($P<0.05$ corrected for multiple comparison, Figure. 11B), while humans' preference occurred in two stages and later in time, with a first short-lived preference emerging at 1800ms and a more stable one at 3615ms ($P<0.01$, Figure. 11E). To provide information on the spatial distribution of visual exploration (see Methods), heat maps and barycenter of eye fixations were generated (Figure. 11 C-F). Monkeys preferentially allocated their attention in the region surrounding the nose (Figure. 11C), while humans eye gazed mostly around the eye region (Fig. 11F).

### 2.3.2.Correlation between looking time and trustworthiness' predictors: FWHR, happiness and femininity scores

Considering the role of facial width-to-height ratio (FWHR) in human judgments of trustworthiness(Stirrat & Perrett, 2010), we tested whether this character might have a specific role in driving the viewing preference for trustworthiness in monkeys and humans. FWHR was calculated as width divided by height using standard landmarks (Carré & McCormick, 2008). To compute FWHR, two independent raters measured the distance between the lip and brow (upper facial height) and the left and right zygion (bizygomatic width) of each face from the entire image database. Inter-rater reliability was high for all measures (all Rs>.79, all Ps<0.001). In agreement with Stirrat and Perrett's findings, we found that faces that have been judged by humans as trustworthy had a lower FWHR than untrustworthy ones ($F_{1,\ 24}$=116.97, P<0.05; $M_T$=0.02; SD=0.019, $M_{UT}$=2.15, SD=0.02). The classification of face stimuli obtained was then regressed against monkeys' viewing preferences (Fig.2). Interestingly, total viewing time on a given face was negatively correlated to its FWHR in monkeys ($r(48)=-0.35$, $P<0.05$) and humans ($r(48)=-0.46$, $P<0.01$) (Figure 13). Thus, longer and narrower faces (lower FWHR) were looked longer in both species.

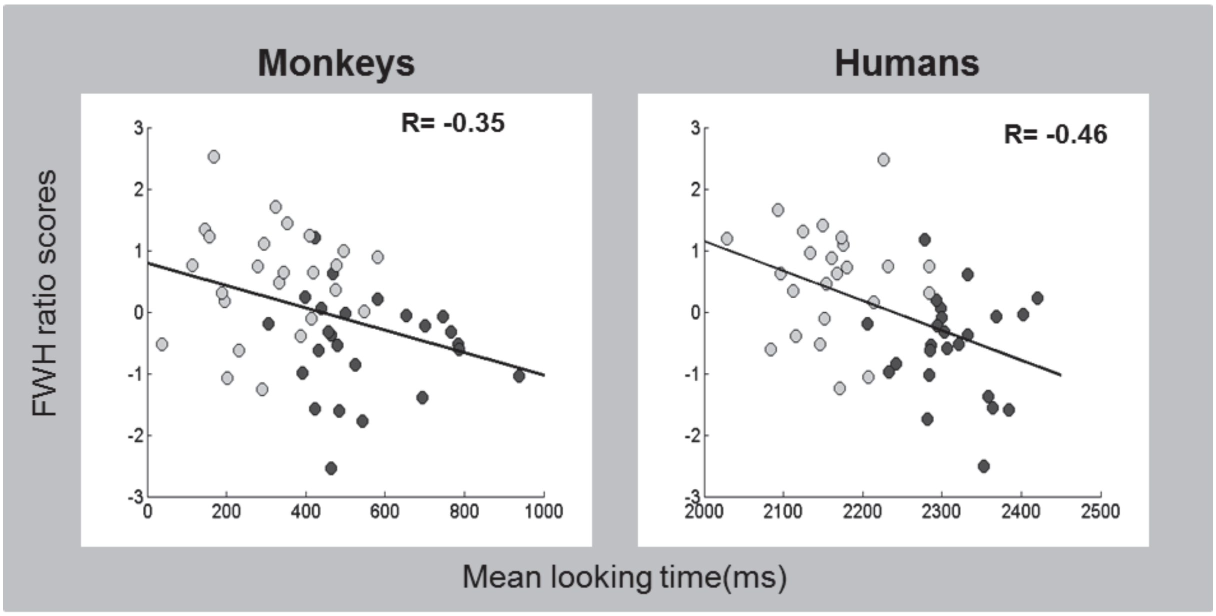



*Figure 13.* **Monkeys and humans'correlation between mean looking time and facial width height ratio (FWHR) score for each face.** Mean looking time of monkeys (left graph, $r(48)=-0.35$, $P<0.05$) and humans (right graph, $r(48)=-0.46$, $P<0.001$) are negatively correlated to the FWHR scores, i.e., longer and narrower faces were looked at longer by both species. Light grey points correspond to trustworthy faces and dark grey points to untrustworthy ones.

As shown by Todorov and colleagues (Nikolaas Oosterhof & Todorov, 2008, 2009) other facial parameters such as emotional valence and femininity also matter in the first impression of trust. To establish whether monkeys' looking preference was affected by these attributes we performed further correlation analyses. First, independent human observers rated each facial identity for their degree of apparent happiness (happy/unhappy) ($n=7$) and femininity (feminine/masculine) ($n=7$). Scores for each face were given by the percentage of instances the stimulus was selected as the happiest and the most feminine, respectively. The obtained scores were then correlated with subjects' looking time. We found that monkeys' total viewing time on a face was positively correlated to the emotion score ($r(48)=0.56$, $P<0.0005$) and to the face femininity score ($r(48)=0.58$, $P<0.0001$). The same pattern of correlation was observed in humans (happiness: $r(48)=0.731$, $P<0.000001$; femininity: ($r(48)=0.738$, $P<0.000001$) (for discussion about correlation analyses see Methods). Here we confirm that trustworthiness faces are perceived as more happy and feminine. Moreover, we showed that this intrinsic connection can be confirmed using an implicit measure as it is the looking time. Hence monkeys and humans were more attracted to faces judged ad more happy and feminine.

## 2.4. Discussion

Our results show in monkeys and humans a preferential attention to trustworthiness-associated facial cues, suggesting that certain facial features and trait-related information might be selected together through evolution.

The establishment of this visual preference differed in time in monkeys and humans: monkeys settled on the preferred face early on (510 ms) whereas humans first explored both faces equally for about 2s before exhibiting a preference for the trustworthy face. In addition to exposure time to the face stimuli (2s and 5s for monkeys and humans, respectively), the different temporal profiles may be related to the fact that monkeys were rewarded only for maintaining gaze within the limits of the screen, not for exploring the two faces, whereas human subjects may have, wittingly or not, construed the task as requiring exploration and comparison of both images. In humans, explicit judgments of trustworthiness are made in less than 33ms(Todorov et al., 2009). Differently, in this study we found a tardive onset of the preference. This result is not surprising considering that two faces have been simultaneously presented on the screen and that no explicit instructions were given to participants.

Spatial distribution of eye fixations confirms previous reports that monkeys attending to human faces do not explore the eye region as they do when looking at monkey faces (Katalin M Gothard et al., 2009) and, more generally, that both monkeys and humans make more eye fixations toward conspecifics than no conspecifics (Dahl, Wallraven, Bülthoff, & Logothetis, 2009; Katalin M Gothard et al., 2009; Leonard et al., 2012). Overall, despite different temporal and spatial strategies across species, the main finding of the present study is that monkeys and humans were significantly attracted to trustworthy faces.

Obviously, the main question that this finding raises is: what are the features of trustworthy human faces that attract monkeys' attention?"

The question can be addressed at two different levels. We can ask whether monkeys have an abstract representation of human trustworthiness. The answer to this ultimate question is bound to be somewhat speculative at this stage. We can also take one step back and ask about proximal mechanisms, i.e. whether the same underlying facial features drive monkeys' and humans' preference for trustworthy faces. The correlation analyses, showing that in both species, longer looking time is associated with femininity, positive emotional valence, and a low facial width to height ratio", seem to suggest that common facial features determine monkeys' and humans' preferences. This is not entirely surprising as all those dimensions are correlated with trustworthiness judgments and through these correlations - with one another.

The importance of femininity and emotional cues to attribution of trustworthiness has been pointed out in different studies (Nikolaas Oosterhof & Todorov, 2008; Todorov, 2008a). These subjective dimensions are derived from human subjective judgments, and their ecological relevance to monkeys is largely unknown. At 9 months of age, nursery-raised macaque monkeys distinguish the gender of human faces: while not showing a visual preference, they have been found to produce significantly more lip-smacking, which is an affiliative gesture, toward female but not male human faces (Paukner, Huntsberry, & Suomi, 2009). A similar preference towards female faces has been found in 3–4 months human infants (P C Quinn et al., 2002). To our knowledge, no study has investigated whether monkey recognize human emotional cues of sadness/happiness. However, we know that human newborns prefer looking at faces with happy as compared to fearful expressions (Farroni et al., 2007). According to the emotion overgeneralization hypothesis, resemblance of neutral faces to emotional expressions is perceived as indicating the trait attributes associated with these emotions (Knutson, 1996; Montepare & Dobish, 2003; Todorov, 2008b).

An emerging explanation for monkeys' preference for trustworthy human faces is that expertise with human faces enables them to detect gender and, possibly, general emotional valence of the face, hence to perceive trustworthy faces as more positive and approachable than untrustworthy ones.

Social traits inferences are constructed from multiple sources of information. In humans, in addition to the physical facial features contributing to perceived femininity and emotional valence, it has been shown that facial structure is another underlying dimension of trustworthiness. Particularly, faces with lower FWHR are more likely to be judged as trustworthy(Stirrat & Perrett, 2010). We confirm here the findings of Stirrat and Perrett (Stirrat & Perrett, 2010) in humans, using an implicit measure (visual preference) rather than explicit judgments, and we further showed that monkeys have a similar preference for faces with lower FWHR.

Interestingly, recent studies have demonstrated a link between FWHR and social dominance in Capuchin Monkeys (Borgi & Majolo, 2016; Carmen Emilia Lefevre et al., 2014). This raises the possibility that this species-typical facial trait may be used by monkeys to infer social personality of conspecifics, and recycled for making similar inferences about human faces. To our knowledge, no such data is available for macaque monkeys and an interesting further study would be to manipulate fWHR of monkey face pictures and assess looking preference of monkey observers. However, this facial metric correlates with different social judgments (Geniole, Denson, Dixson, Carré, & McCormick, 2015). Thus, this modification may not be sufficient to generate "trustworthy" monkey's faces.

We lack in fact of a monkey's database where personality traits and facial feature are connected, as it is done in the database of human faces. However this database may be realized thank to a combined effort of ethological and engineering studies. A possibility may be to observe monkey's behavior to detect personality traits and correlate these ones with facial features. This attempt may allow understanding how social traits are express by monkey's faces and successively test what these facial features trigger in monkeys' observers. Our results exhort to continue in this direction since nonhuman primates detected certain morphological cues that are associated with trustworthiness.

As previous studies have shown, monkeys are able to observe and interpret human social cooperation, by choosing to interact with individuals who demonstrate reciprocity with peers (Anderson, Kuroshima, et al., 2013; Anderson, Takimoto, et al., 2013). Monkeys also have a preference for human imitators compared to non-imitators (Paukner, Suomi, et al., 2009) in terms of time spent looking at them and interacting with them. The fact that monkeys can distinguish positive and negative social attitudes from humans' non-verbal behavior is consistent with the hypothesis that they are attentive to the visual social cues emitted by our species (Dahl et al., 2009; Katalin M Gothard et al., 2009; Leonard et al., 2012). Such comprehension of human social behavior might also be the basis of monkeys' ability to form human-like "first impression" of human faces.

Darwin proposed that facial displays of emotions serve to predict an individual's current intentions (Darwin, 1872). Social trait inference would extend the prediction to future behavior (Knutson, 1996). Considering the present finding, it is reasonable to assume that the implicit visual preference that monkeys and humans displayed is made possible thanks to a strong predisposition to use not only emotional cues but also stable face characteristics. Invariant and morphological aspects of the face have a fundamental role in making these inferences. The present study suggests that the combined effect of FWHR, happiness and femininity cues induced humans and monkeys' preference for trustworthy faces. Future studies may investigate if individual morphological features alone can induce visual preferences.

Physiognomy is the ancient art of connecting facial features with underlying character. It is unlikely and unexpected that judgments on social traits based on facial features are always accurate; however there might be a reason why evolution is keeping the mechanisms necessary to be sensitive to trustworthiness facial features. Detecting fast who can be approached and whom should be avoided may constitute a basic reflex-like mechanism intrinsically tied with all primates' social survival.

# Appendix 1

## Monkey's preference towards attractiveness human's faces

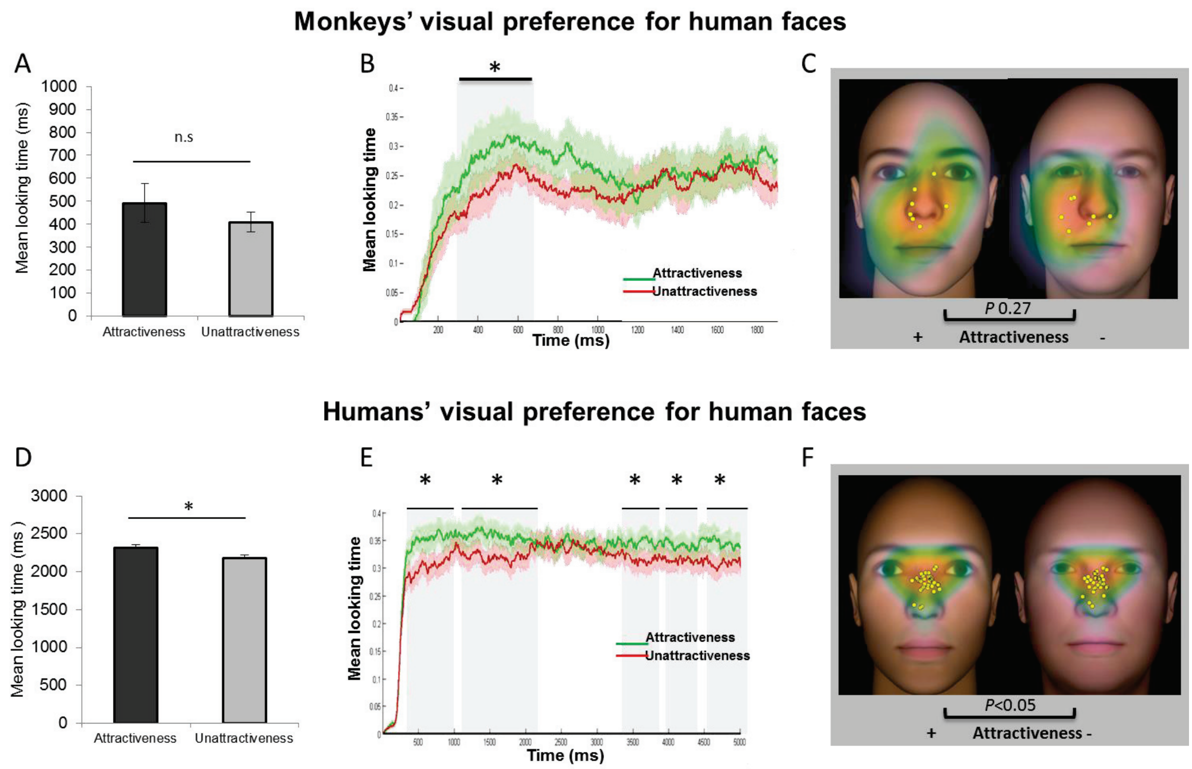
In the previous experiment we observed that monkeys, like humans, are sensitive to trustworthy human modulated faces. To test whether this preference is selective for trustworthiness or it can be extended to other social dimension I investigated preference towards attractive modulated faces selected from Todorov's database of Attractiveness.

Monkeys (N=7) and an independent group of humans (N=29) performed the implicit preferential looking paradigm (the same used for the trustworthiness experiment). The choice of the database of attractiveness relies on the fact that trustworthiness judgments are positively correlated with attractiveness judgments(Nikolaas Oosterhof & Todorov, 2008).

Humans tend to judge trustworthy faces as more attractive than untrustworthy ones. Both dimensions are positively correlated to the general dimension of valence. However, trustworthiness is the dimension that more explains valence judgments. Based on this assumption I expected that monkeys may show the same preference looking more attractive faces than unattractive ones.

***Results***

To evaluate monkeys' looking time from chance level we performed an analysis that revealed that monkeys looked significantly longer at the two faces than predicted by chance (chance level represented with dotted line was 160ms for each face region of interest) ($M_{attractive}$=492.47, SD=227.58, $T_6$=3.86, $p$=0.008; $M_{unattractive}$=409.48, SD=113.5, $T_6$= 5.81, $p$= 0.0011). Overall, monkeys looked more at attractive than unattractive faces, however the preference did not reach significance $p$=0.27. Differently, humans looked significantly longer at attractive than unattractive faces ($T_{28}$= 2.10; $p$=0.04) (Appendix Figure 1, A-D). An effect of preference was observed when testing the temporal dynamic of the preference at each trial. A cluster-based permutation test showed that monkey's preference for the attractive faces (green line) was significant between 300 ms and 560ms ($P$<0.05 corrected for multiple comparison). In humans, the temporal dynamic of the preference revealed five blocks of significant preference for attractive faces (480 to 1000, 1100 to 2200, 3400 to 3800, 4000 to 4400 and from 4600 to the end of the trial)(B-E). Spatial distribution of fixation confirms the same pattern of exploration already observed during the trustworthiness experiment. As for the trustworthiness dimension, average barycenter of fixation was located in the region surrounding the nose in monkeys whereas it is around the eye region in humans (C, F).




*Appendix Figure 1*. ***Looking preference for attractive vs. unattractive faces by rhesus macaques and human subjects.***

**MONKEYS (n=7): (A)** *Mean looking time in milliseconds* (ms) for the most attractive (+3SD of the neutral face) and the least attractive (-3SD of the neutral face) versions of the same facial identities. The error bars denote standard error of the mean. *$P$<0.05. **(B)** *Time course of looking preference*. Mean viewing time on each facial prototype plotted each ms. Green bar represent the period within the trial where the preference was significant (cluster permutation analysis).**(C)** *Gaze heat maps for attractive and unattractive faces* averaged across subjects (attractive face on the left by convention, facial prototype spatial location was counterbalanced within and between subjects). Yellow dots show fixation centers of gravity for each subject.

**HUMANS (n=29) (D)** *Mean looking time in milliseconds* (ms) for the most attractive (+3SD of the neutral face) and the least attractive (-3SD of the neutral face) versions of the same facial identities. The error bars denote standard error of the mean. *$P$<0.05. Humans looked significantly longer at attractive than unattractive faces ($T_{28}$= 2.10; $P$=0.04) **(E)** *Time course of looking preference*. Significant cluster are represented in green. **(F)** *Gaze heat maps for attractive and unattractive faces* averaged across subjects. Yellow dots show fixation centers of gravity for each subject.

***Perceived happiness and femininity for trustworthiness and attractiveness database***

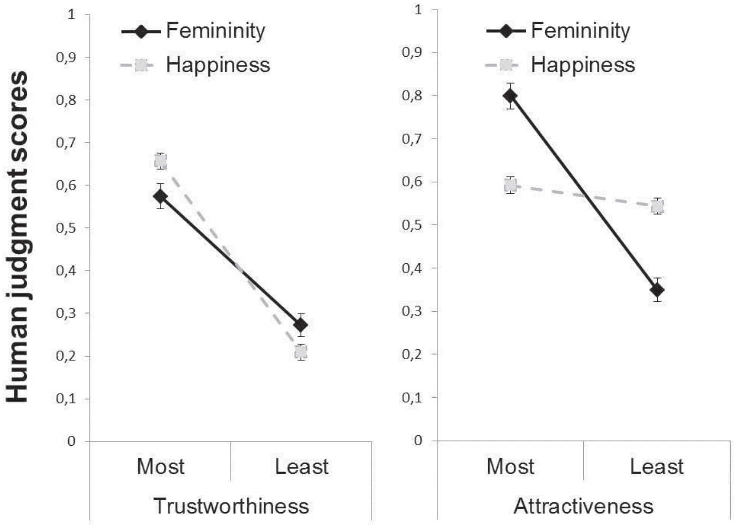
Happiness and femininity judgments are strongly correlated in the trustworthiness database ($R$= 0.83 $P$ =0.0001). These facial features are also correlated in the attractiveness database but significantly less (Fisher test: Z=4.098, $p$ <0.0001; $R$= 0.33, $p$ =0.016).

To assess weather monkeys's attention towards trustworthy and attractive faces may have been biased by these facial attributes I asked an independent group of participants to judge the level of happiness and femininity of Todorov's modulated faces. Subjects' choices have been then correlated to monkeys and humans' viewing time over the trustworthy and attractive faces database.

Scores of happiness and femininity judgments for each face of the "Trust" and "Attractive" database were obtained from an independent sample of 12 human subjects. All possible pairs of faces (from both databases) were presented in a random order. Half participants judged which of the two faces was the happiest. The other half judged which one was the most feminine. Each of the 50 faces was evaluated 16 times. Inter-subjects reliability in pairs evaluation was high for both, happiness ($R$=0.74) and femininity ($R$=0.65).

To provide further evidence that femininity cues are not sufficient, and that both cues are necessary to form a preference in monkeys, we assessed whether judgments of emotion and femininity from the two databases were different across trust and attractiveness. To that aim, we performed an ANOVA for human judgments of happiness and of femininity of faces with the valence of the face (negative, positive) and the database of the face (attractiveness, trustworthiness). Human judgments scores are reported in Figure 3.

Coherently with a previous study (Nikolaas Oosterhof & Todorov, 2008), these analyses showed a main effect of the valence for each social dimension (Left and right graph of Figure 3): faces judged as more trustworthy or more attractive were also judged as more happy ($F_{1,5}$= 10.94; $P$<0.05) and as more feminine ($F_{1,5}$= 44.20; $P$ <0.001) than less trustworthy or less attractive ones. More interestingly, for judgments of happiness, there was a significant interaction between valence and database type ($F_{1,5}$= 66.10; $P$ <0.001): Post hoc Duncan tests showed that faces judged as more trustworthy were also judged as more happy ($P$<0.001) than untrustworthy ones (Appendix Figure 2) but that faces judged as very attractive or unattractive were not judged to have different emotions ($P$= 0.23, right graph of Figure 3). When participants judged femininity, no interaction appeared between the valence and the database of the face ($F_{1,5}$= 2.87; $P$> 0.15). Hence, the results suggest that when both happiness and femininity are available monkeys are able to show a preference. In other words, to perform judgments of trustworthiness, both happiness and femininity could be used as cues, but to perform judgments of attractiveness, only femininity may subjectively distinguish the faces between them.




*Appendix Figure 2*: On the left average of human subjects' judgment on happiness (grey dotted line) and femininity (black full line) for the more trustworthy and the less trustworthy faces of the "Trustworthiness" database. On the right average of human subject's judgment on happiness (grey dotted line) and femininity (black full line) for the more attractive and the less attractive faces of the "Attractiveness" database. Faces judged as more trustworthy or more attractive were also judged as more happy and as more feminine than less trustworthy or less attractive ones. The interaction effect when participants judged happiness means that faces judged as more trustworthy were also judged as more happy than untrustworthy ones (left graph) but that faces judged as very attractive or unattractive were not judged to have different emotions (P= 0.23, right graph). When participants judged femininity, no interaction appeared between the valence and the database of the face. The results show that happiness is an important facial cue that allow to disciminate between trustworthiness and attractiveness database. This may suggest a reason for the lack of looking preference in monkeys.

### *Monkeys and human's correlation between looking time and facial attributes (happiness and femininity)*

Monkeys' total viewing time on a face from the trustworthiness database was positively correlated to the emotion score ($R$=0.50 $P$< 0.001) and to the face femininity score ($R$=0.57 P< 0.001). Hence, in the trustworthiness database, monkeys looked longer at faces that were both happier and more feminine. On the contrary, in the attractiveness database, no correlation was observed between monkeys' total viewing time on faces and the emotion ($R$=0.09; $P$= 0.33) or femininity scores ($R$=0.22, $P$= 0.15). In the trust database, humans like monkeys, looked longer at faces that were both happier ($R$=0.65 $P$< 0.001) and more feminine ($R$=0.77 $P$< 0.001). In the attractiveness database, humans' looking time like that of monkeys, were not correlated to the emotion score (R =0.09; *P*= 0.33). However, humans who showed a preference for attractive human faces, unlike monkeys, also looked longer at faces judged as more feminine (R=0.41, *P*<0.05). These results suggest that the preference for attractive faces in humans may be driven by femininity cues of the face for humans only.

To identify which factor was more predictive of the viewing time behavior in monkeys and in humans, we performed a multiple regression analysis using the three facial cues (FWHR, femininity judgments and happiness judgments) to predict viewing time for the two social dimensions: trustworthiness and attractiveness. We observed the following beta scores: for Monkeys, in Trust: $\beta_{FEMININITY}= 0.584$, $T_{46}=2.4$, $P<0.05$; $\beta_{HAPPINESS}= -0.066$, $T_{46}<1$; $\beta_{FWHR}= -0.133$, $T_{46}<1$; and in attractiveness: $\beta_{FEMININITY}= 0.242$, $T_{46}=1.6$, $P=0.12$; $\beta_{HAPPINESS}= -0.060$, $T_{46}<1$; $\beta_{FWHR} = 0.152$, $T_{46}<1$; for humans, $\beta_{FEMININITY}= 0.687$, $T_{44}=3.6$, $P<0.001$; $\beta_{HAPPINESS}= -0.051$, $T_{44}<1$; $\beta_{FWHR}= -0.08$, $T_{44}<1$ in Trust and $\beta_{FEMININITY}= 0.424$, $T_{44}=2.9$, $P<0.01$; $\beta_{HAPPINESS}= -0.030$, $T_{44}<1$; $\beta_{FWHR}= 0.030$, $T_{44}<1$ in attractiveness. Hence, this multiple regression analysis suggests that in monkeys, femininity is a crucial cue in detecting trustworthiness but not in detecting attractiveness. Moreover, it suggests that femininity is a crucial cue in both trustworthiness and attractiveness preference in humans. Because happiness and femininity judgments were strongly correlated in the trustworthiness database (*R*= 0.83 *P* =0.0001; they were also correlated in the attractiveness database but significantly less - Fisher test: Z=4.098, *P* <0.0001; *R*= 0.33, *P* =0.016), it is possible that both subjective cues are nonetheless involved in this detection.

Together these results suggest that for the attractiveness database, even if the faces can be subjectively distinguished in terms of femininity, this cue alone does not seem sufficient or relevant for monkeys, as they do not display a preference for attractive human faces, unlike human participants. Hence, it is possible that femininity cues are not processed in the same way across species to infer social judgments. Pascalis and colleagues (under review) showed that monkeys do have a preference for attractive monkey faces, suggesting that monkeys can discriminate such cues within their own species. However, in monkeys, attractiveness may play a fundamental role in mating selection (Brighina et al., 2002), which could explain why these cues are not relevant across species as such social interactions does not generally occur across species.

# Appendix 2

## Preference towards trustworthy faces: an innate ability or expertise-based

It is reasonable to assume that monkeys' preference is possible thanks to a strong predisposition to use facial cues for controlling approach/avoidance behavior. However, it is unclear whether the preference toward trustworthy human faces reflects an innate capacity shaped by experience or whether it arises only from expertise.

A previous study has shown that differences in serotonin transporter (5-HTTLPR) can predict scanpath differences (Gibboni et al., 2009). We then performed a genotyping analysis in order to find out whether genetic differences between monkeys may explain the different pattern of visual explorations.

A genotyping analysis for the promoter region of the serotonin transporter regulatory gene (5-HTTLPR) was performed using the same method of (Gibboni et al., 2009).

Whole blood was collected from the saphenous vein and the ear capillary of monkeys as previously described (A. Lefevre, 2015), blood samples treated with heparin are stored at–20°C. Genomic DNA was isolated from the samples using QIAmp DNA micro Kit following the manufacturer's instructions (Qiagen, Courtaboeuf, France).This protocol is for isolation of genomic DNA from small volume, in our experiment only 90 µl of whole blood was used. Samples were quantitated using The QuantiFluor® dsDNA System with the Quantus™ Fluorometer. (Promega Corporation 2800 Woods Hollow Road Madison, WI 53711-5399 USA). The rh-5-HTTLPR was amplified from 25ng of genomic DNA with flanking oligonucleotide primers (forward, 5'- GGCGTTGCCGCTCTGAATGC; reverse, 5'-CAGGGGAGATCCTGGGAGGG)(Barr et al., 2004) in 15 µL reactions with Platinum Taq and the PCRX Enhancer System kit, according to the manufacturer's protocol (Invitrogen, Carlsbad, California). Amplifications were performed on LightCycler® 2.0 Instrument ((RocheDiagnostics, GmbH, Mannheim, Germany) with one cycle at 94°C for 5 min followed by 30 cycles of 94°C for 15 sec, 62°C for 10 sec, 68°C for 30 sec, and a final 3-min extension at 72°C. Amplicons were separated by electrophoresis on the Agilent 2100 Bioanalyzer system using High Sensitivity DNA Analysis Kits. Short (s, 388 bp) and long (l, 419 bp) alleles of the rh5-HTTLPR were identified as in previous studies(Dobson & Brent, 2013; Lesch et al., 1997).

Results revealed that the monkeys that performed trustworthiness experiment had different genotypes. Three monkeys (O-Y-Z) were homozygous for the short allele (S/S) of the serotonine transporter gene; four monkeys (T-E-V-S) were homozygous for the long allele (L/L), and one monkey (D) was heterozygous (L/S) (Appendix Table 1).

Previous studies have reported that the presence of a short allele of the serotonin transporter is a pattern of anxiety and stress sensitivity in humans(Caspi et al., 2003; Gotlib, Joormann, Minor, & Hallmayer, 2008; Karg, Burmeister, Shedden, & Sen, 2011; Mueller, Brocke, Fries, Lesch, & Kirschbaum, 2010) and monkeys(Barr et al., 2004; Champoux et al., 2002; Watson, Ghodasra, & Platt, 2009). This effect is higher when individuals with short allele are exposed to stressful life experiences(Gotlib et al., 2008; Mueller et al., 2010).

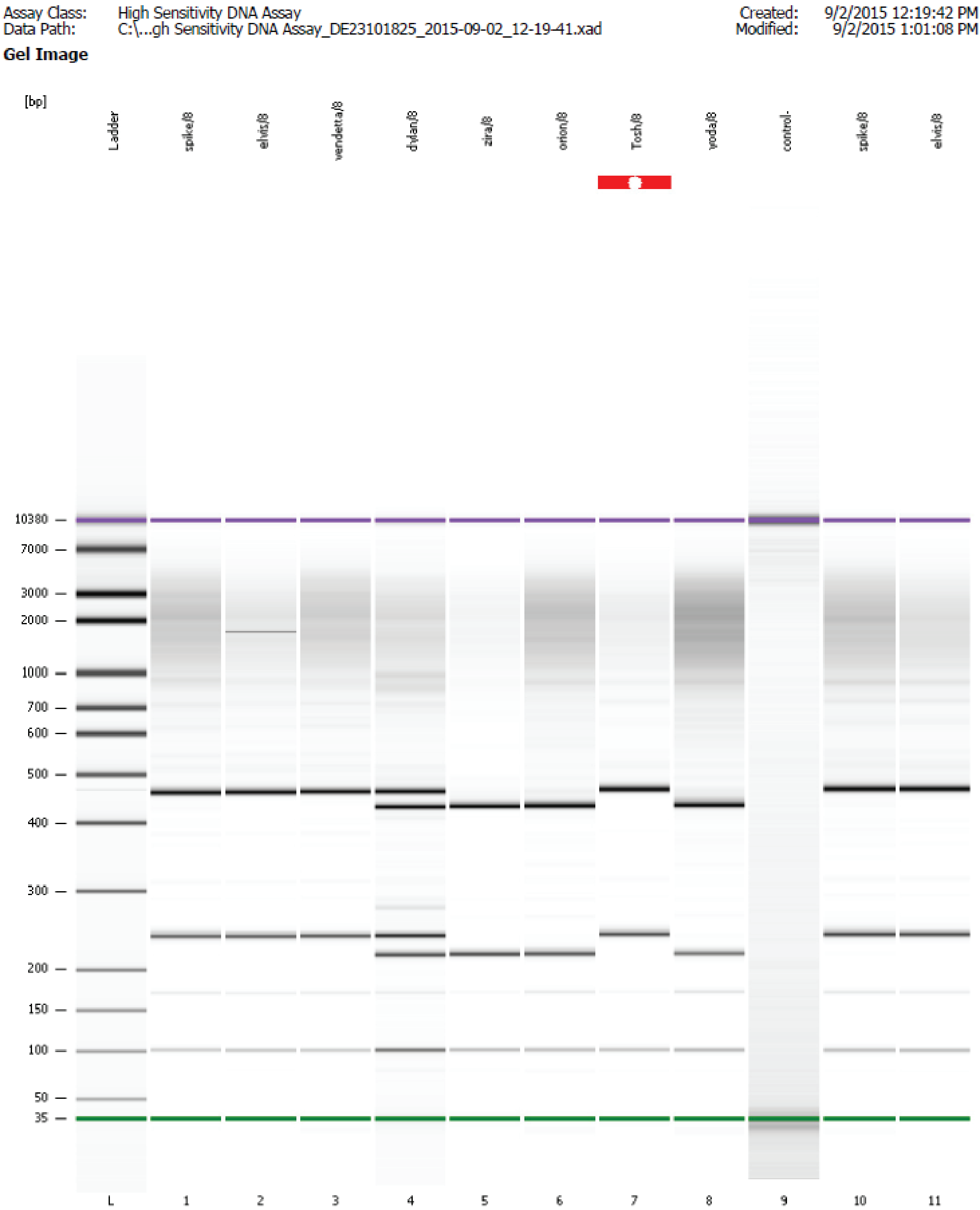
Differences in monkeys' scan paths match with differences in genotype(Gibboni et al., 2009). I considered whether the serotonin transporter regulatory gene (5-HTTLPR) may have an impact on the monkey's preference towards trustworthy faces correlating looking preference of individual monkeys and the genotype. Since untrustworthy faces are more threatening than trustworthy faces, I hypothesized that S/S monkeys may had a strongest preference towards trustworthy faces. Results did not confirm this hypothesis. The correlation performed between looking time and serotonin transporter regulatory gene (5-HTTLPR) was not significant (R=-0.34 P=0,401). Nevertheless a trend was observed when we controlled for sex (R=-0.74 P=0,057).

Face processing ability in monkeys can be affected by stimulus exposition acquired with age (Sugita, 2008). To test the impact of learning on monkey's behavior during the task, we further performed a correlation between looking time and monkey's age. Age was selected as an indicator of monkey's expertise in interacting with humans. Neither in this case the result reached significance (R=0.53 P=0,169). Still, a trend was observed when controlling for sex (R=0.74 P=0.053).

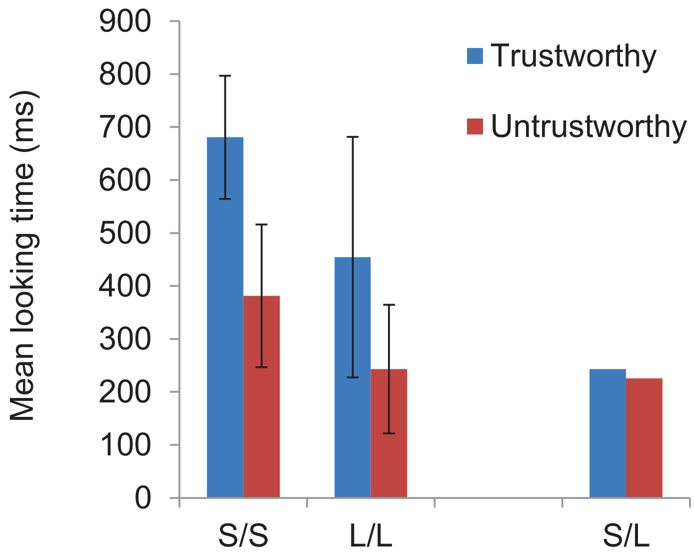
Finally, I performed a partial correlation controlling the influence of genotype in order to understand the effect of age on looking time (R=0.874, P<0.05 corrected for multiple comparison).

In the light of these findings we cannot exclude that monkey's experience may generate a difference in the expression of genetic predisposition. A possible explanation is that young S/S individuals do not have a clear preference towards trustworthy human faces compared to L/L individuals. Faced with the opportunity to interact with a trustworthy and untrustworthy person they may be more prone to shift their attention towards both positive and negative valence of stimuli (Belsky et al., 2009; Homberg & Lesch, 2011; Uher, 2008). However this behavior may change as an effect of positive interaction with humans. In other words, SS individuals may need more experiences and interaction to look at trustworthy faces

(Appendix Figure 4). This interpretation needs to be taken with caution because of the small sample and the distribution (age and genotype in our sample are strongly correlated (R=-0,96 P<0.05), note that this correlation do not exist in real life). Thus perhaps with a larger monkey sample sizes more robust results might have been found.

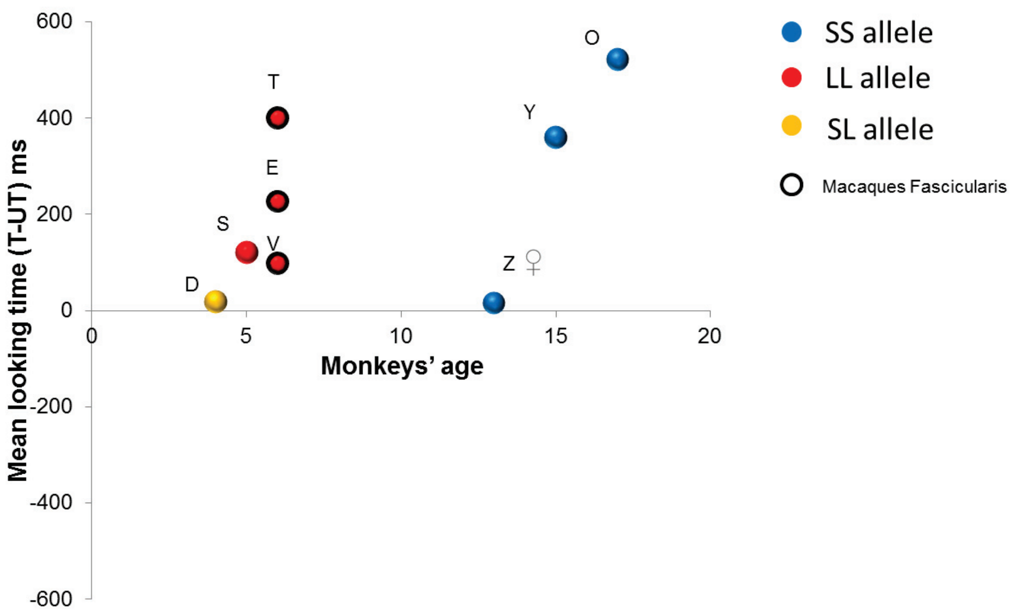



*Appendix Figure 3.* **Genotyping analysis.** Classification of the 8 monkeys for the promoter region of the serotonin transporter regulatory gene (5-HTTLPR).




*Appendix Figure 4.* **Looking time towards trustworthy and untrustworthy faces following the** classification for the promoter region of the serotonin transporter regulatory gene (5-HTTLPR). Based on the genotyping analysis monkey O-Y-Z were are SS. Monkeys T-E-S-V are LL. Monkey D is SL.

## Genotype and age to explain monkeys' preference




*Appendix Figure 5.* **Correlation between monkeys' age and genotypes and looking preference towards trustworthy faces**

**Appendix Table 1.** *Compilation of biographical, behavioral and genetic characteristics of the eight monkeys for trustworthiness and attractiveness experiment.*

| | **Age at the time of the experiment** | **Genotype** | **Sex** | **Specie** | **Preference towards trustworthy faces** | **Total Looking time at the faces** |
|---|---|---|---|---|---|---|
| O | 17 years | S/S | M | Rhesus Macaque | 522,39 | 1282,60 |
| Y | 15 years | S/S | M | Rhesus Macaque | 360,00 | 657,29 |
| S | 6 years | L/L | M | Macaque fascicularis | 98,15 | 308,15 |
| T | 6 years | L/L | M | Macaque fascicularis | 400,39 | 870,13 |
| V | 5 years | L/L | M | Rhesus Macaque | 121,4 | 885,20 |
| Z | 13 years | S/S | F | Rhesus Macaque | 15,74 | 1245,95 |
| E | 6 years | L/L | M | Macaque fascicularis | 226,44 | 726,18 |
| D | 4 years | S/L | M | Rhesus Macaque | 17,81 | 468,43 |

| | **Age at the time of the experiment** | **Genotype** | **Sex** | **Specie** | **Preference towards attractive faces** | **Total Looking time at the faces** |
|---|---|---|---|---|---|---|
| O | 17 years | S/S | M | Rhesus Macaque | -239,9 | 569,5 |
| Y | 15 years | S/S | M | Rhesus Macaque | 326 | 1071,5 |
| S | 6 years | L/L | M | Macaque fascicularis | 17,7 | 886,7 |
| T | 6 years | L/L | M | Macaque fascicularis | 228,2 | 1273,4 |
| V | 5 years | L/L | M | Rhesus Macaque | 64,6 | 784 |
| Z | 13 years | S/S | F | Rhesus Macaque | 140,2 | 1240 |
| E | 6 years | L/L | M | Macaque fascicularis | 44 | 461,6 |

# Chapter 3

# How components of facial width to height ratio differently contribute to the perception of social traits

This chapter is a modified version of:

Manuela Costa, Guillaume Lio, Alice Gomez, Angela Sirigu.
How components of facial width to height ratio differently contribute to the perception of social traits.
Submitted to *Psychological Science*.

## Abstract

*Aim:* Perception of social traits is largely influenced by morphological and stable facial features. Facial width to height ratio (fWHR) is a reliable morphological cue that correlates with sexual dimorphism and social traits like trustworthiness, aggressiveness and dominance. Still, it is currently unclear how vertical and horizontal morphological traits the two components of fWHR, distinctly capture faces' social information.

*Methods:* Using a new methodology, we orthogonally manipulated the upper facial height (vertical component) and the bizygomatic width (horizontal component) in order to test the selective and the combined effect of fWHR in the formation of face first impressions.

Subjects (*N*=90) saw on the screen pair of faces and were requested to select the face expressing better the social trait being tested (trustworthiness, aggressiveness and femininity). Using a female/male databased we further investigated how sex and fWHR components interact in the formation of these judgements.

*Results:* In all experiments and conditions, the vertical component always better predicted participants' judgements than either the horizontal component or their combined effect. Perception of trustworthiness changed as a function of upper height manipulation: faces with smaller height were perceived as less trustworthy, less feminine and more aggressive than medium or higher upper height faces values. Judgements of aggressiveness and femininity but not of trustworthiness were affected also by the horizontal.

*Discussion:* By distinctly testing the effect of horizontal and vertical components from the effect of the ratio, we show that upper facial height and byzigomatic width weigh differently in the formation of social impression. Our findings demonstrated that by dissociating fWHR into two dissociable components we can obtain a much powerful and discriminative measure of how facial morphology biases social judgements.

## 3.1. Introduction

Facial perception is largely influenced by detection of emotions such as smiling, frowning, fearfulness. However, it is also influenced by some morphological and stable factors such as gender, skin color and facial width to height ratio (fWHR). Among these factors fWHR has recently received great attention (Geniole et al., 2015; Michael P. Haselhuhn, Ormiston, & Wong, 2015).

Researches in social psychology showed that fWHR is used implicitly to form social judgments from facial appearance. Male faces with higher fWHR are more likely to be judged as untrustworthy (Stirrat & Perrett, 2010, 2012), dominant ( Hehman, Leitner, Deegan, & Gaertner, 2015; Mileva, Cowan, Cobey, Knowles, & Little, 2014), more powerful and competent (Hehman, Leitner, & Freeman, 2014). Strikingly, this measure of facial appearance can have strong impact on real life since it has been recently considered in the context of sentencing decisions where the prisoners' fWHR has contributed to the jury's decision (Hehman, Leitner, Deegan, & Gaertner, 2013).

A link between fWHR and behavioral tendencies has also been established (Carré & McCormick, 2008; M. P. Haselhuhn & Wong, 2012; Stirrat & Perrett, 2010). For example, it has been suggested that man with higher fWHR have a higher propensity to aggression (Carré, McCormick, & Mondloch, 2009; Carré & McCormick, 2008), they are more likely to show unethical behavior such as deception, cheating (Haselhuhn & Wong, 2012) self-interest (Haselhuhn, Wong, & Ormiston, 2013) or little propensity to trust others (Stirrat & Perrett, 2010).

Recent studies however have not replicated these findings (Deaner, Goetz, Shattuck, & Schnotala, 2012; Gómez-Valdés et al., 2013; Kramer, Jones, & Ward, 2012; Özener, 2012), especially when trying to correlate fWHR with individuals' real behavior. For instance, Deaner et al. (2012) found that body weight, and not fWHR, predicts aggression in hockey player. Along the same line, Gómez-Valdés et al. (2013) did not find any link between fWHR and bellicose tendencies in male mexican prisoners.

Moreover, as pointed by Geniole and colleagues (2012) data are lacking concerning the generalizability of this effect for female faces. In line with this, Carré and McCormick (2008) found a relationship between fWHR and aggressiveness in man only. The lack of fWHR effect for woman' faces has been also reported by Stirrat and Perret (2010) when studying trustworthiness or by Haselhuhn and Wong when investigating unethical behavior (2012).

Such contrasting findings may suggest that fWHR isn't perhaps a reliable facial dimension strong enough to influence social perception. Others may argue that disparity of results could also reflect differences in the method used among the different studies given that no strict consensus exist on how evaluate fWHR. In other words, we might still lack a validated fWHR method as a standard of face metric.

FWHR is a combined measure obtained by dividing byzigomatic width, the distance between the left and right zygion of the face, by upper facial height, distance between the upper lip and mid-brown. No study has yet investigated how the two components distinctly contribute to the formation of social impression.

Perhaps the only exception is the study conducted by Weston, who employed this method to search for a morphological cue of sexual dimorphism. He showed that, in male, facial width grows proportionally with body size but not upper facial height. Hence, for the first time the two components were measured separately and then combined to account for the growth of body size.

Following studies have systematically employed fWHR as a whole measure, mainly by modifying the entire proportion of the face's features, without controlling the effect of each single component on social judgments. Indeed, variations on byzigomatic width or variations of upper facial height can modify fWHR and produce different percepts. There are evidence supporting a link between upper facial height and some human characteristics (behavior, sex), independently from facial width.

Upper facial height, and not facial breadth, is a potential target of selection during evolution, (Weston et al., 2007). In fact, Weston and colleagues (2007) reported that facial height can unambiguously distinguish an adult male from a female face, whereas the facial width may fluctuate with variation in body size. Despite signaling sex differences, upper facial height may also reveal other characteristics. In fact, faces with smaller upper height have been shown to display more bite force, a trait that may play a crucial role in survival (Proffit, Fields, & Nixon, 1983; Raadsheer, van Eijden, van Ginkel, & Prahl-Andersen, 1999).

The objective of this study is to show that fWHR can become a more powerful tool when considering how upper facial height (vertical component) and byzigomatic width (horizontal component) are used by people to infer others' personality or to form fast social impressions. More precisely, the present study aims to investigate the independent and combined function of vertical and horizontal components of fWHR during the formation of different social judgements like trustworthiness, femininity and aggressiveness.

Additionally, given the lack of effect for female faces as previously reported (Carré & McCormick, 2008; Geniole et al., 2012; Michael P. Haselhuhn et al., 2013; Stirrat & Perrett, 2012), we investigated the role of fWHR for trustworthiness, femininity and aggressiveness judgements using a female database. If gender bias social perception, we expect differences in vertical and horizontal manipulation when using female faces. To test this effect, we built two databases, male and female faces, where all modifications have been done using identical methods (Figure 14).

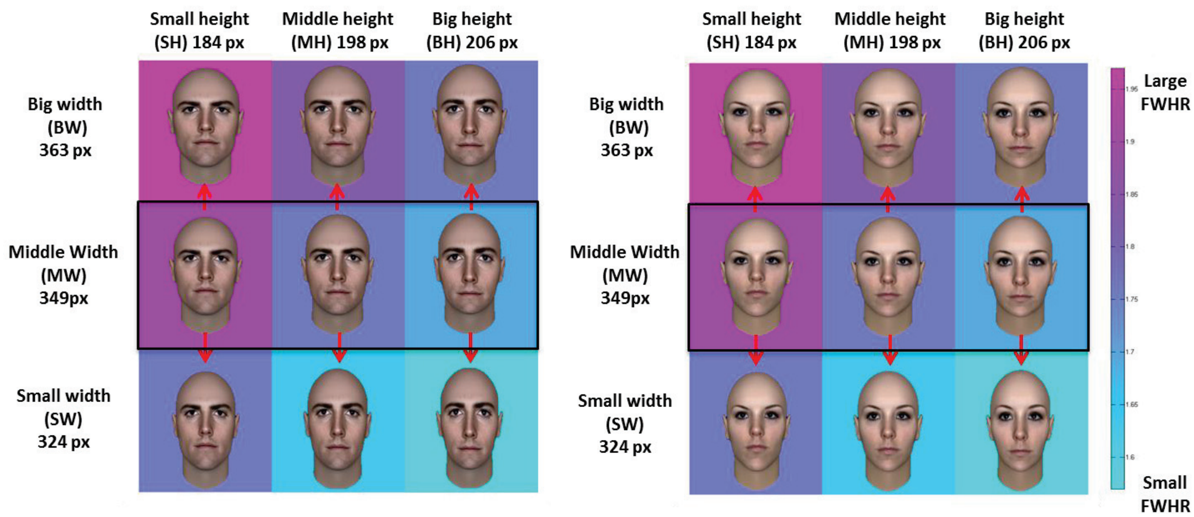



*Figure 14.* **Categories of male and female stimuli.** Typologies of faces from male (on the left) and female (on the right) dataset for the three visual conditions according to vertical modification (Small Height, Middle Height, Big Height) and to horizontal modification (Big Wide; Middle Wide; Small Wide). Color bar represents the FWHR. Big Wide Height, Small Wide Height and Middle Wide Height faces have the same FWHR.

Overall, we performed six experiments to investigate whether both components - vertical and horizontal - are significant predictors of social judgments for trustworthiness (Experiment 1, 2), aggressiveness (Experiment 3-4) and femininity (Experiment 5-6).

## 3.2.Methods

To test the effect of each dimension, we orthogonally manipulated the upper facial height (vertical component) and the bizygomatic width (horizontal component) to disentangle their contribution to face impression. Consider for few seconds the three faces in Figure 12, do they represent the same level of trustworthiness or aggressiveness or femininity? These faces have exactly the same fWHR but the size of each component (upper facial height, bizygomatic width) differs. As Figure 15 should suggest face with identical fWHR but with different combination of the two components (vertical and horizontal) can trigger different social judgments.

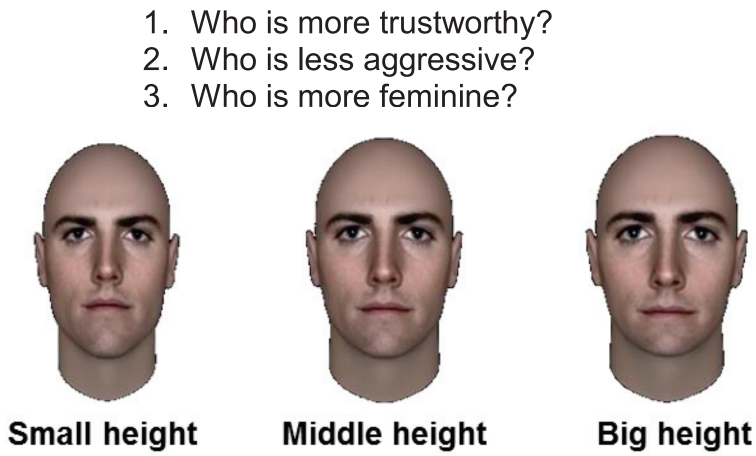



*Figure 15.* **Can faces with the same fWHR trigger different social perception?** Example of three faces selected from the database used in the present study. These faces have the same facial width to height ratio (fWHR) obtained by differently combined upper facial height and bizygomatic width.

### Participants

Six independent groups of French-speaking subjects with normal vision (*N*=15 for each experiment) participated in this study. Participants were randomly selected to attend the trustworthiness, femininity or aggressiveness experiment, using either male or female database (3 social judgments X 2 sex of database used). In each group we recruited 7 males and 8 females (age between 21 and 43 years, M=27.6 and SD=5.8).

### Face stimuli

The original dataset was composed of three male and three female Caucasian faces. All stimuli were oriented straight and with a neutral expression. Eyes positions were aligned.

To generate the vertical modification, the upper facial height of each face was manipulated using the "face-brow-nose-chin-ratio" of FaceGen Modeller 3.5. Three categories of vertical faces were created (Figure 16): 1) the small height faces (SH; 184 px), corresponding to +3 in FaceGen Modeller; 2) the middle height faces (MH; 198 px) corresponding to 0 in FaceGen Modeller; 3) the big height faces (BH; 206 px) corresponding to -3 in FaceGen Modeller.

To generate the horizontal modification, the bizygomatic width of each category of vertical faces (SH, MH, BH) was modulated using Gimp (Version 2.8, http://gimp.org). Three categories of horizontal faces were created (Figure 16): 1) the small wide faces (SW); 2) the middle wide faces (MW); 3) the big wide faces (BW). The degree of horizontal modification was determined using the fWHR (1.76) of the baseline face (MWxMH) and the value of vertical modification previously generated. Thus, big width faces (BW) have a width calculated in a way that the face with the biggest height (BWxBH) has an fWHR of 1.76. Likewise, the small width face has a width (SW) calculated in a way that the face with the smallest height (SWxSH) has an fWHR of 1.76 (Appendix Figure 5). The same methods have been used to modulate female faces.

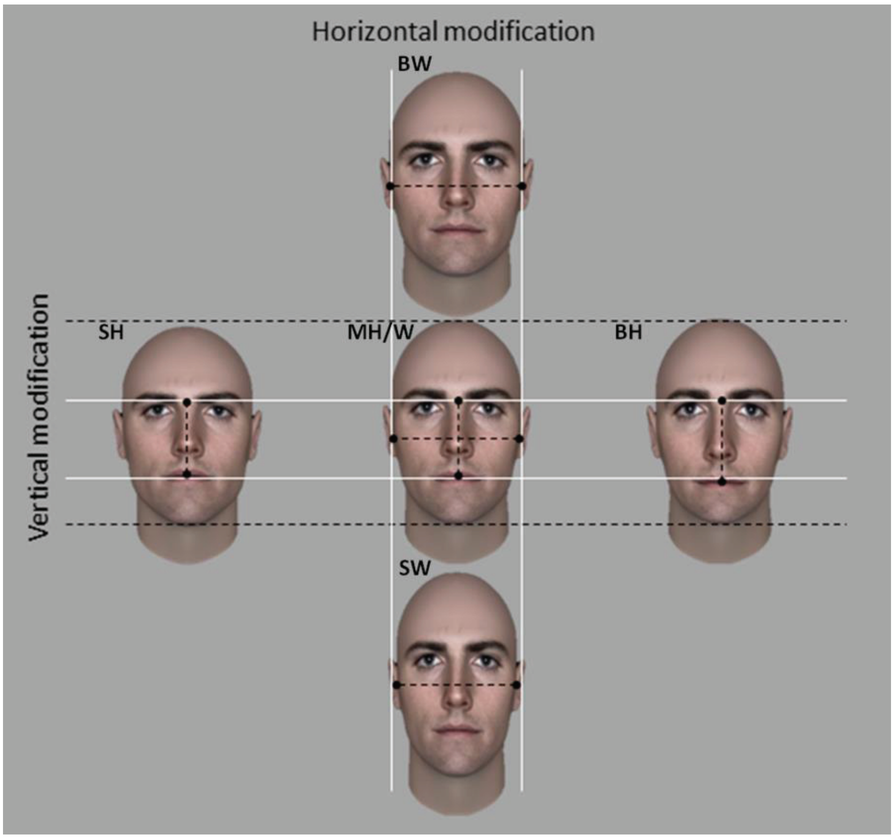



*Figure 16.* **Vertical and Horizontal modification.** Examples of computer-generated identities modulated for the upper facial height (vertical modification) and byzigomatic width (horizontal modification). From left to right, vertical modification: Small Height face (SH) corresponds to +3 in FaceGen Modeller (face-brown-chin ratio); Middle Height face (MH), corresponds to 0 zero and Big Height face (BH) corresponds to -3. From top to bottom, horizontal modification: Big Wide (BW), Middle Wide (MW), Small Wide faces (SW).

Following these metrics, each face was modified in 9 different ways, resulting in a final database composed of 54 stimuli (3 vertical x 3 horizontal modifications x 3 identities x 2 conditions, female or male) (Figure 14). The same dataset has been used for the six experiments.

Importantly, middle faces have been chosen to represent exactly the median value for faces in the distribution of real population. We then used middle faces as starting point for all modifications and we further assess that also modified faces of our database were still within the same distribution by comparing upper facial height and byzigomatic width of our stimuli with a large sample of real measures. The metric of faces for upper facial height and byzigomatic width for real faces is provided by "FaceBase" database (Weinberg et al., 2015). Z-scores were computed for each category of stimuli generated with to respect to FaceBase data (Appendix figure 7-8).

**Task procedure**

Faces have been presented side by side in a screen of 1920 X 1200 pixels using Matlab, the image resolution was 757 X 820 pixels. Rating of faces consisted in multiple presentations of two randomly selected stimuli from the 27 of the entire Database. All possible couples have been presented for a total number of 351 trials per participant. At each trial, participants (*N*=15) had to choose the most 'trustworthy' or the most 'feminine' or the most 'aggressive' face, by pressing the corresponding key on the keyboard. Participants knew that they were attending a study on first impression where they were encouraged to respond with their "gut feelings". The score for each face was calculated as the average of all scores obtained from subjects' choices (face selected=1; otherwise 0). Because aggressiveness judgments are negatively correlated to trustworthiness and femininity judgments, values reported were then transformed as follow: Used value=1-observed value to clarify the effect across experiments. This transformation does not modify any of the statistical analysis.

For each experiment, to control variability, the averages across identities have been performed in accord with the vertical component (SH/MH/BH) and the horizontal component (SW/MW/BW). For each experiment a mixed ANOVA was run on the average rating score for the vertical (SH/MH/BH) and horizontal components (SW/MW/BW) as within subject-variable. To test the effect of stimulus type (male or female dataset), we included "type of stimulus" (male/female) as categorical factor. Effect of linearity has been tested using planned comparison. All post-hoc analyses were done using Bonferroni correction.

## 3.3. Results

### 3.3.1. Trustworthiness judgments

A mixed ANOVA showed that perceived trustworthiness of faces was significantly modulated by the vertical component, $F(2,56)=13.59$, $p<.001$, $\mu^2=.33$. Planned comparisons showed a linear effect $F(1,28)=7.025$, $p<.01$. Faces with smaller height were judged as less trustworthy ($M_{SH}=0.43$, SD=0.018 < $M_{MH}=0.55$, SD=0.11 < $M_{BH}=0.51$, SD=0.02) (Figure 17-A, left graph). This effect was independent from sex, (male/female database used), as there was no interaction ($F<1$) between sex and the vertical component.

The horizontal component also slightly biased participants perceived trustworthiness of faces, $F(2,56)=4.61$, $p<.05$, $\mu^2=0.14$. However, planned comparisons did not show a linear effect $F(1,28)=1.77$, $p=0.19$. Post hoc analysis showed that faces with smaller width were judged as less trustworthy ($M_{SW}=0.47$, SD=0.01 < $M_{MW}=0.52$, SD=0.01 = $M_{BW}=0.50$, SD=0.01) (Figure 17-A, right graph). This effect was independent from sex, (male/female database used), as there was no interaction ($F<1$) between sex and the vertical component.

For trustworthiness judgments, the ANOVA showed a significant interaction between the vertical and the horizontal components: $F(4,112)=5.90$, $p<.001$, $\mu^2=0.17$. A triple interaction with sex and components (vertical X horizontal X sex), $F(2,112)=3.15$, $p<.05$, $\mu^2=0.10$ was also observed. Planned comparisons showed that vertical/horizontal interaction was significant for male $F(4,56)=7.29$, $p<.001$, $\mu^2=0.34$ but not for female faces ($F<1$).

### 3.3.2. Aggressiveness judgments

A mixed ANOVA showed that perceived aggressiveness of faces was strongly modulated by the vertical component, $F(2,56)=83.1$, $p<.0001$, $\mu^2=0.75$. Planned comparison showed a linearity effect $F(1,28)=90.82$, p<.001. Faces with smaller height were judged as more aggressive ($M_{SH}=0.35$, SD=0.018 < $M_{MH}=0.51$, SD=0.012 < $M_{BH}=0.62$, SD=0.019) (Figure 17-B left graph). There was a significant interaction between sex and vertical component, $F(2,56)=7.70$, $p<.005$, $\mu^2=0.21$. Planned comparison showed a significant linear effect for female $F(1,28)=21.85$, $p<.001$ and male faces $F(1,28)=77.50$, $p<.001$. The linear difference between the two sex was also significant $F(1,28)=8.5$, $p<.01$.

The horizontal component also biased participants perceived aggressiveness of faces, $F(2,56)=18.5$, $p<.001$, $\mu^2=0.39$. Planned comparison showed a linear effect $F(1,28)=20.90$,

$p<.001$, faces with smaller width were judged as less aggressive ($M_{SW}=0.55$, $SD=0.018$ > $M_{MW}=0.49$, $SD=0.015$ > $M_{BW}=0.45$, $SD=0.016$) (Figure 17-B right graph). This effect was independent from sex, i.e., database used, as there was no interaction between sex and horizontal component, $F(2,56)=2.26$, $p=.11$.

Finally, there was no significant interaction between vertical and horizontal component $F(4,112)=1.33$, $p=.26$, and no triple interaction between components and sex (vertical X Horizontal X sex) $F(4,112)=1.41$, $p=.23$.

### 3.3.3. Femininity judgments

A mixed ANOVA showed that perceived femininity of faces was strongly modulated by the vertical component, $F(2,56)=46.9$, $p<.001$, $\mu^2=0.62$. Planned comparisons showed a linear effect $F(1,28)=49.99$, $p<.001$: faces with smaller height were judged as less feminine ($M_{SH}=0.42$, $SD=0.015$ < $M_{MH}=0.51$, $SD=0.009$ < $M_{BH}=0.56$, $SD=0.013$) (Figure 17-C left graph). There was a significant interaction between sex and vertical component, $F(2,56)=3.39$, $p<.05$, $\mu^2=0.10$. Planned comparison, showed, as in the aggressiveness experiment, a linear effect when using both female $F(1,28)=13.19$, $p<.01$ and male database $F(1,28)=40.54$, $p<.001$. Finally, using the difference between values obtained for the male and female database we observed only a trend for the linear effect $F(1,28)=3.75$, $p=.063$.

The horizontal component also biased participants perceived femininity of faces, $F(2,56)=15.8$, $p<.001$, $\mu^2=0.36$. Planned comparison showed a linear effect $F(1,28)=17.97$, $p<.001$. Faces with smaller width were judged as more feminine ($M_{SW}=0.55$, $SD=0.013$ < $M_{MW}=0.49$, $SD=0.010$ < $M_{BW}=0.45$, $SD=0.014$) (Figure 17-C right graph). This effect was independent from the sex, (male/female database), as there was no interaction ($F(2,56)=2.18$, $p=.12$) between sex and horizontal component.

Finally, the interaction between vertical and horizontal components was not significant $F(4,112)=2.14$, $p=.08$, but there was a significant triple interaction (vertical X horizontal X sex) $F(4,112)=2.47$, $p<.05$, $\mu^2=0.08$. Planned comparison showed that the vertical and horizontal interaction was significant for male faces $F(4,56)=2.80$, $p<.05$, $\mu^2=0.16$, but not for female faces $F(4,56)=1.71$, $p=.16$. In sum, the effect of vertical component in all experiments and conditions was higher compared to the horizontal component effect or compared to the interaction effects: Trustworthiness experiment ($\mu^2$Vert=.32 > $\mu^2$Inter=.17 > $\mu^2$Hor=.14);

Aggressiveness experiment ($\mu^2$Vert=.74 > $\mu^2$Hor=.21; Inter=n.s); Femininity experiment ($\mu^2$Vert=.62 > $\mu^2$Hor=.10; Inter=n.s).

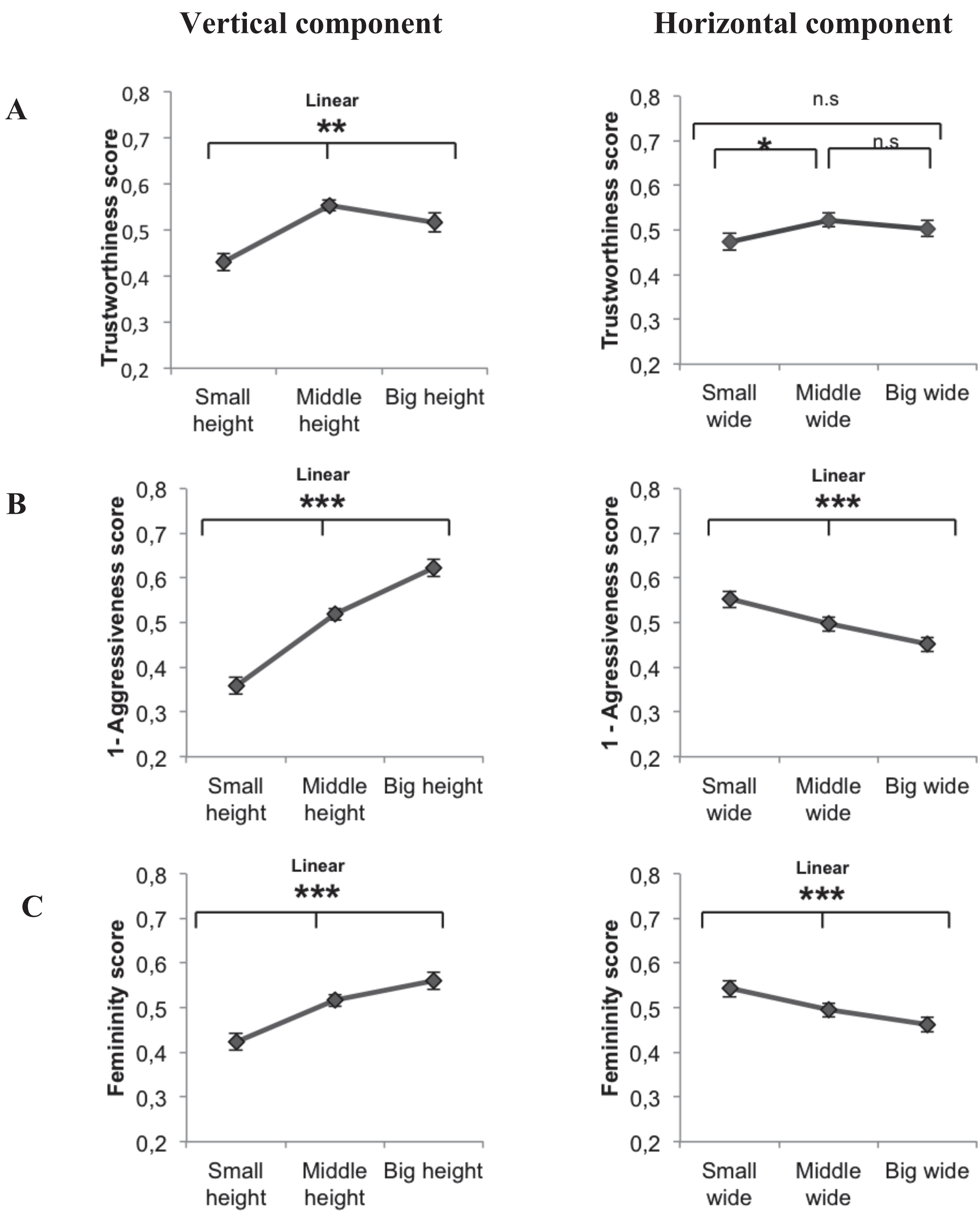



*Figure 17.* **(A-B-C)** Average score for trustworthiness, femininity and aggressiveness judgments in all visual conditions. Left panels: result for the vertical component (Small-Middle-Big Height faces); right panels: results for the horizontal component (Small-Middle-Big Wide faces). Because aggressiveness judgments are negatively correlated to trustworthiness and femininity judgments, values reported are 1-Aggressiveness values to clarify the effect across experiments. This transformation does not make statistical changes.

### 3.3.4. Inter-subject reliability

To provide further evidence of the relative bias that each component added to social judgments, we assessed the inter-subject-rating reliability. For each participant, ratings of the vertical component (SH, MH, BH) have been correlated with the ratings of the other participant (*N*=15) that performed the same experiment. For each experiment (trustworthiness, aggressiveness or femininity) and for each type of data based used (male or female). Results are presented as a correlation matrix for each experiment (trustworthiness, aggressiveness or femininity) and each database used (male or female) (Figure 19, left graph). The same analysis was conducted with participants' ratings for the horizontal component (SW, MW, BW), (Figure 19, right graph).

We then computed the average of participants' agreement for the vertical and the horizontal component for each experiment (trustworthiness, aggressiveness, femininity). Because data didn't follow a normal distribution we performed a non-parametric test (sign test) to compare agreement among participants while judging faces modified for the vertical component (SH-MH-BH) against judgments of faces modified for the horizontal component (SW-MW-BW). The sign test showed a strong effect signaling that the highest agreement was reached for the vertical component $p<7.78 \times 10^{-14}$ (Figure 18).

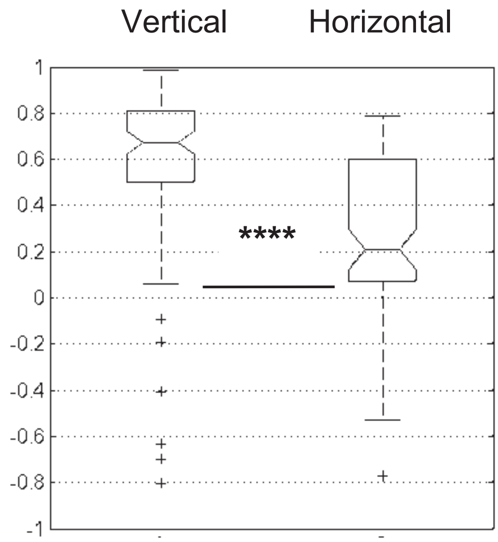



*Figure 18.* **Vertical effect:** average of participants's agreement for the vertical component (modulation of upper facial height) and the horizonthal component (modulation of the bizygomatic width). Sign test showed a strong significant difference $p< 7.78 \times 10^{-14}$

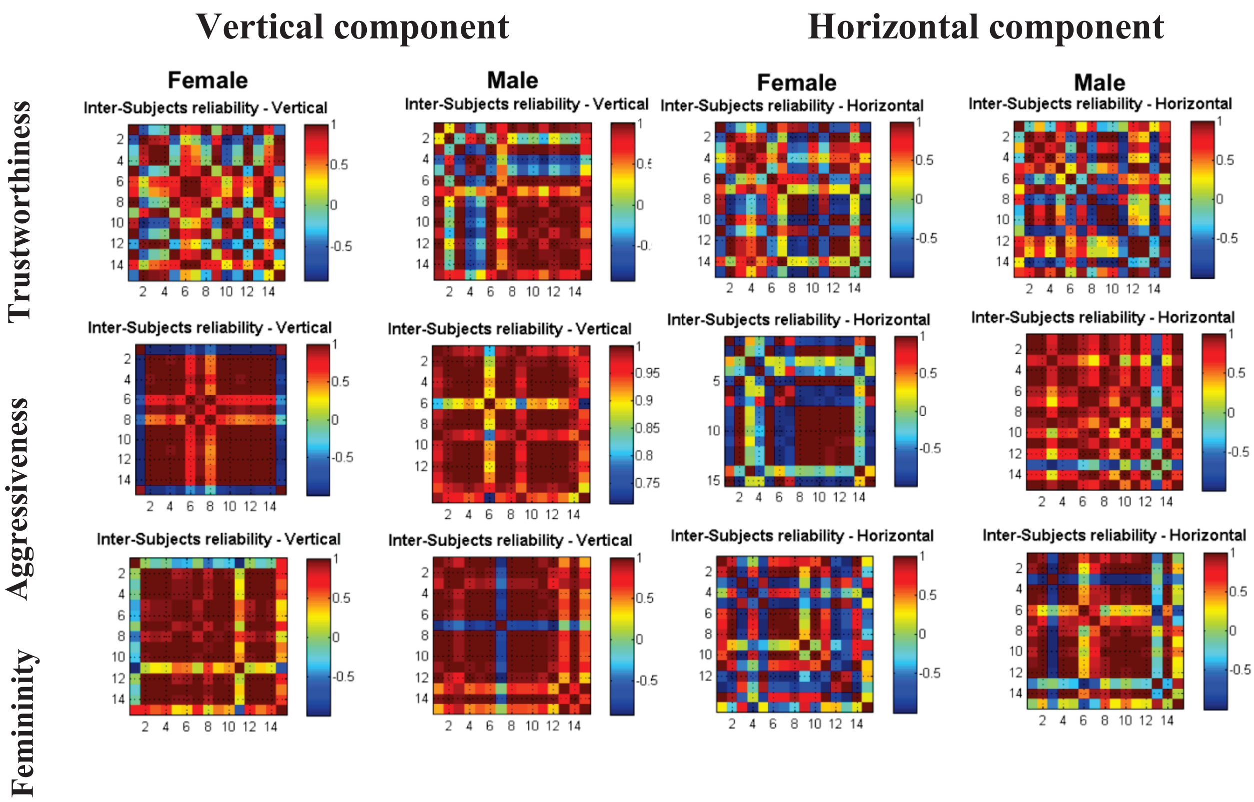



*Figure 19.* **Inter-subjects reliability.** From left to right, matrixes of correlation for female and male database used for the vertical and the horizontal component. From top to bottom the three social judgments evaluated during the experiments: trustworthiness, aggressiveness and femininity. Each graph presents a matrix of correlation across the participants (*N*=15 for each experiment). For each participant, ratings of the vertical component (SH, MH, BH) have been correlated to ratings of the three categories of every other participant. The same analysis have been performed using scores of the horizontal component (SW,MW,BW). Colorbar represent the degree of correlation using R from blue (negative correlation, -1) to red (positive correlation, +1). Participants reached higher agreement while judging faces modulated by the vertical component.

## 3.4. Discussion

Our results show that vertical and horizontal components of fWHR play different roles in the formation of social impression. The methodology employed in this study allowed us to disentangle the facial impression induced by the vertical component from that produced by the horizontal one. We also measured the whole effect of fWHR to assess if the contribution of one component (i.e., vertical) was dependent or independent from the other component (i.e., horizontal).

For judgments of trust, the vertical component strongly affected the attribution of trustworthiness for both male and female faces. Changes in perceived trustworthiness were function of the amount of vertical manipulation: faces with smaller height were perceived as less trustworthy, less feminine and more aggressive. This effect on trust was not observed following horizontal modification. The pattern of results differed when participants judged aggressiveness and femininity. Judgment of femininity and aggressiveness were perceptually affected by both the vertical and the horizontal components, and this was found for both female/male dataset, although the effect of the vertical component for male faces yielded more significance.

Hence, these results demonstrate that the modulation of the upper facial height is a relevant cue affecting several types of social judgments. In fact, as shown by the size of the eta square, across all experiments and conditions, the vertical effect always better explained participants' judgments than either the horizontal effect or their combined effect. This result was also confirmed by the inter-subject reliability analysis, which showed a higher agreement across participants while judging vertically modulated faces.

Why would the vertical component play a more significant role in social judgments?

One explanation may be that upper facial height (and not facial breadth) is a potential target of selection during evolution, as previously argued (Weston et al., 2007). In fact, Weston and colleagues (2007) reported that the relationship between bizygomatic width and the usual skull size does not differ between males and females whereas the relationship between upper facial height and skull size significantly differs between the sexes. Therefore, facial height can unambiguously distinguish an adult male from a female faces, whereas the facial width may fluctuate with variation in body size. Therefore, this component may be crucial for judging the face femininity. Despite signaling sex differences, this cue may also reveal other characteristics. In fact, faces with smaller upper height have been shown to display more bite force which may play a crucial role in survival (Proffit et al., 1983; Raadsheer et al., 1999). As a consequence, it is possible that faces with such characteristics may be perceived and judged as more aggressive as well. Here, coherently with this literature, we found that participants strongly relied on the vertical dimension and that faces with small upper facial height have been judged as more aggressive and less feminine but also less trustworthy compared to all other stimuli.

Another possible explanation for the advantage of the vertical component over the horizontal one may be that the upper facial height is less variable than facial width in humans (Franciscus, Long, 1991; Bastir, Rosas, 2004). In fact, facial width but not upper facial height may greatly vary with change in skin quality related to oldness, body size or fat variation. Indeed, it has been already argued that the presence of fat facial tissue in cheekbones makes fWHR difficult to measure (Kramer et al., 2012). Hence, upper facial height would be a less variable feature and thus easier to perceive from a face than facial width.

In agreement with previous studies, we also found that faces with larger width were judged as more aggressive and less feminine, regardless the sex identity. These results are coherent with previous literature showing that during puberty under the influence of testosterone, males would get larger facial width (Penton-Voak & Chen, 2004 Enlow & Hans, 1996, cited in Weston, Friday, Johnstone, Schrenk, 2004); and that, in return, the faces with larger width would be perceived as more aggressive (Carmen E. Lefevre, Lewis, Perrett, & Penke, 2013). Hence, testosterone can be considered as a potential modulator of both physical (width of the face) and behavioural aspects. This step forward was important to clarify the role of biological constraints exert on facial metrics which are relevant for femininity and aggressiveness judgments. Following this reasoning, while taking advantage of our results on trust, future studies may assess the influence of neuromodulators relevant for trustworthiness, such as level of oxytocin (Lambert, Declerck, & Boone, 2014) and/or serotonin (Simonsen et al., 2014) on upper facial length.

In all three experiments we did not observe a significant interaction between the vertical and horizontal components when judging female faces on trust, aggressiveness and femininity. This lack of interaction was also observed when judging male faces on aggressiveness. In other words, in most conditions, the effect of the vertical dimension was completely independent from the effect of the horizontal component. Again, this observation is coherent with the hypothesis that selection pressure exerted on facial height is independent from facial width (Weston et al., 2007).

As a consequence for future studies, these results strongly favor a methodology where the measure of the vertical component *per se* is favored over the complex and less fine-grained measure of fWHR. In fact, there are different ways to modulate fWHR, but as it has been shown in this work, it is important to determine which of these two variants enable faces' first impressions to occur. Based on a previous literature, faces with higher fWHR *tout court* are judged as less trustworthy, more aggressive and less feminine (Carré & McCormick, 2008; Stirrat & Perrett, 2012). However, our findings rather demonstrate that it is possible to determine which component of fWHR is more relevant in the formation of these social impressions. Future studies may continue to use our methodology to investigate other differences in social traits that are not been considered in the present work. Nonetheless, our study draws attention on the need to control for these two components when discussing the impact of the fWHR as an integrated measure.

Finally, contrary to previous results that threaten the validity of fWHR as they did not report significant effects using female faces (Carré & McCormick, 2008; Geniole et al., 2012;

Michael P. Haselhuhn et al., 2013; Stirrat & Perrett, 2012), we observed that the vertical effect for trust and the horizontal effect for aggressiveness and femininity were sex-independent.

The objective of this study was not to undermine the interest of the fWHR *per se*. Rather we wanted to show that each component of this measure could become a more powerful tool when used to catch differences between social perceptions and differences among male and female faces. Altogether our findings suggest the use of clearer methodology of the fWHR where the contribution of the vertical and horizontal components should be tested independently. This may limit the ambiguity in measuring fWHR.

# Appendix 3

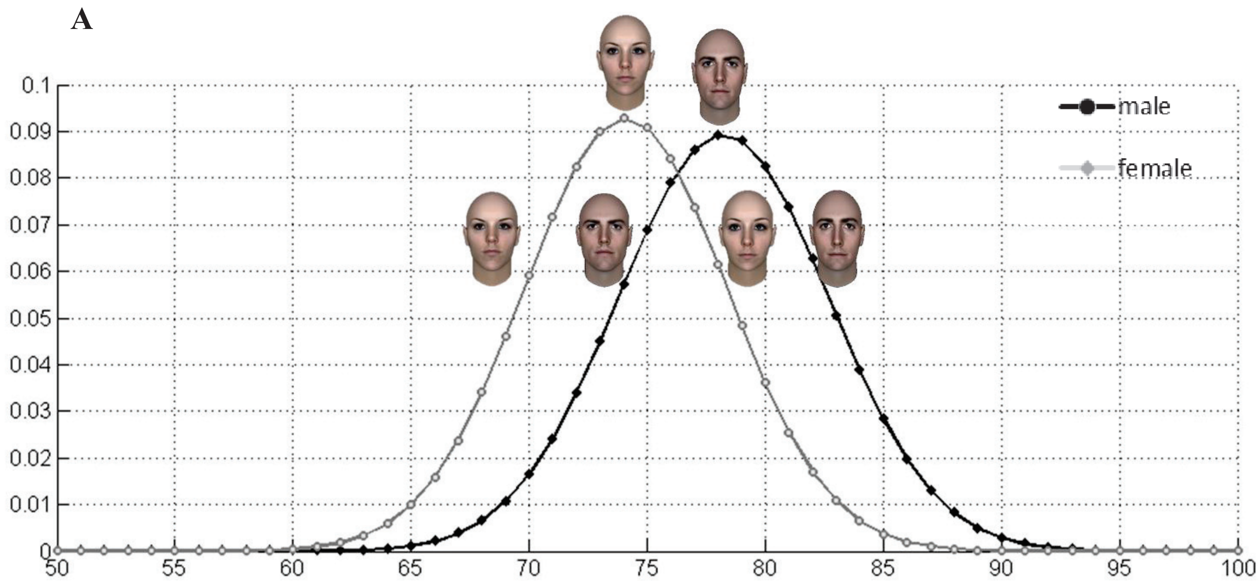
## Face stimuli and facial measures in normal population

To insure that stimuli from our dataset were realistic samples from human real faces, our measures of bizygomatic width and upper facial height were compared to the metrics obtained in a normal Caucasian male population provided in the database "FaceBase" (Weinberg et al., 2015). In male, the values selected in Facebase data were the following: (a) upper facial height, $N$=655; M=78.23, SD=4.47; (b) facial width, $N$=614; M=137.76 mm, SD=6.28. The values selected in FaceBase for female: (a) upper facial height, N=1214; M=74.09, SD=4.30; (b) facial width, N=1080; M=129.89, SD=5.37.

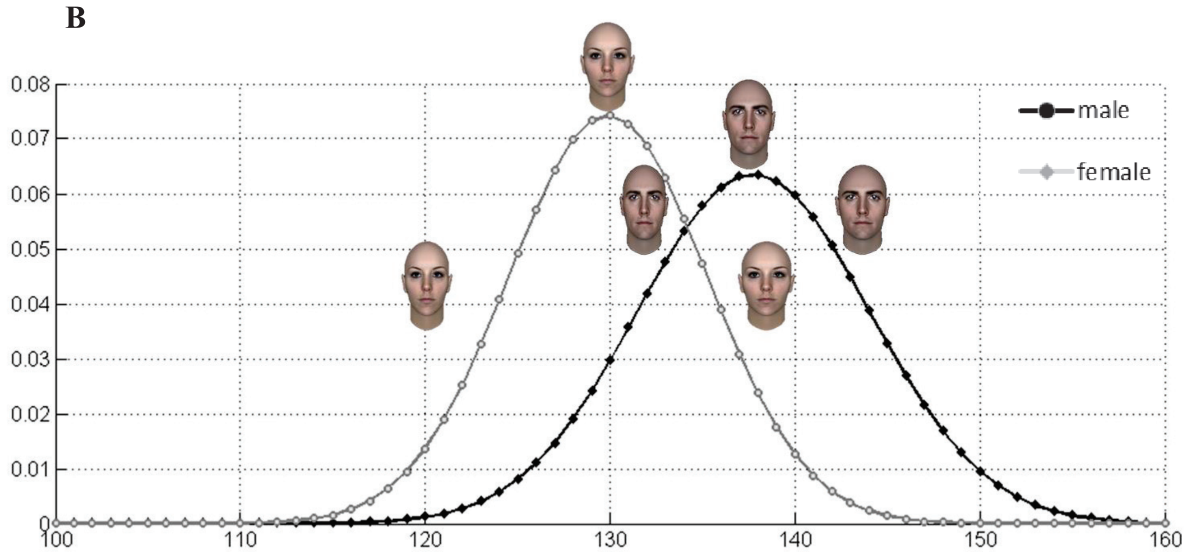
We then calculated Z-scores for stimuli in each face category of our database and transformed the value in percentile. We obtained the following results: (a) vertical component, SH=72.7(11%), MH=78.23 (50%), BH=81.42 (76%), (b) horizontal component, SW=128.06 (6%), MW=137.76 (50%), zBW=143.47 (81%).

The same analysis has been performed for female faces: (a) vertical component, SH=68.91 (11%), MH=74.15(50%), BH=77.15 (76%); (b) horizontal component, SW=120.5 (4%), MW=129.88 (50%), BW=135.09 (83%). This analysis confirmed that all stimuli used were within the range of the normal population.

## Upper facial height: normal distribution in real population




## Bizygomatic width: normal distribution in real population



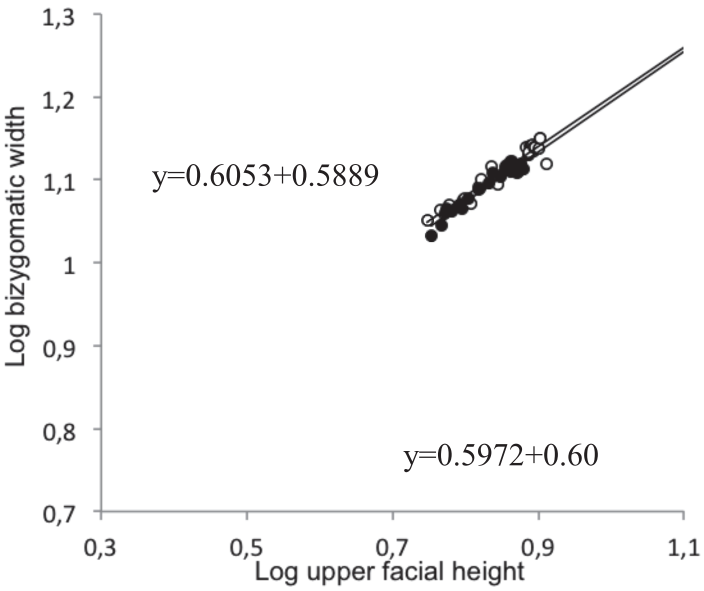

*Appendix Figure 6.* **(A)** Normal distribution of upper facial height in male population (dark grey) and female population (light grey). From left, sample of Small Height faces (SH), Middle Height faces (MH) and Big Height faces (BH); **(B)** normal distribution of bizygomatic width in male population (dark grey) and female population (light grey). From left, sample of Small Wide faces (SW), Middle Wide faces (MW) and Big Wide faces (BW).

## Appendix 4

### Test of sexual dimorphism using FaceBase data

Standing from the analysis performed in the previous section we further questioned whether the sexual dimorphism pointed out by Weston (2007) exists and could be replicated using FaceBase database. As argued in the main text Weston and colleagues (2007) reported that the relationship between bizygomatic width and the usual skull size does not differ between males and females, whereas the relationship between upper facial height and skull size significantly differs between the sexes.

However, different studies do not found the sexual dimorphism stated by Weston (Kramer et al., 2012; Carmen E. Lefevre et al., 2013; Özener, 2012). Coherently with these findings, here we did not find a sign of sexual dimorphism (Appendix Figure 7).




*Appendix Figure 7.* **Male and Female trajectories byzigomatic width (BZW) and of upper facial height (FHT).** Values of upper facial height and bizygomatic width provided by FaceBase database. Weston found a significantly difference in the intercept between male and female. The graph show the absence of this difference using this values.

We observed in fact that values of FaceBase for male and female after puberty differed both for their width (Mmale=137.76, SD=6.28; Mfemale=129.89, SD=5.37; $T_{18}=16.21$, $p<.001$), and height (Mmale=78.23, SD=4.47; Mfemale=74.09, SD=4.30; $T_{18}=14.03$, $p<.001$). As a consequence we did not observe a significant difference ($p=0.31$) in the fWHR across sex: Mmale=1.76, SD=0.019; Mfemale=1.75, SD= 0.025. On the contrary difference between variance was significantly different when consider byzigomatic width only ($p<0.5$), upper facial height variance ($p>.05$). A different variance between female and male for the byzygomatic width may be one of the factors that could explain the sexual dimorphism measured by Weston. Another methodological difference was that Weston obtained the measures from skulls.

In line with Ozener (2012) we agree that if fWHR is an important characteristic that emerged as a result of sexual selection in the evolution, evidences for sexual dimorphism should be perceptible from the face (Özener, 2012). In this work, using a new method we found that independently from the existence of sexual dimorphism, modulation of fWHR due to changes in the vertical or in the horizontal component can bias attribution of social judgments from faces both in female and male. We thus strongly suggest further investigating about the existence of sexual dimorphism and the presence of effect in female population as two independent but related topics.

# Chapter 4

## Processing of facial information in patients with Williams syndrome: a behavioral and neuronal investigation


### Abstract

Trust is strongly involved in human social interactions. When faced with unfamiliar individuals, humans make judgments of trust based on a fast and automatic processing of facial features such as shape of the eyes, eyebrows and mouth. The way individuals explore faces, especially over the eyes, constitute an extraordinary source of social information. The ability to detect and process such information from faces is altered in pathologies characterized by atypical social behavior. Williams-Beuren syndrome (WS) is a rare genetic disorder with mental retardation but preserved linguistic competence and an exaggerated social appetence. In the current work we choose to examine patients suffering from this pathology as a model to investigate perception of social information from faces. Specifically we were interested in examining how facial features important for forming first impression of trustworthiness were represented in WS.

WS patients (N=12) and healthy participants (N=12) with typical development (TD), matched for age, participated in three different experiments.

The first experiment investigates trustworthiness' detection from faces using eye-tracking recording. The novelty of this study is to compare eye-tracking data during an implicit and explicit task while participants were seeing computer generated trustworthy and untrustworthy faces. During the implicit task WS patients spontaneously looked at both trustworthy and untrustworthy faces confirming their behavioral aptitude to approach everyone regardless of perceived facial traits. Differently, during the explicit task the effect of preference towards trustworthy faces was present in both WS and TD. The data suggest dissociation between the implicit and explicit task.

In the second experiment, using a different method, we aimed at investigating how a trustworthy face is built from the perspective of WS patients, a procedure thereby allowing exploring patients' internal representation of trustworthiness. More precisely, I asked whether WS patients might be able to create a trustworthy face from noisy images of faces and weather this representation has similar properties to the ones created by a TD group. For that aim we used a paradigm of reverse correlation (RC). This new technique allows generating a

mean image that reflects participants' internal representation of faces, computed from their choices of noise images. Differently from experiment 1, where trustworthy-modulated faces were presented, here no assumption exists about the features that are considered relevant for this social judgment. Representations of 9 WS have been then compared to the representations of 24 TD subjects obtained from noise. As expected, all controls by modulating specific facial features created trustworthy faces that were similar across subjects. Cluster analysis showed that eyes color, eyebrows and a smiling mouth are relevant facial cues for trustworthy faces in controls. In contrast, representation of trustworthy faces in WS patients seems to differ from the control group. These results suggest that comparing to healthy subjects the representation of trustworthiness is differently processed in WS.

Results from experiment 1 and 2 confirmed the presence of different social behavior in WS patients. Standing from these results, in the third experiment I asked whether brain network could explain patients' representational variability. Looking at the electrophysiological brain sources, with particular attention to the source localized in the superior temporal sulcus (STS), I asked whether this activity was differently modulated compared to a control group. I used a spatial filter over STS using blind source separation. In a previous experiment performed in our team we found that when looking at different parts of a face (cheek, nose, eyes, mouth etc.) healthy subjects displayed a source activity in STS which peaked at 240ms after stimulus onset. Interestingly, this STS activity was not found in patients with autism spectrum disorders (ASD) when performing the same task. In WS we found the same source in the STS at 240 ms thus showing that neural activity in these patients is similar to that found in TD. The results suggest that this source, that allows to discriminate ASD from TD, at this specific time (240 ms) seems not be the cause of the atypical social behavior in WS.

## 4.1. Introduction

Trust is taking the risk of putting one's own fate in someone else's hands, hence the importance of trustworthiness assessment to minimize this risk. A possible neurobiological model with marked social deficits is Williams-Beuren syndrome (WS). In brief, these patients have cognitive disturbances such as visuo-spatial deficit (Deruelle, Rondan, Mancini, & Livet, 2006) and inhibition problems(Little et al., 2013). Importantly, they display abnormal social functioning with overexpression of approachability towards others (Barak & Feng, 2016; Bellugi, Adolphs, Cassady, & Chiles, 1999; Doyle, Bellugi, Korenberg, & Graham, 2004;

Wendy Jones et al., 2000; Mimura et al., 2010). This overexpression of kindness and positive social interactions can appear at first sight a positive competence. However, WSP families greatly complain about this atypical behavior, which often leads WS patients to fail to build strong interpersonal relationships during adulthood and lead to loneliness and anxiety and risk of suicide (Binelli et al., 2014)

Thus, understanding the neurocognitive mechanisms responsible for this exaggerated social behavior is an important practical health problem. Nonetheless, studying these patients will contribute to the understanding of brain-behavior correlations in social functioning while processing faces and while making trustworthiness judgments based on facial appearance.

Williams-Beuren syndrome (WS) is a rare neurodevelopmental disorder (1:20.000) caused by a deletion of approximately 26 genes on the long arm of the chromosome 7 (7q11.23). Children affected by this pathology present special facial appearance, complex clinical dysfunctions and distinctive cognitive profile (Bellugi, Wang, & Jernigan, 1994). The clinical spectrum is characterized by cardiovascular and gastrointestinal problems but also psychomotor problems that impact coordination and walk. The cognitive profile, defined with peaks and valleys, reveals impaired cognitive functions such as visuo-spatial construction and attentional deficit. Differently, language skills are preserved and sometimes above average (Losh, Bellugi, Reilly, & Anderson, 2000; Reilly, Klima, & Bellugi, 1990).

In the field of face processing, the debate is still open. Some findings show that face processing is preserved in this pathology (Bellugi et al., 1994; Riby, Doherty-Sneddon, & Bruce, 2009) and others reveal abnormal skills such as deficit in holistic or configural face processing (Deruelle et al., 2006; Karmiloff-Smith et al., 2004).

Generally, WSP display a preference in processing local elements and features compared to global stimuli, being more focused on details. This special behavior is confirmed using an inversion face tasks (i.e. faces presented up-side down) (Karmiloff-Smith et al 2004) a task where healthy subjects show difficulties but not WS patients. Facial emotions recognition is disrupted in this pathology as observed in recognition of emotional expression tasks (Gagliardi et al., 2003; Skwerer et al., 2011). Last, WS show difficulties in detecting fear and they show less arousal when looking at angry faces (Barak & Feng, 2016).

This impaired ability to detect facial expressions may be linked to patient's altered social behavior. WS are characterized by overfriendliness and appetence to create social links and interactions, with most them displaying an exaggerated gregarious personality. Several explanations may be provided for this special behavior. One may refer to the absence of behavioral inhibition which may be the origin of this abnormal pattern of social behavior(Little et al., 2013). Another view suggests the presence of atypical neural activity in the amygdala that may disrupt patients´ social skills, particularly while looking at faces.

Differential amygdala activity, in fact, has been found while patients were looking at emotional scenes and faces. An fMRI study showed that amygdala activity in WS only increased while looking at threatening scenes but not when processing threatening faces (Meyer-Linderberg 2005). These findings have been confirmed in another fMRI study (Haas 2009) where atypical amygdala activation was triggered during processing of fearful faces (amygdala activity WS<controls perceiving fearful scenes and WS>controls while perceiving happy faces). Moreover, fMRI data supports this explanation showing absent connectivity between frontal cortex (OFC), and the amygdala (Meyer-Lindenberg et al., 2005, 2006).

Importantly, WSP while performing explicit judgments from faces they tend to rate unfamiliar and negative faces as more approachable compared to controls (Frigerio et al., 2006). Finally, WS display a positive bias for positive expression, with ratings higher for positive faces compared to controls.

Previous findings have shown a link between perception of emotion from faces and detection of social traits in healthy subjects. The hypothesis of overgeneralization of emotion is in fact one mechanism that has been argued to be the base of perception of social traits. According to the emotion overgeneralization hypothesis, resemblance of neutral faces to emotional expressions is perceived as indicating the trait attributes associated with these emotions (Knutson, 1996; Montepare & Dobish, 2003; Todorov, 2008b). The link with emotional perception suggested by the overgeneralization hypothesis can be tested in this pathology. In fact if WS have a disrupt ability to perceive facial expression from faces they might be impaired in detect social traits. Nevertheless, the two abilities, detection of emotion and detection of social traits, may be dissociate with WS preserved in the ability to perform social judgments.

Standing from these findings that confirm the outstanding social ability of WSP, in the first experiment I asked how WS would explore and possibly interact with faces displaying different level of trustworthiness. I use the measure of looking time towards trustworthy versus untrustworthy faces as measure of approachability.

## 4.2. Experiment 1

## Implicit and explicit detection of trustworthy faces in WSP

WS and typically developing (TD) participants matched for age and sex, attended an eye movement experiment. Pairs of trustworthy and untrustworthy computer generated faces selected from Todorov's database were presented side by side. In a first session, participants looked at faces (implicit task) without any explicit instruction, they were told to look at faces as they prefer. The objective of this task was to determine the spontaneous looking preference towards faces a priori judged as trustworthy and untrustworthy without elicit the concept. In a second session, participants were asked to explicitly select the trustworthy face.

### 4.2.1. Methods

**Participants**

12 participants (6 female, 6 male) with WS were individually matched for sex and chronological age (mean=11.9, SD=4.01) with 12 healthy controls.
Control participants were recruited through advertisements in local newspapers, or local schools. The inclusion criteria for healthy participants were age, no current or past history of organic disease, neurological or psychiatric disorders, learning disorders, mental retardation or prematurity (<39 weeks of gestation), and no medication.

Patients were recruited through the GENOPSY center of Dr Demily and through advertisements in Williams' syndrome associations. Moreover, in WS group, children were included only if they were already diagnosed with WS by a medical doctor with a genetic assessment including a report of deletion of the gene at 7q11.23 (American Psychiatric Association, 2000).

**Neuropsychological evaluation**

A short neuropsychological evaluation of the intellectual abilities was performed to insure the typical phenotype of WS. We used two sub-tests of the WISC-IV which are known to be the most correlated to the general IQ score: the similarities sub-test which is most correlated to the verbal intellectual quotient and the matrix sub-tests which is most correlated to the IQ performance. Neuropsychological results showed that in the WS population the typical intellectual retardation in verbal reasoning (VR, Similarities, $F(1,19)=129.74$, *p*<.0.001: WS mean =3.27; SD=2.41; TD mean=15.44; SD=2.5) and visuo-spatial reasoning (matrix: $F(1,19)=46.7$, $p<.0.001$ (WS mean=3.4 ; SD=2.16; TD mean=12.00; SD=3.30).

Because we hypothesized that the impairments in the trust detection test could be explained by attentional, visuospatial or emotion sensitivity abilities we also included sub-tests from the NEPSY to assess these abilities.

In the auditory attention sub-test of the NEPSY, participants are facing a paper sheet with 4 colored circles (red, yellow, black and blue). They are asked to put their hand on the table and touch the red circle whenever they are told "red" and to put back their hand on the table afterwards to be ready for the next words. The number of commission, omission and correct detection are scored and used to assess the standard score.

In the Arrows sub-test of the NEPSY, participants are facing a target with several arrows directed toward it. They have to indicate which arrows will directly hit the center of the target.

In the emotion recognition task of the NEPSY, participants are shown a child expressing an emotion, and then they are asked to select among other children the one expressing the first child emotional status.

**Stimuli and setup**

The stimuli used in the experiment were 48 computer-generated male faces created with the FaceGen software development kit (Singular Inversions, Toronto, Canada) selected from the Todorov's well-controlled quantitatively validated stimulus repertoire of faces. All faces were bald and Caucasian. The Trustworthiness database is composed of facial identities varying on 7 levels of trustworthiness (Todorov et al., 2013). Todorov et al' work showed that human explicit judgments of trustworthiness match with the model's prediction (Todorov et al., 2013) For the current study, we selected from the 24 identities the two most extreme versions (-3 SD and +3 SD), resulting in 24 couples

Stimuli were displayed on a 17-inch computer screen at a resolution of 1280 x 1024 pixels using Presentation® software (Version 14.9, www.neurobs.com). The viewing distance from the participants' eyes and the screen on which stimuli (subtending 7.8° x 12.5° of visual angle) (377 x 604 pixels) were displayed was 73cm. Subjects' eye positions were recorded using an infrared video-based tracker (Tobii 1750) at a 60-Hz sampling rate and Clearview 2.7.0 allowed online recording of eye-gaze data. The two systems were synchronized using the Tobii extension for Presentation.

**Task procedure**

The implicit task was always performed before the explicit task. During the implicit task, participants were instructed to look at pairs of faces during 5 s. This session was composed of 48 trials, as each face of the 24 couples was presented twice to counterbalance side position. During the explicit task, participants were asked to explicitly select the face that matched their answer to the question by pointing at the face.

All questions were chosen to assess something related to the concept of trustworthiness. The questions allowed us to insure that patients and children did not fail to choose the correct face due to misunderstanding of the concept of trustworthiness. Therefore, the following three questions were pseudo-randomly presented at the onset of the face stimuli presentation: Q1. To whom would you say a secret? Q2. To whom would you ask help? Q3. Who would you trust? (Figure 20). This session was also composed of 48 trials to counterbalance for the side of presentation. The response given by the participant was recorded by the experimenter. Prior to the experiment, participants underwent a 5 point-calibration task. During both sessions participant's eye gaze was monitored on-line with a hidden second screen to insure that the posture was correct for data acquisition.

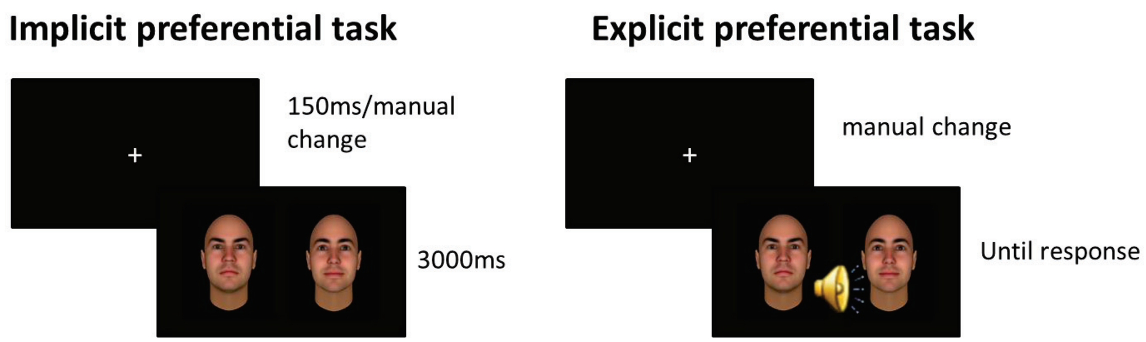



*Figure 20.* **Task paradigm. A.** Implicit detection of trustworthy, extreme variation of faces selected from the Todorov database of trustworthy faces. In the first trial on the left the trustworthy face (+3SD), on the right the untrustworthy one (-3SD). **B.** Explicit task and the three type of questions that participants listened.

## Statistical Analysis

*Pre-processing and data analysis*

### Eye-movement

ClearView fixation filter was used to filter the data (with a visual angle of 1° and duration of 100ms). An in-house Matlab script was used to process eye-tracking data. In order to quantify allocation of attention to faces, regions of interest (ROI) delimiting each face were defined manually. The mean looking time was calculated as the average of the total time spent within each ROI during a trial. For the main statistical analysis, mean looking times on each face were calculated for each participant and for each trial. The results from the different trials in each condition were averaged for each participant.

### Spatial distribution of fixations

To provide information on the spatial distribution of the fixations, the barycenter of fixations and a heat map representation were calculated for each face at the subject level. Heat maps were calculated using Gaussian kernel density mapping of the fixations, weighted by the fixations' duration (Caldara & Miellet, 2011). Then, at the group-level, individual heat-maps were normalized and averaged to visualize the spatial distribution of the fixations of the studied population.

### Statistical analysis

As first analysis we wanted to insure whether WSP and TD had a difference preference towards faces. An ANOVA using difference of looking time during the implicit task was performed and the group (TD and WS) was used as categorical factor.

ANOVA for implicit on the looking time difference and a t-test to 0 to define whether there was a preference in both group. The same analysis has been performed for the explicit task.

### Behavioral measures during the explicit task

As first measure I calculate the percentage of trustworthy face choice in both groups. This measure reveals the participant's agreement with Todorov's database. Another measures computed during the explicit task was the self-coherence. The same paired of face was presented twice to counterbalance for the position of the target face. The self-coherence looked whether each participant select as trustworthy the same face in both case. The ANOVA was performed using the score of agreement and the group (WS and TD) as categorical factor. The same analyses have been performed using score of self-coherence.

### 4.2.2. Results

*Neuropsychological results*

As expected, WS patients were impaired in verbal reasoning (VR, Similitude) and visuo-spatial reasoning (matrices). A group effect was significant for both test $F_{VR}(1,19)=129.74$, $p<.0.001$; ($WS_{VR}=3.27$; SD=2.41; $TD_{VR}$ mean=15.44; SD=2.5) and visuo-spatial reasoning (VSR, Matrice) $F_{VSR}(1,19)=46.7$, $p<.0.001$ ($WS_{VSR}=3.4$ ; $SD_{VSR}=2.16$; TD mean=12.00; SD=3.30).

Visuospatial skills (Flesh test-NEPSY) showed a similar deficit in patients as compared to TD $F_{VS}(1,16)=66.66$, $p<.0.001$; ($WS_{VS}=2$; SD=2.29; $TD_{VS}=12.62$; SD=2.87). Attention and executive function (attention auditive) $F_{EF}(1,16)=1.82$, $p=.19$, ($WS_{EF}=6.66$; SD=5.89; $TD_{EF}=10.75$; SD=4.47). Synthesis of the results is reported in the Table 3.

**Table 3**. *Sample demographical and cognitive details*

| | WSP | | TD | | |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SE | p value |
| **Age** | 11.9 | 4.01 | 12.2 | 3.9 | n.s |
| **Verbal reasoning** | 3.27 | 2.41 | 15.44 | 2.5 | 0.001 |
| **Visuo-spatial reasoning** | 3.4 | 2.16 | 12 | 3.30 | 0.001 |
| **Visuo-spatial skill** | 2 | 2.29 | 12.62 | 2.87 | 0.001 |
| **Emotion recognition** | 6.54 | 2.46 | 10.75 | 4.55 | 0.01 |
| **Auditory attention** | 6.5 | 1.98 | 10.75 | 1.31 | n.s |

An altered recognition of emotional information was detected in WS patients as compared to controls $F_{ER}(1,18)=9.01$, $p<.0.01$, showing a reduced ability to discriminate across different emotion ($WS_{ER}=6.54$; SD=2.46; $TD_{ER}=10.75$; SD=4.55).

Emotional recognition task was also used for fulfilling the aim of the study. Todorov's and colleagues mentioned the overgeneralization of emotion as the mechanism underlying perception of trustworthiness. Although trustworthy and untrustworthy faces are neutral they are in fact perceived as resembling respectively happiness and anger (Engell et al., 2010).

Hence, an ANOVA was performed using the scores for happiness and anger and the group as categorical factor. A main effect of group was reported $F(1,18)=7.38$, $p<0.05$, a main effect of emotion was also observed $F(1,18)=8.43$, $p<.0.01$. The interaction between emotion and group was not significant p=0.25. Score obtained from the emotion recognition test are reported in Figure 21.

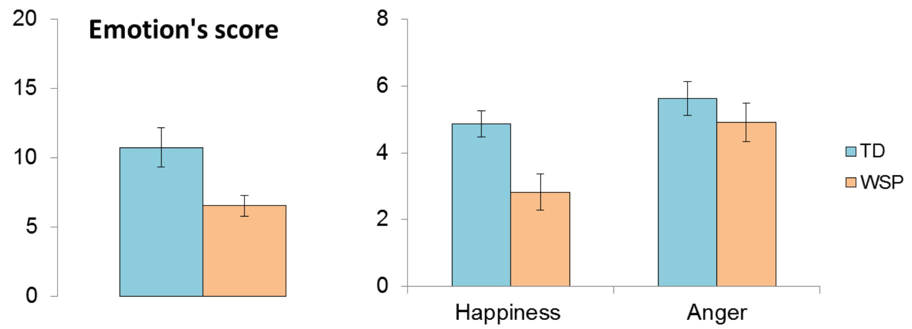



*Figure 21*. **A.** Difference of emotion's score in WS and TD. **B.** Score of happiness and anger detection for TD and WSP.

*Eye tracking measure. Looking time in the AOI during the implicit and the explicit task*

In order to quantify gaze allocation, regions of interest (ROIs) encompassing the trustworthy and untrustworthy faces were defined. Ocular fixations within and outside these ROI were recorded during each trial (see Methods). The mean looking time was calculated as the average of the total time spent within trustworthy and untrustworthy faces for all stimulus pairs presented. As first analysis we wanted to insure whether WSP and TD had a difference preference towards faces.

The ANOVA performed to test the preference during the implicit task reveal the following group effect $F(1,20)= 3.31$, $p=0.08$. During the explicit task we observed the following values $F(1,20)= 3.31$, $p=0.083$. We then performed a t-test to 0 to define whether there was a preference in WSP and TD. Interestingly WSP did not show any preference during the implicit task $T_{11}=-0.43$ $p=0.66$; differently TD spontaneously looked more the trustworthy faces $T_{11}=2.29$, $p<0.05$. During the explicit task, WSP showed a tendency to look more at the trustworthy faces, $T_{11}=1.93$ $p=0.081$. For TD the preferential looking time towards trustworthy faces was highly significant, $T_{11}=5.43$ $p<0.001$.

In sum, during the implicit task, TD group looked longer at trustworthy faces, while WSP do not show any preference. In other words, WSP showed an abnormal positive bias towards untrustworthy faces look more than TD the untrustworthy faces. During the explicit task this different approach behavior towards faces disappeared. In fact, when WSP were asked to choose the most trustworthy faces; they used the facial features to perform their choice. The eye tracking collected during the explicit task are supported and confirmed by the behavioral data presented in the next section.

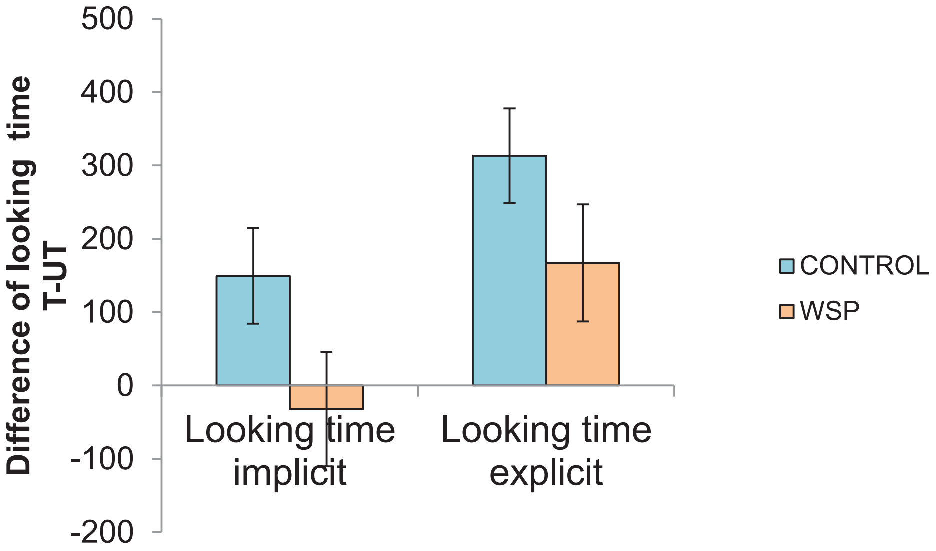



*Figure 22.* **Results of WS and TD for the Implicit and Explicit task.** In y axe I present the difference between looking time towards trustworthy and untrustworthy faces (T-UT) for TD and WS during the implicit and explicit task.

*Behavioral measures during the explicit task*

As first measure I calculate the percentage of trustworthy face choice in both groups. This measure reveals the participant's agreement with Todorov's database. WSP selected trustworthy faces (76%) as TD group (90%).  T-test between TD and WS showed that the agreement was not significantly different. Importantly, this choice do not differed across type of questions neither in TD and WS. Another important measures computed during the explicit task were the self-coherence. The same paired of face was presented twice to counterbalance for the position of the target face. The coherence looked whether each participant select as trustworthy the same face in both case. As shown in Figure 23 TD and WSP were both coherent (TD=20.27, SD=3.58; WS=16.9, SD=5.3). This result shows that when WSP were asked to detect trustworthy faces they are able to perceive and use the relevant information from a computer generated face.

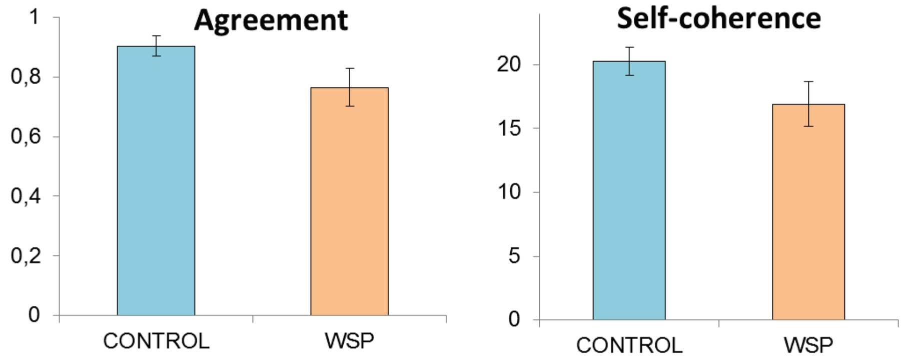



*Figure 23.* **Behavioral measures during the explicit task for TD and WS. A.** Agreement**. B.** Self-coherence**.**

*Exploration of face during the implicit task and the explicit*

Previous studies claimed a strong preference in WSP for the eye regions (M. a Porter, Shaw, & Marsh, 2010). We looked at the pattern of exploration between one face and the other across trials. In TD we found a typical holistic exploration with the barycenter of fixation over the regions of the eyes nose and mouth. Interestingly in WSP we found at least two different pattern of exploration. One group of patients mainly displayed a barycenter of fixation similar to the one found in TD (represented by the yellow characters that are within the curve of the distribution, see Fig 24). Nevertheless, another group of patients showed the barycenter of fixation over the mouth.

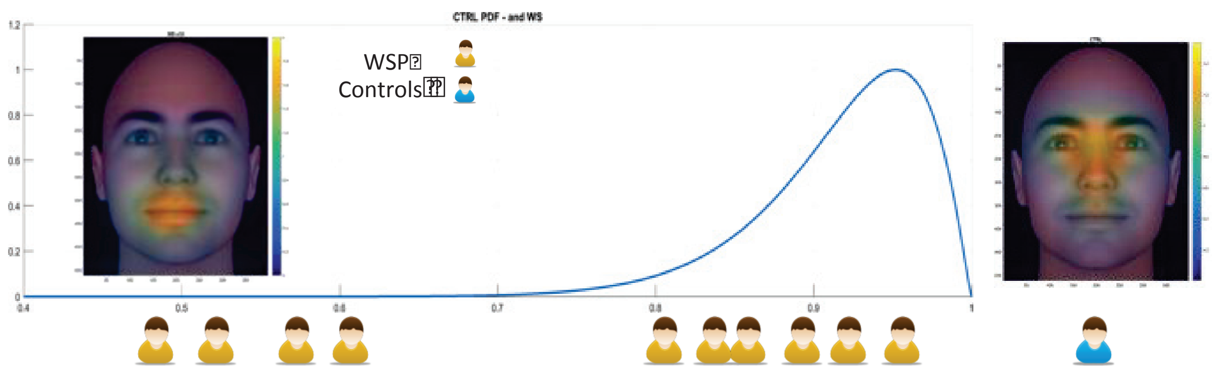



*Figure 24.* **Distribution of the probability based on normed data from TD group**. Blue line represents the distribution of fixation maps for TD group. Right image show the pattern of exploration of the TD participants that was the median value of the curve. Yellow characters referred to the WS that attended the trustworthiness experiment. Four of them had a pattern of exploration that significantly differed to the TD group. Image on the left represents the map of the WS that was most extreme in the curve.

### 4.2.3. Discussion

During the implicit task, the control group spontaneously looked longer at trustworthy faces. Differently WSP looked at both faces signaling an approachable behavior towards both trustworthy and untrustworthy faces. This result is in line with previous research showing an exaggerated and inappropriate approaching behavior in WSP which extends even to strangers (Doyle et al., 2004; Wendy Jones et al., 2000; M. A. Porter, Coltheart, & Langdon, 2007). Moreover, their inability to distinguish the trusty from the untrusty is also in line with the reported disrupted mechanisms in detecting the emotional valence of faces (Haas et al., 2009; Meyer-Lindenberg et al., 2005). Interestingly here we found similar results using an implicit preferential looking paradigm.

During the explicit preference task, when participants were forced to choose the most trustworthy face between the two stimuli, WSP choice did not differ from the one performed by TD group. In agreement with this, looking time during the explicit task was not different from controls. Moreover, neuropsychological results on emotion detection tasks showed that WSP had difficulties in detecting happiness and anger compared to the control group. From this, we can infer that the emotional deficit impacted perception of both trustworthy and untrustworthy faces. Thus the ability to detect trustworthiness may concurrently require both, detection of the emotional valence and a processing of facial features. Further testing should be performed to confirmed whether emotional detection and detection of social traits can be dissociable or not in this pathology.

Changes of preference between implicit and explicit task support two different conclusions in TD and WSP. Results observed in TD suggest that eye movements can predict explicit preference. We observed in fact that TD during the implicit task looked more at faces that they choose afterword as more trustworthy during the explicit task. In contrast, data from WS patients suggest that spontaneous/implicit detection of trustworthiness may be dissociable from explicit decisional mechanisms.

## 4.3. Experiment 2

## Representation of trustworthiness in WS

In exp. 1, WS patients spontaneously did not show preference for the trustworthy faces. Moreover, I also found different patterns of face exploration across patients, thus suggesting a higher variability in patients' visual strategies compared to controls. Indeed, one group explored trustworthy faces as control while a second displayed a different gaze pattern where saccades were mainly directed toward the mouth region.

Mental representation shapes perception. If experiment 1 explored the detection of trustworthiness using computer generated faces, in a second experiment I explored the internal representation of trustworthiness in WS patients to search for difference and similarities with respect to the detection task. We asked the following questions: do WS patients have the representation of trustworthiness as the TD group? Can they generate a trustworthy face from noise? To address these questions, I used a different paradigm know as neurophysiological reverse correlation (RC) (Kontsevich & Tyler, 2004; Mangini & Biederman, 2004). This technique allows generating images reflecting subjects' endogenous representation of faces without prior assumption about the features that contribute to the participant choices. In other words I tested how top-down mechanisms in WSP generate trust to faces.

Recently, reverse correlation has been used in social cognitive research to investigate social behavior (see, e.g., Dotsch,Wigboldus, Langner, & van Knippenberg, 2008; Dotsch, Wigboldus,& van Knippenberg, 2011; Jack, Caldara, & Schyns,2011; Karremans, Dotsch, & Corneille, 2012). Particularly, Dotsch and Todorov (2012), used this technique to investigate the representation of trustworthiness using faces in normal population. Here the interest is to use this technique as a as a quantifiable objective measures to disclosure in a social pathological condition internal representations of trust, to therefore define how patients imagine or see others faces.

**Participants**

A first group of 15 typically developing (TD) controls were recruited (8 male and 7 female) age from 18 to 42 (M=24.8, SD=7.88). 9 of the 12 WSP that performed the eye tracking task participated to this study. We also included the performance of the 9 participants of the control group (TD) that were matched in chronological age for the eye tracking experiment. These participants were different from the first group of control. Thus TD group was finally constituted by 24 individuals.

## Stimuli

Stimuli were built using a baseline facial image, average of three male identities selected from the Nimstim database (Tottenham et al. 2009) (Figure 25) and a randomly generated noise pattern (Figure 25) superimposed on the face. Two types of noise were created: high frequency (HFn) and low frequency (LFn). High frequency noise (HFn) was generated by randomizing the image spatial high frequencies while low frequency noise (LFn) was generated by randomizing the spatial low frequencies of images. These two types of noise have been then embedded into the original face image. Because for each trial a diverse random noise pattern was created, on each trial faces looked differently as the noise differently distort the original face.

For each HF and LF trials, two images displaying the original noise and the negative version of the same random noise were presented side by side. The negative pattern was mathematically constructed with each dark pixel converted into white one (Figure 25).

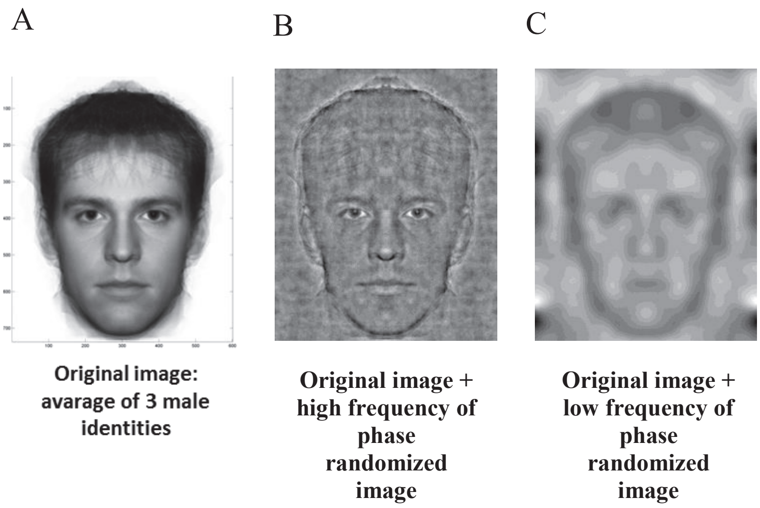



*Figure 25*. **A. Face base average of three male identities**. **B. High frequency stimulus**. The stimulus was created using the original image and the randomization of the high frequency of the image 'phase. **C. Low frequency stimulus.** The stimulus was created using the original image and the randomization of the low frequency of the image 'phase.

## Task procedure

Participants performed a typical reverse correlation (RC) image classification task. Based on their first impression they were asked to choose the face that best resembles a trustworthy face as in Dotsch et al. (2012). WSP and TD performed 800 trials; the same numbers of HF and LF stimuli were presented in a random order (Figure 26).

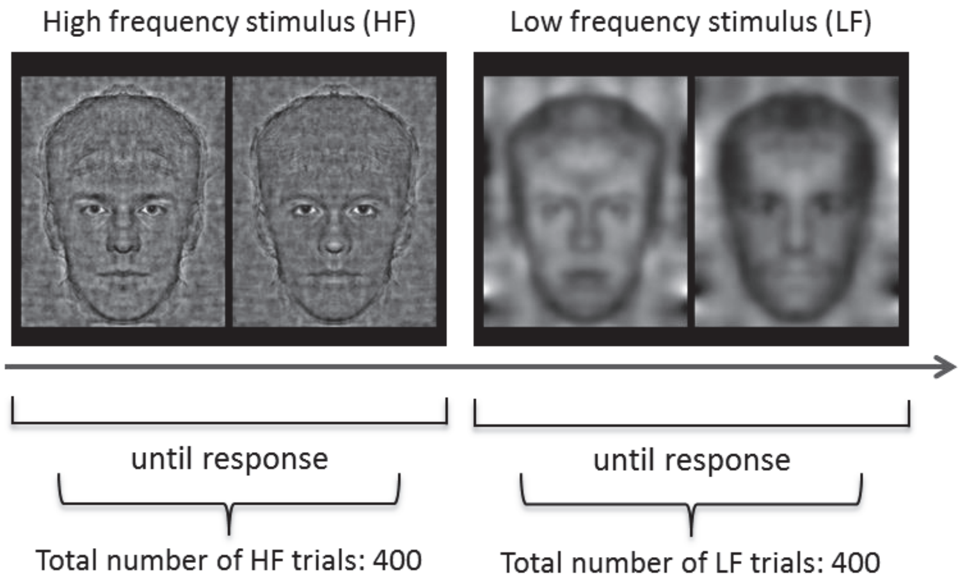



*Figure 26.* **Task paradigm.** Example of original and negative random noise for high and low frequency.

*Processing of subjects choice*

Analysis was first performed at an individual level. To generate the trustworthy face (Ci, classification image) for each individual, an average of the HF and LF noise selected as most trustworthy was done and then overlaid on the original face (baseline) Average of the mean parameters across participants was performed to visualize the prototypical trustworthy face for WSP and TD. The noise chosen by all patients was further averaged to visualize the image at group level. The same was done using the noise chosen by TD group to create average image of TD. Mean chosen noise for HF and LF for both WSP and TD is presented in Figure 27. In each trial, the noise that was not chosen by participants was used to create the anti-trustworthy face. The result represents the internal representation of a trustworthy face for each individual.

To test which clusters have been selected as significantly relevant for the trustworthy face reconstruction, a cluster permutation test has been done. Clusters presented in Figure 28 are the significant clusters for the TD group after correction (cluster correction and maxT). The clusters represent the face part where all the TD subjects significantly modulated pixel luminance. Red color represents increase of pixel luminance (white), while blue color represents decrease of pixel luminance (black).

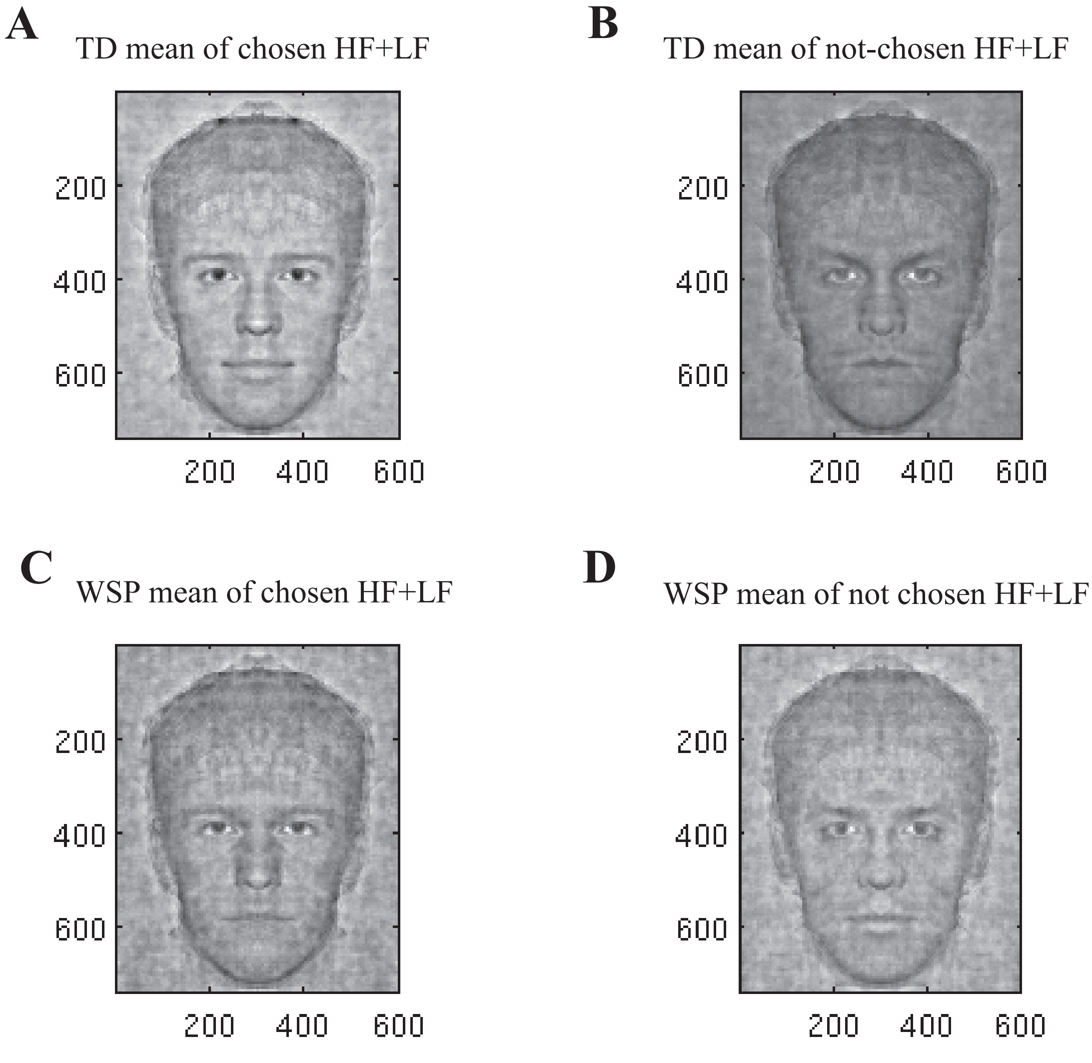



*Figure 27.* **Average image chosen by TD and WSP**. **A.** average of the noise chosen by TD group and overlie on the baseline image. This represents the internal representation of trustworthy face as it was generated by 22 TD **B.** Average of the not-chosen noise that by construction represents the anti-trustworthy face**. C.** average of the noise chosen by WS and overlie on the baseline image, the image display the representation of trustworthy face created by our group of WS patients **D.** average of the noise not-chosen by WSP.

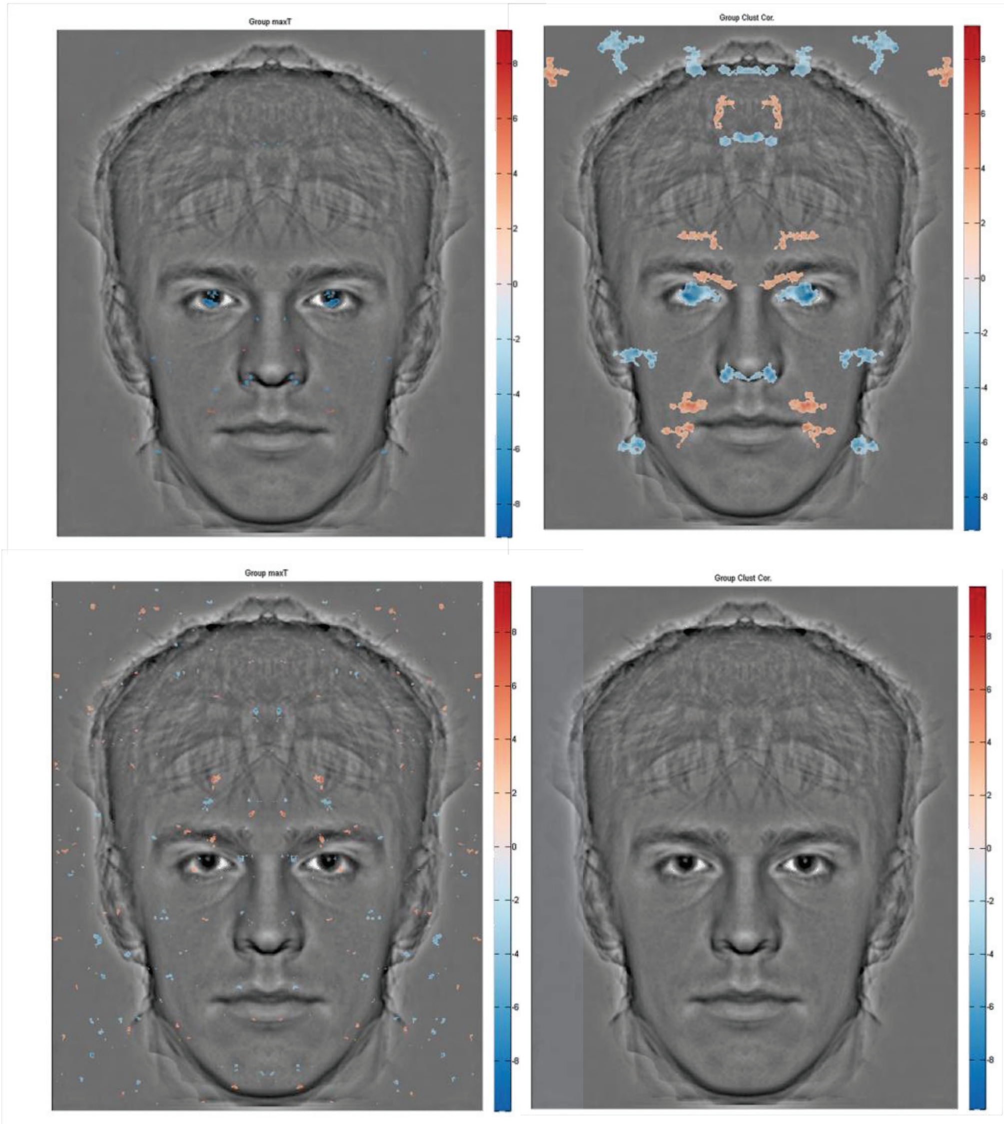



*Figure 28.* **Cluster analysis and MaxT** showing which regions are significantly chosen by TD (upper graph) and WSP when building the trustworthy face.

Because only 9 WS patients completed the task, the same analyses do not reveal any significant cluster. However MaxT analysis showed the presence of small regions that are modulated by WS group. Individual map of the 8 WS patients are presented, and statistical comparison to random choice is reported.

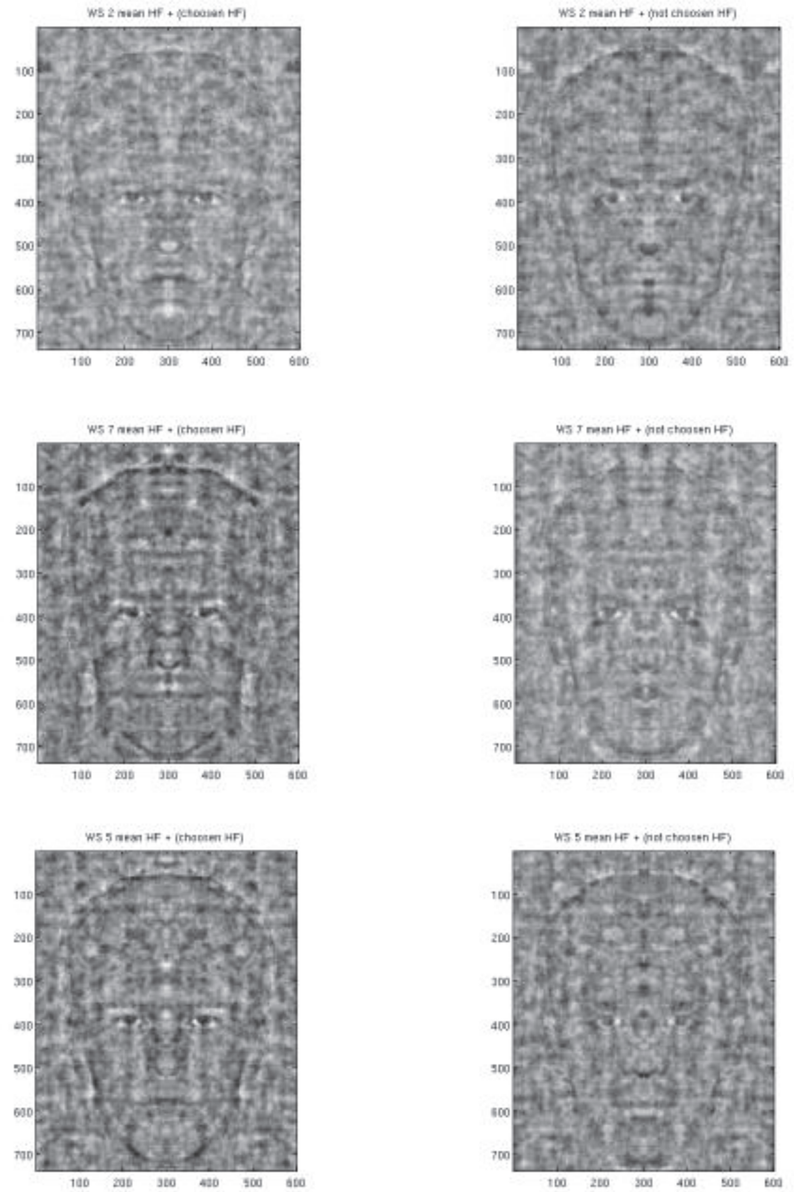



*Figure 29*. Individuals subjects maps showing the trustworthy face (left image) and the anti-trustworthy one (right) build up by WSP.

The absence of cluster significance in WSP may be due to a limited sample. I further compare WSP representations to the ones developed by TD subjects. The probability distribution of TD group was created using TD data as the normative distribution. Then, the representation of each WS patient was set in the curve of the normal distribution to show where each patient is placed compared to the norm.  As present in Figure 30, representation of WSP differs significantly from the representation of TD.

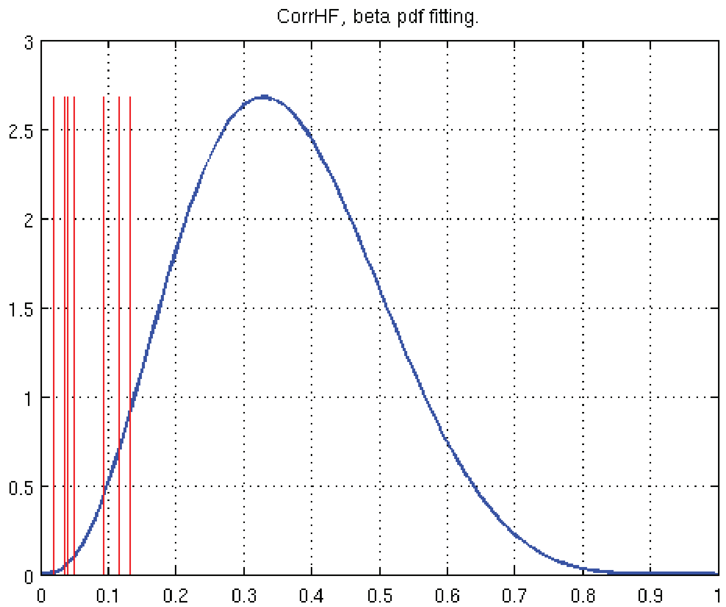



*Figure 30.* Blue line is the curve of probability's distribution of TD group created using TD data as norm. Red bars represent the 9 WS patients.

But do WS patients share the same representation among themselves? To assess agreement within and between groups (WS and TD), a matrix of correlation across subjects was created. Each cell represents the correlation between pixels selected by each participant and all the other (Figure 31). Plots present the mean of correlation score for each group. Score of correlation t-test to 0 showed that the scores of correlation of both WSP and TD were significantly different from 0, Ps>0.05. Importantly, the analysis reveals that two different representations exist between TD and WS as there was not agreement between TD and WS.

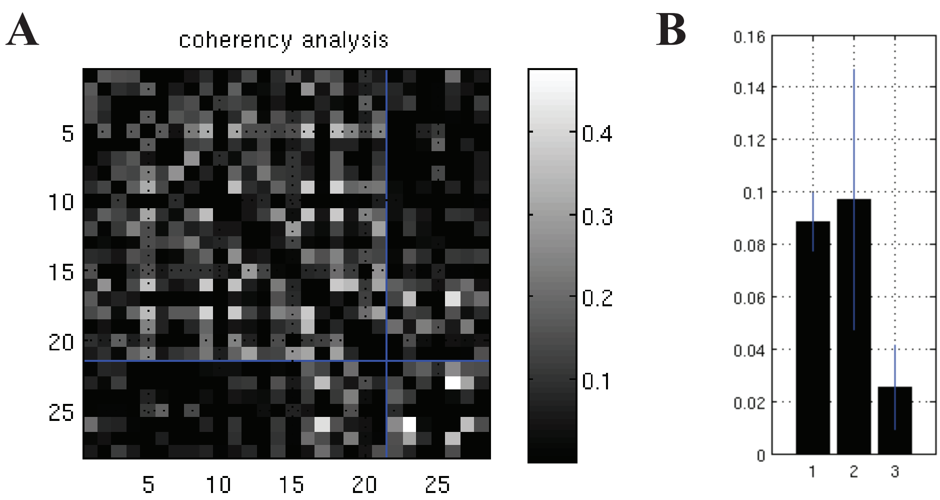



*Figure 31.* **A. Matrix of correlation**. From 1 to 21 are TD subjects. WS are from 22 to 28. **B. Graph bar report** the mean for each group. Graph bar as average of the agreement for each group. 1 is the agreement of TD, 2 is the agreement across WS. 3 is the agreement between TD and WS.

As expected, all controls create a similar trustworthy face modulating specific facial features. Cluster analysis showed that eyes color, eyebrows and a smiling mouth are relevant facial cues for trustworthy faces in controls. Interestingly, the representation of trustworthy faces in WS patients significantly differed from the control group. Moreover, while TD subjects built two faces that seem to differ for their trustworthiness, this was not the case of WSP who generates two faces that looked pretty much the same. To determine how the two faces stimuli generated by the two groups are endowed in trustworthiness level I am currently testing an independent group of subjects requested to rate these images for trust.  Overall, these results suggest that the “imprinted” representation of trust in doesn’t meet the normative requirements in a neurodevelopmental and genetic syndrome as WS.

## 4.4. Experiment 3

*Background*

Bilateral superior temporal regions are highly implicated in social cognition. Neural dysfunctions of this region may thus be the cause of atypical social behaviors in ASD or WSP(Allison et al., 2000; Zilbovicius et al., 2006).

In a previous work, atypical high-level perceptual processing of faces has been found in autistic patients (ASD) compared to a control group (Lio et al under submission). The study demonstrated that evoked activity in the superior temporal sulcus (STS), around 240 ms post stimulus onset, is critically impaired in the ASD population compared to a control group as reported in Figure 32. This result has been obtained using advanced linear decomposition techniques (Parra, Spence, Gerson, & Sajda, 2005), multi-variate pattern classification (MVPC – e.g.(Haynes & Rees, 2006)) and group blind source separation (gBSS – e.g. Lio & Boulinguez, 2013).

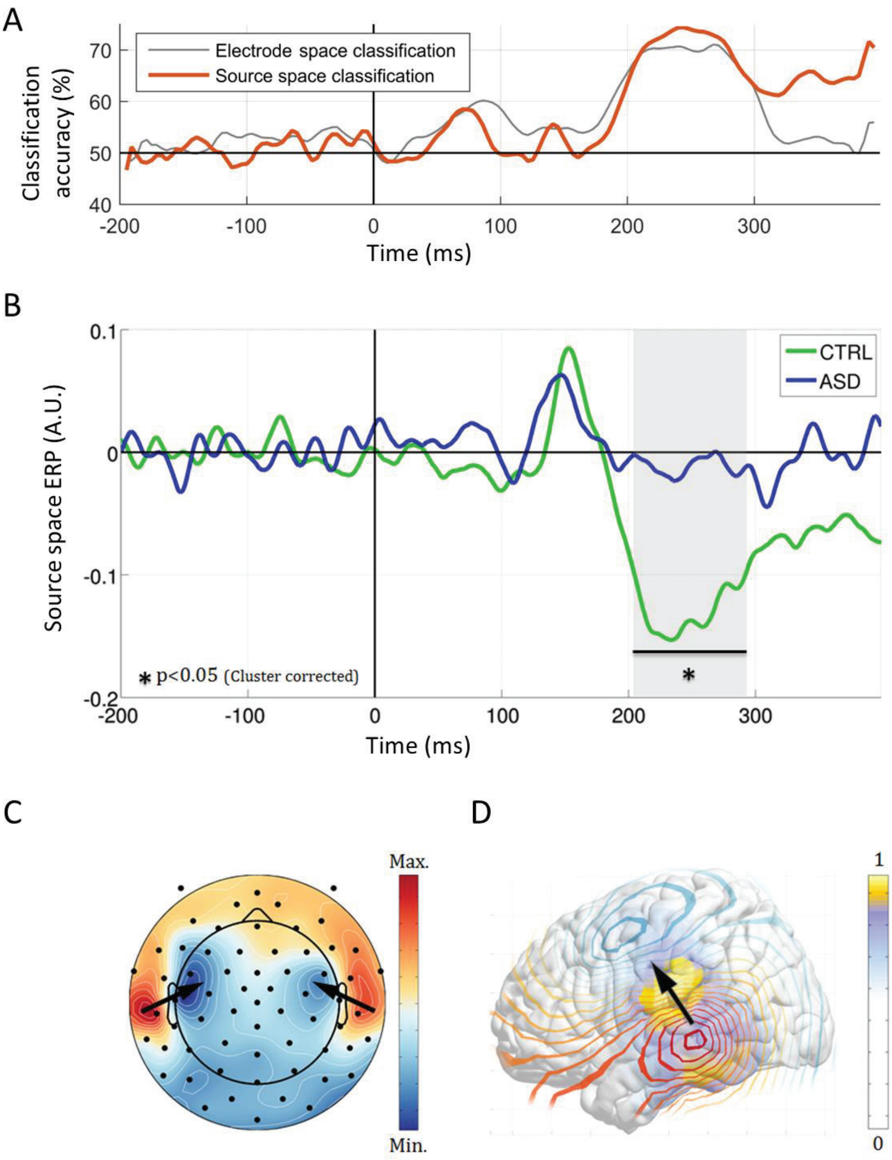



*Figure 32.* **Autism Spectrum Disorder (ASD) impaired evoked activity during neutral faces presentation, detected after group blind source separation** (Independent Component 15). **A-** Time-course of accuracy of the classification of ASD patients and control subjects at the source level, compared to the time-course of classification accuracy estimated in the electrode space. Classification accuracy reached a peak of 74%, ~240ms after the presentation of the face stimulus. **B-** Evoked group activity at the source level after group blind source separation. Control subjects show a prominent evoked activity between 200 and 300ms after the stimulus onset that is not observable in ASD patients ($p<0.05$ corrected. Maximum ~240ms after the stimulus onset). **C-** Scalp topography of the detected independent component (IC). The considered source presents a characteristic bilateral topography. **D-** sLORETA distributed sources localization of the IC. This estimated tomography, located bilaterally on the middle part of the temporal cortex, has a maximum around the lateral fissure and a local maximum around the inferior temporal sulcus. Combined with the directions of the two dipoles on the scalp topography and the latency of the evoked activity in the region, this information highly suggests a location of the component in superior temporal regions. **Data and Figure from Lio et al, under submission.**

In the same study, the authors investigated weather this source activity may be modulated by the face region where subject's visual attention is drawn on. The prediction was in fact that the observed negative activity at 240 ms, absent in ASD patients, may be highly linked to the face eye region (simulation of eye-contact). In a second experiment Lio et al., (under submission) addressed the following question: is STS activity modulated by participants' attention to the eyes or is this just elicited by perception of the face? To answer this question, a second study was designed to extract single trial dynamic of the activity of the STS evoked by drawing subjects' attention at different sectors of the face (Experiment 1). Differently from experiment 1, in this paradigm the subject was instructed to look constantly to the fixation cross. Crucially, face location differed across trials with respect to the fixation cross. Thus, subject's foveal vision was directed towards a specific part of the face (25 ROI may be overlapped to the fixation cross).

*The evoked activity, in superior temporal regions, is eye-sensitive*

Since non-parametric testing on Gaussian Kernel Density maps can be overly conservatives, a second analysis based on face regions of interest analysis was realized (Figure 33). Three region of interest (ROI) were built: One 'Rectangular Eyes region' (5° height, 12° width) encompassing the eye's region, one 'Triangular Nose/Mouth region' (10° height, 12° width) encompassing the lower facial features and one 'no-facial feature region' encompassing the other tested areas. For each subject, the mean evoked activities extracted from the cortical source of interest, during the [200ms, 300ms] time-cluster, were calculated at each ROI and Z-transformed. Then, a group analysis was processed using the Kruskal-Wallis non parametric test with post-hoc multiple comparisons FWER corrected by the Tukey-Kramer method. Statistical non-parametric mapping of the cortical evoked activity (permutation test, $p<0.05$ FWER corrected) showed that only the eyes region remains significant, with a maximum sensitivity when subject's foveal vision is located in the Eyes/Eyebrows region. Differently the two minimal activities are observed when subject's fovea was outside the face picture.

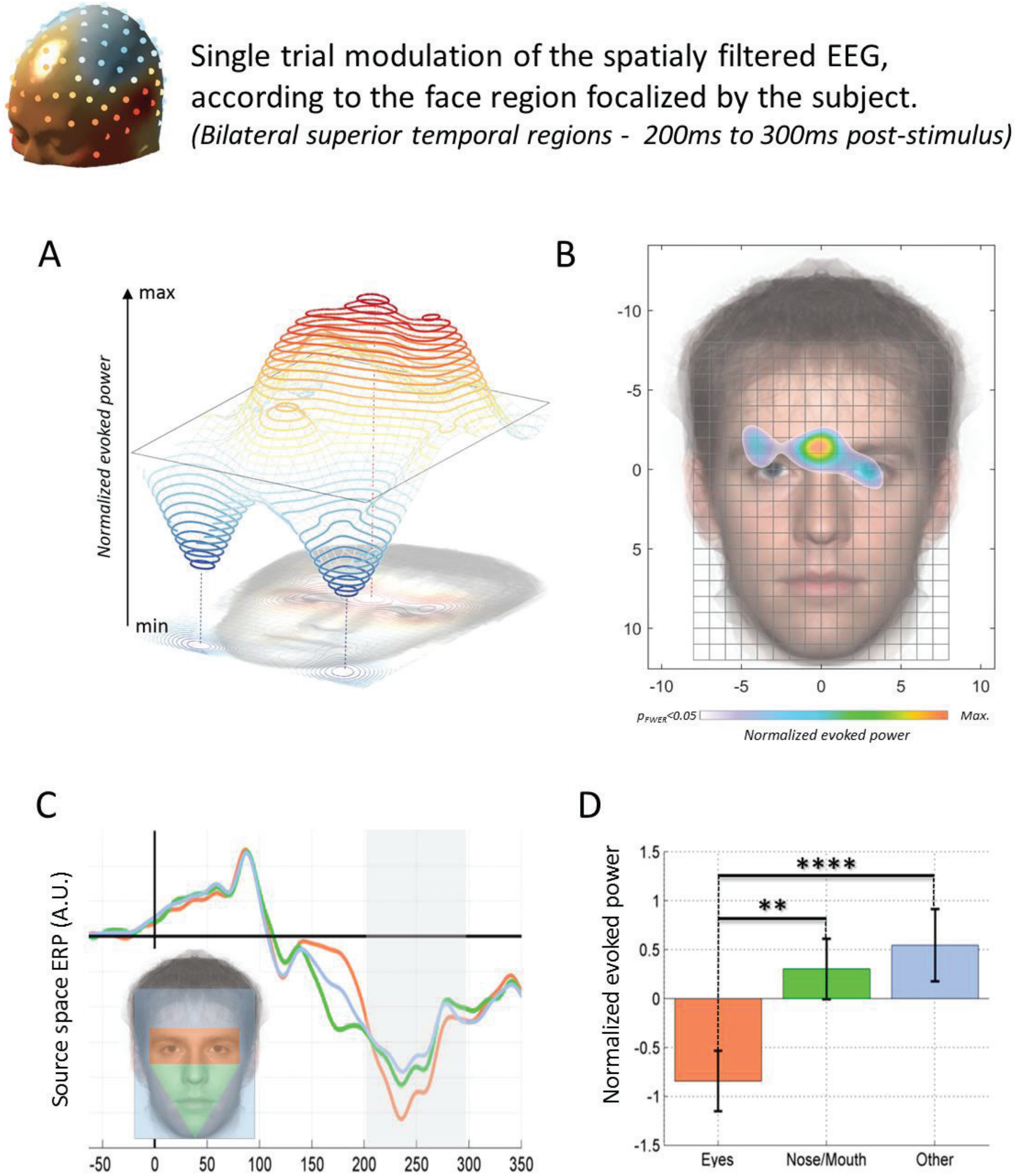



*Figure 33*. **Single trial modulation of the spatially filtered EEG, according to the face region focalized by the subjects. A**- Group average of the Normalized Gaussian Kernel Density mapping of the evoked activity in the superior temporal regions**.** Evoked power is maximal when subject's foveal vision is located in the Eyes/Eyebrows region while two local minima can be observed in the lower part of the tested area, when the line of sight is pointing outside the face picture. **B**- Statistical non-parametric mapping of the cortical evoked activity (permutation test, $p<0.05$ FWER corrected). Only the eyes region remains significant, with a maximum sensitivity when subjects are focusing in a particular region, between the two eyes. **C**- Regions of interest (ROI) analysis: Three ROIs were selected for the analysis. The recovered source presents the same negativity of the source recovered in the first experiment, in the 200ms – 300ms time period, reaching a maximum approximately 240ms after the stimulus onset. **D**- The eye region shows a marked higher activity relative to the nose/mouth area ($p<0.01$, FWER corrected) and relative to the other face regions ($p<0.0001$, FWER corrected). No significant differences were found between the nose/mouth area and the 'other / not internal face features' region ($p>0.05$, FWER corrected). Data and figures from (Lio et al under submission.

*Method*

**Single trial spatial filtering**

In order to measure the single trial behavior of the component identified in the first experiment, a spatial filter was calculated for each trial using minimum variance beamformer techniques (Van Veen et al. 1997, Robinson and Vrba 1996) in combination with the spatial information estimated at the group level with gBSS as described elsewhere (Albares et al. 2014). First, the mixing vector $A_i$ of the IC of interest, estimated in the first experiment, was interpolated to the 128 electrodes layout using the scalp surface of the MNI152 template (Mazziotta et al. 2001).

Then, considering the measured signal $x_t$ at trial $t$, the source signal $\widehat{s_t}$ was estimated by:

$$\widehat{s_t} = \mathrm{W_t}\, x_t$$

where the spatial filter $W_t$ is estimated by:

$$W_t = Cr_t^{-1} A_i \, (A_i^T Cr_t^{-1} A_i)^{-1}$$

and the regularized noise covariance matrix $Cr_t$ by:

$$Cr_t = C_t + \mu \, diag(C_t)$$

where $C_t$ is the data covariance matrix computed for the trial $t$, $\mu$ the Backus-Gilbert regularization parameter (=10) and $diag(C_t)$ the matrix of the diagonal elements of $C_t$ (the diagonal matrix of sensor noise).

**Density Mapping**

For each subject, a source spatial sensitivity map was built using Gaussian kernel density mapping. Gaussian Kernel Density mapping is a non-parametric way to estimate the probability density functions of a random variable, weighted here by the intensity of the evoked activity estimated at the source level. The method is similar to 'heat map' representations used in eye-tracking studies (see e.g. Caldara and Miellet, 2011). At each trial t, were the focalized area is located at the coordinates [xt, yt] on the face pictures, the mean evoked potential mt extracted from the cortical source of interest $\widehat{s_t}$, during the [200ms, 300ms] previously identified time-cluster, was multiplied by a two dimensional Gaussian kernel function with a mean value of [xt, yt] and a Full Width at Half Maximum (FWHM) of ~3.53° of visual angle (standard deviation = 1.5°). Then, the subject-level source spatial sensitivity map was built by averaging all Gaussian kernel functions. Finally, to highlight the 'most positive' and the 'most negative' areas, the mean value of the map was removed.

To highlight the face areas that evoked the 'most negative' activities at the subject level, a threshold was defined using non-parametric random-permutation test ($p<0.05$) to visualize each individual map.

Finally, the most sensitive face regions were identified at the group level using the 10 maps estimated at the subject level. Each ROI was tested using the non-parametric, one tailed, sign test (306405 tests, $p<0.05$) while the Family Wise Error Rate (FWER) was controlled using the maxT/minP multiple testing procedure (Westfall and Young, 1993), leading to a statistical non-parametric mapping of the evoked activity in the superior temporal regions. For all permutation-based tests the permuted values were, at the subject-level, the mt values.

## Modulation of STS activity in Williams's syndrome patients: data analysis using EEG separation source

Williams-Beuren Syndrome patients have been reported to be strongly interested by faces(Wendy Jones et al., 2000; Mervis et al., 2003; Paul, Stiles, Passarotti, Bavar, & Bellugi, 2002), with special interest in the eyes(M. a Porter et al., 2010). The interest for faces has been considered an expression of the high social appetence in this pathology but its neural mechanisms are largely unknown. Here, I test whether the abnormal social behavior during face processing may be due to an abnormal modulation of the superior temporal source using the paradigm used in (Lio et al. under submission).

### 4.4.1. Methods

**Participants**

The same 12 WSP (6 female, 6 male) and the 12 healthy controls (TD) individually matched for sex and chronological age (mean=11.9, SD=4.01) participated to EEG recording paradigm (128 electrodes). TD subjects had normal or corrected-to-normal vision and no history of psychiatric or neurological disease.

### Stimuli

12 real face pictures, built from photography selected from the Nimstim database (Tottenham et al. 2009) were used in this study (3 men, 3 women and their horizontal-flip counterparts). The proportions of the faces were slightly modified with the Gimp software (http://www.gimp.org/) in order to control the distances between facial features for each picture: interpupillary distance= 6.5°, eyes/nose distance= 5°, nose/mouth distance= 2°, mouth/chin distance= 5° - screen resolution ~31 pixels/degree of visual angle (Figure 34). The faces have been presented with an angular size that corresponds to a relatively close interpersonal distance, around 56 cm as it is reported in the personal space regulation scale (Hall 1963, Kennedy et al. 2009).

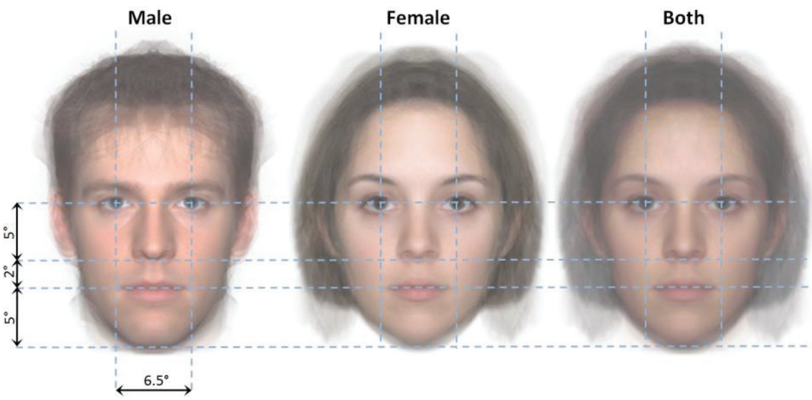



*Figure 34.* **Stimuli**. Average of respectively male/female/all stimuli used in this study.

### Task procedure

WS patients and control performed three blocks of 1500 trials; each block's duration was around 17 minutes. WS and controls performed the task in a Faraday's cage with always one experimenter inside. The subjects were seated in a darkened, shielded room with the head position controlled by an ophthalmic chin-rest device. The eyes at the same level of the fixation cross. The subject was instructed to look constantly to the fixation cross. Faces with same luminance distribution were presented before a tag question appeared. Number of face presented before the question was randomly determined and may vary form 7 ± 3. Crucially, location of the face with respect the fixation cross differed across trials. Thus, subject's foveal vision was directed towards a specific part of the face (25 ROI may be overlapped to the fixation cross as shown in Figure 35). When the question tag appeared, the subject was instructed to determine the gender of the last seen face stimulus. To avoid artefacts in the EEG signals due to movements or coordination problems, WS and control gave a vocal response and the experimenter press the right or left bottom accordingly to their response. Correctness of the response was recorded.

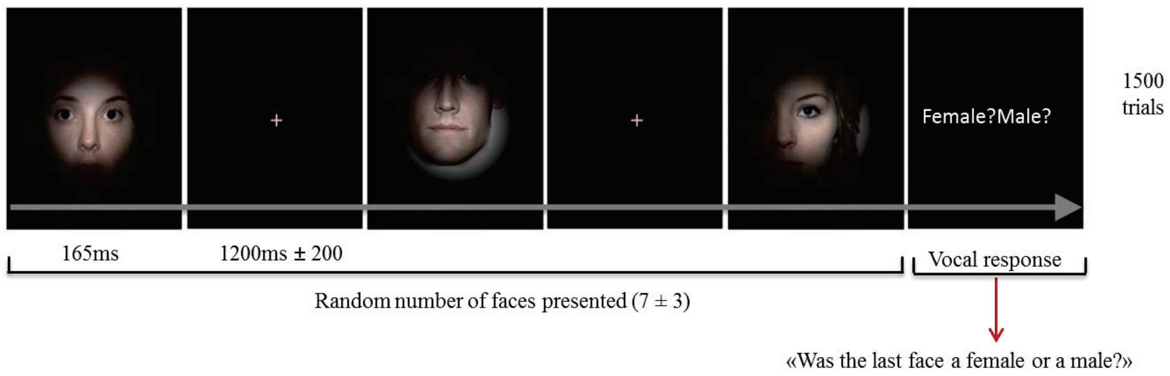



*Figure 35*. **Time course of a trial.** A random number of faces were presented (7 ± 3). Faces differed for their identity and sex (3x2) and for the facial feature's location to respect to the fixation cross. The subject was instructed to look constantly to the fixation cross. Stimulus has the same luminance distribution. Randomly (every 7±3 trials), a question tag appears and the subject was instructed to determine the gender of the last seen face stimulus (vocal response).

### EEG recording and preprocessing

The Brain Product™ actiCHamp system was used to record the electroencephalographic signal from 128 active electrodes (actiCAP 128Ch Standard-2) mounted in an elastic cap at 10-10 and 10-5 system standard locations (Oostenveld and Praamstra. 2001). All electrode impedances were kept below 50 kOhms. EEG data were recorded at a sampling rate of 5000 Hz with an online reference at the Fz electrode. Offline, data were band pass filtered using zero-phase Chebychev type II filters (Low pass - cutting frequency: 45 Hz, transition band width: 2 Hz, attenuation: 80 dB; order: 35, sections: 18 | High pass – cutting frequency: 0.3 Hz, transition band width: 0.2 Hz, attenuation: 80 dB; order: 9, sections: 5) and re-referenced to common average. Then, data were epoched from 200 ms before to 400 ms after the stimulus onset.

### Face regions of interest analysis

In order to measure the face region sensitivity of the superior temporal source we assessed single trial evoked response of this cortical region as function of the face part focalized by participants.

25 region of interest (ROI) were defined. For each subject, the mean evoked activities extracted from the cortical source of interest, during the [200ms, 300ms] time-cluster, were calculated at each ROI and Z-transformed. Then, a group analysis was processed using the

Kruskal-Wallis non parametric test with post-hoc multiple comparisons FWER corrected by the Tukey-Kramer method. An ANOVA was also performed to test whether the modulation of the source differed between TD and WSP.

### 4.4.2. Results

*Modulation of the STS source in WSP and TD*

The objective of this study was to replicate the results from (Lio et al, under submission) and extend this result considering a different pathology (WS) caractherized by an abnormal social behavior. More precisely, the main analysis aimed to look at the single trial modulation of the spatially filtered EEG, according to the face region (25) presented in the foveal field. For each subject the evoked activity in superior temporal regions, 200ms to 300ms post-stimulus, is normalized and mapped as probability density function of cortical activation. The ANOVA showed a strong significant effect of modulation of the source across ROI $F(1,24)=7.67$, $p<0.0001$. No effect of group was reported neither an interaction between group WSP/TD and face part focalized by the participants.

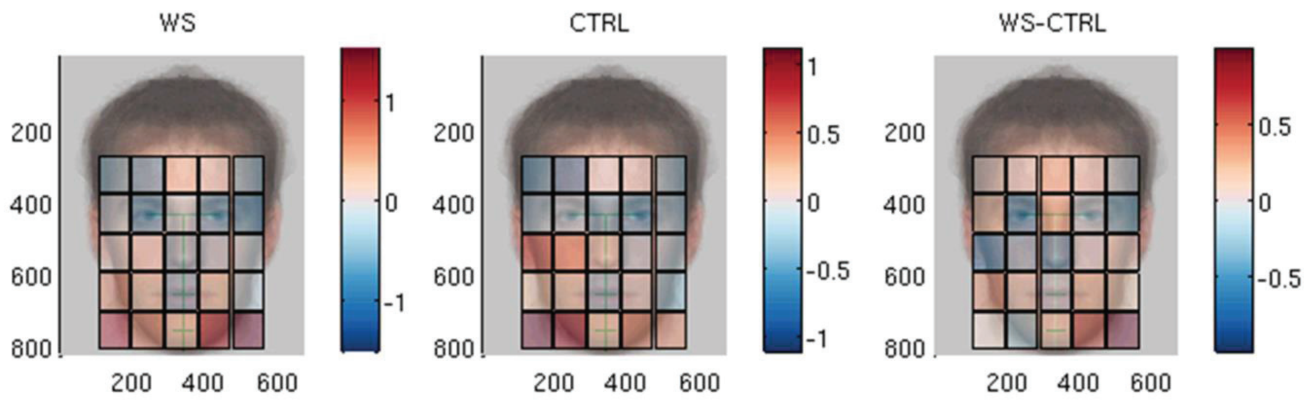



*Figure 36* **A.**WSP average of the Normalized Gaussian Kernel Density mapping of the evoked activity in the superior temporal regions. Evoked power is maximal when subject's foveal vision is directed towards the eyes/eyebrows region while minimal power is observed when subjects' foveal vision was directed towards parts outside the face. **B.** Control group. **C.** Difference between WSP and TD subjects.

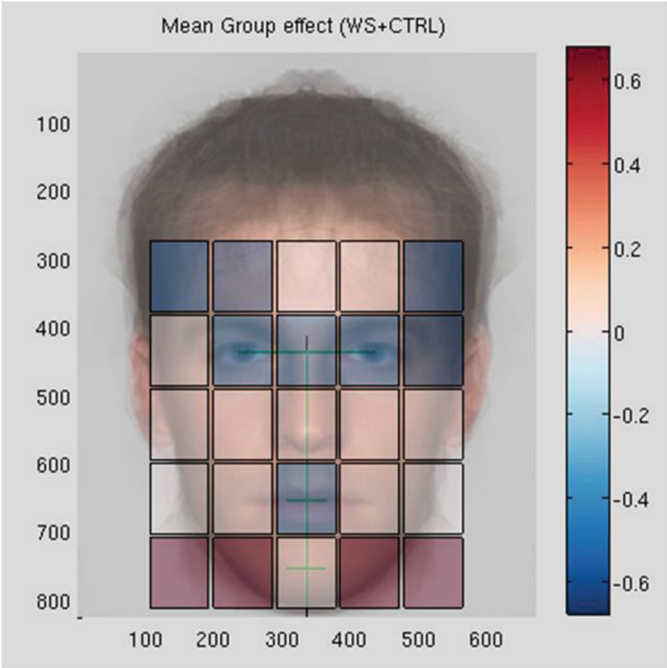



*Figure 37* **Evoked activity in the superior temporal regions by ROI**. Main effect of source modulation in both WSP and TD. Evoked power is maximal when subject's foveal vision is directed towards the eyes/eyebrows region and the mouths. Minimal power is observed when subject's foveal vision was directed towards the part out of the face. Activity reported is corrected using FDR.

### 4.4.3. Discussion

When looking at the eyes of natural face stimulus, a particular evoked activity was observed for neurotypical subjects using EEG and spatial filtering techniques. This source presents a pronounced bilateral evoked activity in the superior temporal regions to reach a maximum 240ms after the stimulus onset (Lio et al. under submission). Moreover, this activity is modulated by the region of the face where the subject's fovea is attracted on, with a maximum of activity elicited by the eyes area.

fMRI studies have identified tree cortical areas that appear to be the core of the adult face processing system: the inferior occipital gyrus ('OFA' – occipital face area), the middle fusiform gyrus (FFA – Fusiform Face Area) and the superior temporal sulcus (STS) (Haxby et al. 2000, O'Toole et al. 2002, Pascalis et al. 2011, Bernstein and Yovel. 2015). The current dominant neural models suggest a division of these structures in two neural pathways of face processing. The ventral stream, including the FFA, more dedicated to the treatment of face traits (the invariant facial aspects such as gender, age, identity, etc.). The dorsal pathway, including the STS, more specialized in the processing of the face state (the changeable facial aspects such as speech, implied motion, emotional expressions, attention and intentions) and also considered as the gateway of an extended system for social perception encompassing the orbitofrontal cortex and the amygdala (Allison et al. 2000). Results of Lio et al (under submission) suggested that this second pathway may be the cause of the abnormal social behavior in ASD population, especially of an averted gaze-contact behavior (abnormal evoked activity observed bilaterally in the STS region). We reasoned that this second pathway may also explain the atypical social behavior of WSP characterized by a strong will to create social relation, hyper-approachability and high interest for the eyes. For the moment the results showed that the modulation of the STS at 240ms has the same pattern in both WSP and TD. This investigation has shown that this region at that precise time is not the region that can explain difference in social perception between WSP and TD. However it is a good technique to investigate whether the time course of activity differ among these two groups.

# General Discussion

The research presented in this thesis focused on how facial information and, especially trust, is spontaneously and implicitly detected from faces.

In the first study, this topic has been approached from an evolutionary perspective. The objective was to test whether monkeys (Macaca Mulatta, Macaca Fascicularis) have a spontaneous preference towards trustworthy human faces, thus suggesting a capacity to detect facial cues similar to those used by humans when making fast judgments of trust. Monkeys and humans' eye movements were recorded while looking at pairs of trustworthy modulated faces (same identity and different level of trustworthiness). Despite different temporal and spatial strategies across species, the main finding was that monkeys and humans are significantly attracted to trustworthy human faces. Both monkeys and humans, in fact, showed a preferential attention to trustworthiness-associated facial cues.

Human data also revealed a correlation between looking time in the implicit and in the explicit task. In other words, I observed that during the implicit task ("look at faces as you want"), humans looked more those faces chosen as more trustworthy during the explicit task. These results clarify the implicit behavior of humans, although the question remains open for monkeys: what are the features of trustworthy human faces that attracted monkeys' attention? Perception of trustworthiness is constructed from multiple sources of information such as emotional valence, femininity and morphological facial features. In both species, I observed (study 1) that longer looking time was associated with faces judged as more feminine and with positive emotional valence. Moreover, monkeys and humans have a similar preference for faces with lower facial width to height ratio (FWHR). In humans, the facial metric of FWHR has been shown underlying the dimension of trustworthiness. Particularly, faces with lower FWHR are more likely to be judged as trustworthy(Stirrat & Perrett, 2010). Results of study 1 extend previous findings in humans and monkeys showing correlation between implicit looking time and the FWHR. Overall, these findings seem to suggest that certain facial features and trait-related information might be selected together at a particular point in the evolution development as they captured the attention of both species. This behavioral data thus exhort to further investigate the evolutionary origin of the mechanisms underlying the ability to detect social information from faces.

Overgeneralization of emotions may be the shared mechanism for trust detection used by monkeys and humans. This mechanism has been already pointed out as a potential explanation of perception of social traits in humans (Knutson, 1996; Montepare & Dobish, 2003; Todorov, 2008b). According to this hypothesis, neutral faces may resemble to emotional expressions. Thus an observer can infer trait attributes or social behavioral tendencies of neutral faces exploiting the emotional information. One explanation of monkeys and humans behavior in my results is that they both perceived the emotional positive valence from trustworthy faces, preferring to approach positive rather than negative valence faces. A strong claim may be that facial appearance of social traits does not really exist and that inferences of social traits are only the results of emotional evaluation of faces. The correlation found between happiness score and looking time may thus be interpreted in this direction. The data recording in the attractiveness experiment also go in the same direction. At the implicit level this mechanism may be useful, in fact when we meet an unknown person, overgeneralization of emotion will allow us to quickly decide whether this individual deserve our friendship or should be avoided.

Nevertheless, we may be able to explain if our choice to approach someone was more driven by a perception of attractiveness, trustworthiness, or extraversion. The results of study 2 showed that human ability to perform explicit judgments could be indeed performed without the influence of emotional information. The facial metric used to reach this result was the facial width to height ratio.

More precisely, the objective of study 2 was to investigate the independent and combined function of vertical and horizontal components of fWHR to disentangle how these facial features bias social judgements (trustworthiness, femininity and aggressiveness). Based on a previous literature, faces with higher fWHR tout court are judged as less trustworthy, more aggressive and less feminine (Carré & McCormick, 2008; Stirrat & Perrett, 2012). However, our findings rather demonstrate that it is possible to determine which component of fWHR is more relevant in the formation of these social impressions. Overall we observed that in all experiment and conditions the vertical component always better predicted participants'' judgments than the either the horizontal component or their combined contribution.

In agreement with previous studies, we also found that faces with larger width were judged as more aggressive and less feminine, regardless of sex identity. These results are coherent with previous literature showing that during puberty under the influence of testosterone, males would get larger facial width (Penton-Voak & Chen, 2004 Enlow & Hans, 1996, cited in Weston, Friday, Johnstone, Schrenk, 2004); and that, in return, the faces with larger width would be perceived as more aggressive (Carmen E. Lefevre et al., 2013). Hence, testosterone can be considered as a potential modulator of both physical (width of the face) and behavioral aspects. This step forward was important to clarify the role of biological constraints exert on facial metrics which are relevant for femininity and aggressiveness judgments. Following this reasoning, while taking advantage of our results on trust, future studies may assess the influence of neuromodulators relevant for trustworthiness, such as level of oxytocin (Lambert et al., 2014) and/or serotonin (Simonsen et al., 2014) on upper facial length.

As a consequence for future studies, these results strongly favor a methodology where the measure of the vertical component *per se* is favored over the complex and less fine-grained measure of fWHR. In fact, there are different ways to modulate fWHR, but as it has been shown in this work, it is important to determine which of these two variants enable faces' first impressions to occur. Future studies may continue to use our methodology to investigate other differences in social traits that are not been considered in the present work. Nonetheless, our study draws attention on the need to control for these two components when discussing the impact of the fWHR as an integrated measure.

The results of study 2 are interesting for better interpreting the results of study 1 and open new direction of research. Recent studies have demonstrated a link between FWHR and social dominance in Capuchin Monkeys (Borgi & Majolo, 2016; Carmen Emilia Lefevre et al., 2014). This raises the possibility that this species-typical facial trait may be used by monkeys to infer social personality of conspecifics, and recycled for making similar inferences about human faces. Modulation of upper facial height may be used to generate "trustworthy" monkeys' faces and show those faces to monkeys' observers. Another database of social traits in monkeys faces may be realized thank to a combined effort of ethological and engineering studies. A possibility may be to observe monkeys' behavior to detect personality traits and correlate these ones with facial features. This attempt may allow understanding how social traits are express by monkey's faces and successively test what these facial features trigger in monkeys' observers.

The second part of the manuscript presented three studies conducted in patients affected by Williams-Beuren syndrome (WS). This pathology has been considered a neurobiological human model for the atypical social behavior, as for instance overexpression of approachability towards others (Bellugi 1999, Jones 2000, Doyle 2004). This overexpression of kindness and positive social interactions strongly affect WSP life, since these patients fail to build strong interpersonal relationships during adulthood, a situation that leads in the long time to loneliness and anxiety. In fact WSP have high expectations about others behavior. The three studies conducted in this works shed light on the neurocognitive mechanisms responsible for this exaggerated social behavior and may be used as a basis for a cognitive treatment of this disease. Experiment 1 showed dissociation between the implicit and the explicit task. During the implicit task, WSP did not show any preference for the trustworthy or untrustworthy face; they spend the same time looking at both faces. This result is coherent with previous studies reporting a spontaneous tendency to approach strangers. Differently, during the explicit task the effect of preference towards trustworthy faces was present in both WSP and TD subjects. Moreover our results showed that when we forced subjects to find the trustworthy face between two faces, they select the same face as TD and they are coherent with their choice. The results of the explicit task show that WSP can perceive the modulated facial cues of trustworthy faces and they can use those features to perform coherent and correct social judgments from facial appearance. This dissociation between the implicit and explicit task also suggest that WSP can compensate their spontaneous atypical social behavior. This result is thus relevant for future cognitive remediation targeting their exaggerated social behavior.

The second experiment investigated the cognitive mechanisms at the root of their spontaneous atypical preference observed in study 1. Based on their first impression they were asked to choose among two noisy faces, the face that best resembles a trustworthy individual. This technique is interesting because it allows generating an averaged face image that reflects participant's internal representation of a trustworthy face (computed from their choices of noise). However for the first time it has been used in the clinical domain. As expected, all controls created a trustworthy face, which shared specific facial features modifications across control participants. Cluster analyses showed that eyes color, eyebrows and a smiling mouth are relevant facial cues for trustworthy faces in controls, a result that confirm and extend previous results. Interestingly, the modification of facial features performed to represent a trustworthy in WSP patients significantly differed from that of the control group. The cluster correlations across patients suggest however that WS patients use facial cues that are coherent. These results suggest that representation of trustworthiness may be different in a neurodevelopmental and genetic syndrome as WS. The existence of coherence across patients in the selection of facial features however, suggests that rather than lacking a stereotypic representation of trustworthiness, WS patients may use different sets of facial cues. In a further step of this project we are asking an independent group of WSP and TD subjects to judge the image created by the participants of the experiment. We expect in fact that TD will agree with the image create by his/her own group, but to disagree with WSP choices. Likewise, we also WSP should detect trustworthy faces prior created by WS but not chosen by TD subjects.

Results from experiment 1 and 2 confirmed the presence of an atypical pattern of social behavior in WS patients when they process facial features. In experiment 3, using EEG, we hypothesized that WSP atypical choices of trust faces may be linked to a dysfunctional activity of STS, a region, important for face processing. To analyze the EEG signal, we used a spatial filter which allowed in a previous study of our team to track a brain signal over the STS occurring at 240 ms in a population of healthy subjects when looking at human faces (Lio et al., under submission). Interestingly this signal was found absent in ASD patients (Lio et al., under submission). This last approach aimed to provide the neural correlate behind the deficient processing of social information from faces in WSP. As expected, control participants showed a modulation of the STS response at 240ms depending of the region of the face that was presented in the fovea: the eye region of the face responding more greatly than the rest of the face. Remarkably, eye contact also modulated activity of STS in WSP, suggesting that the same neural pattern exist between TD and WS. The results suggest that this source, that allows to discriminate ASD from TD, at this specific time (240 ms) seems not be differently modulated in WS. Further analysis, of these data, will consider the presence of different other sources that may allow to discriminate between WS and TD.

In conclusion, because trust is generally expressed as the product of an individual behavior, accurate judgments of trustworthiness must be based on real life interactions. This thesis showed however that two different stages of this process exist. Specifically, I demonstrated that bottom up mechanisms are employed to make first impression judgments of trustworthiness based on the perception of facial features. As describe in the present work, this fast route plays the fundamental role of modulating approach/avoidance behavior and may thus serve as a screening process of long-term social relationships. This low level mechanism is probably neurally written in the primates' brain since humans and monkeys showed similar tacit choices of trusty faces. . Finally, this work has also shown that judgment of trust is a multidimensional process. At a second stage, cognitive mechanisms take place for integrating information used to create a fast first impression and then generate a more accurate social decision of trust. Learning and experience further shape people's impressions of trust, as shown by the different results that I observed in WSP and controls. Future studies may continue on this path to better clarify all the steps of these two different stages (bottom-up and top-down mechanism), to learn more about this fascinating and fundamental process that underlies basic social interaction. Research on this topic will be also help to further our understanding of atypical social behavior in different pathologies.

## References


Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, *393*(June), 470–474. doi:10.1038/30982

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278. doi:10.1016/S1364-6613(00)01501-1

American Psychiatric Association. (2000). *DSM-IV-TR. Diagnostic and Statistical Manual of Mental Disorders.Fourth Edition. Text Revision*. (A. P. Association, Ed.). Washington, DC.

Anderson, J. R., Kuroshima, H., Takimoto, A., & Fujita, K. (2013). Third-party social evaluation of humans by monkeys. *Nature Communications*, *4*, 1561. doi:10.1038/ncomms2495

Anderson, J. R., Takimoto, A., Kuroshima, H., & Fujita, K. (2013). Capuchin monkeys judge third-party reciprocity. *Cognition*, *127*(1), 140–6. doi:10.1016/j.cognition.2012.12.007

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion (Washington, D.C.)*, *6*(2), 269–78. doi:10.1037/1528-3542.6.2.269

Barak, B., & Feng, G. (2016). Neurobiology of social behavior abnormalities in autism and Williams syndrome. *Nature Neuroscience*, *19*(6), 647–655. doi:10.1038/nn.4276

Barr, C. S., Newman, T. K., Shannon, C., Parker, C., Dvoskin, R. L., Becker, M. L., … Higley, J. D. (2004). Rearing condition and rh5-HTTLPR interact to influence limbic-hypothalamic- pituitary-adrenal axis response to stress in infant macaques. *Biological Psychiatry*, *55*, 733–738. doi:10.1016/j.biopsych.2003.12.008

Barraclough, N. E., & Perrett, D. I. (2011). From single cells to social perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1571), 1739–52. doi:10.1098/rstb.2010.0352

Barraclough, N. E., Xiao, D. K., Baker, C. I., Oram, M. W., & Perrett, D. I. (2005). Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *Journal of Cognitive Neuroscience*, *17*(3), 377–391. doi:0898929053279586

Bauer, R. M. (1984). Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the guilty knowledge test. *Neuropsychologia*, *22*(4), 457–469. doi:10.1016/0028-3932(84)90040-X

Beauchamp, M. S., Argall, B. D., Bodurka, J., Duyn, J. H., & Martin, A. (2004). Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci*, *7*(11), 1190–1192. Retrieved from http://dx.doi.org/10.1038/nn1333

Bellugi, U., Adolphs, R., Cassady, C., & Chiles, M. (1999). Towards the neural basis for hypersociability in a genetic syndrome. *NeuroReport*, *10*(8), 1653–1657. doi:10.1097/00001756-199906030-00006

Bellugi, U., Bihrle, A., Jernigan, T. L., Trauner, D., & Doherty, S. (1990). Neuropsychological, neurological, and neuroanatomical profile of Williams syndrome. *American Journal of Medical Genetics. Supplement*, *6*, 115–125. doi:10.1002/ajmg.1320370621

Bellugi, U., Wang, P. P., & Jernigan, T. L. (1994). Williams Syndrome: an Unusual Neuropsychological Profile. *Atypical Cognitive Deficits in Developmental Disorders:*

*Implications for Brain Function*, 1–59. Retrieved from http://www.wsfamilyofhope.org/uploads/An_Unusual_Neuropsychological_Profile.pdf

Belsky, J., Jonassaint, C., Pluess, M., Stanton, M., Brummett, B., & Williams, R. (2009). Vulnerability genes or plasticity genes? *Molecular Psychiatry*, *14*, 746–754. doi:10.1038/mp.2009.44

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience and Biobehavioral Reviews*, *55*, 536–546. doi:10.1016/j.neubiorev.2015.06.010

Binelli, C., Subirà, S., Batalla, A., Muniz, A., Sugranyés, G., Crippa, J. A.,Martin-Santos, R. (2014). Common and distinct neural correlates of facial emotion processing in social anxiety disorder and Williams syndrome: A systematic review and voxel-based meta-analysis of functional resonance imaging studies. *Neuropsychologia*, *64*, 205–217. doi:10.1016/j.neuropsychologia.2014.08.027

Borgi, M., & Majolo, B. (2016). Facial width-to-height ratio relates to dominance style in the genus *Macaca*. *PeerJ*, *4*, e1775. doi:10.7717/peerj.1775

Braccini, S. N., Lambeth, S. P., Schapiro, S. J., & Fitch, W. T. (2012). Eye preferences in captive chimpanzees. *Animal Cognition*, *15*(5), 971–978. doi:10.1007/s10071-012-0523-9

Brighina, F., Bisiach, E., Piazza, A., Oliveri, M., La Bua, V., Daniele, O., & Fierro, B. (2002). Perceptual and response bias in visuospatial neglect due to frontal and parietal repetitive transcranial magnetic stimulation in normal subjects. *Neuroreport*, *13*(18), 2571–5. doi:10.1097/01.wnr.0000052321.62862.7e

Bruce, V., & Young, a. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. doi:10.1111/j.2044-8295.1986.tb02199.x

Caldara, R., & Miellet, S. (2011). i Map : a novel method for statistical fixation mapping of eye movement data. *Behaviour Research Methods*. doi:10.3758/s13428-011-0092-x

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews. Neuroscience*, *6*(8), 641–51. doi:10.1038/nrn1724

Carré, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings. Biological Sciences / The Royal Society*, *275*(1651), 2651–6. doi:10.1098/rspb.2008.0873

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior: Research report. *Psychological Science*, *20*(10), 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., Poulton, R. (2003). Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science (New York, N.Y.)*, *301*(July), 386–389. doi:10.1126/science.1083968

Champoux, M., Bennett, a, Shannon, C., Higley, J. D., Lesch, K. P., & Suomi, S. J. (2002). Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. *Molecular Psychiatry*, *7*, 1058–1063. doi:10.1038/sj.mp.4001157

Corkum, V., & Moore, C. (1998). The origins of joint visual attention in infants. *Developmental Psychology*, *34*(1), 28–38. doi:10.1037/0012-1649.34.1.28

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In and J. E. T.

Barkow, Jerome H., Leda Ed Cosmides (Ed.), *The adapted mind: Evolutionary psychology and the generation of culture* (Oxford Uni., pp. 163–228). New York.

Dahl, C. D., Wallraven, C., Bülthoff, H. H., & Logothetis, N. K. (2009). Humans and macaques employ similar face-processing strategies. *Current Biology : CB*, *19*(6), 509–13. doi:10.1016/j.cub.2009.01.061

Darwin, C. (1872). *Expression of the Emotions*. (M. John & A. Street, Eds.)*Man and Animals*. London.

Deaner, R. O., Goetz, S. M. M., Shattuck, K., & Schnotala, T. (2012). Body weight, not facial width-to-height ratio, predicts aggression in pro hockey players. *Journal of Research in Personality*, *46*(APRIL), 235–238. doi:10.1016/j.jrp.2012.01.005

Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus macaques. *Current Biology*, *15*(6), 543–548. doi:10.1016/j.cub.2005.01.044

DeBruine, L. M. (2005). Trustworthy but not lust-worthy: context-specific effects of facial resemblance. *Proceedings. Biological Sciences / The Royal Society*, *272*(1566), 919–22. doi:10.1098/rspb.2004.3003

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015a). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex (New York, N.Y. : 1991)*, 1–14. doi:10.1093/cercor/bhv111

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015b). Functional Organization of Social Perception and Cognition in the Superior Temporal Sulcus. *Cerebral Cortex (New York, N.Y. : 1991)*, 1–14. doi:10.1093/cercor/bhv111

Deruelle, C., Rondan, C., Mancini, J., & Livet, M. O. (2006). Do children with Williams syndrome fail to process visual configural information? *Research in Developmental Disabilities*, *27*(3), 243–253. doi:10.1016/j.ridd.2005.03.002

Dobson, S. D., & Brent, L. J. N. (2013). On the evolution of the serotonin transporter linked polymorphic region (5-HTTLPR) in primates. *Frontiers in Human Neuroscience*, *7*(November), 588. doi:10.3389/fnhum.2013.00588

Dotsch, R., & Todorov,  a. (2012). Reverse Correlating Social Face Perception. *Social Psychological and Personality Science*, *3*, 562–571. doi:10.1177/1948550611430272

Doyle, T. F., Bellugi, U., Korenberg, J. R., & Graham, J. (2004). "Everybody in the world is my friend" hypersociability in young children with Williams syndrome. *American Journal of Medical Genetics. Part A*, *124A*(3), 263–273. doi:10.1002/ajmg.a.20416

Dufour, V., Pascalis, O., & Petit, O. (2006). Face processing limitation to own species in primates: A comparative study in brown capuchins, Tonkean macaques and humans. *Behavioural Processes*, *73*(1), 107–113. doi:10.1016/j.beproc.2006.04.006

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.

Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Common neural mechanisms for the evaluation of facial trustworthiness and emotional expressions as revealed by behavioral adaptation. *Perception*, *39*(7), 931–941. doi:10.1068/p6633

Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *The European Journal of Developmental Psychology*, *4*(1), 2–13. doi:10.1080/17405620601046832

Ferrari, P., Kohler, E., Fogassi, L., & Gallese, V. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13997–4002. doi:10.1073/pnas.250241197

Ferrari, P., Paukner, A., Ionica, C., & Suomi, S. J. (2009). Reciprocal Face-to-Face Communication between Rhesus Macaque Mothers and Their Newborn Infants. *Current Biology*, *19*(20), 1768–1772. doi:10.1016/j.cub.2009.08.055

Ferrari, P., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A., & Suomi, S. J. (2006). Neonatal imitation in rhesus macaques. *PLoS Biology*, *4*(9), 1501–1508. doi:10.1371/journal.pbio.0040302

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. S. (2009). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *NeuroImage*, *44*(2), 569–580. doi:10.1016/j.neuroimage.2008.09.011.The

Frigerio, E., Burt, D. M., Gagliardi, C., Cioffi, G., Martelli, S., Perrett, D. I., & Borgatti, R. (2006). Is everybody always my friend? Perception of approachability in Williams syndrome. *Neuropsychologia*, *44*(2), 254–259. doi:10.1016/j.neuropsychologia.2005.05.008

Gagliardi, C., Frigerio, E., Burt, D. M., Cazzaniga, I., Perrett, D. I., & Borgatti, R. (2003). Facial expression recognition in Williams syndrome. *Neuropsychologia*, *41*(6), 733–738. doi:10.1016/S0028-3932(02)00178-1

Geniole, S. N., Denson, T. F., Dixson, B. J., Carré, J. M., & McCormick, C. M. (2015). Evidence from meta-analyses of the facial width-to-height ratio as an evolved cue of threat. *PLoS ONE*, *10*(7), 1–18. doi:10.1371/journal.pone.0132726

Geniole, S. N., Keyes, A. E., Mondloch, C. J., Carré, J. M., & McCormick, C. M. (2012). Facing aggression: Cues differ for female versus male faces. *PLoS ONE*, *7*(1), 1–12. doi:10.1371/journal.pone.0030366

Germine, L. T., Duchaine, B., & Nakayama, K. (1AD). Where cognitive development and aging meet: Face learning ability peaks after age 30 (English). *Cognition*, *118*(2), 201–210. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=fcs&AN=23790036&lang=fr&site=eds-live

Getov, S., Kanai, R., Bahrami, B., & Rees, G. (2014). Human brain structure predicts individual differences in preconscious evaluation of facial dominance and trustworthiness. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsu103

Gibboni, R. R., Zimmerman, P. E., & Gothard, K. M. (2009). Individual differences in Scanpaths correspond with serotonin transporter genotype and behavioral phenotype in rhesus monkeys (Macaca mulatta). *Frontiers in Behavioral Neuroscience*, *3*(November), 50. doi:10.3389/neuro.08.050.2009

Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41. doi:10.1016/j.neuropsychologia.2006.04.015

Gómez-Valdés, J., Hünemeier, T., Quinto-Sánchez, M., Paschetta, C., de Azevedo, S., González, M. F., González-José, R. (2013). Lack of Support for the Association between Facial Shape and Aggression: A Reappraisal Based on a Worldwide Population Genetics Perspective. *PLoS ONE*, *8*(1). doi:10.1371/journal.pone.0052317

Gothard, K. M., Battaglia, F. P., Erickson, C. a, Spitler, K. M., & Amaral, D. G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal* *of Neurophysiology*, *97*, 1671–1683. doi:10.1152/jn.00714.2006

Gothard, K. M., Brooks, K. N., & Peterson, M. a. (2009). Multiple perceptual strategies used by macaque monkeys for face recognition. *Animal Cognition*, *12*(1), 155–67. doi:10.1007/s10071-008-0179-7

Gotlib, I. H., Joormann, J., Minor, K. L., & Hallmayer, J. (2008). HPA Axis Reactivity: A Mechanism Underlying the Associations Among 5-HTTLPR, Stress, and Depression. *Biological Psychiatry*, *63*, 847–851. doi:10.1016/j.biopsych.2007.10.008

Haas, B. W., Mills, D., Yam, A., Hoeft, F., Bellugi, U., & Reiss, A. (2009). Genetic influences on sociability: Heightened amygdala reactivity and event-related responses to positive social stimuli in Williams syndrome. United States, North America. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsbas&AN=edsbas.ftpubmed.oai.pubmedcentral.nih.gov.2754840&lang=fr&site=eds-live

Hadj-Bouziane, F., Liu, N., Bell, A. H., Gothard, K. M., Luh, W.-M., Tootell, R. B. H., … Ungerleider, L. G. (2012). Amygdala lesions disrupt modulation of functional MRI activity evoked by facial expression in the monkey inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), E3640–8. doi:10.1073/pnas.1218406109

Haselhuhn, M. P., Ormiston, M. E., & Wong, E. M. (2015). Men's facial width-to-height ratio predicts aggression: A meta-analysis. *PLoS ONE*, *10*(4), 1–10. doi:10.1371/journal.pone.0122637

Haselhuhn, M. P., & Wong, E. M. (2012). Bad to the bone: facial structure predicts unethical behaviour. *Proceedings of the Royal Society B: Biological Sciences*, *279*(July 2011), 571–576. doi:10.1098/rspb.2011.1193

Haselhuhn, M. P., Wong, E. M., & Ormiston, M. E. (2013). Self-Fulfilling Prophecies as a Link between Men's Facial Width-to-Height Ratio and Behavior. *PLoS ONE*, *8*(8). doi:10.1371/journal.pone.0072259

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, *32*, 203–218. doi:10.1016/S0166-4328(89)80054-3

Haxby, J. V., Hoffman, E. a., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. doi:10.1016/S1364-6613(00)01482-0

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature reviews.Neuroscience*, *7*(1471-003X (Print)), 523–534. doi:10.1038/nrn1931

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2013). Facial Structure Is Indicative of Explicit Support for Prejudicial Beliefs. *Psychological Science*, *24*, 289–296. doi:10.1177/0956797612451467

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015). Picking Teams: When Dominant Facial Structure is Preferred. *Journal of Experimental Social Psychology*, *59*, 51–59. doi:10.1016/j.jesp.2015.03.007

Hehman, E., Leitner, J. B., & Freeman, J. B. (2014). The Face-Time Continuum: Lifespan Changes in Facial Width-to-Height Ratio Impact Aging-Associated Perceptions. *Personality and Social Psychology Bulletin*, *40*, 1624–1636. doi:10.1177/0146167214552791

Homberg, J. R., & Lesch, K. P. (2011). Looking on the bright side of serotonin transporter gene variation. *Biological Psychiatry*, *69*(6), 513–519. doi:10.1016/j.biopsych.2010.09.024

Infants, Y. H., Yu, C., Smith, L. B., Yu, C., & Smith, L. B. (2016). The Social Origins of Sustained Attention in One- Report The Social Origins of Sustained Attention in One-Year-Old Human Infants. *Current Biology*, 1–6. doi:10.1016/j.cub.2016.03.026

Johnson, M. H. (2005). Subcortical Face Processing, *6*(October), 766–774. doi:10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1-2), 1–19. doi:10.1016/0010-0277(91)90045-6

Jones, W., Bellugi, U., Lai, Z., Chiles, M., Reilly, J., Lincoln, A., & Adolphs, R. (2000). II. Hypersociability in Williams Syndrome. *Journal of Cognitive Neuroscience*, *12*(supplement 1), 30–46. doi:10.1162/089892900561968

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*. doi:10.1038/nature12715

Kanwisher, N., McDermott, J., & Chun, M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.*, *17*, 4302–4311.

Karg, K., Burmeister, M., Shedden, K., & Sen, S. (2011). The serotonin transporter promoter variant (5-httlpr), stress, and depression meta-analysis revisited: Evidence of genetic moderation. *Archives of General Psychiatry*, *68*(5), 444–454. Retrieved from http://dx.doi.org/10.1001/archgenpsychiatry.2010.189

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N.,Campbell, R. (2004). Exploring the Williams syndrome face-processing debate: The importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(7), 1258–1274. doi:10.1111/j.1469-7610.2004.00322.x

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior*, *20*(3), 165–182.

Kramer, R. S. S., Jones, A. L., & Ward, R. (2012). A lack of sexual dimorphism in width-to-height ratio in white european faces using 2D photographs, 3D scans, and anthropometry. *PLoS ONE*, *7*(8). doi:10.1371/journal.pone.0042705

Ku, S. P., Tolias, A. S., Logothetis, N. K., & Goense, J. (2011). FMRI of the Face-Processing Network in the Ventral Temporal Lobe of Awake and Anesthetized Macaques. *Neuron*, *70*(2), 352–362. doi:10.1016/j.neuron.2011.02.048

Lambert, B., Declerck, C. H., & Boone, C. (2014). Oxytocin does not make a face appear more trustworthy but improves the accuracy of trustworthiness judgments. *Psychoneuroendocrinology*, *40*(1), 60–68. doi:10.1016/j.psyneuen.2013.10.015

Lander, K., & Butcher, N. (2015). Independence of face identity and expression processing: Exploring the role of motion. *Frontiers in Psychology*, *6*(MAR), 1–6. doi:10.3389/fpsyg.2015.00255

Lander, K., Humphreys, G., & Bruce, V. (2004). Exploring the role of motion in prosopagnosia: recognizing, learning and matching faces. *Neurocase*, *10*(6), 462–470. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=15788286&lang=fr&

site=eds-live

Latham, M. (2012). Cosmetic Surgery. *Encyclopedia Of Applied Ethics*, *1*, 667–673. doi:10.1016/B978-0-12-373932-2.00325-2

Lavater, J. K. (1880). *Essays on physiognomy; for the promotion of the knowledge and the love of mankind*.

Lefevre, A. (2015). Blood microsampling from the ear capillary in non-human primates. *Laboratory Animals*. doi:doi:10.1177/0023677215586911

Lefevre, C. E., Lewis, G. J., Perrett, D. I., & Penke, L. (2013). Telling facial metrics: Facial width is associated with testosterone levels in men. *Evolution and Human Behavior*, *34*, 273–279. doi:10.1016/j.evolhumbehav.2013.03.005

Lefevre, C. E., Wilson, V. a D., Morton, F. B., Brosnan, S. F., Paukner, A., & Bates, T. C. (2014). Facial width-to-height ratio relates to alpha status and assertive personality in capuchin monkeys. *PloS One*, *9*(4), e93369. doi:10.1371/journal.pone.0093369

Leonard, T. K., Blumenthal, G., Gothard, K. M., & Hoffman, K. L. (2012). How macaques view familiarity and gaze in conspecific faces. *Behavioral Neuroscience*, *126*(6), 781–91. doi:10.1037/a0030348

Lesch, K. P., Meyer, J., Glatz, K., Fliigge, G., Hinney, a, Hebebrand, J., … Heils, a. (1997). Neural Transmission Rapid Communication. *Journal of Neural Transmission*, 1259–1266.

Lio, G., & Boulinguez, P. (2013). Greater robustness of second order statistics than higher order statistics algorithms to distortions of the mixing matrix in blind source separation of human EEG: Implications for single-subject and group analyses. *NeuroImage*, *67*, 137–152. doi:10.1016/j.neuroimage.2012.11.015

Little, K., Riby, D. M., Janes, E., Clark, F., Fleck, R., & Rodgers, J. (2013). Heterogeneity of social approach behaviour in Williams syndrome: The role of response inhibition. *Research in Developmental Disabilities*, *34*(3), 959–967. doi:10.1016/j.ridd.2012.11.020

Losh, M., Bellugi, U., Reilly, J., & Anderson, D. (2000). Narrative as a social engagement tool: The excessive use of evaluation in narratives from children with Williams syndrome. *Narrative Inquiry*, *10*(2), 265–290. doi:10.1075/ni.10.2.01los

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190. doi:10.1016/j.jneumeth.2007.03.024

Maurer, D., & Salapatek, P. (1976). Developmental changes in the scanning of faces by young infants. *Child Development*, *47*(2), 523–527. doi:10.2307/1128813

Méary, D., Li, Z., Li, W., Guo, K., & Pascalis, O. (2014). Seeing two faces together: preference formation in humans and rhesus macaques. *Animal Cognition*, *17*(5), 1107–19. doi:10.1007/s10071-014-0742-3

Meltzoff, A. N., & Moore, K. M. (1977). Imitation_of_facial_and_manual_gestures.pdf. *Science*, *198*(4312), 75–78.

Mende-Siedlecki, P., Said, C. P., & Todorov, A. (2013). The social evaluation of faces: a meta-analysis of functional neuroimaging studies. *Social Cognitive and …*, *8*(3), 285–299.

Mende-Siedlecki, P., Said, C. P., & Todorov, A. (2013). The social evaluation of faces: a meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience*, *8*(3), 285–99. doi:10.1093/scan/nsr090

Mervis, C. B., Morris, C. a, Klein-Tasman, B. P., Bertrand, J., Kwitny, S., Appelbaum, L. G., & Rice, C. E. (2003). Attentional characteristics of infants and toddlers with Williams syndrome during triadic interactions. *Developmental Neuropsychology*, *23*(1-2), 243–68. doi:10.1080/87565641.2003.9651894

Meyer-Lindenberg, A., Hariri, A. R., Munoz, K. E., Mervis, C. B., Mattay, V. S., Morris, C. a, & Berman, K. F. (2005). Neural correlates of genetically abnormal social cognition in Williams syndrome. *Nature Neuroscience*, *8*(8), 991–3. doi:10.1038/nn1494

Meyer-Lindenberg, A., Mervis, C. B., & Berman, K. F. (2006). Neural mechanisms in Williams syndrome: A unique window to genetic influences on cognition and behaviour. *Nature Reviews Neuroscience*, *7*(5), 380–393. doi:10.1038/nrn1906

Mileva, V. R., Cowan, M. L., Cobey, K. D., Knowles, K. K., & Little, a. C. (2014). In the face of dominance: Self-perceived and other-perceived dominance are positively associated with facial-width-to-height ratio in men. *Personality and Individual Differences*, *69*, 115–118. doi:10.1016/j.paid.2014.05.019

Mimura, M., Hoeft, F., Kato, M., Kobayashi, N., Sheau, K., Piggot, J.,Reiss, A. L. (2010). A preliminary study of orbitofrontal activation and hypersociability in Williams Syndrome. *Journal of Neurodevelopmental Disorders*, *2*(2), 93–98. doi:10.1007/s11689-009-9041-8

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*(4), 237–254.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. *Advances in Experimental Social Psychology*, *30*, 93–161.

Mormann, F., Kornblith, S., Quiroga, R. Q., Kraskov, A., Cerf, M., Fried, I., & Koch, C. (2008). Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *The Journal of Neuroscience*, *28*(36), 8865–8872. doi:10.1523/JNEUROSCI.1640-08.2008

Mosher, C. P., Zimmerman, P. E., & Gothard, K. M. (2011). Videos of conspecifics elicit interactive looking patterns and facial expressions in monkeys. *Behavioral Neuroscience*, *125*(4), 639–652. doi:10.1037/a0024264

Mosher, C. P., Zimmerman, P. E., & Gothard, K. M. (2014). Facial Expressions in Monkeys, *125*(4), 639–652. doi:10.1037/a0024264.Videos

Mueller, A., Brocke, B., Fries, E., Lesch, K. P., & Kirschbaum, C. (2010). The role of the serotonin transporter polymorphism for the endocrine stress response in newborns. *Psychoneuroendocrinology*, *35*, 289–296. doi:10.1016/j.psyneuen.2009.07.002

Natu, V., & O'Toole, A. J. (2011). The neural processing of familiar and unfamiliar faces: A review and synopsis. *British Journal of Psychology*, *102*(4), 726–747. doi:10.1111/j.2044-8295.2011.02053.x

O'Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, *6*(6), 261–266. doi:10.1016/S1364-6613(02)01908-3

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, 1–5. doi:10.1016/j.tics.2014.09.007

Olivola, C. Y., & Todorov, A. (2010a). Elected in 100 milliseconds : Appearance-Based Trait Inferences and Voting. *Journal of Nonverbal Behavior*, 83–110. doi:10.1007/s10919-009-0082-1

Olivola, C. Y., & Todorov, A. (2010b). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315–324. doi:10.1016/j.jesp.2009.12.002

Oosterhof, N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092.

Oosterhof, N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(32), 11087–11092.

Oosterhof, N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion (Washington, D.C.)*, *9*(1), 128–133. doi:10.1037/a0014520

Özener, B. (2012). Facial width-to-height ratio in a Turkish population is not sexually dimorphic and is unrelated to aggressive behavior. *Evolution and Human Behavior*, *33*, 169–173. doi:10.1016/j.evolhumbehav.2011.08.001

Parr, L. a. (2011). The evolution of face processing in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *366*(1571), 1764–1777. doi:10.1098/rstb.2010.0358

Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2005). Recipes for the linear analysis of EEG. *NeuroImage*, *28*(2), 326–341. doi:10.1016/j.neuroimage.2005.05.032

Pascalis, O., de Haan, M., & Nelson, C. a. (2002). Is face processing species-specific during the first year of life? *Science (New York, N.Y.)*, *296*(5571), 1321–3. doi:10.1126/science.1070223

Pascalis, O., & Kelly, D. J. (2009). The Origins of Face Processing in Humans, *4*(2), 200–209.

Paukner, A., Huntsberry, M. E., & Suomi, S. J. (2009). Visual discrimination of male and female faces by infant rhesus macaques. *Developmental Psychobiology*, *52*(1), 54–61. doi:10.1002/dev.20412

Paukner, A., Suomi, S. J., Visalberghi, E., & Ferrari, P. F. (2009). Capuchin monkeys display affiliation toward humans who imitate them. *Science (New York, N.Y.)*, *325*, 880–883. doi:10.1126/science.1176269

Paul, B. M., Stiles, J., Passarotti, A., Bavar, N., & Bellugi, U. (2002). Face and place processing in Williams syndrome: evidence for a dorsal-ventral dissociation. *Neuroreport*, *13*(9), 1115–9. doi:10.1097/00001756-200207020-00009

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, *25*(4), 229–241. doi:10.1016/j.evolhumbehav.2004.04.003

Perrett, D. I., Smith, & Potter. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings Royal Society of London B*, *223*, 293–317.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology*, *24*(17), 2066–2070. doi:10.1016/j.cub.2014.07.060

Porter, M. a, Shaw, T. a, & Marsh, P. J. (2010). An unusual attraction to the eyes in Williams-Beuren syndrome: a manipulation of facial affect while measuring face scanpaths. *Cognitive Neuropsychiatry*, *15*(6), 505–30. doi:10.1080/13546801003644486

Porter, M. A., Coltheart, M., & Langdon, R. (2007). The neuropsychological basis of hypersociability in Williams and Down syndrome. *Neuropsychologia*, *45*(12), 2839–2849. doi:10.1016/j.neuropsychologia.2007.05.006

Porter, S., ten Brinke, L., & Gustaw, C. (2010). Dangerous decisions: the impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, *16*, 477–491. doi:10.1080/10683160902926141

Proffit, W., Fields, H., & Nixon, W. (1983). Occlusal forces in normal-and long-face adults. *Journal of Dental Research*. doi:10.1177/00220345830620051201

Quinn, P. C., Anzures, G., Izard, C. E., Lee, K., Pascalis, O., Slater, A. M., & Tanaka, J. W. (2011). Looking across domains to understand infant representation of emotion. *Emotion Review*, *3*(2), 197–206. doi:10.1177/1754073910387941.Looking

Quinn, P. C., Conforto, A., Lee, K., O'Toole, A. J., Pascalis, O., & Slater, A. M. (2010). Infant preference for individual women's faces extends to girl prototype faces. *Infant Behavior and Development*, *33*, 357–360. doi:10.1016/j.infbeh.2010.03.001

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*(9), 1109–1121.

Raadsheer, M. C., van Eijden, T. M., van Ginkel, F. C., & Prahl-Andersen, B. (1999). Contribution of jaw muscle size and craniofacial morphology to human bite force magnitude. *Journal of Dental Research*, *78*(1), 31–42. doi:10.1177/00220345990780010301

Reilly, J., Klima, E. S., & Bellugi, U. (1990). Once more with feeling: Affect and language in atypical populations. *Development and Psychopathology*, *2*(4), 367–391. doi:10.1017/S0954579400005782

Riby, D. M., Doherty-Sneddon, G., & Bruce, V. (2009). The eyes or the mouth? Feature salience and unfamiliar face processing in Williams syndrome and autism. *Quarterly Journal of Experimental Psychology*, *62*(1), 189–203. doi:10.1017/CBO9781107415324.004

Ross, W., & Lacroix, J. (1996). Multiple meanings of trust in negotiation theory and research: a literature review and integrative model. *The International Journal of Conflict Management*, *7*(4), 314–360.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the temporo-parietal junction in "theory of mind." *NeuroImage*, *19*(4), 1835–1842. doi:10.1016/S1053-8119(03)00230-1

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others : Neural correlates of social interaction, *44*, 718–730. doi:10.1016/j.neuropsychologia.2005.07.017

Sereno, M. I., & Tootell, R. B. H. (2005). From monkeys to humans: What do we now know about brain homologies? *Current Opinion in Neurobiology*, *15*(2), 135–144. doi:10.1016/j.conb.2005.03.014

Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional Neuroanatomy of Face and Object Processing. *Brain*, *115*, 15–36. doi:10.1093/brain/115.1.15

Sigala, R., Logothetis, N. K., & Rainer, G. (2011). Own-species bias in the representations of monkey and human face categories in the primate temporal lobe. *Journal of Neurophysiology*, *105*(6), 2740–52. doi:10.1152/jn.00882.2010

Simons, D. J., & Levin, D. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, *5*(4), 644–649. doi:10.3758/BF03208840

Simonsen, A., Scheel-Kruger, J., Jensen, M., Roepstorff, A., Moller, A., Frith, C. D., & Campbell-Meiklejohn, D. (2014). Serotoninergic effects on judgments and social learning of trustworthiness. *Psychopharmacology*, *231*(14), 2759–2769. doi:10.1007/s00213-014-3444-2

Skwerer, D. P., Ammerman, E., André, M. C., Ciciolla, L., Fine, A. B., & Tager-Flusberg, H. (2011). A multimeasure approach to investigating affective appraisal of social information in Williams syndrome. *Journal of Neurodevelopmental Disorders*, *3*(4), 325–334. doi:10.1007/s11689-011-9100-9

Sliwa, J., Duhamel, J.-R., Pascalis, O., & Wirth, S. (2011). Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(4), 1735–1740. doi:10.1073/pnas.1008169108

Sliwa, J., Plante, A., Duhamel, J. R., & Wirth, S. (2016). Independent Neuronal Representation of Facial and Vocal Identity in the Monkey Hippocampus and Inferotemporal Cortex. *Cerebral Cortex*, *26*(3), 950–966. doi:10.1093/cercor/bhu257

Steeves, J., Dricot, L., & Rossion, B. (2004). Durham Research Online. *Journal of Business Ethics*, *44*(April), 0–103. doi:10.1063/1.2756072

Stewart, L. H., Ajina, S., Getov, S., Bahrami, B., Todorov, A., & Rees, G. (2012). Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology. General*, *141*(4), 715–27. doi:10.1037/a0027950

Stirrat, M., & Perrett, D. I. (2010). Valid facial cues to cooperation and trust: male facial width and trustworthiness. *Psychological Science*, *21*(3), 349–54. doi:10.1177/0956797610362647

Stirrat, M., & Perrett, D. I. (2012). Face structure predicts cooperation: men with wider faces are more generous to their in-group when out-group competition is salient. *Psychological Science*, *23*(7), 718–22. doi:10.1177/0956797611435133

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences of the United States of America*. doi:10.1073/pnas.0706079105

Todorov, A. (2008a). Evaluating faces on trustworthiness. *Annals of the New York Academy of Sciences*.

Todorov, A. (2008b). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*, 208–224. doi:10.1196/annals.1440.012

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience*, *3*(2), 119–127. doi:10.1093/scan/nsn009

Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion (Washington, D.C.)*, *13*(4), 724–38. doi:10.1037/a0032335

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*, *25*(3), 395–410. doi:10.1080/02643290802044996

Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*(1), 163–173. doi:10.1016/j.neuropsychologia.2006.04.018

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, (August 2014), 1–46. doi:10.1146/annurev-psych-113011-143831

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating Faces on Trustworthiness After Minimal Time Exposure. *Social Cognition*, *27*(6), 813–833.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. H. (2003). Faces and objects in macaque cerebral cortex. *Nat Neurosci*, *6*(9), 989–995. Retrieved from http://dx.doi.org/10.1038/nn1111

Tsao, D. Y., Moeller, S., & Freiwald, W. a. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 19514–19519. doi:10.1073/pnas.0809662105

Uher, R. (2008). The implications of gene-environment interactions in depression: will cause inform cure? *Molecular Psychiatry*, *13*(August), 1070–1078. doi:10.1038/mp.2008.92

Van Essen, D. C. (2004). Organization of visual areas in macaque and human cerebral cortex. *The Visual Neurosciences*, *1*, 507–521.

Watson, K. K., Ghodasra, J. H., & Platt, M. L. (2009). Serotonin transporter genotype modulates social reward and punishment in rhesus macaques. *PLoS ONE*, *4*(1). doi:10.1371/journal.pone.0004156

Weinberg, S. M., Raffensperger, Z. D., Kesterke, M. J., Heike, C. L., Cunningham, M. L., Hecht, J. T., … Marazita, M. L. (2015). The 3D Facial Norms Database: Part 1. A Web-Based Craniofacial Anthropometric and Image Repository for the Clinical and Research Community. *The Cleft Palate-Craniofacial Journal*, *00*(August), 15–199. doi:10.1597/15-199

Weston, E. M., Friday, A. E., & Liò, P. (2007). Biometric evidence that sexual selection has shaped the hominin face. *PLoS ONE*, *2*(8). doi:10.1371/journal.pone.0000710

Willis, J., & Todorov, A. (2006). First Impression Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science*, *17*(7), 592–598.

Winston, J. S., Strange, B. A., O’Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat Neurosci*, *5*(3), 277–283. Retrieved from http://dx.doi.org/10.1038/nn816

Yovel, G., & Freiwald, W. a. (2013). Face recognition systems in monkey and human: are they the same thing? *F1000prime Reports*, *5*(April), 10. doi:10.12703/P5-10

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in Neurosciences*, *29*(7), 359–366. doi:10.1016/j.tins.2006.06.004