



**HAL**  
open science

# Neural mechanisms of contextual influences during social perceptual decisions

Marwa El Zein

► **To cite this version:**

Marwa El Zein. Neural mechanisms of contextual influences during social perceptual decisions. *Neurons and Cognition [q-bio.NC]*. Université Pierre et Marie Curie - Paris VI, 2015. English. NNT: 2015PA066436 . tel-01630679

**HAL Id: tel-01630679**

**<https://theses.hal.science/tel-01630679v1>**

Submitted on 8 Nov 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THESE DE DOCTORAT DE  
L'UNIVERSITE PIERRE ET MARIE CURIE**

Spécialité

Neurosciences cognitives

Ecole doctorale Cerveau Cognition Comportement

Présentée par

**Marwa El Zein**

Pour obtenir le grade de

**DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE**

# **Neural mechanisms of contextual influences during social perceptual decisions.**

Soutenue le 6 Novembre 2015 devant le jury composé de :

Dr. Julie GREZES, Directrice de thèse, DR

Dr. Valentin WYART, Co-directeur de thèse, CR

Pr. Philippe SCHYNS, Rapporteur, Professor

Dr. Floris DE LANGE, Rapporteur, Professor

Dr. Mathias PESSIGLIONE, Examineur, DR

Dr. Nathalie GEORGE, Examinatrice, DR

Pr. Beatrice DE GELDER, Examinatrice, Professor

Dr. Jérémie MATTOU, Examineur, CR

## Table of contents

<b>Abstract</b> .....	8
<b>Résumé</b> .....	9
<b>Abbreviations</b> .....	10
<b>INTRODUCTION</b> .....	11
1-Social perception: the case of faces and facial expressions .....	12
1.1- Why the face and what are facial expressions of emotions? .....	12
1.2- How are faces and facial expressions represented in the brain?.....	16
1.3- When are faces and facial expressions represented in the brain? .....	18
1.4- Where are faces and facial expressions represented in the brain? .....	19
2-Contextual effects in social perception .....	24
2.1- Contextual elements characteristic of the observer .....	25
2.2- Contextual signals related to the environment.....	28
2.3- Contextual signals from the emitter .....	32
3-Perceptual decision-making.....	42
3.1- Principal decision-making models .....	42
3.2- Neural findings in perceptual decision-making .....	48
4-Contextual effects in perceptual decision-making.....	57
4.1- Expectation/probability cues .....	58
4.2- Attention/relevance cues .....	60
5-Informing contextual effects in emotion perception with decision models .....	62
<b>Part 1: Attention as a contextual factor during the perception of social cues</b> (Experiment 1).....	65
<b>Part 2: Gaze direction as a contextual factor during emotion categorization</b> (Experiment 2).....	74

<b>Part 3: Gaze direction influence on emotion categorization: a few unanswered questions</b> (follow-up behavioral studies).....	116
3.1- General approach in behavioral analyses.....	116
3.2- Question 1: Characterization of the effect: contribution of high and low spatial frequencies during Fear/Angry categorization (Experiment 3) .....	117
3.3- Question 2 Specificity of the effect: Categorization of Happy/Angry expressions (Experiment 4) .....	124
3.4- Question 3 Reciprocity of the effect: Emotion influence on gaze categorization (Experiment 5) .....	130
3.5- Question 4 Generalization of the effect to a representative pool of the population (Experiment 6).....	137
<b>Part 4: Expectation about the emotions of individual identities as a contextual factor</b> (Experiment 7).....	145
<b>General discussion</b> .....	174
The evolutionary advantage of threat detection .....	176
Sensory and motor regions in emotion processing: functional roles? .....	178
Selective representation of threat signals: is the amygdala the starting point? ...	182
Anger and Fear emotions .....	184
Contextual effects related to congruence vs. significance to the observer.....	187
Decision-making in social contexts .....	191
<b>Conclusion</b> .....	194
<b>References</b> .....	196

*To my parents*

## Acknowledgments

First and foremost, I would like to thank my supervisor Julie Grèzes. Through your enthusiastic approach, you gave me the taste for research from the very first months we worked together. You taught me all the basics of a scientific approach, how to address questions, operate them, interpret findings and write papers, by challenging me at various levels, always pushing me further in my thoughts, and guiding me at the appropriate moments. I cannot thank you enough for all your support, for your trust, and for giving me so many opportunities to present my work abroad and to collaborate in numerous projects. Working with you was above all a real pleasure on a daily basis.

I am extremely grateful to my co-supervisor Valentin Wyart. Thank you for sharing your positive enthusiasm and your prolific ideas, and for communicating your scientific rigour and expertise with great pedagogy, things seem so simple when you explain them! I consider myself very lucky to have worked with you along this road.

Thanks to Etienne Koechlin for welcoming me in the LNC. Thanks to Laura Chadufeu and Marine Magne for making life so easy at the lab. Thanks to all the members of the LNC for the pleasant atmosphere at work every day. Particular thanks to Anne-Do Lodeho, for being so helpful and being a very good listener. Special thanks to Marianna, Marion and Florence for your friendship and support. Marion it was great to have you on my side in the writing process, your optimism and positive advices really helped.

Thanks to the social team (past and current), my lab 'family': Guillaume *GuiGui*, thanks for pushing me to work in the social team in the first place, discussions with you were and always are very enjoyable moments. Michèle *MichMich*, thanks for your constant support and encouragement, and for always bringing laughter into the lab. Lou, thanks for the very pleasant moments of collaboration, discussions, and laughs. Terry *Chérie*, thanks for wonderful advices, those we aim to follow. Mathias and Amelie, it was always nice to see you when you popped in the lab. Christina and Hannah, your presence brought a positive dynamism into the team, and it is really nice working with you too, thanks for your trust. Thanks to all those who helped with my experiments, Lena, Florie, Hélène, Lydia, and Flavie. Thanks to Lise, for a nice year of collaboration. Thanks to Lucile, for the great push you gave me at the very beginning and for being there all along. Thanks to Laurence Conty and Coralie Chevallier, it was great and fruitful to collaborate with you.

And a particular IMMENSE thanks to Emma *querida*, who shared almost every minute I spent in the lab since the beginning, thanks for the laughs, the hugs, the advices, the complicity, the friendship... You were there every single day, always in my perspective, listening to any (really any) thing I had to say at any moment, thanks for your impressive patience and for always cheering me up. Thank you for making work a warm place.

Thanks to all the friends that were around all these years, since the cogmaster, Antoine, Aureliane, and Juliane, since the 'licence', Lea and Stephanie, and since childhood Maria, Mia and Mira.

Turning to family, I warmly thank my parents to whom I dedicate this PhD, you both are my cognitive and emotional strengths and drives. Thanks a million Diala for your continual friendship, and for just being there all these years, making Paris our home. Thanks to Azza, when you told me that doing a PhD was 'a life experience', I didn't realize how much you were right until now, thanks for being there with all your encouraging and helpful words despite the oceans between us. Thanks to Birendra, the discussions with you are always very pleasant, and your comments very helpful. Thanks to my little niece Meera, you made the writing process such an easy one by your simple day-to-day presence. Thanks to Hyam, for being a family in Paris, and such a good listener.

Thanks to my best friend Alizee, *and yes ! we did it !* we realized our undergraduate dream, the one we shared since the very first day we met at Jussieu, '*étudier les neurosciences*'. We really walked along this road together, from the biology undergrad at UPMC when we already travelled to Gif sur Yvette to listen to incomprehensible neuroscience conferences, to the Cogmaster when *oh finally* the brain was the main topic, to our PhDs, sharing ALL the good and bad moments - I cannot even imagine how the past years would have been without you and our weekly cafe lea.

Finally, thanks to Kleanthi, well, for so many things, everything, and I am unable to write everything down. Thank you for encouraging and listening to my work and ideas and for all our intense discussions around science. Thank you for knowing exactly how to handle every difficult moment I went through and for your perpetual optimism, you keep me going.

*“...expression in itself, or the language of the emotion, as it has sometimes been called, is certainly of importance for the welfare of mankind”*

Charles Darwin, 1872 *The expression of the emotions in man and animals*. London: John Murray. 1st edition.



## Abstract

Everyday social decisions require the combination of multiple sources of information and therefore build upon abundant contextual elements such as the social cues of emitters (e.g., gaze direction, emotion, gesture), the attentional focus of observers, their mood and their past experience. The work conducted during this Ph.D. (including three main studies in healthy human subjects) aimed at characterizing the cognitive and neural mechanisms of contextual influences in social settings.

In the first Electroencephalography (EEG) study, evoked potential and source reconstruction analysis revealed that unattended social signals are integrated in the human brain within 170 ms after stimulus onset as long as they are contextually relevant to the observer. However, selective attention increased the contribution to this integration of neural regions specific to the attended feature. In the second study, model-based behavioral and single-trial EEG analyses demonstrated that the mechanisms underlying the integration of multiple social cues from faces implicate an enhancement of the perceptual sensitivity to observer-relevant threat signals. This increased sensitivity to threat-signaling emotions emerged in ventral temporal and dorsal motor regions in parallel, starting 200 ms after stimulus onset until response. Moreover, individual characteristics of the participants, here their anxiety, influenced the regions where threat-signaling emotions were selectively encoded in a doubly dissociable fashion: preferentially in face selective temporal regions for low anxious individuals but in motor regions for high anxious individuals, consistently with the adaptive role of anxiety in action preparation under threat. Finally, in the third study, model-based behavioral and pupillometric analyses showed that experimentally induced priors as well as participants' idiosyncratic priors about the emotion (mood) of individual identities – i.e., identity/emotion associations, modulated decision biases rather than the sensitivity to displayed emotions. This was reflected in pupillary responses, which were increased to effortful decisions against biases on unexpected stimuli.

While co-emitted social cues interact by boosting bottom-up processing of relevant threat signals, prior experience enacts as a top-down contextual factor biasing decisions toward expected options, and attention and individual traits modulate the relative contribution to social processing of relevant neural regions. Altogether, these findings shed light on the distinct cognitive mechanisms underlying the influence of different contextual factors during perceptual decisions in social settings.

## Résumé

Les décisions que nous prenons au quotidien nécessitent le traitement de plusieurs sources d'information, et dépendent par conséquent de nombreux éléments contextuels tels que les indices sociaux provenant d'un émetteur (par exemple, le regard, l'émotion et le geste), ainsi que le centre de l'attention, l'humeur et l'expérience passée de l'observateur. Le travail réalisé durant cette thèse (qui inclue trois expériences principales chez des sujets humains sains) a eu pour but de caractériser les mécanismes cognitifs et neuraux sous-tendant l'impact de ces éléments contextuels sur la prise de décision dans un environnement social.

La première étude d'électroencéphalographie (EEG), basée sur des analyses de potentiels évoqués et de reconstruction de source, a montré que les indices sociaux sont intégrés dans le cerveau dès 170 ms après l'apparition du stimulus, en fonction de leur pertinence pour l'observateur et indépendamment de l'attention qui leur était portée. Toutefois, l'attention augmente l'implication des régions cérébrales spécifiques au traitement de l'indice social sur lequel était portée l'attention. Grâce à des analyses comportementales et EEG basées sur la modélisation, la deuxième étude a démontré que les mécanismes qui sous-tendent l'intégration de différents indices sociaux extraits d'un visage induisent une augmentation de la sensibilité perceptive aux signaux de menaces pertinents pour l'observateur. Une telle augmentation de la sensibilité perceptive aux signaux de menace est apparue en parallèle au sein de régions temporales ventrales et motrices dorsales, dès 200 ms après l'apparition du stimulus. De plus, les analyses ont révélé que certaines caractéristiques propres à l'observateur, ici son niveau d'anxiété, ont influencé les régions dans lesquelles les signaux de menaces étaient représentés: chez les sujets peu anxieux, cette représentation s'est produite préférentiellement dans les régions temporales propres à l'encodage du visage, tandis qu'elle a impliqué les régions motrices chez les sujets plus anxieux, en accord avec la fonction adaptative de l'anxiété pour la préparation à l'action face à la menace. Enfin, en conjuguant modélisation comportementale et enregistrements pupillaires, la troisième étude a montré que les aprioris induits par l'expérimentateur ainsi que les aprioris propres à chaque individu concernant l'émotion exprimée par une identité particulière, influencent le biais décisionnel plutôt que la sensibilité à l'émotion. Cette influence a été mise en évidence par une augmentation des signaux pupillaires en réponse aux stimuli inattendus.

En conclusion, alors que les indices sociaux émis simultanément interagissent en renforçant le traitement ascendant 'bottom-up' des signaux pertinents de menace, l'expérience apriori agit comme un facteur contextuel descendant 'top-down' qui biaise les décisions vers les options attendues, tandis que l'attention et les caractéristiques de l'observateur modulent la contribution relative de régions pertinentes dans le traitement de stimuli sociaux. Les résultats pris dans leur ensemble mettent en lumière les différents mécanismes cognitifs qui sous-tendent l'influence de facteurs contextuels distincts lors de la prise de décision perceptive dans un environnement social.

## Abbreviations

EEG: Electroencephalography

MEG: Magnetoencephalography

fMRI: functional Magnetic Resonance Imagery

RDM: Random dot motion

P100: Positive evoked potential at 100 ms at occipital electrodes.

N170: Negative evoked potential at 170 ms at temporal electrodes.

P300: Positive starting at 300 ms at parietal electrodes.

CPP: Centro-parietal positivity in EEG signals.

OFA: Occipital Face Area

FG: Fusiform Gyrus

STS: Superior Temporal Sulcus

FFA: Fusiform Face Area

AMG: Amygdala

DLPFC: Dorso-Lateral Prefrontal Cortex

FEF: Frontal Eye Field

PM: Premotor

LC-NE: Locus Coeruleus-Norepinephrine system

## INTRODUCTION

Humans are animals that are distinguished by their very elaborate social interactions. Every day, we communicate with several individuals, and the ability to correctly read and interpret social cues from others is critical for proper social integration and survival. When such abilities are modified, individuals suffer from social dysfunctioning reflected in psychiatric disorders often associated with altered social behaviors such as clinical anxiety disorders, autism spectrum disorders and schizophrenia. Investigating the functioning of the human brain when dealing with social situations to establish models of healthy brain functioning is a key step toward a better understanding of what exactly is different in such disorders. The domain of social cognition has expanded through the last years, and different levels of social processing have been extensively explored, from low level social perception to more complex collective decisions and ‘theory of mind’ judgments.

The work conducted in my Ph.D. focused on the visual processing of social information and how perceptual decisions are formed based on this information. In social contexts, perceptual decisions represent the first step in social interactions and often occur in an implicit fashion. For example, if you perceive a social individual as dominant because of visual characteristics in his/her face, you might avoid that person without explicitly forming the decision that he/she looks dominant. However, understanding what actually led you to (implicitly or not) judge that person as dominant (i.e., which visual features) and therefore avoid him/her, is a crucial step to understanding the reasons/causes for your behavior. The ecological theory of perception extended to social perception (McArthur and Baron, 1983) indeed proposes that perception itself is adaptive – as *“the external world must provide information to guide biologically and socially functional behaviors”*. This theory emphasizes the necessity of studying social perception to comprehend social functioning. Therefore, understanding the neural mechanisms of how simple perceptual decisions are constructed based on social stimuli can bring important insights as to how individuals behave adaptively in social contexts. Importantly, social environments are rich with abundant contextual information, as social cues are never perceived in isolation. In addition, perceivers themselves represent a source of variability (in terms of their internal

state – e.g., anxious or not) as to how the social environment is perceived. Taking into account the scenery and the characteristics of the emitter and perceiver is therefore pivotal when investigating social perception.

In cognitive neuroscience, taking advantage of the powerful advantage of modeling when trying to understand the underlying computations of brain functioning is more and more put forward. However, computational modeling has been used more frequently outside of the field of social cognition, in domains such as decision-making and visual perception. As the computational approach provides a compelling tool to better understand the mechanisms of brain functioning, most of the work conducted in this Ph.D. took advantage of behavioral and neural modeling to characterize the way individuals behave and perceive social stimuli. For this reason, the literature that investigated perceptual decision-making in the non-social domain and the leading models that attempt to explain non-social decision-making will be reviewed, as well as the influence of context and priors on these decisions.

The introduction will be organized as follow: first, I will attempt to summarize the work that has been conducted in the field of social perception in the particular case of facial and emotion perception; second I will describe the findings of the studies that addressed contextual influence in emotion/face perception; third I will review the studies of perceptual decision-making in the non-social domain; fourth I will depict the influence of priors and contexts in the non-social domain and finally discuss the importance of bridging the gap between the social and non-social domains to address the question of perceptual decision-making on stimuli of emotional and social significance.

## **1-Social perception: the case of faces and facial expressions**

### **1.1- Why the face and what are facial expressions of emotions?**

The word 'social' derives from the latin 'socii' or 'socialis' which refers to partners, comrades, allies. Although there is no single definition of the word 'social', it clearly

refers to interactions with others. In the human species, when you think of other individuals, the first thing that comes to mind is most probably their face. We identify others by their faces, we decode their intentions and moods in their facial reactions. It is therefore not surprising that research on facial perception is very prominent in the domain of social perception. Faces probably represent the category of visual stimuli that is most salient biologically and socially in humans. Emotional signals are primarily expressed through faces. Charles Darwin has been a pioneer in the development of the idea proposing that facial expressions have an evolutionary role, serving as communication signals (Darwin, 1872). In his book 'The Expression of the Emotions in Man and Animal' the parallel is drawn between animals and humans both expressing emotions in similar ways and using them as signals for adaptive purposes. This idea was later developed by Jaak Panksepp who defended that emotions stem from the challenges species have to face in the environment driving behaviors for survival (Panksepp, 1998, 2005). Further corroborating this idea, emotions are suggested to represent brain circuits that contribute to behaviors related to survival functions (see review/perspective by LeDoux, 2012). Appraisal theories of emotions (Frijda, 1986; Scherer, 1984) suggest that motivations and intentions are inherent to the emotion emergence process. In this framework, it becomes important to take into account the behavioral relevance of an expressed emotion. Within evolutionary accounts of emotions, they have been defined by their role of possessing the adaptive function of adjusting relevant behaviors (Nesse, 1990):

*“The emotions are specialized modes of operation shaped by natural selection to adjust the physiological, psychological, and behavioral parameters of the organism in ways that increase its capacity and tendency to respond adaptively to the threats and opportunities characteristic of specific kinds of situations.”*

Why is this important for emotion perception? Just as objects in the environment provide action possibilities (Gibson, 1979), social stimuli induce action tendencies in observers, and thereby their perception is already influenced by their functional significance (McArthur and Baron, 1983). If emotions have the role of driving adaptive behaviors, do these behaviors motivate their perception? Enhanced processing of socially significant stimuli is intuitive within this framework and greatly depends on their emotional implication (Compton, 2003). For example, fear, anger,

snakes and fires can be considered of high emotional significance as they signal that the environment emanates threat, but also dominance and aggressiveness that would require adaptive action responses (McArthur and Baron, 1983). This implies that the perception of faces and facial expressions, among other socially meaningful stimuli, is already influenced by their behavioral significance.

Influential work of Ekman and colleagues (Ekman, 1972, Ekman and Friesen 1978) emphasizes the existence of innate muscle programs allowing facial expressions to be recognized universally. This research concentrated on prototypical expressions of what have been referred to since as the basic emotions: Anger, Fear, Disgust, Surprise, Happiness and Sadness (Figure 1).



Figure 1. From Ekman 1999. Pictures of the 7 basic emotions. From left to right, Anger, Fear, Disgust, Surprise, Happiness and Sadness.

Although using 'prototypical' emotional facial expressions is a matter of discussion as they are very rarely encountered in real life situations (Hess and Thibault, 2009), the idea of basic emotions governs today most of the research on emotion perception. Ekman further proposed that the expression of basic emotions is the product of combinations of a set of facial action units (Facial Action coding system FACS, Ekman and Friesen, 1978). For example, the combination of lowering eyebrows (AU 4), raising the upper eyelid (AU 5), tightening the eyelid (AU 7) and pressing and tightening the lips (AU23 and AU24) mediated by specific facial muscles (such as the corrugator for the brow lowerer) give rise to an angry expression. This suggests that a hardwired program of emotional expressions exists, and justifies the recognition of basic emotions across cultures (Ekman and Friesen, 1971). This view has been repeatedly challenged since (e.g., Ortony and Turner, 1990; Barrett et al., 2007). Nev-

ertheless, most of the existing faces databases include the 6 basic emotions and a neutral condition (NimStim, Tottenham et al., 2009, KDEF, Lundqvist et al. 1998) and some of them are based on the FACS (Radboud database, Langner et al., 2010, JACFEE/JACNEUF Matsumoto & Ekman 1988).

Even though the recognition of emotional expressions seems to be highly accurate with prototypical expressions, backward masking that is a way to reduce the detectability of visual stimuli in general through short time presentations reduces this ability in emotion recognition (Sweeny et al., 2013). Importantly, this diminished recognition is especially noticeable if the categorization concerns emotions that share characteristics, for example negative emotions such as fear and anger (Sweeny et al., 2013). Such observations question the basic component of visual features pushing individuals to recognize a face as emotional and which emotion particularly it is expressing, i.e. why is it more difficult to categorize anger and fear expressions and what are the visual features that allow their discrimination? It therefore seems important to integrate visual and dynamic components of facial expressions, and using finer methodological analyses can open a window into the mechanisms of face/emotion perception. Recently, to identify how emotion signals are transmitted over time, face stimuli were created by a random combination of action units and categorized by participants on the basis of the 6 basic emotions (Jack et al., 2014). Using Bayesian classifiers and information theory, findings revealed a first early stage of emotion processing where only four basic emotional expressions were perceptually represented as a function of their behavioral significance: happy, sad, fear/surprise and disgust/anger. This first stage is suggested to involve biologically rooted signals such as raising the upper lid in both fear and surprise to enhance sensory processing in possible approaching danger situations (fearful and surprise expressions can signal danger in the near environment). A second late stage of information processing is marked by diagnostic information for each emotion, dissociating the 4 categories into the socially ingrained 6 basic emotions. These findings are in line with the ecological theory of perception and appraisal theories as they show that early perceptual recognition of emotion is based on the biological relevance of transmitted signals. As later stages dissociated the basic emotions, these results also reconcile both accounts stated above by shedding light on the dynamics



of the signal transmission: 1) behavioral/biological relevance as inherent to the emotion recognition process, involving an early stage of processing, 2) hardwired programs for each emotion, involving a later stage of processing.

## 1.2- How are faces and facial expressions represented in the brain?

The question of how faces are perceived can be meticulously addressed through data-driven methods that map facial information into their subjective perception. Reverse correlation techniques indeed help characterizing the precise contribution of different visual information in the perception mechanism. These methods are successful in identifying the precise features of the face that help determining its belonging to a specific social or emotional category. The idea behind is to present visual inputs embedded in noise while participants make specific judgments that are used to reconstruct classification images showing which visual input drove the decisions. For example, using this method to visualize trustworthiness and dominance social traits, the social judgment was reported to involve visual information from the mouth, eye, eyebrow and hair (Dotsch and Todorov, 2012). Furthermore, when judging the gender of a face, information from eyes, eyebrows and mouth regions was necessary for accurate categorization, whereas judging if a face is happy or neutral only required information from the mouth (Gosselin and Schyns, 2001).

In this last study, the specific method used, referred to as the “Bubbles” technique, consists in adding noise to faces by creating random Gaussian windows revealing different portions of the face. Afterwards, participants’ responses mapped into correct and incorrect trials depending on the task are used to create classification images representing diagnostic information processed by observers’ brain to correctly perform the task. Different facial features were involved in the mapping of emotions when using Bubbles to categorize basic emotions: for example, while recognition of fear requires information from the eyes, happiness recognition requires information from the mouth (Smith et al., 2005). In Figure 2, the diagnostic information that differed significantly from an emotion to the other (uncorrelated across emotions) proves that the brain transmits information about different emotional ex-

pressions with minimal overlap possibly to optimize transmission as interpreted by the authors.

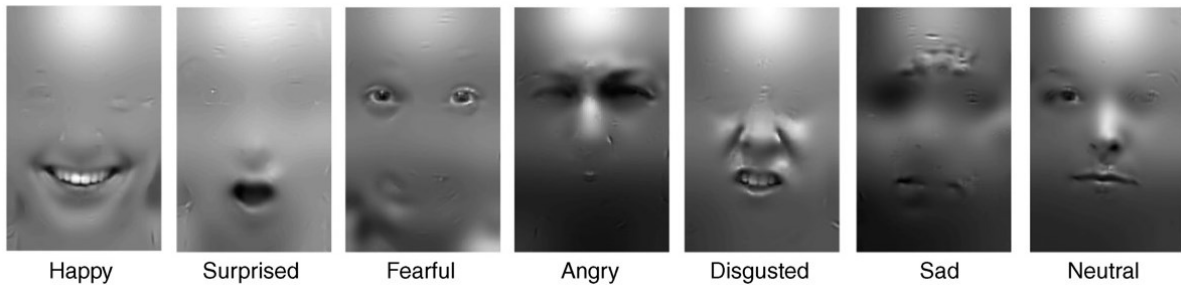


Figure 2. From Schyns et al 2009. Diagnostic information of 6 basic emotions and neutral face.

These diagnostic features trigger increased attention even during a gender categorization task as eyes were more attended in fearful faces and mouths were more attended in happy facial expressions (Scheller et al., 2012).

The modulation of spatial frequency conveyed by the faces reveals that different frequency bands contribute differentially to the recognition of emotions (Smith and Schyns, 2009). As it is adaptive to recognize information signaling threat even if it is ambiguous or far away, one can imagine that emotions signaling threat are transmitted through low spatial frequencies that only provide coarse information. Accordingly, low spatial frequencies in fearful stimuli increase the activity of emotion selective brain regions (Vuilleumier et al., 2003) and brain waves (Vlamings et al., 2009). However, reverse correlation techniques during emotion recognition (Smith and Schyns, 2009) reveal that happiness and surprise are conveyed by low spatial frequency bands rather than fear and anger (two specific signals of threat).

To conclude, perceptual features constituting facial expression contribute differently to the representation of emotions and understanding how they do so sheds light on the role of the transmission of these signals.

### 1.3- When are faces and facial expressions represented in the brain?

Time-resolved techniques such as electroencephalography (EEG) and magnetoencephalography (MEG) offer the possibility to determine when the visual information is encoded in the brain at the level of the millisecond. These techniques, combined with reverse correlation methods, provide a window into the algorithms of information processing in the brain (Schyns et al., 2009). Before going into the results of this powerful combination, let us first take a step back and look at primary EEG/MEG studies which investigated emotion and face processing.

Emotional expressions, particularly fearful expressions, trigger increased brain responses within 100 ms after stimulus onset on central electrodes (Pourtois et al., 2004), fronto-central electrodes (Eimer and Holmes, 2002), and in the amygdala (Luo et al., 2007; Hung et al., 2010; Liu and Ioannides, 2010). While this early involvement in the processing of fear can be interpreted as increased attentional resources toward threatening stimuli (Vuilleumier and Pourtois, 2007), this claim is contradicted by the fact that other emotions including joy (Dubal et al., 2011), anger, disgust, sadness and surprise (Batty and Taylor, 2003) show enhanced electrical positivity at occipital electrodes P100 when compared with neutral expressions. At 170 ms after stimulus onset, a negative occipito-temporal activity, the N170, responds specifically to perceived faces (George et al., 2005). This face selective N170 is also influenced by facial expressions as it is increased for fear as compared to other basic emotions (Batty and Taylor, 2003), peaks earlier for fearful expressions as compared to other basic expressions (Schyns et al., 2007) and varies as a function of the emotional intensity of angry, disgusted and fearful facial expressions (Sprengelmeyer and Jentzsch, 2006). After 300 ms, positive electrical activity at centro-parietal electrodes is modulated by negative (Schupp et al., 2004) and positive emotions as compared to neutral also for emotional scenes (Cuthbert et al., 2000; Liu et al., 2012).

While these studies show that emotional information modulates brain activity as soon as 100 ms after stimulus onset, they fail to characterize the specific contribution of visual information in the emotion recognition process over time. By combining electrical brain recordings with reverse correlation methods, the curvature and

the peak of the N170 were related to information processing showing that visual information is integrated through a pattern (Schyns et al., 2007). First, visual representation started by the eyes and went downward to the nose and mouth, thus explaining the earliest peak of the N170 to fearful expression (diagnostic information is in the eyes). Second, the peak of the N170 actually coincided with the time where the diagnostic information for each emotion was integrated (e.g., nose for disgust, mouth for happiness). Using MEG, this type of analyses is transferred from single electrode (occipito-temporal for the N170) to the source level during an emotion categorization as well as gender categorization (Smith et al., 2009). In addition to showing that maximal visual information is processed 170 ms after stimulus onset (corresponding to the N170), this study revealed different patterns of integration dependent on the task at this same latency. After 300 ms, only task specific information was represented in cortical activity, (for example, mouth information if categorizing happy faces), consistently with a decision related signal at this later latency.

While in the last two sections, we addressed how and when facial expressions are visually represented in the brain, the next section will deal with the neural sources of emotion processing.

#### 1.4- Where are faces and facial expressions represented in the brain?

Functional magnetic resonance imagery (fMRI) that has a great spatial resolution, allows for the isolation of brain regions involved in face processing. The most influential model in face processing is the model of the distributed human neural system for face perception (Haxby et al., 2000) suggesting that after structural encoding of visual features in occipital areas (occipital face area OFA), invariant aspects of faces such as identity are processed in the fusiform gyrus (FG) while variant aspect of faces such as gaze and emotion are processed in the posterior superior temporal sulcus (pSTS) (Figure 3). According to this model, it is only after the visual analysis in the 'core system' which comprises the OFA, FFA and STS, that facial information will be integrated with other dimensions such as attention and emotion in the 'extended system'.

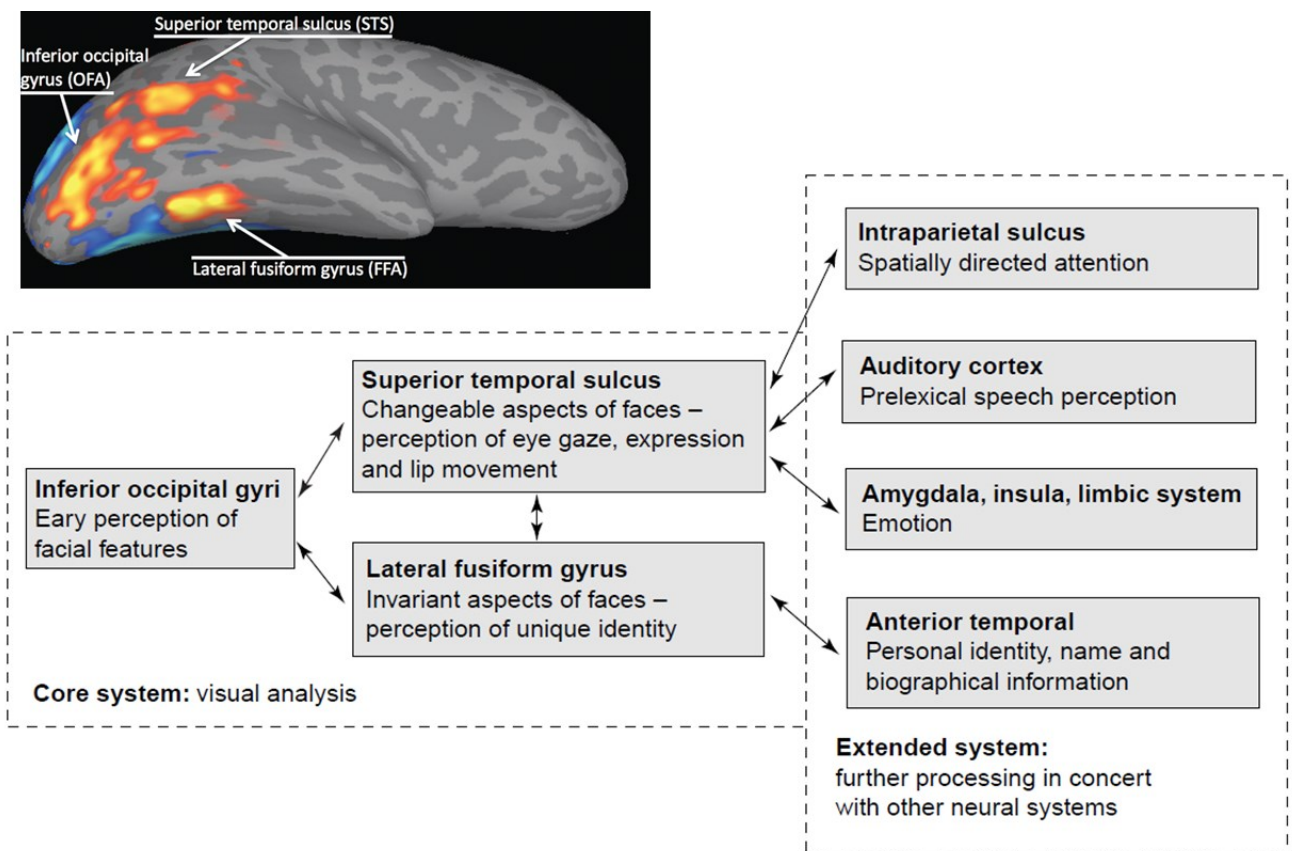


Figure 3. From Haxby et al. 2000. Model of the distributed human neural system for face perception.

As will be described next, evidence from the literature clearly highlights the role of FG and STS in face processing. Other core regions of face processing are also often reported, including the amygdala (AMG), the insula and the pulvinar for emotion processing, but also the orbitofrontal cortex, and the medial prefrontal cortex for inference processes (theory of mind). Here I will concentrate on the areas involved in the sensory processing of faces and facial expressions: the FG and STS. Then I will also describe the role of the AMG because of its involvement in threat and social relevance processing, and finally draw attention to the less acknowledged involvement of motor regions in emotion/threat processing.

## **The Fusiform Gyrus (FG)**

The FG responds preferentially to faces shown in different viewpoints when compared to houses, objects animals and other body parts (Kanwisher et al., 1997). The specific region in the FG responding to faces is referred to as the fusiform face area (FFA). Even when subliminally presented, the identification and detection of faces was associated with modulation in FFA while the identification of other subliminally presented objects such as houses, guitars and cars modulated other regions in the ventral temporal cortex (Grill-Spector et al., 2004). FG has been persistently shown to be the source of face-selective N170 mentioned earlier (Itier and Taylor, 2002; Rossion et al., 2003; Conty et al., 2007). In accordance with the core system model, the FG seems to be specialized in identity and familiarity processing (George et al., 1999; Rotshtein et al., 2005). The FG also responds to emotional expressions, preferentially to high spatial frequencies (Vuilleumier et al., 2003).

## **The superior temporal sulcus (STS)**

The STS responds to variant aspects of the face such as moving eyes and moving mouth (Puce et al., 1998). The STS seems to have a general role in the perception of movement (even from other body parts), and its activation to static face and body is interpreted as a role in the detection of implied motion and stimuli signaling actions (Allison et al., 2000). Accordingly to its role in the encoding of variable features in the faces, the STS also co-varies with the emotional intensity of a face (Said et al., 2010) and is involved in continuous, rather than categorical, representations of facial expressions (Harris et al., 2012). In a recent review about the field of social neuroscience, the pSTS is suggested to have an integrative role in social processing as it is involved in three key neural systems: social perception, action observation and theory of mind (Yang et al., 2015). In line with the integrative role of the STS occurring at multiple levels of social processing, single neuron recordings in monkeys show that the STS encodes different types of social cues including identity in a hierarchical manner (Morin et al., 2014). Therefore the thorough dissociation between invariant and variant features of the face is questionable, and the fact that STS and FFA inter-

act during face processing (Baseler et al., 2014) rather suggests that these regions act in concert to process socially relevant stimuli.

### **The amygdala (AMG)**

The AMG is undoubtedly a central structure in emotion perception, its role in threat detection being essentially put forward. Indeed, AMG activity to fearful expressions when compared with neutral or positively valenced stimuli is increased (e.g., Whalen et al., 2004). Moreover, a patient that has a bilateral AMG damage (patient SM) is unable to recognize static fearful expressions (Adolphs et al., 1994). Accordingly, in the monkey brain, AMG responds more strongly to threat expressions as compared to neutral and appeasing (Hoffman et al., 2007).

However, although threatening expressions are processed within 100 ms after stimulus onset in the AMG (Luo et al., 2007; Hung et al., 2010; Liu and Ioannides, 2010), there is no clear evidence that this processing occurs earlier than other emotions or uniquely involves a subcortical pathway (Palermo and Rhodes, 2007). Indeed, many studies point toward the involvement of the AMG in the perception of all emotions including positively valenced (Breiter et al., 1996; Winston et al., 2003).

The role of the AMG as a “relevance detector” comprising but not restricted to threat detection has been proposed (Sander et al., 2003) and could reconcile the different findings. In contrast to continuous representations of emotions in STS, the AMG categorically represents emotion expressions including happiness, disgust, fear and anger (Harris et al., 2012), in accordance with the role of the AMG in the detection of relevant information (here, which emotion is expressed, independently of its intensity). The sensory processing of emotions of both appease and threat in facial expression selective temporal regions was impaired in monkeys with damaged AMG, in contrast to intact processing of neutral faces (Hadj-Bouziane et al., 2012). Once more, this finding highlights the involvement of the AMG in relevance detection by showing that it influences the sensory processing of salient social signals. Finally, the impaired recognition of threat in patient SM seems to be restricted to static facial expressions of fear. Indeed, when patient SM with bilateral AMG damage is instruct-

ed to pay attention to the eyes to recognize fearful expressions (diagnostic information in fear recognition), she is able to do the task accurately (Adolphs et al., 2005). This shows that it is not the recognition of fear emotion *per se* that is impaired but rather the allocation of attention to the relevant feature of the face, consistently with the role of the AMG in orienting to salient information. Moreover, patient S.M is perfectly able to recognize fearful expressions from static and dynamic body expressions (Atkinson et al., 2007), highlighting the possible involvement of somatosensory regions bypassing the amygdala in threat processing.

### **Motor-related regions**

Although the implication of motor related regions is not yet firmly rooted in emotion processing, its implication in threat perception in tandem with the AMG is suggested (Grèzes et al., 2015). This stems from experimental evidence showing that premotor regions are co-activated with the AMG during the perception of threatening emotions of anger and fear conveyed by body expressions (de Gelder et al., 2004; Grèzes et al., 2007; Pichon et al., 2009, 2012). The involvement of motor regions in emotion perception extends to natural emotional scenes and facial expressions. Indeed, disrupting the motor cortex using TMS impaired emotion categorization of anger and fear expressions specifically, and not happy expressions (Balconi and Bortolotti, 2012, 2013). Moreover, the perception of natural emotional scenes engages the motor cortex at very early stages (150 ms after stimulus onset) when the emotional valence of the scene is negative (Borgomaneri et al., 2014). Finally, evidence for a structural connection between amygdala and motor regions (Grèzes et al., 2014) supports the existence of a strong link between emotion and motor systems both representing threat signals to prepare individuals for action in threatening contexts (Ohman and Mineka, 2001; Frijda, 2009).

### **Networks of regions rather than specialized regions in emotion representation?**



Recent meta-analyses (Lindquist et al., 2012; Wager et al., 2015) argue for a constructionist approach entailing the involvement of a general brain network rather than discrete regions in the representation of emotions. They support the activation of general networks for different categories of emotions. Importantly and together with other meta-analyses and reviews on emotion processing (e.g Fusar-Poli et al., 2009; George, 2013), these integrative studies pinpoint the numerous areas of the brain involved in emotion perception, sometimes shared between different emotions or emotion modalities (faces/scenes/bodies), sometimes distinct, and also highly influenced by experimental settings and task demands. It therefore becomes crucial to not only assess where the emotion is represented in the brain but rather understand the mechanisms of such representation, in terms of networks such as suggested here, or detailed information processing such as described in the previous 2 sections.

## **Chapter 1 conclusion**

Research on the perception of face and facial expressions, stimuli of high biological and social significance, has been extensively investigated in the last years showing the involvement of a broad brain network in facial processing starting from 100 ms after stimulus onset. Although the idea of facial expressions perceived in categories of basic emotions dominates the field of affective neuroscience, recent evidence points toward the perception of emotions as a function of their significance in fewer categories. Importantly, social significance is also provided by the abundant contextual factors present in social settings, therefore rendering the consideration of these factors essential when studying social perception.

## **2-Contextual effects in social perception**

While many of the studies cited in the previous section investigated the processing of emotional faces without manipulating the context, it is crucial to take into account the different contextual factors that can interact with emotion perception. It is obvious

that facial expressions encountered in our everyday life do not appear on their own, but are rather always embedded in a certain setting. Moreover, the perceiver who has to judge a certain emotion has specific characteristics and past experiences that might influence the way he/she will judge and react to facial expressions. The nature of the contextual information therefore varies from signals provided by the environment to signals coming from the sender of the emotional expression, to signals characteristic of the observer. In this section I will review some of the papers that have manipulated such contingent elements to emotions. First, I will present the contextual elements that are characteristic of the observer/perceiver such as his/her traits or origins, second the contextual elements related to the environment, and finally the contextual modulators belonging to the emitter such as features and body.

### 2.1- Contextual elements characteristic of the observer

When we think of an observer (represented by the participants in the studies we conduct), we can think of all the possible states we can be in depending on the day, on how much we slept, on our general character, on our knowledge of the word, on our past experience etc. It seems more than obvious that the person we are and our personal experiences - shape our perception of the world. This is particularly true for social perception as successful interactions with other people sometimes require social efforts that might be altered as a function of personal factors. The first example that comes to mind when mentioning personal difference is that of psychiatric disorders as social skills alteration is one the most prominent deficit in these disorders. As this literature is very broad (see review Fett et al., 2015) and involves a large number psychiatric disorders (such as autism, anxiety, schizophrenia, compulsive obsessive disorder, depression...), I will only give a few examples demonstrating the strong impact of individual differences on social perception. Psychiatric dysfunctions are known to lie on a continuum that explains individual differences even in healthy individuals (Kosslyn et al., 2002). A useful way to investigate these individual differences is provided by personality scales that have long been used in experimental psychology and neuroscience and that clearly demonstrate differences in social perception within healthy individuals.

A vast number of studies show that highly anxious individuals (on personality scales, the most common for anxiety being the State-Trait Anxiety Inventory STAI (Spielberger et al. 1983)) and clinically anxious patients are strongly biased toward the interpretation of events and stimuli as threatening (for meta-analytic study see Bar-Haim et al., 2007). In addition, this bias is reflected in the brain by an increased activation of the AMG in highly anxious people when they are in a threatening context (e.g Bishop et al., 2004) (Figure 4). AMG activation to fearful vs neutral faces correlated with anxiety scores. Further, attention modulated less strongly this activation in high anxious individuals, showing that they encode threatening emotions even if their attention is not engaged in the processing of facial expressions.

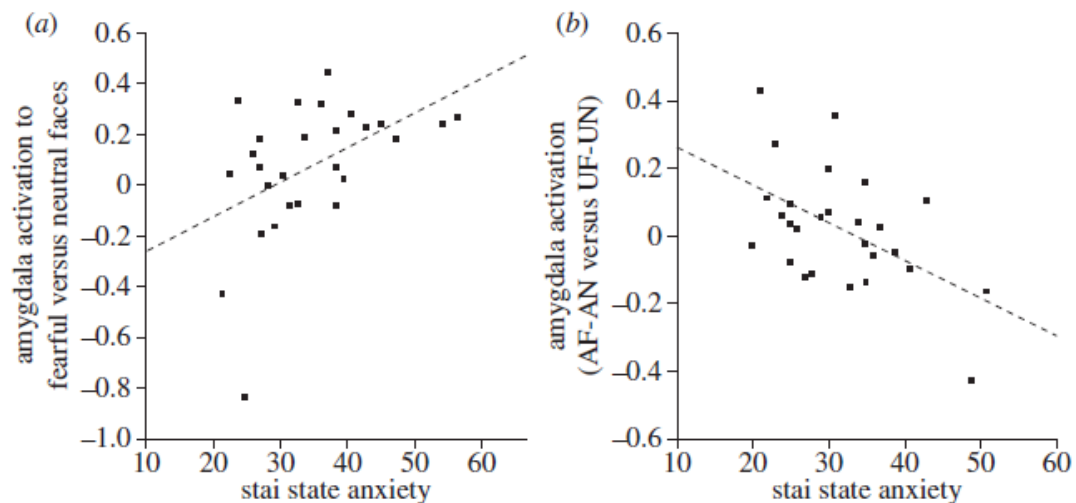


Figure 4. From Bishop et al. 2004. Influence of state anxiety on the perception of threatening emotions. a) AMG activation increase to fearful vs neutral with anxiety scores b) Attentional modulation decrease with anxiety scores in the perception of threatening emotions. AF=Attended Fear, AN= Attended Neutral, UF=Unattended Fear, UN= Unattended Neutral.

Another example of AMG modulation by individual trait concerns extroversion and the perception of happy faces: the more individuals were extroverted and therefore susceptible to positive mood, the greater was AMG activity in response to happy faces, specifically in contrast to neutral faces (Canli and Amin, 2002). These findings therefore suggest that the ways emotional expressions are encoded in the brain are susceptible to individual traits.

Individual traits linked to psychiatric disorders aside, another example of individual elements influencing social perception is that of race: the perception of faces from different races is influenced by the perceiver's own race as reflected in an increased AMG to dissimilar race and is related to the implicit measure of race bias (Phelps et al., 2000). Moreover, N170 EEG response to faces from a different race as compared to similar race was increased (Ofan et al., 2011) and finally activity in FG only predicted the race of faces of dissimilar race in people who showed an important implicit race bias (Brosch et al., 2013).

More generally, the belonging to a group can greatly influence the way out-group and in-group individuals are perceived. Group belonging can be manipulated experimentally quite easily, by using minimal group paradigms (Tajfel, 1970) for example, where people are assigned to groups on the basis of trivial criteria. Belonging to a group can change how the participant perceives in-group as compared to out-group members relatively automatically as it is independent of the task. This is reflected in an increased activity in face selective regions such as the AMG and the FG to in-group members (Van Bavel et al., 2008, 2011). Behaviorally, in-group members are perceived as more trustworthy, caring and intelligent as compared to out-group members (e.g Ratner et al., 2014).

Finally, cultural differences influence the perception of facial expressions as Eastern and Western cultures differ in the way they represent basic emotions (Jack et al., 2012a, 2012b). For example, contrarily to Westerners, Easterners do not represent the 6 basic emotions with their distinct action unit set and they seem to be particularly attentive to eye region in the decoding of expressive information.

To conclude on this part, personality traits have a role in changing the way social stimuli are perceived, mainly by biasing interpretations toward what is considered relevant to a certain individual (i.e. threat for an anxious person), and enhancing the processing of such relevant social inputs. Although internal elements of the perceiver were considered in a separate section, all the contextual information that will be discussed in forthcoming sections cannot be dissociated from their significance to the observer.

## 2.2- Contextual signals related to the environment

Social environments are complex settings that include a variety of elements which can influence our perception of people. It is actually well acknowledged that any visual feature is influenced by the surrounding context (Bar, 2004), as there is no ecological situation where an object is perceived on its own. It is also important to keep in mind that the observer might have previous experience in a given environment, such as the people encountered just before, or the elements observed preceding a specific categorization, or the specific aims driving his/her behavior in this setting.

A clear example of how contextual settings change the categorization of emotional expressions comes from studies that manipulated the congruency between a facial expression and the concomitant scene. They show that when emotional faces (disgust, happiness, fear) are embedded in emotionally congruent scenes (Figure 5), their emotion categorization is facilitated even under high load (Righart and de Gelder, 2008a).



Figure 5. From Righart & DeGelder 2008. Example of stimuli: left, disgusted face in a congruent disgusting environment. Right, disgusted face in an incongruent pleasant environment.

This is also reflected in the brain where scene information and emotions seem to be combined as early as 170 ms after stimulus onset, as the evoked potential N170 was larger for fearful faces in a fearful context both when the task was explicit and implicit to the emotion (Righart and de Gelder, 2006, 2008b). The integration between emotional scene and emotion information involved the fusiform gyrus (Van den Stock et al., 2014). While the last mentioned studies manipulated the congruency of concurrent scene and face, the effect of positive and negative context induced by short videos on the rating of subsequent images of facial expressions was also investigated (Mobbs et al., 2006).

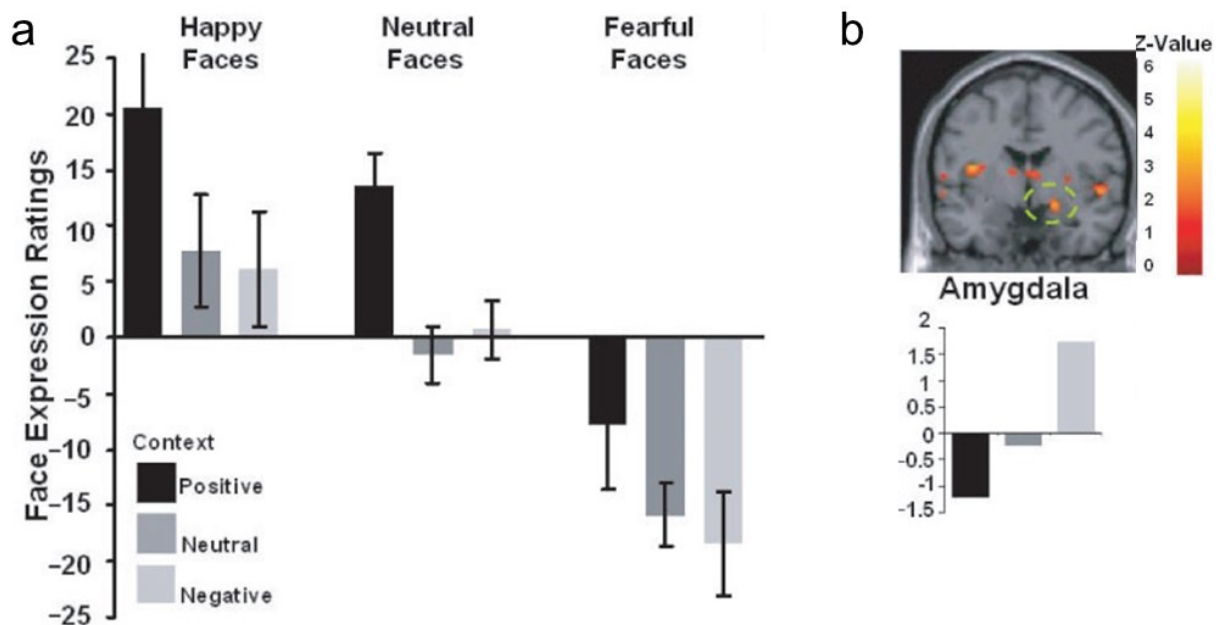


Figure 6 From Mobbs et al. 2006. Influence of context and neural correlates of the interaction. a) ratings of happy, neutral and fearful faces in positive, neutral and negative context. b) Amygdala activity reflecting the interaction of context with facial expression (Negative context Fear – Neutral) – (Neutral context Fear – Neutral), and betas for fearful faces in positive, negative and neutral contexts.

In line with the previous studies, congruency effects were evidenced in biased ratings: positive ratings for happy faces increased in positive contexts, and negative ratings for fearful faces increased in negative contexts (Figure 6). Interactions in the AMG, the FG and temporal pole, reflected these effects, with increased activity to fearful vs neutral preceded by negative context as compared to neutral context (interaction in the AMG shown in Figure 6b).

Aside from natural contexts, social settings often involve more than one person, and therefore other faces present in the setting can also influence the appraisal of emotional expressions. A study that investigated that issue showed that the presence of a fearful face in the periphery enhances the recognition of anger, however only an angry face looking at a fearful face increases the recognition of a fearful face (Mumenthaler and Sander, 2012). As fearful faces signal potential danger in the surrounding, they can increase the perception of other elements in the scene (Phelps et al., 2006). However in the case of anger, this emotion signals a threat only if it is looking at its target, which explains why the fearful face is better recognized only if looked at. The distinction between these two emotions and their association with gaze direction will further be discussed in forthcoming sections.

Just like faces concurrent with a to-be-judged face can change the categorization of emotional expressions, previously encountered faces or judgments on one type of emotion before another can influence decisions about facial expressions. It has been demonstrated since early studies that judging an emotion before judging another strongly shifts decision about the second (Russell and Fehr, 1987). For example, a neutral face will be perceived as sadder if the decision about that neutral face was preceded by the judgement of a happy face. The idea defended by the authors is that the perception of one emotion will displace your perception of the next one in the opposite direction. A more recent MEG study also highlights the effect of preceding facial expressions by showing that participants were more likely to categorize an ambiguous face as “fearful” if it is preceded by an ambiguous rather than a prototypical fearful face (Hsu, 2015). This effect was coupled with increased suppression of alpha range in frontal electrodes 120 ms before the onset of an ambiguous fearful face when the preceding context consisted of ambiguous fearful faces.

Also in line with these findings, however during the judgment of faces but not facial expressions, the perception of facial identities seems to be highly modified by the preceding encountered identities (Apps and Tsakiris, 2013; Liberman et al., 2014). In an fMRI study (Apps and Tsakiris, 2013), contextual familiarity to faces measured as a direct effect of identities seen on previous trials predicted participants

choice and this effect was mediated by the STS that tracks moment to moment fluctuations. View-independent familiarity that updated the familiarity of each identity (how many times the same identity was seen throughout the experiment) also influenced decision and co-varied with fusiform face area activity, consistent with the role of the FG in detecting static information in the face such as identity. In another study (Lieberman et al., 2014), again with neutral faces only, participants were asked to adjust the identity of a face to a target stimulus. Results showed that the perceptual error is highly influenced by the preceding trial, biasing responses toward the identity seen in the previous trial. This effect survives changes in the viewpoint of the identity but also task modality (if forced choice and not adjustment, but also if no required response at the previous trial) suggesting that the perception of faces is serially dependent. The authors argue that this serially dependent perception of visual objects might serve for visual stability decreasing the neural computations needed to identify objects over short periods of time. This might be useful to our daily interactions with other individuals that constantly move by changing their facial expressions, head orientations, and body movements. The described results of the last two studies converge by showing that viewing independent identities encountered over short periods of time highly predicts the perception of subsequent faces, and therefore that face perception is remarkably influenced by preceding context.

Information about others can also constitute a contextual element during their perception/judgment. For example, faces paired with the sound “stupid” elicit increased skin conductance (Iidaka et al., 2010). Faces associated with an insult toward an observer trigger increased activity in the AMG (Davis et al., 2010). Similarly, when specific behaviors were attributed to identities, those assigned a negative behavior triggered increased activity in STS (Todorov et al., 2007). Furthermore, within negative behaviors, disgusting versus aggressive behaviors triggered increased insula activity independently of explicit memory.

Together, the findings provided in this section demonstrate that signals present in the environment are part of facial perception even though not relevant to the task a participant is engaged in. As our actions and intents in real life situations are



most often interrupted or unwarily altered by all sorts of external stimuli such as backgrounds, people and events, these results describe how they possibly do so and shed light on the neural basis of these indisputable influences.

### 2.3- Contextual signals from the emitter

Let's now consider the contextual elements intertwined with the central element to be perceived, the emitter. These elements can be *within-face features* such as dynamics in the face or gaze direction, or *within-sender features* such as body postures and voice (Wieser and Brosch, 2012). When we perceive people every day, we probably perceive them holistically taking into account their movements, their voice etc rather than focusing on specific features. Therefore, all the visual and non-visual signals from an emitter should be combined to form a social signal that has a meaning to the observer (Figure 7).

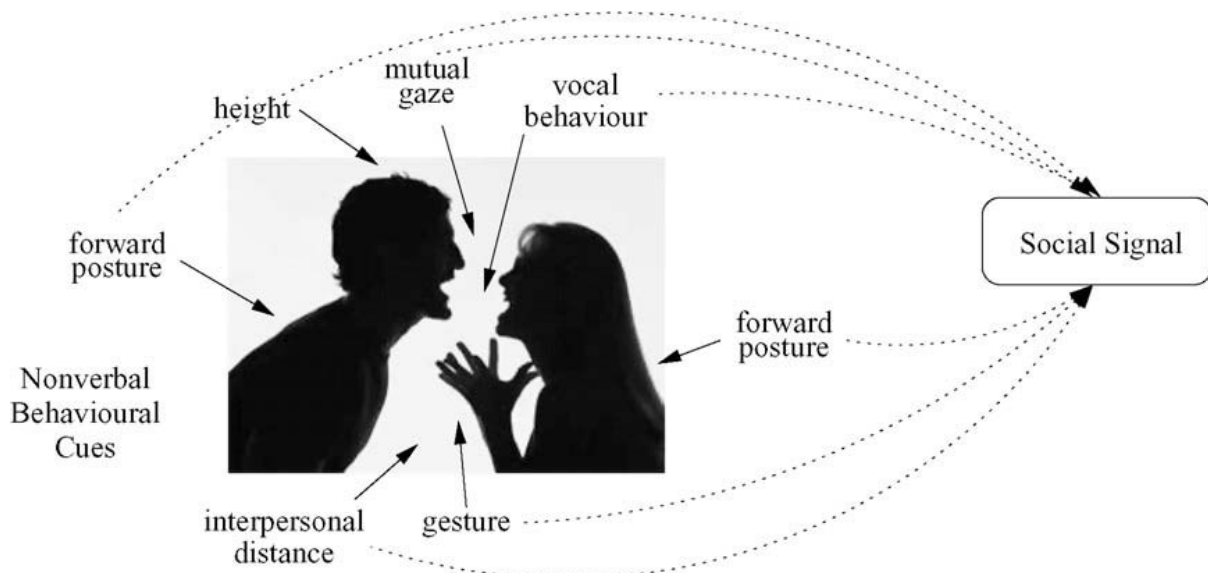


Figure 7. From Vinciarelli, Pantic & Bourlard, 2009. Social cues such as posture, voice and gaze are combined to form a social signal.

#### ***Within face***

Perceptual determinants within the face can highly influence perceptual or social judgments about faces (Todorov et al., 2015). Many examples can illustrate this influence such as baby-faced looking individuals judged as weak and submissive (Zebrowitz and Montepare, 2008), masculine faces and angry faces perceived as more dominant and angry faces perceived as more aggressive. Reverse correlation methods can be helpful to visualize how people expect a face in terms of a certain characteristic. When using this technique to visualize the internal representation of trustworthy and dominant individuals (Dotsch and Todorov, 2012), mouth, eye, eye-brows and hair regions mostly influenced decisions. The resulting classification images clearly show how people have a bias in representing a certain characteristic which implies that certain traits in the face will influence their judgment of people as dominant or trustworthy (Figure 8).

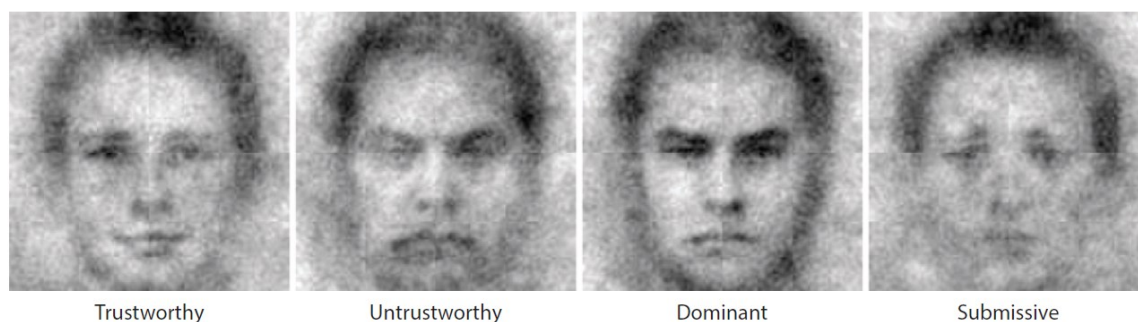


Figure 8. From Todorov et al. 2015. Classification images of trustworthy and dominant individuals.

Gaze is the most explored cue when faces are processed (Yarbus, 1967; Pelphrey et al., 2002; Spezio et al., 2007) and therefore undoubtedly processed and very likely influencing judgments about faces. This part about gaze as a contextual cue will be elaborated as most of the experiments in this Ph.D. included gaze as a contextual factor and addressed its influence on emotion categorization.

When faces are neutral, direct gaze seems to be advantageously processed, as even 4 months old infants show an increased processing of faces with direct gaze as compared to averted gaze 290 ms after face onset (Farroni et al., 2002). In adults, faces with a direct gaze trigger enhanced N170 as compared to averted gaze, and this effect involves the medial prefrontal cortex, the orbito-frontal cortex

and the STS (Conty et al., 2007). In addition, under uncertainty, individuals expect a direct gaze as they are biased toward the interpretation of an ambiguous gaze direction as directed toward them (Mareschal et al., 2013). When perceiving a face, as long as it is neutral, direct gaze seems to be most relevant (Senju and Hasegawa, 2005). This can be explained by the signal provided by a direct gaze, a signal of a potential social interaction. Infants who already process direct gaze advantageously might do that as they know they will learn from their parents or other adults who interact with them by looking at them in the eye. This strong sensitivity to direct gaze has been suggested to come from the fact that it is an ostensive signal, used as communicative tool since birth (Senju and Csibra, 2008).

When faces express emotions, gaze direction can impact the perception of the emotion. A happy person with a direct gaze toward an observer probably wishes engaging an interaction with the observer. Therefore in the case of happy expressions, direct gaze should be processed more easily. The shared-signal theory (Adams and Kleck, 2005) posits that emotions sharing an approach tendency similarly to direct gaze should be better processed when accompanied by a direct gaze. This involves emotion of happiness, just mentioned, but also anger that reflects the motivation of the expresser to engage with the perceiver for a fight for example. Contrarily, emotions such as fear and sadness that rather share with averted gaze avoidance tendencies, should be better perceived with an averted gaze. This theory is corroborated by increased emotional ratings and faster reaction times for angry and happy expressions displaying a direct gaze and for fearful and sad expressions with an averted gaze, but also increased ratings on anger and happiness scales for direct gaze neutral stimuli and on sadness and fear scales for averted gaze neutral faces (Adams and Kleck, 2003, 2005). Participants also rated feeling more repulsed and more anxious when they saw an angry face looking at them as compared to an angry person looking away, more repulsed for averted fear as compared to direct, and more anxious for a an averted gaze sad face as compared to direct gaze (Hess et al., 2007). The latter results suggest that the effects are linked to the feelings of the observer. When angry, fearful and happy expressions were manipulated on schematic faces controlling for action units, ratings of anger and fear and the correspond-

ing emotional intensity were again increased respectively for direct and averted gaze while no effect of gaze emerged for happiness (Sander et al., 2007) (Figure 9).

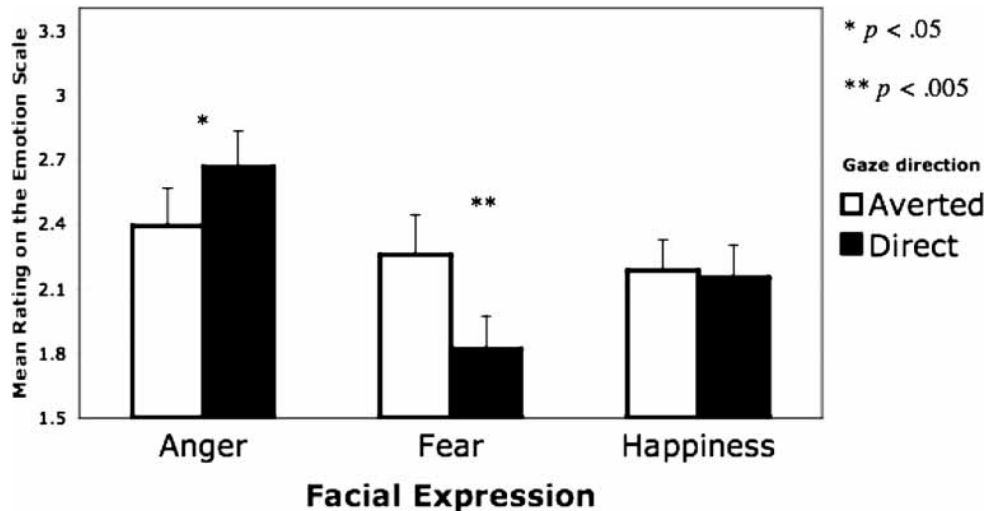


Figure 9. From Sander et al. 2007. Ratings of Anger and Fear are influenced by gaze direction.

The authors argue that this effect stems from appraisal theories of emotion suggesting that direct anger and averted fear both signal a threat to the observer: while an angry person gazing at an observer makes the observer the target of the threat, a fearful face gazing at the environment away from the observer might signal a nearby threat in the environment. Therefore, observers perceive more anger in direct anger and more fear in averted fear because their evaluation of the displays as being relevant to them (here threatening).

Reciprocally, when judging gaze direction, anger and happiness trigger in observers the tendency to judge stimuli as looking at them (Lobmaier et al., 2008; Adams and Franklin, 2009; Ewbank et al., 2009) The interaction between emotions of fear and anger and gaze direction was also evidenced when judging gaze, with faster reaction times for direct anger and averted fear (Adams and Franklin, 2009).

Although the results are not always very clear, and for this purpose summarized in table 1, the processing of direct anger and averted fear seems to be facilitated quite consistently, and this appears to be the case since childhood (Akechi et al., 2010; Rhodes et al., 2012).

The neural basis of the interaction between emotion and gaze cues has also been investigated mostly during implicit tasks, drawing attention away from the social cues of interest. The activity in the AMG consistently varies as a function of this interaction during gender categorization (Adams et al., 2003; Sato et al., 2004, 2010), passive viewing (Hadjikhani et al., 2008) and emotional intensity rating (N'Diaye et al., 2009). Its activity varies as a function of stimulus duration, responding more strongly to signals of threat (direct anger and averted fear) when faces are presented for a short time (300 ms), and more strongly to threat ambiguity (averted anger and direct fear) if stimuli are presented for 1 second (Adams et al., 2012). Other brain areas such as FG, medial prefrontal cortex (mPFC) and hypothalamus have been reported in integrating gaze and emotion (see table 1). Moreover, when the relevance related to threat signals was parametrically manipulated by adding a third social cue to gaze and angry emotion, pointing gesture, the premotor cortex mediated the integration of social cues as a function of their relevance to the observer (Conty et al., 2012). This integration was revealed within 200 ms after stimulus onset. The involvement of the premotor at 200 ms implied an increased motor preparation to socially relevant stimuli. Only one other study in children had shown a similar early signature of social cues interaction (Akechi et al., 2010) as the N170 was enhanced for averted fear and direct anger as compared to direct fear and averted anger. In contrast, another EEG study found a later interaction between emotion and gaze, around 300-400 ms after stimulus onset, and argued for a first stage of independent processing followed by an interaction (Klucharev and Sams, 2004). Interestingly, even 7 months infants showed an increased negativity (Nc) to direct as compared to averted anger (Hoehl and Striano, 2008) which suggests that the interactions between gaze and emotion might arise very early on, because of their important social significance.

Study	Participants	Stimuli	Task	Face duration	Behavioral effect	Neural basis
Adams&Kleck 2003	Healthy adults	Study 1: A/F, prototypical and ambiguous; study 2: H/S, prototypical. Both studies: Gaze D/Av	1) Categorization A/F; 2) categorization H/S	until response	1) Emotion*gaze interaction on RT: faster for prototypical AD and FAv- More F labels than A for Av on ambiguous expressions; 2) interaction on RT: faster for HD and SAV	-
Adams et al. 2005	Healthy adults	Photographs: Study 1: Neutral faces; Study 2: A/F, prototypical and ambiguous; Study 3: A/F/S/H; All studies: Gaze D/Av	1) Rating of neutral faces on scales of A/F/S/H ; 2) choose emotion label between the 4, then intensity rating; 3) same as study 1-2 faces presented one with D and one with Av gaze.	until response	1) Emotion*gaze interaction: H+A higher rating for D, F+S for Av; 2) more labels A and higher intensity rating of A for D /same for for FAv; 3) interaction, ratings higher for H+A in D gaze, and for F+S in Av.	-
Sander et al. 2007	Healthy adults	Dynamic schematic faces controlled for AU: H/F/A, Gaze D/Av	Rating on 7 scales: general intensity + 6 basic emotions	1.1 s	Emotion scales: more A for D, more F for Av - Intensity scale A more intense with D, F with Av	-
Hess&Adams 2007	Healthy adults	Emotions H/S/F/A, Gaze D/Av	1) Rating on 6 basic emotion scales; 2) rating of emotional reactions: cheerful irritated anxious miserable surprised and repulsed	until response	1) D gaze drives better recognition, opposite pattern for F, 2) anxiousness and feeling of repulsion higher for AD vs AAv, for F only feeling repulsed higher, anxious feeling increased for Sav	-
Adams et al. 2003	Healthy adults	Emotions A/F/H, Gaze D/Av	Gender categorization	2 s	-	Aav + FD > AD + FAv : Left AMG
Sato et al 2004	Healthy adults	A/N, Head and gaze Av or D	Gender categorization during fMRI; post test: intensity of experienced emotion and recognized emotion	0.7 s	Interaction face direction*emotion for experienced emotion only, anger direct elicits more negative feelings, no effect on gender task	AD>Aav : left amygdala + correlation with experienced emotion
Hadjikani et al. 2008	Healthy adults	F/Neutral, Gaze D/Av	Passive viewing	0.3 s	-	Fav>FD: STS, IPS , FG, OFA , AMG, HYP, PaI, PM, SPL

N'Diaye et al. 2009	Healthy adults	A/F/H mild and intense, Gaze D/Av	Rating of emotional intensity	Gaze shift: 0.1 s/ dynamic emotion 0.6s	Emotion*gaze interaction: AD judged more intense than AAv , opposite for fear	AD + FAv > Aav + FD: AMG, FG, mPFC only for mild intensities
Sato et al. 2010	Healthy adults	H/A Dynamic and static, Gaze D/Av	Gender categorisation task during fMRI; post test: valence and arousal rating	1.5 s	Greater arousal rating for direct gaze under dynamic and not static, no effect on gender	AD + HD > Aav + Hav: left AMG only in dynamic stimuli
Adams et al. 2011	Healthy adults	Fear with D/Av gaze	Passive viewing during fMRI	0.3 or 1 sec	–	Fav>FD left AMG when 0.3 s ; FD>Fav left AMG when 1 sec
Klucharev & Sams 2004	Healthy adults	H/A, gaze D/Av	Detect repetition of gaze during EEG	0.3 s	–	AD+Hav > Aav + HD: interaction at 270-450 ms in parietal electrodes
Rigato et al. 2010	Healthy adults and infants	Neutral/ F/H, gaze D/Av	Passive viewing of cartoons, with sounds to keep attention- during EEG	1 s	–	Adults: HD+FAv>HAv+FD interaction on occ channels around 250 ms; infants: HD>HA>A frontocentral Nc
Akechi et al. 2010	Typically developing and autistic children (9-17 years)	A/F, gaze D/Av	Categorization anger/ fear during EEG	1.2 s	emo gaze interaction for typical only with better performance anger D and fear A	AD + FAv > Aav + FD: N170 for typical
Hoehl & Striano 2008	7 months infants	A/F, gaze D/Av	Passive viewing	1 s	–	AD>Aav: Nc - F>A : N290 P400: A>F

Cristinzio et al. 2010	patients with amygdala damage and healthy controls	A/F/H, gaze D/Av	7 scales intensity than the 6 emotions	2 s	Emo* gaze interaction for controls only: Emotion scales: more A for D, more F for Av - Intensity scale A more intense with D, F with Av	Right AMG damage impairs gaze emotion interactions on behavior
Lobmaier et al. 2008	Healthy adults	A/F/H, gaze D/Av	Gaze direction judgement: Face looking at you or not?	1 s	More answers "looking at you" for H and A	–
Ewbank et al. 2009	Healthy adults	A/F, gaze D/Av	Gaze judgement: face looking left right or directly at them	0.2 s	Tendency to judge A expressions as looking at the observer	–
Adams&Franklin 2009	Healthy adults	A/F, gaze D/Av	Gaze judgment: direct or averted gaze	until response	Interaction emo*gaze on accuracy and RT, quicker FAv/ AD- correlation with the main effect of gaze (faster for Av than D)	–
Perett et al 2011	Healthy adults	H/ F/ A, gaze D/Av, eyes and no eyes condition	Gaze judgement: Attending to you or not ?	0.15 s	More answers attending to you for H expressions, even in the no eyes condition	–
Calder et al. 2012	Healthy adults and 8 years old	A/F, gaze D/Av	Gaze judgement: face looking left right or directly at them	0.4 s	Tendency to judge A expressions as looking at both adults and children observers	–
Conty et al. 2012	Healthy adults	A, gaze and head D/Av, Pointing gesture or not	Gaze judgement: is the person addressing you or another	1.3 s	Only behavioral pretest : more self involvement for anger direct pointing	Interaction emotion*gaze*gesture at 200 ms in PM, AD pointing gesture > all other conditions

Table 1: Main findings of studies that investigated emotion by gaze interactions. N/A:Non Applicable. A:Anger, F:Fear, S:Sad H:Happy, D:Direct, Av:Averted. AMG=Amygdala, STS: Superior temporal sulcus, IPS: Inferior temporal sulcus , FG: Fusiform gyrus, OFA: occipital face area , HYP: hypothalamus, Pal: Pallidum, PM: Premotor, SPL: Superior parietal lobule. occ:occipital, Nc: EEG component fronto-central negativity.



## ***Within sender***

Elements *within sender* can influence the perception of an individual such as the body posture. The categorization of facial expressions of anger, fear and disgust is facilitated when they are matched congruently with body posture (such as a fearful face with a withdrawal-like posture) (Meeren et al., 2005; Aviezer et al., 2011). This is reflected in brain activity as soon as 100 ms after stimulus onset, as the evoked potential P100 in occipital electrodes was larger for incongruent as compared to congruent displays (Meeren et al., 2005) (Figure 10).

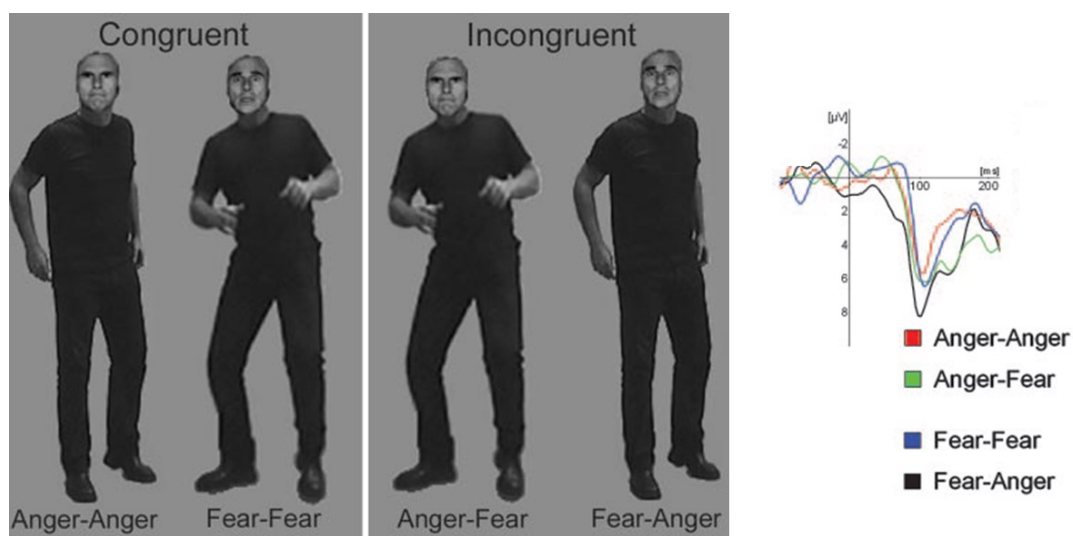


Figure 10. From Meeren et al 2005. Body emotion congruency. Left, example of stimuli. Right, increased P100 for incongruent stimuli.

Body congruency with voice prosody also influences the recognition of emotion in voice by biasing judgments toward concurrently presented body expression (Van den Stock et al., 2007). Voice tonality interacts with facial expressions by biasing interpretation of either facial expression or voice toward task-irrelevant simultaneous voice and facial expression respectively (De Gelder and Vroomen, 2000). A neural signature of this interaction is found in the AMG and FG where activity is increased for congruent fearful faces and voice (Dolan et al., 2001). In the same vein, laughter increases the rated intensity of happy facial expressions (Sherman et al., 2012) while screams increase the rating of fearful stimuli (Müller et al., 2011).

To conclude, concurrent signals from the emitter coming from different modalities interact in the way he/she is processed independently from attentional demands. All the cited studies experimentally prove that when a person is perceived, all the different signals emitted by this person will be combined into a significant social signal.

## **Chapter 2 conclusion**

Signals from the emitter, from the environment and elements encountered in the environment over short periods of time strikingly change the way social stimuli such as faces and facial expression are perceived and categorized. However, there is no unequivocal picture of how all these elements influence perception. Face selective areas involved in the perception of decontextualized faces (STS, FG) are often reported to be involved in these influences. Together all these studies provide evidence to the idea that facial expressions result from underlying appraisals of the social context in which the emotion is expressed (Hess and Thibault, 2009).

In the last two chapters, we saw that the perception of emotional expressions and the influence of contextual factors on emotion processing have been extensively investigated. However, the mechanisms underlying the processing of emotional expressions and the influence of context on their processing remain poorly understood. Indeed, while previous studies could show that emotions in specific contexts modulate behaviors as well as relevant brain regions and brain waves, they did not characterize these influences. One way to understand the computational and neural mechanisms driving these influences is to inform analyses with models capable of explaining the observed effects. The following chapters will thus be devoted to the field of perceptual decision-making that uses the powerful advantage of modeling when trying to understand the underlying computations of brain functioning, as most of the work conducted in this Ph.D. took advantage of behavioral and neural modeling to characterize the way individuals behave and perceive social stimuli.

### 3-Perceptual decision-making

On a day-to-day basis, we take innumerable decisions consciously or unconsciously. Some of these decisions are formed based on sensory evidence, such as whether the traffic light turned green (to continue your way), how close is the sound of the police siren (to be careful if it is close) and what color is the sky today (to avoid the rain). This type of decision is referred to as 'perceptual decision-making', as decisions are taken based on perceptual information. This Ph.D. focused on social perceptual decisions which represent a form of perceptual decisions based on socially meaningful stimuli, here emotional expressions. Perceptual decisions have been studied using psychophysical/quantitative methods which led to influential theoretical models, some of which will be described below.

#### 3.1- Principal decision-making models

##### **Signal detection theory (SDT)**

In everyday life, we often face situations where the sensory information around us is ambiguous, and therefore requires some effort to decide what it is and how to react accordingly. This might happen on a rainy day where rain obstructs our way back home or in an unfamiliar place such as a new city where some elements are more difficult to recognize, or in a big crowd where it becomes harder to identify your friends etc. The Signal Detection theory (Green and Swets, 1966) deals with decisions of individuals when they are faced with such ambiguous or noisy sensory information. This psychological theory undeniably represents the groundwork of most of decision theories acknowledged nowadays, both from theoretical and mathematical point of view. Its most crucial contribution in the understanding of human decisions is that it is able to disjoin the sensory information available to an individual from his/her decision biases. Its first applications involved detection tasks: a faint signal (i.e., a target stimulus) has to be detected by participants, this signal can either be present or absent, leading to 4 possible outcomes of a decision:

- Hits: signal present/ response present
- False alarms: signal absent/ response present
- Misses: signal present/ response absent
- Correct rejection: signal absent/ response absent (Figure 11a).

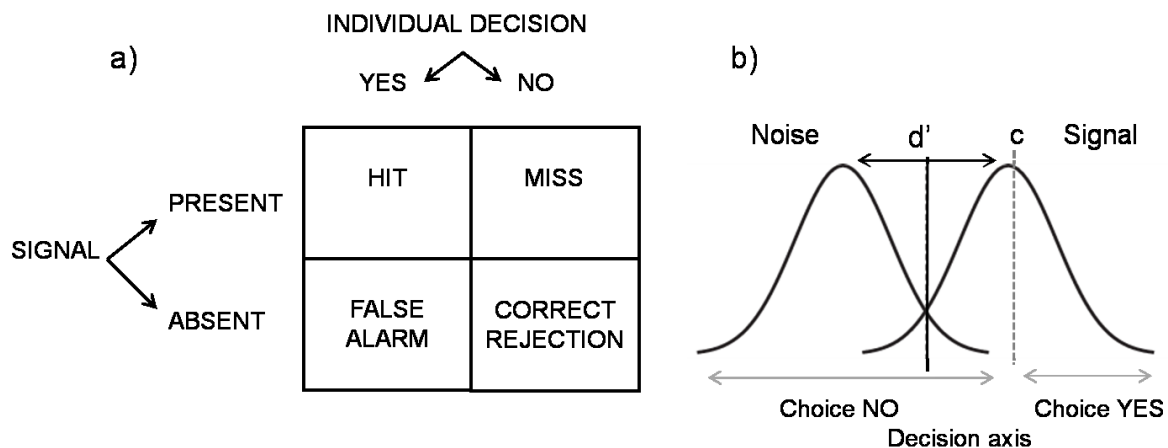


Figure 11. Signal detection theory. a) Individual decision as a function of the presence/absence of the signal leading to 4 possible outcomes. b) Representation of the signal and noise distribution.

The prime hypothesis underlying signal detection theory is that decisions are formed on the basis of a noisy variable which reflects the presence or absence of the signal and corrupted by additive Gaussian noise. Decisions are hypothesized to be taken by comparing the value of the noisy decision variable to a fixed, criterial value: if the decision variable exceeds this value, then the stimulus is classified as containing the signal; otherwise, it is classified as not containing the signal. As the (Gaussian) distribution probabilities of noise and signal overlap (Figure 11b), subjects will necessarily make some mistakes when deciding whether the signal is present or absent. Their behavior can be explained by two parameters: the discriminability or sensitivity parameter  $d'$ , and the criterion  $c$ . The sensitivity  $d'$  corresponds to the signal-to-noise ratio of the decision variable and thus depends on the degree of overlap between the distributions of the signal and noise. In practice, it can be calculated by subtracting  $z$  corrected false alarms from hits. The sensitivity parameter represents how well the decision maker is capable of extracting the sensory information from noise. The less the distributions overlap, the higher will be the  $d'$ , and the better the person can po-

tentially discriminate noise from signal. The second parameter  $c$  represents individuals' decision criterion and can be set at any point on the decision axis to account for idiosyncratic biases in the decision-maker (e.g., conservativeness in reporting the signal): as indicated above, if the signal falls to the left of the criterion, the participant will respond 'NO', if the signal falls at the right of the criterion, he/she will respond 'YES'. Importantly, the criterion can differ from an individual to another: for example conservative decision-makers will tend to shift their criterion to the right (high criterion) needing more evidence to answer YES and thereby increasing misses while reducing false alarms. Liberal decision-makers on the contrary will shift their criterion to the left (low criterion) and respond 'yes' much more often on the cost of making more false alarms. In some cases, decision biases can have useful features. A famous example given to illustrate these differences in decision biases is the one of the radiologist who has to detect a tumour. Some doctors (liberal) will prefer to say that there is a tumour even if uncertain although they increase the rate of false alarms and therefore increase incorrect diagnosis, but it might be better to do a mistake in that case rather than missing a tumour that would likely kill the patient eventually.

Signal detection theory can also be applied to two alternative forced choice tasks, for example in a visual pattern where dots move right or left: the perceptual information to be tracked here would be the movement of the dots, and the decision bias or criterion would be the tendency of each observer to choose more often right or left. In this framework, psychometric functions can describe participants' behavior under uncertainty, by applying general linear models that assign a sensitivity parameter to decision (multiplicative by the sensory evidence), and a decision bias that is an additive parameter depicting the tendency to answer one option over the other. If  $x$  represents the gradual evidence from option 1 to option 2, such as noise to signal, or left to right, the probability of choosing option 2 (such as 'yes' over 'no' in a detection task or 'right' over 'left' in a 2 options tasks) can be described as follow:

$$P(\text{option2}) = \Phi(w*x + b)$$

where  $P(\text{option2})$  corresponds to the probability of choosing option 2 over option 1,  $\Phi$  to the cumulative normal function,  $w$  to the multiplicative gain assigned to the sensory evidence (proportional to the sensitivity parameter  $d'$  in the original formulation of signal detection theory) for option 2 response,  $x$  to the evidence in favor of option 2 response and  $b$  to an additive response bias toward option 2 choice (corresponding to the criterion  $c$  in the original formulation of the theory).

### **Extensions of SDT: sequential-sampling models**

While the above described models do not take into account reaction times, extensions of the SDT do and are useful to have a more complete understanding of decision behavior (Gold and Shadlen, 2007), by describing speed accuracy trade-offs for example. These models, including the drift-diffusion model and the race model, portray the decision over time, assuming that there is an integration of sensory information in time until a decision is made.

The drift-diffusion model (Ratcliff, 1978) proposes that sensory evidence is accumulated until a threshold is reached, at which point the response is selected. The slope of evidence accumulation is equivalent to the sensitivity parameter of SDT and represents the quality of the encoding of the visual information. The threshold at which a response is given corresponds to the criterion, as it changes as a function of individuals: liberal individuals will have a low threshold to respond, and thus need to accumulate less evidence to respond. The race model differs from the drift diffusion model because it considers that two independent populations of neurons will accumulate evidence toward one or the other option. In the drift diffusion model, the symmetric random walk shows accumulated evidence for either hypotheses  $H_1$  or  $H_2$  whereas in the race model, while evidence accumulates for  $H_1$ , it decreases for  $H_2$  (Figure 12).

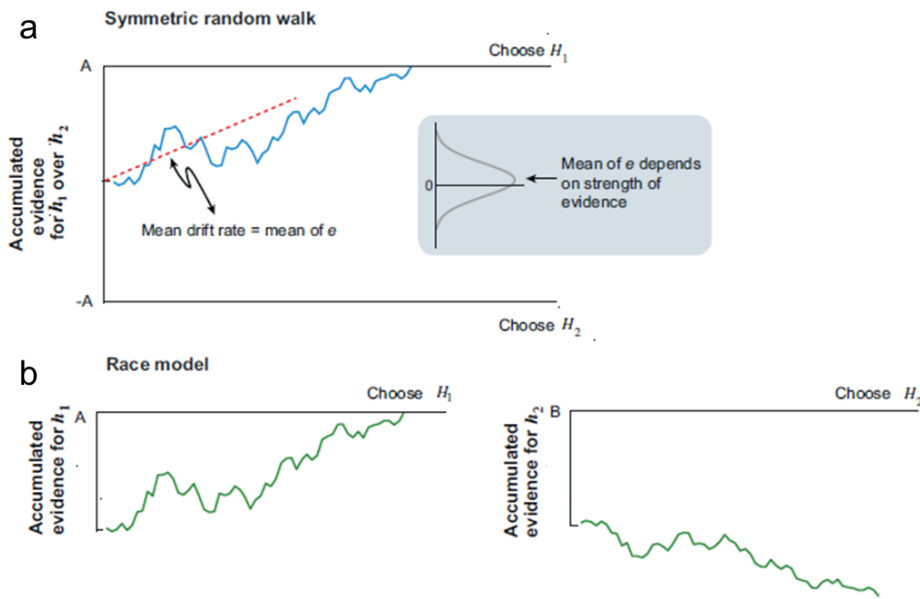


Figure 12. From Gold & Shadlen 2007. Race model (b) in comparison with the drift diffusion model (a).

Neural evidence that corroborate these theoretical accounts of decision-making both in monkeys and humans contributed to their influential impact in the field of decision-making.

Evidence from monkey studies indeed show that sensory areas respond to sensory evidence by firing as a function of evidence phasically after stimulus onset while ‘accumulator’ regions rather fire gradually until response is provided. An example is given in Figure 13 (Freedman et al., 2003): while monkeys categorized morphed pictures of cats and dogs, recordings in the inferior temporal cortex showed a transient neural firing after stimulus onset specific to the evidence to dog/cat stimuli. In contrast, higher-function areas computing the decision, here the prefrontal cortex (PFC), fired in an accumulative fashion, increasing gradually their activity until response. Apart from the PFC, considerable work has focused on a region of the monkey parietal cortex, the posterior parietal cortex (LIP) which shows several properties of evidence accumulation during perceptual decision-making (e.g., Shadlen and Newsome, 2001).

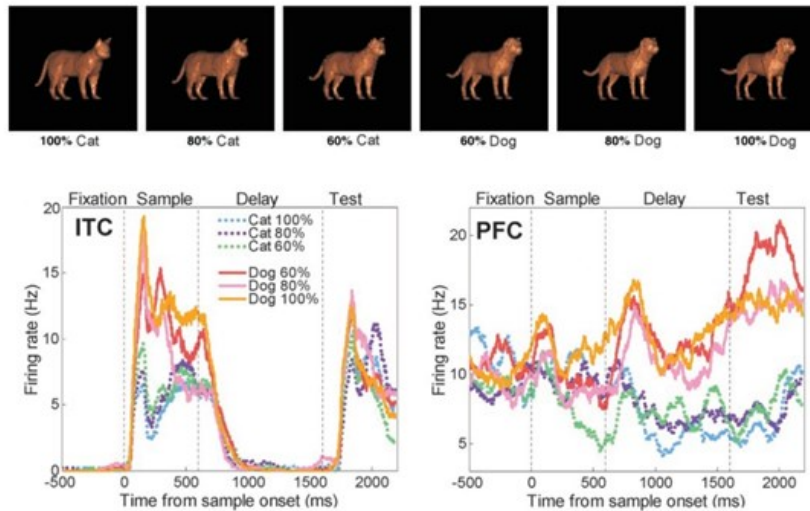


Figure 13. From Freedman et al 2003. Firing of sensory and decision related cells while monkeys categorize cats and dogs stimuli. Here, the ITC cell tracks evidence to dog stimuli. PFC accumulates evidence to dog stimuli until decision.

Neural evidence confirming this model has also been provided in an fMRI experiment in humans involving a face/house categorization task (Heekeren et al., 2004). The available sensory information was degraded by noise added to the images and authors reasoned that two conditions should be fulfilled by higher level cortical regions that integrate decisions: they should show greatest activity for higher perceptual evidence and their activity should be correlated with the difference between face sensory signals and house sensory signals. Sensory responses to faces and houses were computed in face selective (FFA) and house selective areas (parahippocampal place area PPA) in ventral temporal cortex. Only the dorso-lateral prefrontal cortex (DLPFC) fulfilled both conditions as the difference between face selective and house selective regions co-varied with the time course in DLPFC and showed greater response to clearer images. A schematic example of the diffusion model taken from this study is shown in Figure 14, with examples of the faces on the left (more or less degraded information),  $\mu$  corresponds to slope of the evidence accumulation (the sensitivity to visual information), and  $a$  and  $b$  are the thresholds to respond either face or house respectively. In the first example (blue), the accumulation process is much faster as the evidence is clear, and the slope is steeper as compared to the second example where the face is degraded (red) and the accumulation takes more time to reach the face response. This study therefore extends the



drift diffusion model to more complex decisions in humans and suggests that the DLPFC is responsible for accumulation processes.

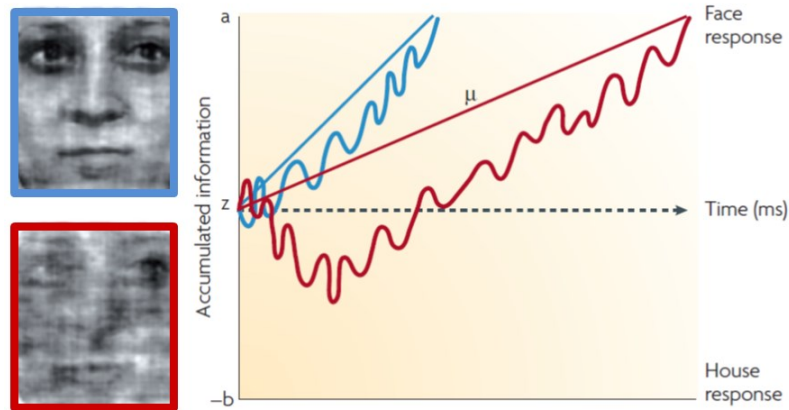


Figure 14. From Heekeren et al. 2004. Accumulation of evidence for a clear (blue) and a degraded face (red) toward face response.

### 3.2- Neural findings in perceptual decision-making

Above, I described the most influential models in perceptual decision-making and illustrated with a few examples from monkey and human studies. Here I will elaborate on detailed recent findings in perceptual decision-making to undertake exposing its underlying neural systems (see extensive reviews, Gold and Shadlen, 2007; Heekeren et al., 2008). While the two examples of stimuli I gave above were house/face and animal morphing, the most common stimuli used in studies that tackled the neural mechanisms of perceptual decision-making are either random dot moving stimuli (RDM) or gabor patches (gratings tilted to the right or left with different contrast levels). These are highly controlled stimuli that can easily be used in both monkey and human studies with different response modalities, such as button press or saccade initiation to respond.

## **Accumulation of sensory evidence**

The accumulation of sensory evidence in the brain, that substantiates the drift-diffusion model, has been frequently reported in perceptual decision-making studies. It involves various brain regions in humans encompassing both sensory and domain-general structures that will be described below.

As reported above, during the categorization of face/house stimuli, evidence accumulation was computed by the DLPFC, suggesting that this area could play the role of a high order decision area in humans (Heekeren et al., 2004). The DLPFC involvement in human perceptual decisions was replicated when the task involved the categorization of the direction of RDM and shown to be independent of response modality (saccade initiation or button press) (Heekeren et al., 2006). The idea that sensory information is integrated in associative processing areas for decision formation has also been corroborated during an object recognition task where parietal and frontal regions represented domain-general regions whose activity mirrored that of a decision accumulator until response (Ploran et al., 2011). Furthermore, when decisions are decoupled from motor response, perceptual choices are decoded in parietal and visual cortices (Hebart et al., 2012).

The combination of EEG and fMRI analysis during a car/face categorization task (Philiastides et al., 2006; Philiastides and Sajda, 2007) evidenced the time course of decision processes (Figure 15): first, early visual perception of face/car stimuli took place in sensory areas such as FFA and the superior temporal gyrus (STG), then at 220 ms task difficulty modulated brain responses in regions such as the anterior cingulate cortex (ACC) and DLPFC, and finally after 300ms, the lateral occipital cortex (LOC) varied with sensory evidence showing that there is a perceptual persistence in post sensory phase, and the ventro-lateral prefrontal cortex (VLPFC) was involved in uncertainty processing. Furthermore, this last component has been shown to be qualitatively different from the early perceptual encoding component as it is sensitive to task demands and reflects post sensory processes directly providing decision evidence (Ratcliff et al., 2009). The difficulty component at 220 ms was however called into question by evidence showing that it is highly dependent on the addition of phase noise to the stimuli thereby increasing their sensory pro-

cessing (Bankó et al., 2011). On the contrary, when overall task difficulty was manipulated by rendering the stimuli ambiguous with morphing techniques, this component was not affected. Nevertheless, these studies importantly show that sensory regions encode evidence strength along with non-sensory associative regions.

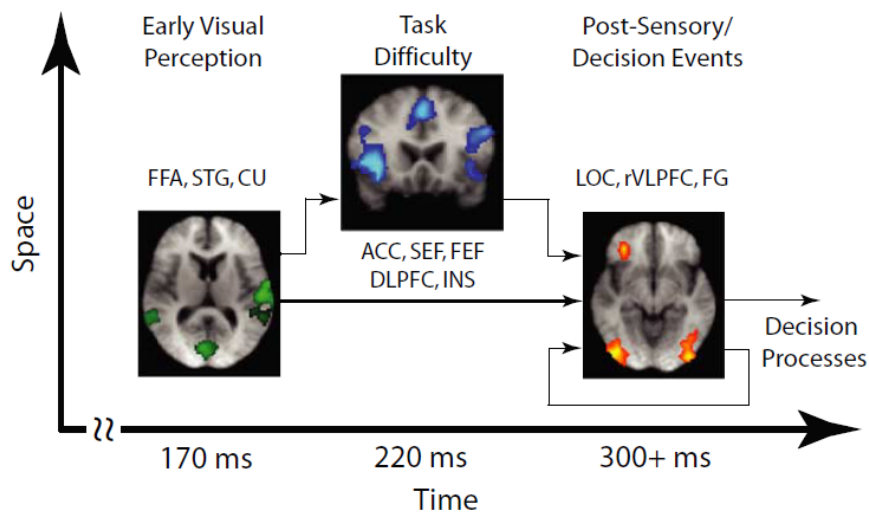


Figure 15. From Philiastides and Sajda, 2007. Model of decision processes in the brain throughout time.

Recently, it has been proposed that the evidence accumulating process involves a domain general component that tracks the decision variable (O’Connell et al., 2012) independently from sensory modality and task. In this study, the flickering of a gradual visual annulus was varied and participants were requested to detect the fading of the stimulus with a key press. In addition to a sensory specific occipital component, a motor selective component and a centro-parietal positivity (CPP) co-varied with the evidence to response. This centro-parietal positivity tracked evidence even when the sensory modality was changed to auditory and the task changed to a mental count of targets. On the contrary, when the visual information was made irrelevant, the CPP did not track the evidence, implying its role in the tracking of only goal-relevant sensory information.

In a different study that used RDM, the CPP showed a buildup rate tracking the sensory evidence to motion coherence (Figure 16) and reflected a bound in the decision process as it stopped building up at the moment of the decision (Kelly and O’Connell, 2013). Once again, these observations validate evidence accumulation

processes and the reach of a threshold in the decision-making procedure. In addition, this component shared all the properties of the parietal evoked potential peaking around 300 ms, the P300, which led to the conclusion that the CPP, and its event-related potential ERP-equivalent P300, constitute a domain general decision variable independent from the motor effector system (Kelly and O’Connell, 2015).

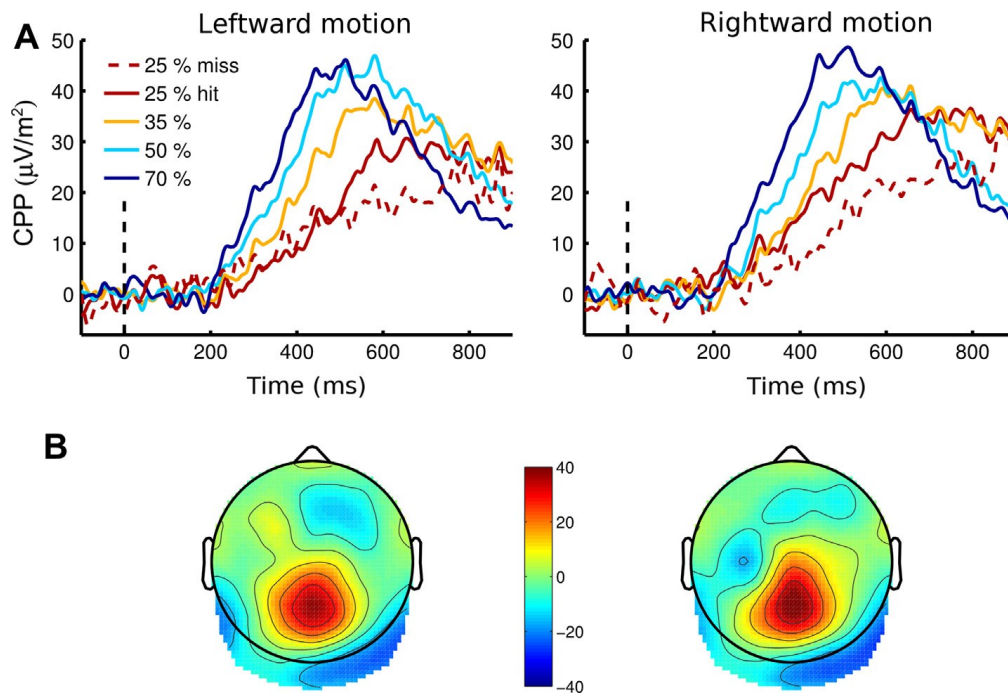


Figure 16. From Kelly & O’Connell 2014 (modified from Kelly & O’Connell 2013). Domain general decision variable tracking leftward and rightward motion. Activity increases with decision evidence.

### Perception/Action in perceptual decision-making

While traditional views of decision-making suppose that the decision process is serial, first requiring the perception of the stimulus on which the decision evidence is based and later the motor-specific selection of the decision, recent studies highlight the involvement of action related areas in the formation of the decision starting very early on. The boundaries between perceptual, decision related and motor systems seem to be less strong than initially considered. Single neuron recordings in monkeys show that motor related regions are involved in the decision formation before the motor execution *per se*. For example, the frontal eye field (FEF) that controls eye movements, encodes both decision and saccade related information when monkeys

select with their eyes whether random dots are moving right or left (Gold and Shadlen, 2003). Likewise, neurons in the premotor cortex coded the entire cascade of perceptual decision-making from sensory processing to motor outcomes during a vibrotactile discrimination task (Hernández et al., 2002; Romo et al., 2004). This is replicated in humans where response selective regions contribute to the decision formation during perceptual tasks, such as the involvement of the FEF in decision formation during an oculomotor decision task (Heinen et al., 2006). However, it has been argued that an effector-general as opposed to an effector-specific mechanism sustains perceptual decision-making: during decisions on random dot movements, evidence accumulation regions including central regions such as FEF and precentral sulcus (Liu and Pleskac, 2011) were involved independently of task modality (saccadic movement or key press) and foreknowledge of the modality to use (cue before or after the visual stimulus).

Although there seem to be regions integrating evidence independently from motor planning, it does not exclude the fact that motor regions, including effector specific regions, are clearly involved in the decision formation process. During a signal detection task that required human subjects to detect movement coherence in random dot movements, effector selective measures built up choice predictive activity starting long before button press (Donner et al., 2009). Effector selective activity was computed by subtracting time frequency decompositions in the alpha-beta bands (8-36Hz) and gamma bands (64-100Hz) from the ipsilateral to the contralateral hemisphere of the hand used to respond (motor lateralization measure). Subjects were requested to use their left and right hands to provide their yes/no responses. This allowed computing when the motor lateralization started to predict choice after stimulus onset. While beta bands are suppressed in the hemisphere contralateral to the response, gamma bands are enhanced. This results in opposed lateralization for beta and gamma bands such as shown in Figure 17– in the motor/premotor cortex. Generally, lateralization in both bands was able to predict the choice better than chance (marked by the dashed line on the figure) during stimulus presentation (shown for 2 seconds). The prediction was similar for hits and false alarms, as well as for misses and correct rejections, approving the idea of different populations of neurons building up for yes or no responses rather than one popula-

tion, in which case the motor preparation should have started building toward yes responses before going to no responses during a miss.

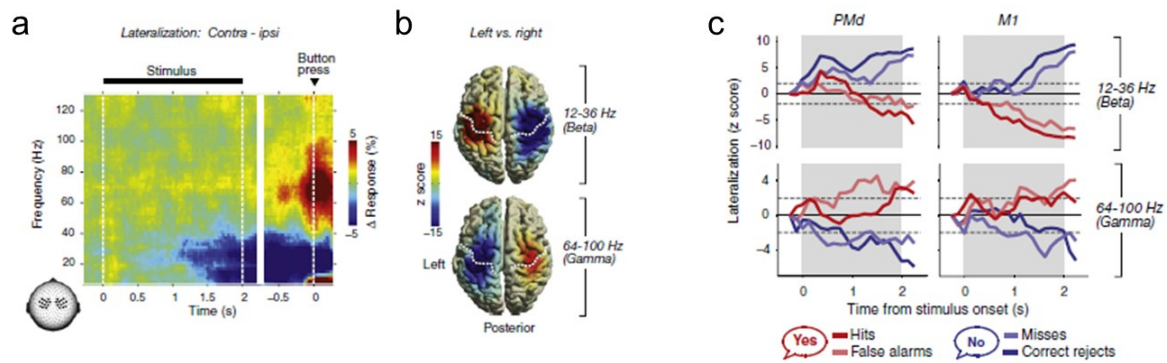


Figure 17. From Donner et al 2009. Effector selective buildup of choice in PM and Motor cortices. a) Time frequency plot showing contralateral minus ipsilateral hemispheres to the hand used for button press. b) Neural sources of the lateralization measure show in a. c) Time course of the lateralization measure in PM and motor regions.

Further, the lateralization measures correlated with the temporal evidence in motion detection area MT (middle temporal area), suggesting that the evidence from sensory areas is integrated into response in motor effector selective structures. Further evidence pinpoints the role of motor beta band activity as a response preparation signal as it co-varied with response updates (i.e., response preparation increments provided by discrete pieces of sensory evidence provided sequentially) during a multisample categorization task (Wyart et al., 2012a). It also had an additive influence on decision-making, as it predicted decision bias rather than the multiplicative sensory evidence. The methods used to draw these conclusions will be described in further details in following sections, as they were applied in the work conducted in this Ph.D. (see Part 2, Experiment 2).

To conclude, growing evidence of a parallel functioning between motor and perceptual systems substantiates the viewpoint of interactive systems rather than segregated serial processing during decision-making (Cisek and Kalaska, 2010). Of course, this does not preclude the importance of having abstract decision mechanisms in effector-independent regions when the motor action is decoupled from

choice, as many studies have shown (e.g., Heekeren et al., 2006; Liu and Pleskac, 2011; Hebart et al., 2012). One way to reconcile the findings supporting the two views (decision in effector related vs common regions) is to consider the distributed consensus model (Cisek, 2012) suggesting that the brain has different types of representations including one for abstract decisions and one related to action selection. Decisions could thus possibly form through simultaneous implications of associative and action-specific brain regions, through a multi-level distributed mechanism. Both representations are necessary depending on whether decisions need a rapid motor reaction in which case the early involvement of motor-related regions might serve to facilitate actions, or whether abstract decisions have to be taken in which case associative domain-general regions could be sufficient to map the decision process. Along these lines, the brain might have evolved to deal with the two types of decisions by concurrently mapping both representations; however how it did so remains an open question.

### **Neuromodulatory systems in decision-making: the role of the coeruleus-norepinephrine system**

Although the role of neuromodulators in the brainstem has been traditionally assigned to basic functions such as reward and arousal, a growing number of recent studies are suggesting their involvement in more complex functions such as decision-making (Aston-Jones and Cohen, 2005). Specifically, mostly substantiated by monkey studies, the Locus Coeruleus-Norepinephrine system (LC-NE) is proposed to have a crucial role in decision-making. According to two modes of activity, LC activity facilitates behaviors increasing performance on decision tasks (phasic activity) and drives the search of alternative behaviors when needed (tonic activity) (Aston-Jones and Cohen, 2005). In fact, phasic firing in LC monkey neurons is observed after a target stimulus but not a distractor therefore responding in a task specific manner while tonic phases are observed when distractibility appears in animals' behaviors. At the level of human studies, the decision-related P300 provides a window into the implication of norepinephrenic systems in decision-making (Kelly and O'Connell, 2015). The P300 is suggested to reflect activity in the LC-NE system that

sustains decision-making and information processing (Nieuwenhuis et al., 2005). More specifically, the phasic activity of LC is linked to the P300, both reflecting the performance and the engagement of subjects performing a task. This framework therefore suggests that a possible measure of LC-NE activity is the P300 recorded at the scalp of the brain. Another physiological measure that directly reflects LC-NE activity is pupil dilation as it highly reflects the rate of LC neurons activity (Figure 18).

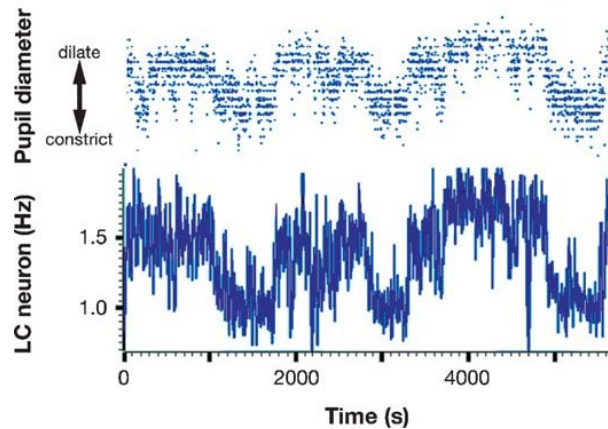


Figure 18. From Aston-Jones and Cohen 2005. Pupil dilation and LC-NE activity in monkeys during signal detection task.

Based on this relation between pupil dilation and LC-NE activity, several studies in humans measure pupil dilation using eye-tracking during behavioral tasks. Pupil dilation is associated with visual detection (Privitera et al., 2010), encodes decision information (e.g Einhäuser et al., 2008; Cheadle et al., 2014) and is increased to uncertain/surprising events (Preuschoff et al., 2011) and effort (Porter et al., 2007). In a recent study with a signal detection task (visual target with noise present or absent), pupil dilation was shown to reflect choices throughout the decision process and to track participants' individual biases (Gee et al., 2014). Pupil dilation was stronger before yes responses (signal present) than before no responses, and this increase correlated with individual criteria (calculated based on SDT). The more participants were conservative (bias toward no responses), the more their pupils dilated when they responded "yes" against their bias (Figure 19). This finding was interpreted as possibly resulting from a surprise signal (yes responses being more infrequent in conservative individuals and therefore more surprising). Further, the authors interpreted this finding within the race model in decision-making, where two



populations of neurons one for “yes” response, one for “no” response accumulate evidence toward opposite thresholds. If conservative people start at a point further away from “yes”, then a stronger neural boost in gain from visual areas is needed to choose the “yes” response and could be reflected in increased pupil dilation .

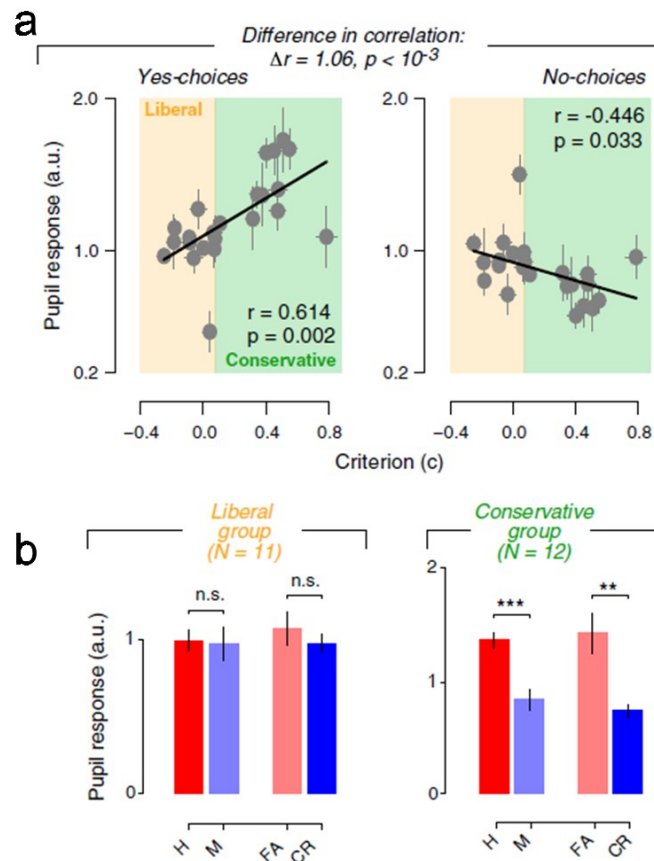


Figure 19. From Gee et al. 2014. Pupil responses increase with yes responses against the bias of conservative individuals. a) Correlation between pupil response and individual biases for yes (left panel) and no (right panel) responses. b) Hits, misses, false alarms and correct rejections for the liberal group (left panel) and the conservative group (right panel).

### Chapter 3 conclusion

Extensive research on the processes underlying decision-making shows that it relies on a set of sensory, associative and motor regions as well as neuromodulatory systems, all seeming to interact rather than working serially and independently. A very recent study in monkeys corroborates this view by showing that sensorimotor decisions are mediated by a flow of information in a network involving visual, temporal, parietal and motor regions (Siegel et al., 2015). This allows the communication

between fronto-parietal networks with sensory visual areas and motor-initiating regions to initiate flexible context-dependent behaviors. Findings thus converge toward the view that sensory, associative and motor regions are concurrently rather than successively involved during the formation of decisions. As suggested earlier, the brain might have evolved in such a way that it is capable of dealing with different sorts of decisions (abstract and action-driven), by forming these decisions in a broad network encompassing sensory, motor and associative regions.

#### **4-Contextual effects in perceptual decision-making**

During perceptual tasks, manipulating prior information and participants' expectation is feasible by adding explicit cues that bias toward a specific response. This has been the question of many perceptual tasks, as stated in section 2, it is indisputable that day to day decisions are influenced by prior information – in particular when the available sensory evidence to reach a decision is noisy or ambiguous. Hence understanding the mechanisms of contextual prior influences is essential to understand decision-making as a whole. In non-social perceptual decision studies, explicit expectation or attentional cues are typically provided before the decision relevant information and are processed consciously by the participant. Contrarily to social stimuli described in section 2, non-social stimuli are not typically encountered in daily situations and thus not particularly associated to specific settings or elements present in natural environments. If we take the example of random dots, there is no reason to think that a background with a blue sky or a grey sky will change the decision toward left or right motion movement. Similarly, the stimulus *per se* does not contain additional features (such as the face that contains many elements: gaze, mouth, etc) that will influence the left/right choice. Therefore what can be modulated is either cues provided before the stimulus that predict the upcoming visual information (also referred to as probability cues) or task relevant information such as the attentional focus or the behavioral relevance of a stimulus.

#### 4.1- Expectation/probability cues

Contextual information can be manipulated by providing cues at the beginning of trials or blocks that will predict the upcoming stimulus with a certain probability. If participants integrate these cues in their decisions although the choice concerns an independent visual target, then it would signify that expectations bias the choices. Normative (ideal) Bayesian accounts of decision-making suggest that conditionally independent sources of information, such as expectation cues and a target stimulus, should be combined additively into a 'posterior' belief about the judgment of the target (Griffiths and Tenenbaum, 2006; Oaksford and Chater, 2007), and this would be reflected in SDT terms as a change in the decision bias. As will be described below, while some studies confirm this statement, recent studies also suggest that expectation cues can change the sensitivity to the visual information. Target brain regions reflecting these effects are those described in the previous section as being involved in the decision process.

To start with, the DLPFC, which has been identified as a decision accumulator region in humans, also takes into account prior information. Expectation was manipulated during a task where participants had to judge whether a patch of dots was expanding or constricting (Rahnev et al., 2011a). A shaped cue (triangles and squares) was provided at each trial before the target stimulus and could either be predictive (75% of the trials valid, 25% invalid) or non-predictive (neutral cues) of the forthcoming stimulus, subjects were fully aware of the design and instructed to take the cue into account for their decisions. Subjects performed better at the task when provided with a valid cue and poorer for an invalid cue as compared to neutral. This expectation information was reflected in DLPFC activity and effective connectivity between sensory regions and DLPFC was also increased when participants had prior expectations. DLPFC reflected individual criterion adjustments that were shifted toward the expected option. To conclude on this study, in addition to being a target region that accumulates evidence during decision-making, the DLPFC incorporates prior expectations into the decision process.

Moreover, the motor system seems to be involved in the integration of priors during decisions. In a random dot task, a similar expectation manipulation was used

where predictive (words “left” or “right”, 75% of trials valid and 25% invalid) and non-predictive (word “neutral”) cues were presented before the motion stimulus (de Lange et al., 2013). Here the motor lateralization described previously (Donner et al. 2009) was computed separately for expected and unexpected conditions and locked to the cue presentation to see whether pre-stimulus activity is modified by expectation cues. Motor preparation signals were biased by the expectation as they started building up toward the congruent answer based on the expected cue even when the final choice was the opposite response (Figure 20). Interestingly, even in neutral cue trials, participants showed biases in the pre-stimulus interval that were partly explained by their preceding responses. Behavioral biases showed the same expectation modulations and correlated with the neural effect.

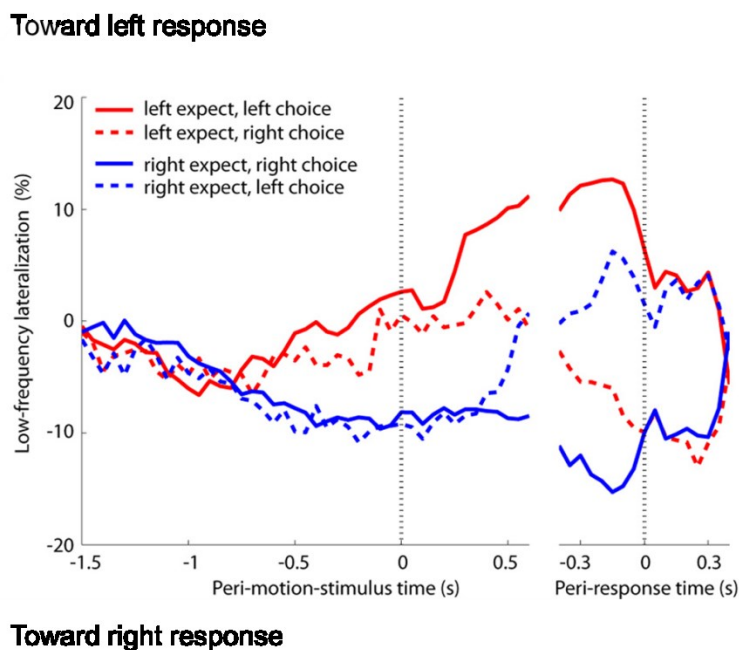


Figure 20. From Delange et al. 2013. Expectations bias motor preparation in pre-stimulus interval.

Finally and strikingly with regards to the Bayesian view described above, prior expectations influence the very early processing of visual stimuli. Indeed, in Bayesian terms, expectations should enact as additive sources of information which should be combined, but should not *a priori* interact, with the processing of the sensory evidence. Effects of expectation were reflected in an interesting dissociation within vis-

ual area V1 in a study where an auditory cue predicted the orientation of an upcoming visual graded stimuli (Kok et al., 2012): while expected stimuli triggered less activity in V1 (expectation suppression (Summerfield et al., 2008)), they were better decoded in this same area using multivariate pattern analysis (MVPA). Expected stimuli being better represented in visual areas, the authors concluded that expectations sharpen the representation of visual information in primary sensory areas. Using a forward modeling approach to check whether reconstructed visual information is influenced by expectations, the same authors showed in a subsequent study (Kok et al., 2013) that prior expectations can modify what is represented in early visual areas. Further substantiating the early effects of prior probability on visual processing, providing information about the probability of stimulus occurrence at the beginning of the block increased sensitivity to weak signals by increasing the baseline activity of threshold-saturating signal-selective units – rendering them more sensitive to weak signals (Wyart et al., 2012b).

Altogether, the findings presented in the last section show that prior expectations are incorporated into the decision process in several brain regions including primary visual areas, domain-general decision areas such as DLPFC, and motor selective regions. Moreover, not only do expectations bias decisions toward expected stimuli by shifting the decision criterion, as predicted by conventional signal detection theory and Bayesian formulations of decision-making, they can also sharpen the sensory representation of expected stimuli.

#### **4.2- Attention/relevance cues**

Operating context in perceptual decision-making studies was most commonly implemented by explicitly shifting the attention of participants to specific information (either in space or among visual features), thereby biasing the processing of one source of sensory information over the other. The results of these studies generally

converge by showing a multiplicative boost in the sensory processing of the relevant information (i.e. attended).

Since early days, manipulating attention in a perceptual task showed that providing participants with a cue that predicts the spatial location of the target stimulus enhances the perceptual sensitivity to that stimulus (H S Bashinski, 1980). As shown in numerous studies since then, attention increases sensitivity to visual inputs by both decreasing the external noise and enhancing the sensory gain of relevant information (see review Carrasco, 2006). Because paying attention to everything in our surroundings is impossible given the limited attentional resources at our disposal, the visual system adapts by increasing visual salience to relevant information. Single neuron recordings in monkeys validate these statements by showing that attention heightens contrast discrimination of attended stimuli in visual area V4 (Reynolds et al., 2000). Similarly, a multiplicative enhancement of V4 monkey neurons ability to discriminate orientation was observed (McAdams and Maunsell, 1999). Task information is also first encoded in early sensory cortices before reaching frontal and parietal regions and finally spreading to other regions from the sensorimotor pathway (Siegel et al., 2015). In humans, task information increases the sensory processing of the attended visual stimulus. As an illustration, during a perceptual categorization task of either face (gender discrimination) or house stimuli (1 or 2 houses), providing a cue that determined whether the houses or faces will appear not only increased performance at the decision stage but also anticipatory BOLD activity in object-selective areas- FFA for faces, PPA for houses (Esterman and Yantis, 2010).

When examining the timing of attentional effects, the early processing of sensory information is enhanced to attended as compared to unattended stimuli at 250 ms. However, the first evident attentional filter only occurs around 500 ms in parietal regions, when the visual information is converted into associative decision signals (Wyart et al., 2015). At this point only task relevant information is encoded. Another study also showed an influence of relevance cues predicting the forthcoming signal within each trial later in the process at the decision moment, by increasing the sensitivity to strong signals through the suppression of the internal noise corrupting the decision process (Wyart et al., 2012b). It is also suggested that this attention-

triggered reduction in internal noise can indirectly modulate the decision bias by diminishing the probability of reaching the criterion – which reduces the probability of a false alarm (Rahnev et al., 2011b).

Although it is acknowledged that attention/relevance cues sharpen the sensory processing of attended/relevant information, evidence from the later mentioned studies show that attention also influences later stages of the decision process: the conversion of sensory signals into decision-relevant evidence, and the apparent decision criterion.

#### **Chapter 4 conclusion**

In the perceptual decision-making domain, contextual cues consist of either varying expectation cues or relevance cues. The use of model based analyses for studying expectation- or relevance-dependent effects on decision-making has allowed for the conclusion to be drawn that both types of manipulations influence sensitivity to perceptual information as well as the decision criterion depending on the experimental design and tasks used. These influences involve sensory, motor and more domain-general regions in the human brain. Extending this model-based approach to contextual effects in the case of social perceptual decisions could thus potentially enable a fine characterization of how decisions are formed on stimuli of both biological and social significance.

#### **5-Informing contextual effects in emotion perception with decision models**

In the last chapters, I separately presented the perception of socially-meaningful stimuli and perceptual decision-making on non-social displays and the impact of contextual information on decisions. Naturally, all social tasks involving biologically rele-

vant stimuli need at their core an adequate processing of the sensory evidence provided by the stimuli. Although countless studies investigated social perception, fewer studies employed methods used in the perceptual decision-making domain (model-based analysis) to address the mechanisms underlying social perception. This partly stems from the fact that social stimuli are harder to control in terms of exact visual components, when compared to highly controlled stimuli such as Gabor patterns or random-dot kinematograms where the experimenter is fully aware of each pixel in the image. Of course, recent studies that used sophisticated methods such as “reverse-correlation” approaches overcame this limitation (e.g., Gosselin and Schyns, 2001; Dotsch and Todorov, 2012). As stated in previous chapters, these studies elegantly dissociate different elements of the face that contribute to perceptual decisions on faces. However, they do not integrate decision theoretical models into the emotion processing mechanism. Using a model-based framework, both behaviorally and neurally, helps clarify the mechanisms mediating perceptual decisions on social displays. In perceptual decision studies, the parametric control over perceptual evidence (how much the Gabor pattern is tilted to the right or left, or the degree of coherence of the random-dot pattern) allows using generative models of choice to predict the performance of human participants given the amount of sensory evidence provided by the stimulus. The way we adopted in this Ph.D. to parametrically control socially-meaningful information when addressing emotion perception is a morphing technique: indeed it permits varying emotional expressions from ambiguous to prototypical displays, thereby rendering the stimuli more ecological and importantly perceptually more difficult to classify as the perceptual information is progressively degraded.

The principal aim of the work conducted in this Ph.D. was to delve further into the depth of contextual influences during the perception of social stimuli. The first conducted study (Experiment 1) did not use model-based analyses, but took advantage of a time-resolved neural recording technique, electroencephalography (EEG), to identify how and when selective attention influences the perception of multi-dimensional social displays. This study explored whether the neural markers of social displays of threat are affected by which social cue the observer is attending (both at the timing and source level). The second and principal study of this work



(Experiment 2) builds on the previous by characterizing how different combinations of social cues are decoded in the brain as a function of how much they portray a threat to the observer. More specifically, this study combined a quantitative model-guided approach with neural recordings (EEG) to address the neural and computational mechanisms of the influence of gaze on emotion perception in the context of threat. Experiments 3, 4, 5 and 6 represent behavioral studies that were designed to answer various questions raised by Experiment 2 by exploiting model-guided analyses based on signal detection theory. In brief, we wanted to further characterize how emotion and gaze were combined by dissociating stimuli into low and high spatial frequencies, test the specificity of the effect to a threat context, check whether emotion reciprocally influences gaze perception, and finally extend our findings to the general population. The last experiment (Experiment 7) aimed at manipulating a different contextual factor during emotion perception: past experience with identities. This last study combined behavioral modeling with pupillometric measures to investigate how identity/emotion associations (i.e., an identity that is angry most of the time) can modify emotion categorization.

## Part 1: Attention as a contextual factor during the perception of social cues (Experiment 1)

In chapter 2 section 3, we saw that emotion and gaze cues are perceived interactively by observers, and that this interaction depends on the significance of the combined signal. Table 1 shows that different types of tasks were used in different studies, these tasks varying from gender categorization to emotion categorization to gaze categorization tasks. Although the findings taken together show that emotion/gaze interactions are persistent across tasks, the regions integrating these social cues and the time at which they occur seem to be variable. This can be due to differences in experimental design and task demands. The question of the influence of the task was only directly addressed in emotion studies that did not vary gaze direction but used both explicit and implicit tasks regarding the emotion (e.g., which emotion, or gender categorization). These investigations show that emotional displays are detected at early stages even if irrelevant to the task (e.g Holmes et al., 2006). Task modulation occurs later, (Krolak-Salmon et al., 2001; Hajcak et al., 2006; Holmes et al., 2006; Van Strien et al., 2010; Rellecke et al., 2012), and often involved the P300 (Hajcak et al., 2006; Van Strien et al., 2010). Here we wanted to address whether the combination of social cues including emotion but not emotional displays alone, is affected by directing attention to one of the social cues manipulated. Importantly, as social cues that consisted of emotion, gaze and gesture were shown to be combined at 200 ms in the PM cortex (Conty et al., 2012) during gaze categorization, our main question in Experiment 1 was whether the timing and source of the integration would vary as a function of the attended cue.

Experiment 1: published in Neuroimage, 2015

***Selective attention effects on early integration of social signals: same timing, modulated neural sources.*** Experiment 1: published in Neuroimage, 2015



## Selective attention effects on early integration of social signals: Same timing, modulated neural sources



Marwa El Zein<sup>a,b,\*</sup>, Lucile Gamond<sup>a,c</sup>, Laurence Conty<sup>a,d</sup>, Julie Grèzes<sup>a,e,\*</sup>

<sup>a</sup> Laboratoire de Neurosciences Cognitives (LNC), INSERM U960, Département d'Etudes Cognitives (DEC), Ecole Normale Supérieure, PSL Research University Paris, France

<sup>b</sup> UPMC Université Paris 06, Paris, France

<sup>c</sup> UFR de psychologie, Université Paris 8, Saint-Denis 93526 cedex, Paris, France

<sup>d</sup> Laboratoire de Psychopathologie and Neuropsychologie (LPN, EA2027), Université Paris 8, Saint-Denis, France

<sup>e</sup> Centre de Neuroimagerie de Recherche, CRICM, Université Pierre et Marie Curie-Paris 6 UMR-S975, Inserm U975, CNRS UMR7225, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

### ARTICLE INFO

#### Article history:

Accepted 28 October 2014

Available online 1 November 2014

#### Keywords:

Attentional influence

EEG

Integration

Self-relevance

Social cues

### ABSTRACT

Humans combine co-emitted social signals to predict other's immediate intentions and prepare an adapted response. However, little is known about whether attending to only one of co-emitted social signals impacts on its combination with other signals. Here, using electroencephalography, we address selective attention effects on early combination of social signals. We manipulated three visual cues: gaze direction, emotional expression, and pointing gesture, while participants performed either emotion or gaze direction judgments. Results showed that a temporal marker of social cues integration emerges 170 ms after the stimulus onset, even if the integration of the three visual cues was not required to perform the task, as only one feature at a time was task relevant. Yet in addition to common temporal regions, the relative contribution of specific neural sources of this integration changed as a function of the attended feature: integration during emotion judgments was mainly implemented in classic limbic areas but in the dorsal pathway during gaze direction judgments. Together, these findings demonstrate that co-emitted social cues are integrated as long as they are relevant to the observer, even when they are irrelevant to the ongoing task.

© 2014 Elsevier Inc. All rights reserved.

### Introduction

The quality of our social interactions depends on correctly detecting co-emitted social signals and adapting our behavior accordingly. Of interest, once merged into a single percept, co-emitted social signals can take on new significance and certain combinations, notably those indicating the presence of threat, become more relevant than others. Typically, angry expressions are perceived as more threatening when associated with a direct gaze than with an averted gaze (Sander et al., 2007; N'Diaye et al., 2009; Sato et al., 2010; Conty et al., 2012). Despite the clear relevance of understanding the spatiotemporal characteristics of the mechanisms underlying the combination of social cues, the question of whether, when and how, directing one's attention towards a specific social signal impacts on its neural integration with other co-emitted social cues remains.

Indeed, the literature exploring the neural sources of gaze and expression integration reports incongruous results (Graham and Labar, 2012). One possible explanation is that the neural sources underlying social cues integration may be task demand-dependent (Graham and Labar, 2012; Dumas et al., 2013). While most of the functional magnetic

resonance imagery (fMRI) experiments revealed that the amygdala integrated emotion and gaze when participants were required to attend to the emotional content of the faces or their gender (Sato et al., 2004, 2010; Sander et al., 2007; N'Diaye et al., 2009; Adams et al., 2012), the premotor cortex was involved in gaze and expression combination when participants were requested to attend to gaze direction (Conty et al., 2012). Electroencephalography (EEG) results also suggest that the temporal marker of such integration is susceptible to task demand influences: it emerges around 170 ms after stimulus onset when participants attend emotional expressions (Akechi et al., 2010), at around 200 ms when participants attend gaze direction (Conty et al., 2012), at 240 ms during passive viewing (Rigato et al., 2010), and later during an oddball task concerning gaze direction (Klucharev and Sams, 2004). Finally, other studies in the literature have reported that when observers orient their attention to only one feature at a time as requested by the task (i.e. judging the emotion or the gaze of emitters), co-emitted facial signals are processed independently (Bindemann et al., 2008) by anatomically and functionally segregated neural structures, at least initially (Klucharev and Sams, 2004; Pourtois et al., 2004).

To directly address how selective attention affects the time course and neural bases of early combination of social signals, we implemented an electroencephalography (EEG) experiment (evoked potentials and source reconstruction analysis). To parametrically modulate the self-relevance of perceived social signals, we manipulated three visual

\* Corresponding authors at: Laboratoire de Neurosciences Cognitives, INSERM U960, DEC, Ecole Normale Supérieure, 29 Rue d'Ulm, 75005 Paris, France. Fax: +33 1 44 32 26 86.

E-mail addresses: [marwa.el.zein@ens.fr](mailto:marwa.el.zein@ens.fr) (M. El Zein), [julie.grezes@ens.fr](mailto:julie.grezes@ens.fr) (J. Grèzes).

cues: emotion, gaze direction, and pointing gesture (as in Conty et al., 2012). Participants performed two tasks: one which requested them to attend only to the gaze directional cue and one which requested them to attend only to the emotional cue. Based on our previous study using the same stimuli (Conty et al., 2012), we expected an early combination (rather than parallel processing) of the social cues that clearly portrayed a relevant threat for the observer within the first 200 ms. Yet how the precise temporality and the sources of this integration were influenced by task demand was addressed in a hypotheses-free manner.

**Methods**

*Participants*

Eighteen healthy subjects (10 females; mean age, 24.0 ± 0.5 years) participated in the EEG experiment. All participants were right-handed, with a normal or corrected vision, and were free of current or past psychiatric or neurological disorders. The subjects gave their written informed consent and were paid for their participation. The study was approved by the local ethics committee and was conducted in accordance with the Declaration of Helsinki.

*Stimuli*

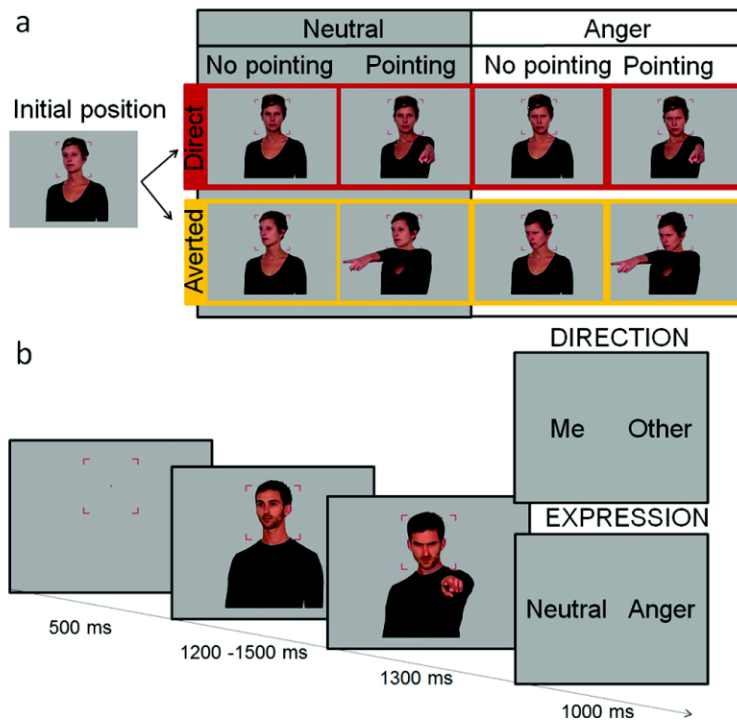
The stimuli used in the present experiment were developed by Conty and Grèzes (2011). They consisted of 192 photographs of 12 actors (six females) seen under 8 conditions: 2 emotions (anger/neutral) × 2 gaze

directions (head and eye gaze directed toward the participant or averted to the right and left side: direct/averted) × 2 gestures (pointing/no pointing) (Fig. 1a). An “initial position” photograph depicts each of the actors with a neutral expression, arms by their sides, and an intermediate head and eye direction of 15°. More detailed description of the stimuli can be found in Conty and Grèzes (2011).

*Procedure*

Each trial started with a fixation screen (500 ms) consisting of a central red fixation point and four red angles, displayed on a uniform gray background. The participant was instructed to fixate the central point and to keep his/her attention inside the fixation area at the level of the central point throughout the trial, avoiding eye blinks and saccades. Because dynamic social displays are more ecologically valid (Sato et al., 2010; Schilbach, 2010), we created an apparent movement by presenting two photographs, one after the other (Conty et al., 2007, 2012). The first photograph always displayed the actor in the initial position during a random period of time, ranging from 1200 to 1500 ms. The second one immediately followed and displayed the same actor in one of the eight conditions of interest (Fig. 1b) for 1300 ms. Throughout the trial, the actor’s face remained within the fixation area. Then, a response screen was presented for 1000 ms, followed by a black screen of 500 ms preceding the next trial. We delayed the responses to ensure that the participants responded at the end of the observation period.

The experiment was divided in 8 experimental blocks of 96 trials each. Two tasks alternated in separate blocks resulting in 4 blocks for each task. Participants were requested to focus either on the nature of



**Fig. 1.** Experimental procedure and stimuli examples. (a) From the initial position, gaze, emotion, and gesture were manipulated: stimuli displayed a direct or an averted gaze, an angry or a neutral expression, a pointing gesture or not. (b) Trial procedure: a central fixation area where the face of the stimuli later appears was presented for 500 ms. In the expression task blocs, subjects had to judge whether the actor displayed an angry or a neutral expression. In the direction task blocs, they had to judge whether the actor was addressing them or another.

the emotional expression (expression task) or the gaze direction of the actor (direction task). We chose to only include those two tasks based on the literature described in the introduction.

The expression task consisted in indicating the content of the actors' facial expressions during a response screen displaying the words "neutral" and "angry." In the direction task, participants had to indicate the gaze direction during a response screen now displaying the words "me" and "other." The location of the two words (on the right or left parts of the screen) was randomized (Fig. 1b). Varying the location of the words allowed preventing the participants from preparing their responses during the observation period. Therefore, our design was not appropriate for detecting any reaction time advantage in making judgments about gaze directions or emotions. During the EEG experiment, the participants perfectly performed the two tasks (expression task: mean correct response =  $96 \pm 0.8\%$  and direction task: mean correct response =  $98 \pm 0.4\%$ ), which assured that participants kept their attention on the stimuli.

To summarize, there were 16 conditions of interest in our experiment: 2 tasks (expression/direction)  $\times$  2 emotions (anger/neutral)  $\times$  2 gaze directions (direct/averted)  $\times$  2 gestures (pointing/no pointing).

#### Post-test

On the same day, just after the EEG experiment, participants had to perform a behavioral post-test. They were exposed to all stimuli using the same procedure described above. After each trial, they had to evaluate their felt self-involvement on a scale ranging from 0 to 9 (0, "not involved"; 9, "highly involved"). A repeated-measures ANOVA was performed on these subjective ratings, with emotion (anger/neutral), gaze direction (direct/averted), and gesture (pointing/no pointing) as within-subjects factors.

#### EEG data acquisition, processing, and analyses

The recording of the EEG activity was held at Centre de Neuroimagerie de Recherche (CENIR at Institut du Cerveau et de la Moelle Epinière; ICM) at a sampling rate of 1000 Hz using a BRAINAMP amplifier (Brain Products, BRAINAMP MR PLUS). Participants were equipped with an EEG cap of 63 sintered Ag/AgCl ring electrodes (EasyCap). The reference channel was placed on their nose, and a forehead ground was used. Impedances were kept under a threshold of 10 k $\Omega$ . EEG signal was low-pass filtered online at 250 Hz.

The raw EEG data were recalculated to average reference, down-sampled to 500 Hz, and low-pass filtered at 30 Hz (using EEGlab version 11 scm.ucsd.edu/eeGLab). Trials containing blinks and muscles artifacts were manually excluded. We computed the ERPs for each condition between 200 ms before and 800 ms after the onset of the second photograph and corrected using a baseline subtraction (from  $-200$  ms to stimulus onset).

Four classic ERP components were studied: the occipital P100, the temporal N170, the central P200 (Ashley et al., 2004; Vlamings et al., 2009) and a later component: the late positive potential (LPP) (Cuthbert et al., 2000; Schupp et al., 2000). For each ERPs of interest, we selected specific electrodes and time range on the peak amplitude observed on mean ERPs computed on all trials. For the P100, average activity on two occipito-parietal electrodes around the peak, between 92 and 116 ms, in each hemisphere was selected (O1/PO7, O2/PO8). For the N170, average activity on three electrodes around the peak, between 162 and 186 ms, in each hemisphere was selected (P7/TP7/TP9, P8/TP8/TP10). For the P200, average activity on six central electrodes around the peak, between 184 and 208 ms, was selected (C1/Cz/C2/CP1/CPZ/CP2). Finally, according to previous literature on the late positive potential (LPP; Schupp et al., 2000; Liu et al., 2012), LPP-related activity was measured by considering the mean activity averaged on five parietal electrodes (on which the LPP activity was maximal) between 300 and 700 ms (P3, P1, Pz, P2, and P4).

Repeated-measures ANOVA was performed on each ERP measure with task (expression/direction), emotion (anger/neutral), gaze direction (direct/averted), gesture (no pointing/pointing), and for bilateral ERP components (P100, N170) hemisphere (right/left), as within-subjects factors.

#### EEG source localization

Brainstorm was used for source reconstruction (Tadel et al., 2011, <http://neuroimage.usc.edu/brainstorm>). A source model that consisted in 15028 current dipoles was used to obtain a cortical current source density mapping per subject and condition. Dipole orientations were loosely constrained to the cortical mantle of a generic brain model taken from the standard Montreal Neurological Institute (MNI) template brain provided in Brainstorm. This head model was warped to the standard geometry of the EEG sensor cap taking into account a mean polhemus of 4 subjects. Computation of the forward model of EEG was done using the OpenMEEG BEM model (Kybic et al., 2005; Gramfort et al., 2010). For each subject and each condition, cortical current maps were computed from the ERPs of interest using the weighted minimum norm current estimate (WMNE).

First, to estimate the source of the integration of emotion, gaze, and gesture observed on the N170, irrespective of the tasks, we averaged the sources data on the time window of the N170 (162–186 ms) for each subject and condition and across tasks. Then, we applied a statistical mask on every vertex of the 15028 vertices of the model that reflects our parametric increase of self-relevance as a function of the number of self-oriented social cues. A vertex was thus considered as significant if (1) the activity of the most self-relevant condition (anger, direct, and pointing) was statistically different (*t*-tests) from the activity in all the other conditions and if (2) the activity in the two conditions with two self-directed cues (anger, direct, no pointing and neutral, direct, pointing) were statically different (*t*-tests) from the activity in the condition with one self-directed cue (neutral, direct, no pointing).

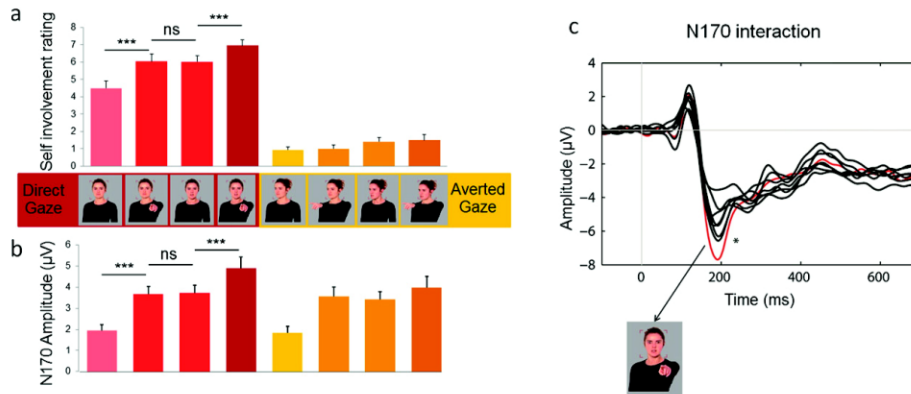
Second, to address selective attention effects on the sources of early social cues integration, we performed the same masked analysis as described above, but for each task separately. We created a conjunction map to identify neural sources that were common to both tasks, and then used this conjunction image as an exclusive mask to identify the sources that were specific to one or the other task.

For both analyses, regions representing a surface area superior to 1.5 cm<sup>2</sup> (and thus included more than 15 contiguous vertices), following the self-relevance pattern are reported and labeled on the basis of the correspondence between the Tzourio-Mazoyer brain atlas (Tzourio-Mazoyer et al., 2002), the brainstorm generic brain and source model implemented in Brainstorm software. For each region, we reported the number of vertices observed. We selected 3 regions: (1) one revealed across tasks, (2) one for the direction task, and (3) one for the expression task, and plotted for comparison the data for one representative vertex of those regions, for each task independently.

## Results

#### Behavioral post-test

As expected from our previous experiment (Conty et al., 2012), stimuli were considered more self-involving in the angry condition as compared with the neutral condition ( $F_{(1,17)} = 32.8, p < 0.001$ ), in the direct gaze condition as compared with the averted gaze condition ( $F_{(1,17)} = 160.9, p < 0.001$ ), and in the pointing condition as compared with the no pointing condition ( $F_{(1,17)} = 69.7, p < 0.001$ ). Interestingly, an interaction between emotion, gaze, and gesture was observed ( $F_{(1,17)} = 10.3, p < 0.01$ ). Post hoc comparisons revealed that the condition that was judged as more self-involving as compared to the other conditions (all  $t_{(1,17)} > 3$ , all  $p < 0.001$ ) was the condition where the 3 relevant cues were co-emitted (angry expression, direct gaze,



**Fig. 2.** Social cues interaction on the N170 as a function of self-relevance. a) Post-test self-involvement ratings on a scale from 0 to 9. Description of the conditions from left to right: 1. neutral, direct gaze, no pointing; 2. neutral, direct gaze, pointing; 3. angry, direct gaze, no pointing; 4. angry, direct gaze, pointing; 5. neutral, averted gaze, no pointing; 6. neutral, averted gaze, pointing; 7. angry, averted gaze, no pointing; 8. angry, averted gaze, pointing. b) Amplitude of the N170 activity averaged between 162 and 186 ms in microvolts. Same order of conditions as above. c) N170 interaction between emotion, gaze and gesture. Neural activity and self-involvement ratings increased with the number of self-oriented social signals. \* $p < 0.05$ ; \*\*\* $p < 0.001$ , ns nonsignificant.

and pointing gesture). Moreover, the self-involvement ratings increased with the number of self-directed social cues: it was higher when two self-directed social cues were present (direct gaze and anger or direct gaze and pointing) as compared to one self-directed cue (direct gaze) (all  $t_{(1,17)} > 5.7$ , all  $p < 0.001$ ), whereas no difference was observed between the two conditions of two self-directed social cues ( $t_{(1,17)} = 0.09$ ,  $p = 0.92$ ) (Fig. 2a).

#### Time course of social cues processing

We addressed when emotion, gaze, and gesture were integrated and whether this integration varied with task demand. On P100 activity, we observed main effects of emotion and gesture. Anger ( $F_{(1,17)} = 17.7$ ,  $p < 0.001$ ) and pointing ( $F_{(1,17)} = 17.3$ ,  $p < 0.001$ ) induced greater positive activity than neutral and no-pointing conditions, respectively (see Fig. S1). No other main effects or interaction between factors were observed (all  $F < 1$ , all  $p > 0.2$ ).

On the temporal N170, main effects of hemisphere, emotion, gaze, and gesture were detected. The activity was more important in the right than in the left hemisphere ( $F_{(1,17)} = 6.69$ ,  $p < 0.05$ ). Anger ( $F_{(1,17)} = 128.04$ ,  $p < 0.001$ ), direct gaze ( $F_{(1,17)} = 6.17$ ,  $p < 0.05$ ), and pointing ( $F_{(1,17)} = 24.99$ ,  $p < 0.001$ ) conditions generated greater negative peak amplitudes than, respectively, neutral, averted gaze, and no-pointing conditions. There was an interaction between emotion and gaze ( $F_{(1,17)} = 6.69$ ,  $p < 0.05$ ), an interaction between emotion and gesture ( $F_{(1,17)} = 16.3$ ,  $p < 0.001$ ), and more importantly, an interaction between emotion, gaze, and gesture ( $F_{(1,17)} = 4.59$ ,  $p < 0.05$ ). This interaction was explained by an interaction between gaze and gesture only in the anger conditions ( $F_{(1,17)} = 4.537$ ,  $p = 0.048$ ) but not in the neutral conditions ( $F_{(1,17)} < 0.001$ ,  $p = 0.991$ ), strongly supporting the idea that there was integration between the three co-emitted cues rather than additive effects only. The peak amplitude of the most self-relevant condition (anger, direct gaze, and pointing) was greater than all other conditions (all  $t_{(1,17)} > 3.3$ , all  $p < 0.01$ ) (Fig. 2b,c). In addition, just as in the post-test ratings, N170 activity increased with the number of self-directed social cues: activity was greater when two self-directed social cues were present (direct gaze and anger or direct gaze and pointing) as compared to one self-directed cue (direct gaze) ( $t_{(1,17)} = 6$ ,  $p < 0.001$ ), whereas no difference was observed between the two conditions of two self-directed social cues ( $t_{(1,17)} = 0.47$ ,  $p = 0.65$ ) (Fig. 2b).

Analysis ran on the central P200 revealed a main effect of emotion and gaze. Anger ( $F_{(1,17)} = 8.9$ ,  $p < 0.01$ ) and averted gaze ( $F_{(1,17)} = 5.1$ ,  $p < 0.05$ ) elicited greater positive activity than, respectively, neutral, and direct gaze conditions (see Fig. S1). No other main effects or interactions between factors were observed (all  $F < 1$ , all  $p > 0.2$ ). Although Conty et al. (2012) found a triple interaction on the P200, our results are not surprising as the triple interaction in this study is already observed on the N170. Greater activity for averted gaze as compared to direct gaze on the P200 replicates the findings of Rigato et al. (2010).

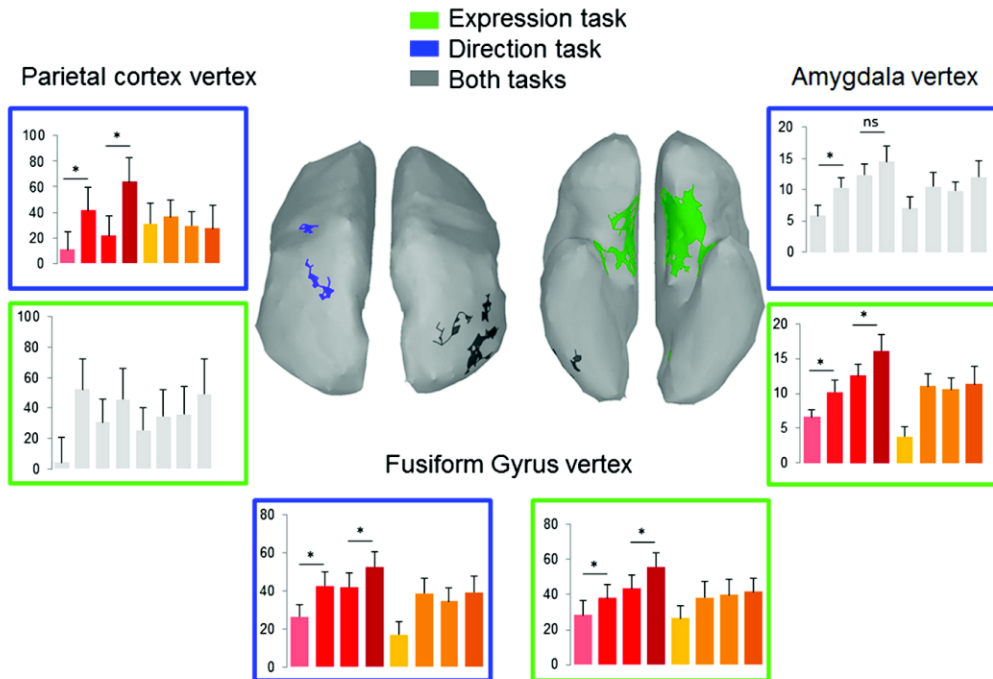
On the late positive potential (LPP), we observed main effects of task and emotion. Anger ( $F_{(1,17)} = 24$ ,  $p < 0.001$ ) induced greater activity as compared to neutral condition. Expression task ( $F_{(1,17)} = 19.4$ ,  $p < 0.001$ ) induced greater activity as compared to direction task (see Fig. S1). An interaction between gaze and task was also observed ( $F_{(1,17)} = 4.7$ ,  $p < 0.05$ ) and was justified by a main effect of gaze only in the direction task ( $F_{(1,17)} = 6$ ,  $p < 0.05$ ). In this task, direct gaze induced greater activity as compared to averted gaze, whereas no such difference was observed in the Expression task (see Fig. S1). These results suggest that, at these later temporal stages, the emotional content of the faces may be always processed in both tasks whereas their gaze direction may only be processed in the direction task. The present main effect of emotion on the late positive potential (LPP) is consistent with previous results that involve the LPP in explicit emotional processing (Cutthbert et al., 2000) and task switching (Rushworth et al., 2002; Astie et al., 2008).

Crucially, there was no main effect of task before the LPP (all  $F < 1.52$ , all  $p > 0.23$ ) nor interaction of the task with the other factors (all  $F < 4.2$ , all  $p > 0.05$ ), suggesting at first view that the integration between gaze, emotion, and gesture during the first 200 ms after stimulus onset are independent of selective attention related to the task.

Finally, we tested whether there were between-subjects correlations between the peak amplitudes of the N170 and the behavioral self-involvement ratings (for all the different conditions separately as well as for the differences of interest between conditions). No significant effects emerged (all  $r < 0.4$ , all  $p > 0.1$ ).

#### Sources reconstruction

We then explored the brain sources of early integration of emotion, gaze, and gesture observed on the N170 (see Fig. 3) and whether the sources of this integration were influenced by task demand.



**Fig. 3.** Neural sources of the N170 interaction. Regions that reflect the parametric increase of self-relevance with the number of self-oriented social signals. Both tasks (grey) involved temporal and occipital regions, including the fusiform gyrus, whereas the Expression task (green) essentially involved regions in the limbic system, including the amygdala, and the direction task (blue) regions in the dorsal pathway, including the parietal. For the three-mentioned example, the activity of a representative vertex is plotted (p.A.m) for the direction task (blue) and for the expression task (green) separately for comparison. The order of the conditions for each plot is similar to the one of Fig. 2. \* $p < 0.05$ ; ns, nonsignificant.

First, we estimated these sources of early social cues integration across tasks. The masked source reconstruction analysis revealed that both structures of the limbic system (such as the amygdala, the hippocampus, the cingulate cortex, and the ventromedial prefrontal cortex), previously reported in experiments using explicit emotional tasks (Sato et al., 2004, 2010; Sander et al., 2007; N'Diaye et al., 2009; Adams et al., 2012), and the dorsal pathway including the parietal and premotor cortex, reported by Conty et al. (2012) that used a gaze direction task, are involved in the early integration of emotion, gaze, and gesture and showed an increase of activity with the increase of self-relevant social cues (same pattern as observed on the N170).

Second, as the literature pointed toward inconsistent neural sources of (a) gaze and emotion combination and (b) of the N170 (Itier and Taylor, 2004; Eimer, 2011), and even though there was no statistical interaction on the N170 activity (evoked potential analysis) between the three manipulated social cues (emotion, gaze, and gesture) and the task factor, we tested whether the sources of the N170 varied as a function of the ongoing task. To do so, we performed the above-mentioned masked analysis for each task separately. When participants attended to the emotional expressions, the amygdala, the hippocampus, and the subgenual cingulate area were involved in the integration of social cues (see Table 1 for a full description of the results), whereas the post-central gyrus and the inferior and superior parietal cortex were detected when participants attended to gaze direction. In addition to these task-specific sources, the fusiform gyrus, middle occipital gyrus, inferior occipital gyrus, and middle temporal gyrus which most probably involved the superior temporal sulcus (STS) were involved in both tasks (as revealed by the conjunction analysis) (Fig. 3).

**Table 1**

Number of vertices ( $> 15$ , surface area  $> 1.5 \text{ cm}^2$ ) for each region reported in source localization analysis based on the correspondence between the Tzourio–Mazoyer brain atlas, the Brainstorm generic brain, and source model implemented in Brainstorm software. R/L: R = right hemisphere, L = left hemisphere. E/D: E = regions involved mainly in the expression task, D = regions involved mainly in the direction task.

Task	Hemisphere	Brain regions	Vertices
E/D	R	Fusiform	21
E/D	R	Middle occipital	36
E/D	R	Middle temporal	103
E/D	R	Inferior occipital	36
D	L	Post-central	18
D	L	Inferior/superior parietal	36
E	R/L	Subgenual cingulate	34/23
E	R/L	Rectus	16/35
E	R/L	Hippocampus	41/38
E	R/L	Amygdala	23/26
E	R/L	Caudate	80/64
E	L	Inferior frontal	23
E	L	Thalamus	16
E	L	Superior temporal pole	16
E	L	Calcarine fissure	16

## Discussion

The aim of the present study was to address whether, when and how, directing one's own attention towards a specific social signal emitted by another person impacts on its neural integration with other co-emitted social signals. Our critical finding was that the early integration of social

visual cues is indexed by a temporal marker (N170) that is independent of the attended feature, yet being dependent on the relevance of the combination. While the underlying neural sources of the N170 included both limbic and dorsal pathway structures, their relative contribution was modulated as a function of the attended feature: integration during emotion judgments was mainly implemented in classic limbic areas but in the dorsal pathway during gaze direction judgments.

The amplitude of the N170 and associated neural sources, as well as of the self-involvement ratings, are maximum when the combination of the co-emitted social cues clearly portrayed a threat for the observer, i.e. when the actor expressed anger, looked, and pointed toward the participant, in agreement with previous findings (Klucharev and Sams, 2004; Akechi et al., 2010; Rigato et al., 2010; Conty et al., 2012). The statistical interactions between the three cues in one hand and gaze and gesture within anger conditions only on the other hand suggest that the three co-emitted cues are merged into a single percept rather than triggering additive effects only. These results substantiate that our sensory system is not globally adapted to the statistics of all natural stimuli but rather evolution has tailored our brain to be specifically optimized to biologically relevant stimuli (Machens et al., 2005).

The temporal marker of the integration of emotion, gaze, and gesture emerges 170 ms after stimulus onset. Importantly, this temporal marker emerges even if the integration of the three visual cues was not required to perform the task, as only one feature at a time was task relevant. To our knowledge, this is the first study that directly compares the impact of two tasks on the spatiotemporal course of early integration of social cues. This allowed to reveal that early integration appears to be temporally task-related attention *independent* and thus contradicts the view that orienting attention to only one feature at a time results in processing co-emitted facial signals independently (Pourtois et al., 2004; Klucharev and Sams, 2004; Bindemann et al., 2008). A general attention effect (rather than a selective attention effect) may explain whether this marker emerges within the first 200 ms or later: indeed, when participants are requested to attend to the stimuli by judging either the emotional expressions (present study, Akechi et al., 2010) or the gaze direction (Conty et al., 2012), the temporal marker of social cues integration emerges earlier (between 170 and 200 ms) than when participants observe passively (Rigato et al., 2010) or perform an oddball task (starting at 190 and lasting till 350 ms (Dumas et al., 2013, but see exception by Klucharev and Sams, 2004).

The integration of social visual cues recruited the posterior occipital cortex, the fusiform gyrus (FG), and the superior temporal sulcus (STS) independently of task demand. In functional imaging studies, those areas have been frequently associated with face perception (Kanwisher et al., 1997; Puce et al., 1998; George et al., 1999), notably, the FG preferentially processes the invariant aspects of the face while the STS processes changeable aspects such as the gaze and the mouth (Haxby et al., 2000; Hoffman and Haxby, 2000). Moreover, EEG source analyses have either reported the STS (Batty and Taylor, 2003; Henson et al., 2003; Itier et al., 2006; Itier and Taylor, 2004; Watanabe et al., 2003) or the FG (Itier and Taylor, 2002; Rossion et al., 2003; Watanabe et al., 2003) as the main source of the N170. Here, we further showed that these classic neural sources of the N170 (Bötzel et al., 1995; Rossion et al., 2003; Conty et al., 2007) are involved in the early integration of facial expression with co-emitted gaze direction and gesture, independently of the attended social cue.

Several limbic areas as well as dorsal pathway structures were also found to be engaged during the early integration of relevant social cues. Yet in contrast to the above-mentioned areas, the contribution of these specific neural sources appears to be modulated as a function of the attended feature: predominantly sustained by classic limbic areas during emotion judgments, the integration of social cues involved the dorsal pathway during gaze direction judgments. Although most of the brain regions were revealed in previous fMRI studies on gaze-emotion integration (Hadjikhani et al., 2008; N'Diaye et al., 2009; Sato et al., 2010), we further demonstrated, using two tasks, that when

participants were requested to attend to the emotional content of the faces, the early integration of emotion, gaze, and gesture took mainly place in the temporal pole, including the amygdala–hippocampal region and the subgenual orbital cortex, brain areas known to be involved in emotional processing (Phillips et al., 2003; Höistad and Barbas, 2008; Fusar-Poli et al., 2009), notably explicit processing (Gur et al., 2002; Keightley et al., 2003; Habel et al., 2007). Nonetheless, these regions were less evidenced in the direction task mainly associated with the post-central and the parietal cortex involved in gaze perception (Puce et al., 1998; Hoffman and Haxby, 2000). Of interest, using fMRI, Zaki et al. (2012) similarly showed the involvement of specific neural networks as a function of the perceiver's goals, i.e., to attend to the gaze direction or to the emotional state of the actor. Here, we further revealed that the contribution of the limbic system to the integration of social cues was enhanced when participants explicitly paid attention to emotional cues but decreased when they paid attention to gaze direction while the contribution of the dorsal pathway increased.

The present study thus reveals that the integration of social cues and not simply the perception of one or the other cues recruits a large brain network around 170 ms after stimulus onset, which activity is partly modulated by task demand. Still, the question remains as to whether and how biologically relevant information flows within the whole network at a more fine-grained time scale. From the data of macaque monkeys studies, one could speculate that the amygdala may initiate the integration and enhance sensory processing of biologically relevant signals thanks to direct projections to the ventral visual pathway including the FG and the STS (Hadj-Bouziane et al., 2012) and premotor cortex in the dorsal pathway (Avendano et al., 1983). To conclude, by addressing selective attention effects on the spatiotemporal course of social cues integration, we were able to show an influence of task demand on the “where” component rather than on the “when” component. Moreover, our results demonstrate that all social cues are encoded even when they are irrelevant to the task, as long as they are relevant to the observer.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2014.10.063>.

#### Acknowledgments

We wish to warmly thank Florence Bouchet and Laurent Hugueville for their useful technical help. The research was supported by ANR-11-EMCO-00902, ANR-11-0001-02 PSL\*, ANR-10-LABX-0087, ANR-10-IDEX-0001-02 PSL and INSERM.

#### References

- Adams Jr., R.B., Franklin Jr., R.G., Kveraga, K., Ambady, N., Kleck, R.E., Whalen, P.J., Hadjikhani, N., Nelson, A.J., 2012. Amygdala responses to averted vs direct gaze fear vary as a function of presentation speed. *Soc. Cogn. Affect. Neurosci.* 7, 568–577.
- Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., Hasegawa, T., 2010. The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: an ERP study. *Neuropsychologia* 48, 2841–2851.
- Ashley, V., Vuilleumier, P., Swick, D., 2004. Time course and specificity of event-related potentials to emotional expressions. *Neuroreport* 15, 211–216.
- Astle, D.E., Jackson, G.M., Swinson, R., 2008. Fractionating the cognitive control required to bring about a change in task: a dense-sensor event-related potential study. *J. Cogn. Neurosci.* 20, 255–267.
- Avendano, C., Price, J.L., Amaral, D.G., 1983. Evidence for an amygdaloid projection to premotor cortex but not to motor cortex in the monkey. *Brain Res.* 264, 111–117.
- Batty, M., Taylor, M.J., 2003. Early processing of the six basic facial emotional expressions. *Brain Res. Cogn. Brain Res.* 17, 613–620.
- Bindemann, M., Mike Burton, A., Langton, S.R.H., 2008. How do eye gaze and facial expression interact? *Vis. Cogn.* 16, 708–733.
- Bötzel, K., Schulze, S., Stodieck, S.R., 1995. Scalp topography and analysis of intracranial sources of face-evoked potentials. *Exp. Brain Res.* 104, 135–143.
- Conty, L., Grèzes, J., 2011. Look at me, I'll remember you: the perception of direct gaze enhances memory and right hippocampal activity. *Hum. Brain Mapp.* 33, 2428–2440.
- Conty, L., N'Diaye, K., Tijus, C., George, N., 2007. When eye creates the contact: ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia* 45, 3024–3037.
- Conty, L., Dezeache, G., Hugueville, L., Grèzes, J., 2012. Early binding of gaze, gesture, and emotion: neural time course and correlates. *J. Neurosci.* 32, 4531–4539.



- Cuthbert, B.N., Schupp, H.T., Bradley, M.M., Birbaumer, N., Lang, P.J., 2000. Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol. Psychol.* 52, 95–111.
- Dumas, T., Dubal, S., Attal, Y., Chupin, M., Jouvent, R., Morel, S., George, N., 2013. MEG evidence for dynamic amygdala modulations by gaze and facial emotions. *PLoS One* 8, e74145.
- Eimer, M., 2011. The face-sensitive N170 component of the event-related brain potential. In: Rhodes, G., Calder, A., Johnson, M., Haxby, J.V. (Eds.), *Oxford Handbook of Face Perception*. Oxford University Press.
- Fusar-Poli, P., Piacentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., Benedetti, F., Abbamonte, M., Gasparotti, R., Barale, F., Perez, J., McGuire, P., Politi, P., 2009. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* 34, 418–432.
- George, N., Dolan, R.J., Fink, G.R., Baylis, G.C., Russell, C., Driver, J., 1999. Contrast polarity and face recognition in the human fusiform gyrus. *Nat. Neurosci.* 2, 574–580.
- Graham, R., Labar, K.S., 2012. Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention. *Neuropsychologia* 50, 553–566.
- Gramfort, A., Papadopoulos, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* 9, 45.
- Gur, R.C., Schroeder, L., Turner, T., McGrath, C., Chan, R.M., Turetsky, B.L., Alsop, D., Maldjian, J., Gur, R.E., 2002. Brain activation during facial emotion processing. *NeuroImage* 16, 651–662.
- Habel, U., Windischberger, C., Derrdt, B., Robinson, S., Kryspin-Exner, I., Gur, R.C., Moser, E., 2007. Amygdala activation and facial expressions: explicit emotion discrimination versus implicit emotion processing. *Neuropsychologia* 45, 2369–2377.
- Hadi-Bouziane, F., Liu, N., Bell, A.H., Gothard, K.M., Luh, W.-M., Tootell, R.B.H., Murray, E.A., Ungerleider, L.G., 2012. Amygdala lesions disrupt modulation of functional MRI activity evoked by facial expression in the monkey inferior temporal cortex. *Proc. Natl. Acad. Sci.* 109, E3640–E3648.
- Hadjikhani, N., Hoge, R., Snyder, J., de Gelder, B., 2008. Pointing with the eyes: the role of gaze in communicating danger. *Brain Cogn.* 68, 1–8.
- Haxby, Hoffman, Gobbini, 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Henson, R.N., Goshen-Gottstein, Y., Ganel, T., Otten, L.J., Quayle, A., Rugg, M.D., 2003. Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb. Cortex* 13, 793–805.
- Hoffman, E.A., Haxby, J.V., 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat. Neurosci.* 3, 80–84.
- Hölstad, M., Barbas, H., 2008. Sequence of information processing for emotions through pathways linking temporal and insular cortices with the amygdala. *NeuroImage* 40, 1016–1033.
- Itier, R.J., Taylor, M.J., 2002. Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *NeuroImage* 15, 353–372.
- Itier, R.J., Taylor, M.J., 2004. Source analysis of the N170 to faces and objects. *Neuroreport* 15, 1261–1265.
- Itier, R.J., Latinus, M., Taylor, M.J., 2006. Face, eye and object early processing: what is the face specificity? *NeuroImage* 29, 667–676.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci. Off. J. Soc. Neurosci.* 17, 4302–4311.
- Keightley, M.L., Winocur, G., Graham, S.J., Mayberg, H.S., Hevenor, S.J., Grady, C.L., 2003. An fMRI study investigating cognitive modulation of brain regions associated with emotional processing of visual stimuli. *Neuropsychologia* 41, 585–596.
- Klucharev, V., Sams, M., 2004. Interaction of gaze direction and facial expressions processing: ERP study. *Neuroreport* 15, 621–625.
- Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R., Papadopoulos, T., 2005. A common formalism for the integral formulations of the forward EEG problem. *IEEE Trans. Med. Imaging* 24, 12–28.
- Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., Ding, M., 2012. Neural substrate of the late positive potential in emotional processing. *J. Neurosci.* 32, 14563–14572.
- Machens, C.K., Gollisch, T., Kolesnikova, O., Herz, A.V.M., 2005. Testing the efficiency of sensory coding with optimal stimulus ensembles. *Neuron* 47, 447–456.
- N'Diaye, K., Sander, D., Vuilleumier, P., 2009. Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emotion* 9, 798–806.
- Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R., 2003. Neurobiology of emotion perception I: the neural basis of normal emotion perception. *Biol. Psychiatry* 54, 504–514.
- Pourtois, G., Sander, D., Andres, M., Grandjean, D., Reveret, L., Olivier, E., Vuilleumier, P., 2004. Dissociable roles of the human somatosensory and superior temporal cortices for processing social face signals. *Eur. J. Neurosci.* 20, 3507–3515.
- Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.
- Rigato, S., Farroni, T., Johnson, M.H., 2010. The shared signal hypothesis and neural responses to expressions and gaze in infants and adults. *Soc. Cogn. Affect. Neurosci.* 5, 88–97.
- Rossion, B., Joyce, C.A., Cottrell, G.W., Tarr, M.J., 2003. Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage* 20, 1609–1624.
- Rustiworth, M.F.S., Passingham, R.E., Nobre, A.C., 2002. Components of switching intentional set. *J. Cogn. Neurosci.* 14, 1139–1150.
- Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., Scherer, K.R., 2007. Interaction effects of perceived gaze direction and dynamic facial expression: evidence for appraisal theories of emotion. *Eur. J. Cogn. Psychol.* 19, 470–480.
- Sato, W., Yoshikawa, S., Kochiyama, T., Matsumura, M., 2004. The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *NeuroImage* 22, 1006–1013.
- Sato, W., Kochiyama, T., Uono, S., Yoshikawa, S., 2010. Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *NeuroImage* 50, 1658–1665.
- Schilbach, L., 2010. A second-person approach to other minds. *Nat. Rev. Neurosci.* 11, 449.
- Schupp, H.T., Cuthbert, B.N., Bradley, M.M., Cacioppo, J.T., Ito, T., Lang, P.J., 2000. Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology* 37, 257–261.
- Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011, 879716.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15, 273–289.
- Vlamings, P.H.J.M., Goffaux, V., Kemner, C., 2009. Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *J. Vis.* 9, 12.1–12.13.
- Watanabe, S., Kalkigi, R., Puce, A., 2003. The spatiotemporal dynamics of the face inversion effect: a magneto- and electro-encephalographic study. *Neuroscience* 116, 879–895.
- Zaki, J., Weber, J., Ochsner, K., 2012. Task-dependent neural bases of perceiving emotionally expressive targets. *Front. Hum. Neurosci.* 6, 228.

The main message of Experiment 1 was that social signals are integrated in the human brain within 170 ms after stimulus onset as long as they are contextually relevant to the observer, even if not relevant to the task at hand. Selective attention to social cues however modulated the relative contribution of feature selective regions: while emotion-related regions were mainly involved during emotion judgments, the dorsal pathway was preferentially recruited during gaze direction judgments. Attention thus enhances the weight of social sensory processing in regions specific to the attended cue. The dissociation between the timing and the regions involved is interesting in regards to the nature of the stimuli used: indeed, while the rapid detection of relevant social information (here threatening) is essential for proper adaptive reactions and survival, the contribution of different brain regions to this integration can be variable.

Classical EEG analyses and source reconstruction showed that there is a neural marker at 170 ms that varies parametrically as a function of relevance to the observers integrating emotion, gaze and gesture independently from attentional demands. While the findings of this study replicate and highlight the fact that different social cues mutually influence each other when social signals are perceived, they do not provide the mechanisms by which they do so. Moreover, although gesture was included in the design in addition to gaze and emotion, to parametrically modulate the relevance of the cues to the observers, only anger emotion was manipulated. As described in chapter 2 section 3, the interaction between emotion and gaze would be more noticeable if both fear and anger are put together, as gaze direction has an opposite effect on their perception (facilitated perception of anger with a direct gaze and fear with an averted gaze). To address these last two issues, the next experiment consisted of stimuli expressing both anger and fear with direct and averted gaze, and the emotional expressions were morphed allowing the use of model-based behavioral and EEG analysis that provide a clearer window into the mechanisms underlying gaze and emotion interactions.

## Part 2: Gaze direction as a contextual factor during emotion categorization (Experiment 2)

In this experiment, we aimed at characterizing the effects underlying the contextual influence of gaze direction on emotion categorization under social threat. Here we took advantage of model-based analyses to distinguish between different accounts of contextual gaze influences, mediated either by a multiplicative change in the sensitivity to the displayed emotion or by an additive change in the decision criterion (bias) toward one or the other emotion. This novel analysis framework allows building on previous studies which reported differences in categorization performance between conditions of interest but could not explain *how* and *why* these differences occurred. To be able to apply quantitative models of decision-making, we created ‘morphed’ emotional expressions in order to parametrically control the amount of evidence toward an emotion or another.

Studies exploring perceptual decisions on ambiguous emotional expressions used conventional analyses described in Figure 21a, by relating either stimulus to choice using psychophysics, stimulus to neural activity using regression analyses, or neural activity to choice using ‘decoding’ techniques. Importantly, none of these studies manipulated contextual information such as gaze direction. For example, multivariate pattern analyses on morphs between fearful and angry avatar facial emotions revealed that the right STS sustains categorical representations of emotional expressions whereas the left STS carries graded representations (Said et al., 2010). In other studies using decoding techniques (Pessoa and Padmala, 2005, 2007), choice-related activity was identified in the prefrontal cortex and insula during a near-threshold fear detection task, and a broader network including temporal and parietal regions using morphs from fear to disgust (Pessoa and Padmala, 2005, 2007; Thielscher and Pessoa, 2007). However, choice-related activity was computed by considering neutral face trials only (a choice decoding method often referred to as of ‘constant stimulus’), and distinguished from stimulus-related activity which was computed for morphs at the highest emotional intensity only. Thus, the intermediate level of morphs that are most informative for the understanding of how the brain uses in-

creasing evidence to make decisions on ambiguous emotional stimuli were not exploited.

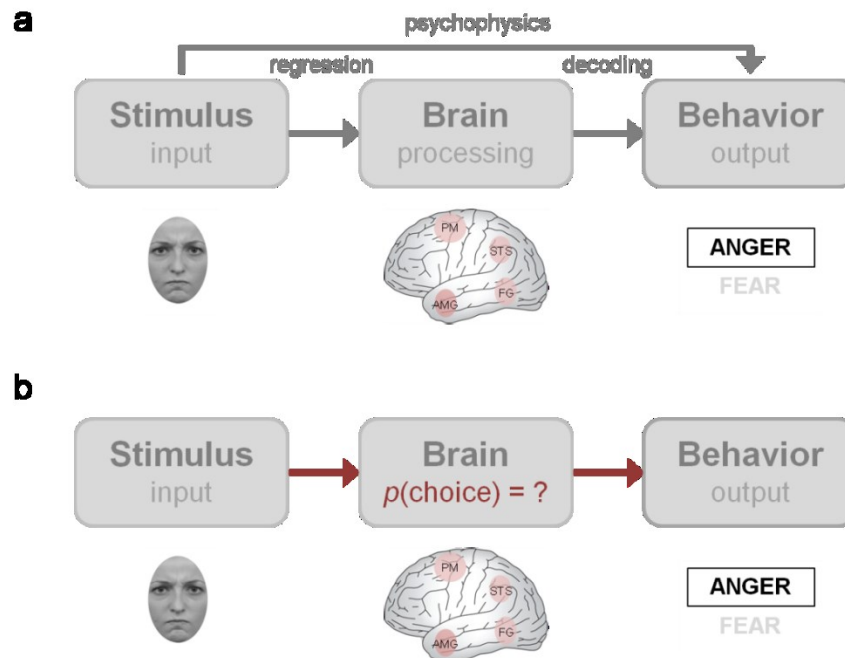


Figure 21. a) Conventional analyses establish pair-wise relations between stimulus, brain and behavior. b) Model-based single trial analyses relate all three sources of variability.

Here we use decision-theoretical models in the aim of reaching a comprehensive mechanistic understanding of how contextual factors influence social perceptual decisions by relating all three sources of variability: variability in stimulation, neural activity and the resulting behavior (Figure 21b). In practice, we apply SDT analyses at the level of brain through the regression of neural signals against model-based variables of interest. Single-trial analyses consider the neural variability as a tool (rather than a nuisance) and permit to relate the trial-to-trial fluctuations in neural activity to behavioral outcomes by assessing whether brain signals co-vary with choice parameters (sensitivity or bias within the SDT framework).

This work will be presented in the form of a manuscript that is now under review, but I will first report the results of pretests performed to calibrate the stimuli that

were less detailed in the manuscript, as these stimuli will be used in all the forthcoming experiments.

### **Stimuli and pretest calibration results**

Stimuli consisted of 36 identities (18 females) adapted from the Radboud Faces Database (Langner et al., 2010). They varied in terms of displayed emotion (neutral, angry or fearful expressions) and gaze direction (direct toward the participant or averted 45° to the left or right). Using Adobe Photoshop CS5.1 (Adobe Systems, San Jose CA), faces were modified to remove any visible hair, resized and repositioned so that eyes, nose and mouth appeared within the same circumference. All images were converted to greyscale and cropped into a 280 x 406 pixel oval centered within a 628 x 429 pixel black rectangle.

To vary the intensity of emotional expressions, faces were morphed from neutral to angry expressions and from neutral to fearful expression using FantaMorph ([www.fantamorph.com](http://www.fantamorph.com)). At first, 7 levels of morphs from neutral to angry expressions and from neutral to fearful expressions (separately for direct and averted gaze stimuli) were created using a simple linear morphing transformation. This resulted in 30 conditions for each identity: 7 levels of morphs \* 2 emotions \* 2 gaze directions = 28 + 2 'neutral' (emotionless stimuli) with direct and averted gaze.

To calibrate the intensity ratings of anger and fear emotions, a pre-test was performed on 19 subjects (9 females, mean age, 24.7 ± 0.9 years). We considered essential to have the same degree of uncertainty at all degrees of emotion intensity for both angry and fearful stimuli. All stimuli were presented for 250 ms and participants rated the emotional intensity perceived on a continuous scale from “not at all intense” to “very intense” using a mouse device (with a maximum of 3 seconds to respond). A main effect of emotion ( $F_{(1,18)}=50.5$ ,  $p<0.001$ ) on perceived intensity ratings was observed, fearful stimuli being judged as more intense than angry stimuli. To adjust for this difference between emotions, we linearized the curve of perceived intensities of anger as a function of the levels of the morphs (Figure 22a). We then extracted the percentage values of morphs of angry and fearful faces needed such

as the perceived intensities fall into this linearized curve. New morphs were correspondingly regenerated using FantaMorph, and another test was carried out on 10 new subjects (4 females, mean age  $24.1 \pm 1.9$ ) to validate the method used. As expected, perceived intensities of the adjusted stimuli did not vary between emotions anymore ( $p > 0.9$ ) (Figure 22b).

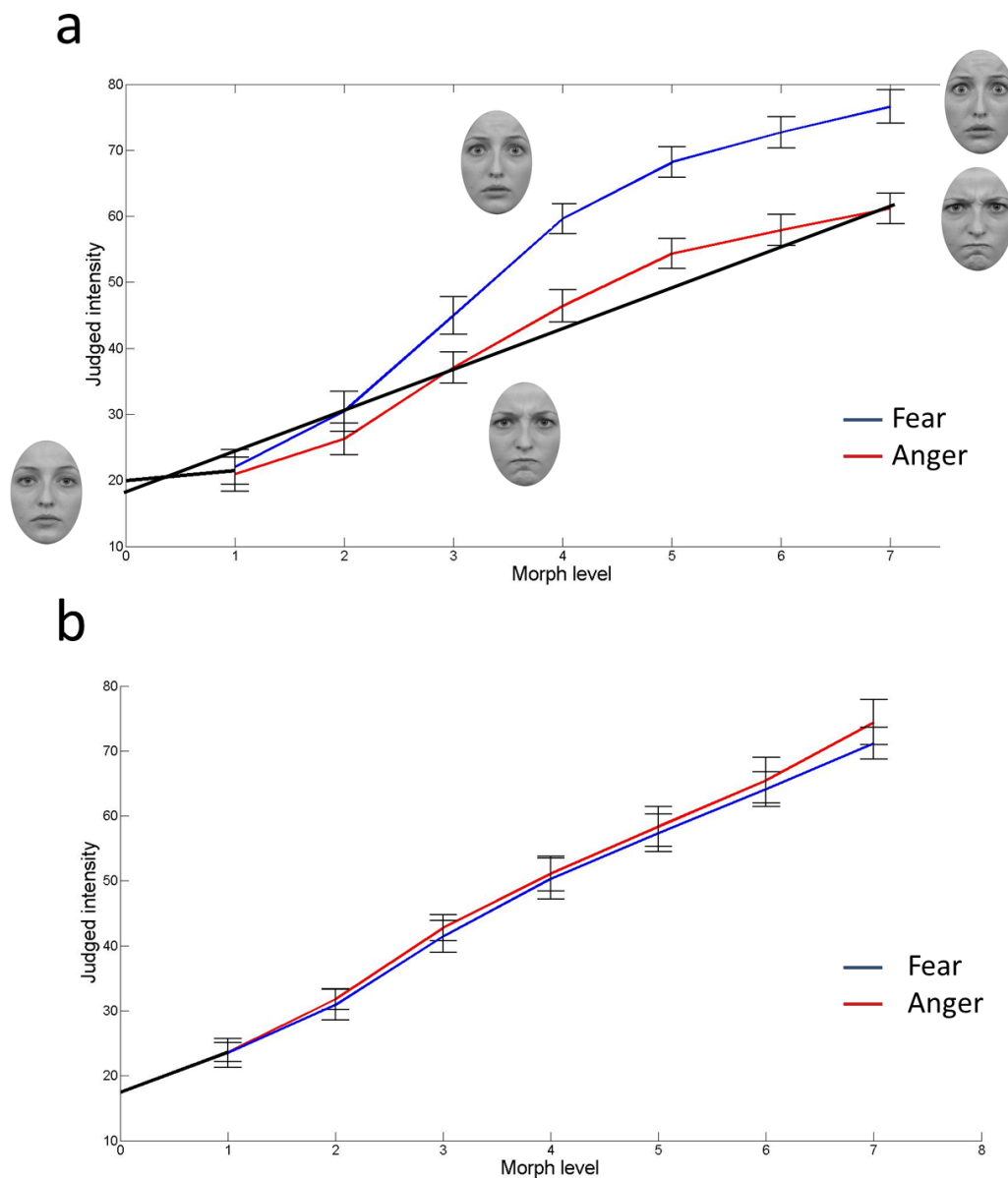


Figure 22 Pre-test of experiment 2: Intensity judgments for angry and fearful stimuli before (a) and after calibration (b). In a, the black line represents the linearized curve based on which morphs of anger and fear were adjusted.

Experiment 2: paper under review

***Anxiety dissociates the adaptive functions of sensory and motor response enhancements to social threats***

# **Anxiety dissociates the adaptive functions of sensory and motor response enhancements to social threats**

Marwa El Zein <sup>1,2,\*</sup>, Valentin Wyart <sup>1,‡</sup>, and Julie Grèzes <sup>1,‡</sup>

<sup>1</sup> Laboratoire de Neurosciences Cognitives, Inserm unit 960, Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, Paris, France

<sup>2</sup> Université Pierre et Marie Curie – Paris 6, Paris, France

\* to whom correspondence should be addressed ([marwa.el.zein@ens.fr](mailto:marwa.el.zein@ens.fr))

‡ shared senior authorship

Keywords: anxiety; threat; emotion; decision-making; computational model

Main text (Introduction, results and discussion): approx. 4600 words.

Methods: approx. 2400 words.

Figures: 7



## **Abstract**

Efficient detection and reaction to negative signals in the environment is essential for survival. In social situations, these signals are often ambiguous and can imply different levels of threat for the observer, thereby making their recognition susceptible to contextual cues – such as gaze direction when judging facial displays of emotion. However, the mechanisms underlying such contextual effects remain poorly understood. By computational modeling of human behavior and electrical brain activity, we demonstrate that gaze direction enhances the perceptual sensitivity to threat-signaling emotions – anger paired with direct gaze, and fear paired with averted gaze. This effect arises simultaneously in ventral face-selective and dorsal motor cortices as a function of individual anxiety at 200ms following face presentation, and does not reflect increased attention to threat-signaling emotions. These findings reveal that threat tunes neural processing in fast, selective, yet attention-independent fashion in both sensory and motor systems, for different adaptive purposes.

## Introduction

Perceptual decisions rely on the combination of weak and/or ambiguous samples of sensory evidence. The accuracy of this decision process is particularly important for the interpretation of negative signals, which require rapid and adaptive responses. In the social domain, identifying the emotional state of a conspecific – e.g., is he/she angry or afraid? – rarely depends solely on facial features, which are usually ambiguous and can imply different levels of threat for the observer. Surrounding cues, such as gaze direction and body posture, are known to act as contextual information during emotion recognition (Righart and de Gelder, 2008; Barrett and Kensinger, 2010; Aviezer et al., 2011). Specifically, the detection of anger represents an immediate threat for the observer when paired with a direct gaze; by contrast, it is when paired with an averted gaze that fear marks the presence (and possibly the localization) of a threat in the environment (Sander et al., 2007). These threat-signaling combinations of gaze direction and emotion have been shown to be better recognized and rated as more intense than other combinations (Adams and Kleck, 2003, 2005; Graham and LaBar, 2007; Sander et al., 2007; Bindemann et al., 2008; N'Diaye et al., 2009), and this as a function of anxiety level of the individuals (Ewbank et al., 2010). However, the computational mechanisms underlying the prioritization of threat-signaling information remain unspecified.

Classical theories distinguish two classes of mechanisms by which contextual information such as gaze direction could influence the recognition of negative emotions. Gaze direction could *bias* the interpretation of negative facial expressions in favor of the emotion signaling higher threat in this context – anger for direct gaze, fear for averted gaze. In signal detection theoretical terms (Green and Swets, 1966; Macmillan and Creelman, 2004), this effect would correspond to a additive shift of the decision criterion as a function of gaze direction. However, gaze direction could also increase the perceptual *sensitivity* to the facial features diagnostic of the emotion signaling higher threat. In contrast to the first account, this effect would correspond to a multiplicative boost of threat-signaling cues in the decision process. While the two accounts predict similar effects of gaze direction on the recognition of threat-

signaling emotions, a bias effect would be maximal for neutral (emotionless) expressions, whereas a sensitivity effect would be maximal at low emotion strengths.

Here we arbitrated between these two competing accounts by recording human electroencephalogram (EEG) signals while participants categorized facial expressions as displaying anger or fear. We manipulated emotion strength by presenting ‘morphed’ facial expressions ranging from neutral to intense anger or fear, and contextual information by pairing facial expressions with direct or averted gaze. The parametric control over emotion strength afforded fitting decision theoretical models to the behavioral and neural data and to arbitrate between bias and sensitivity accounts of threat-dependent effects on emotion recognition. At the neural level, previous studies have reported interactions between emotion and gaze direction from 200 ms following face presentation (Sato et al., 2004; N’Diaye et al., 2009; Adams et al., 2012; Conty et al., 2012), but failed to characterize the computational mechanism responsible for these effects. Here, we applied model-guided regressions of single-trial EEG signals to determine whether the neural ‘encoding’ of threat-signaling emotions is enhanced in ventral face-selective and/or dorsal motor regions (El Zein et al., 2015), and whether this enhancement is mediated by increased top-down attention to threat-signaling signals. As high-anxious individuals show increased sensitivity to threat signals (Bishop, 2007; Cisler and Koster, 2010), we further assessed the neural mechanisms by which anxiety impacts the detection of threat signals as a function of contextual gaze.

## Results

### Behavior

Participants were presented at each trial with a face expressing fear or anger of varying emotion strength and had to categorize the displayed emotion (Fig. 1). Crucially, direction of gaze (direct or averted) was manipulated independently of the displayed emotion in a completely implicit fashion, as it was never mentioned to the subjects nor relevant to the emotion categorization task. Nevertheless, in addition to an expected increase in categorization performance with emotion strength ( $F_{6,138} = 187.3, p < 0.001$ ), gaze direction strongly interacted with the displayed emotion on performance ( $F_{1,23} = 21.2, p < 0.001$ ). Facial displays of anger were better categorized when paired with a direct gaze ( $t_{23} = 4.3, p < 0.001$ ), whereas expressions of fear were better categorized when paired with an averted gaze ( $t_{23} = -3.4, p < 0.01$ ; Fig. 2a). Because these combinations of gaze and emotion are associated with higher threat for the observer (Sander et al., 2007), they will be labeled as THREAT+ as opposed to THREAT- combinations (anger paired with averted gaze, and fear paired with direct gaze).

To characterize the mechanism underlying the improved recognition of threat-signaling emotions, we fitted participants' behavior using a family of nested models of choice which hypothesize that decisions are formed on the basis of a noisy comparison of the displayed emotion to a criterion (see Methods). We compared quantitatively two variants of the model which could account for the improved recognition of THREAT+ combinations: 1. a first variant in which gaze direction *biases* the decision criterion in favor of the emotion signaling higher threat, and 2. a second variant in which gaze direction enhances the *sensitivity* to the emotion signaling higher threat. Bayesian model selection revealed that a sensitivity enhancement for THREAT+ combinations explained decisively better the behavioral data than a criterion shift (Bayes Factor  $\approx 10^8, p_{\text{exc}} > 0.74$ ). As expected, maximum-likelihood sensitivity estimates extracted from the winning model were significantly increased for THREAT+ combinations of gaze and emotion ( $t_{23} = 3.9, p < 0.001$ ; Fig. 2bc).

### Enhanced neural encoding of threat-signaling emotions

To validate the finding of enhanced sensitivity to threat-signaling emotions, and to identify its neural substrates, we then investigated how facial expressions modulated scalp EEG activity recorded during the emotion categorization task. Instead of computing event-related averages, we relied on a parametric regression-based approach consisting in regressing single-trial EEG signals against the strength of the displayed emotion at each electrode and time point following the presentation of the face (Wyart et al., 2012a, 2015). The resulting time course at each electrode represents the degree to which EEG activity ‘encodes’ (co-varies with) the emotion strength provided by morphed facial features.

Parameter estimates of the regression slope revealed significant correlations between emotion strength and EEG activity peaking initially around 290 ms following face presentation at temporal (t-test against zero,  $t_{23} = -12.7$ ,  $p < 0.001$ ) and frontal electrodes ( $t_{23} = 8.7$ ,  $p < 0.001$ ), and then around 500 ms and at response time at centro-parietal ( $t_{23} = 10.2$ ,  $p < 0.001$ ) and frontal electrodes ( $t_{23} = -7.9$ ,  $p < 0.001$ ) (Fig. 3abc). Time points and electrodes where parameter estimates diverge significantly from zero indicate neural encoding of emotion information. The strength of this neural encoding – indexed by the amplitude of the parameter estimate – provides a measure of the *neural sensitivity* to emotion information.

To test for a neural signature of the increased sensitivity to threat-signaling emotions, we compared parameter estimates extracted separately for THREAT+ and THREAT- combinations of gaze and emotion. This contrast revealed increased parameter estimates for THREAT+ combinations first at 170 ms at temporal (paired t-test,  $t_{23} = -2.5$ ,  $p < 0.05$ ) and frontal electrodes ( $t_{23} = 2.2$ ,  $p < 0.05$ ), and then later at 500 ms and at response time at centro-parietal ( $t_{23} = 2.2$ ,  $p < 0.05$ ) and frontal electrodes ( $t_{23} = -2.4$ ,  $p < 0.05$ ) (Fig. 3abc). This finding indicates that the neural gain of emotion encoding was enhanced at these time points and electrodes for threat-signaling emotions. Interestingly, THREAT+ combinations were not associated with increased event-related averages at classical peak latencies (P1, N170, P2, P3: all  $t_{23} < 1.95$ ,  $p > 0.07$ ). We computed the cortical sources of this enhanced encoding of threat-signaling emotions by performing the same regression approach to minimum-norm current estimates distributed across the cortical surface. Parameter estimates at time points of interest (where differences between THREAT+ and THREAT- combina-

tions were observed) were then contrasted between the two conditions (see [Material and Methods](#)). Increases in regression slopes for THREAT+ combinations shifted from ventral visual areas selective to facial expressions of emotion (fusiform gyrus and superior temporal sulcus) around 180 ms, to associative brain regions encompassing parietal, temporal and frontal cortices (superior and middle temporal, temporal pole, and orbitofrontal cortices) at 500 ms, and then to sensorimotor regions around response onset (dorsal central, parietal and frontal regions) ([Fig. 3dfe](#)).

These neural effects converge with behavioral modeling in favor of a sustained enhancement of perceptual sensitivity to threat-signaling emotions, starting 180 ms following face presentation and lasting until after response onset. Additional evidence supports our hypothesis that enhancements in neural sensitivity to THREAT+ combinations are specifically linked to an increase in implied threat for these combinations of gaze and emotion. A separate group of participants rated the identities used in the emotion categorization task in terms of perceived threat and trustworthiness (see [Methods](#)), and the group-level ratings for each identity were regressed against single-trial EEG signals as additional regressors. This regression showed that perceived threat, but not trustworthiness, correlated significantly with temporal and centro-parietal EEG activity at 500 ms following face presentation, in the same direction as the contrast between THREAT+ and THREAT- combinations (threat:  $t_{23} > 3.6$ ,  $p < 0.01$ ; trustworthiness:  $t_{23} < 0.7$ ,  $p > 0.48$ ).

### **Attention-independent enhancement of neural processing by threat**

Analyses of the neural data have so far confirmed the hypothesis that contextual gaze information affects emotion categorization by increasing the perceptual sensitivity to threat-signaling emotions. Such an effect could be mediated by increased top-down attention to threat-signaling emotions – i.e., THREAT+ combinations. To test this possibility, we explored whether residual fluctuations in single-trial EEG signals unexplained by variations in emotion strength (measured by the previous regressions) modulated the accuracy of the subsequent categorical decision – i.e., the perceptual sensitivity to the displayed emotion. This approach is reminiscent of ‘choice probability’ measures applied in electrophysiology to measure correlations between neural activity and choice behavior (Britten et al., 1996; Shadlen et al., 1996; Parker

and Newsome, 1998)– by estimating how much fluctuations in recorded neural signals are ‘read out’ by the subsequent decision (Wyart et al., 2012a, 2015). Stimulus-independent improvements in neural-choice correlations have been classically interpreted as increases in ‘read-out’ weights – i.e., increased top-down attention to these neural signals (Nienborg and Cumming, 2009, 2010). Here, an increased neural modulation of choice for THREAT+ conditions could indicate an increase in top-down attention to threat-signaling emotions, which could in turn explain the observed increase in perceptual and neural sensitivity to these combinations of gaze and emotion.

In practice, we entered EEG residuals from the previous regression against emotion strength as an additional ‘mediation’ predictor of choice – as means to test whether these neural signals co-vary with perceptual sensitivity (see [Methods](#)). The time course and spatial distribution of this neural modulation of perceptual sensitivity followed qualitatively the neural encoding of emotion strength ([Fig. 4abc](#)), with a negative temporal component peaking at 270 ms ( $t_{23} = -4.2$ ,  $p < 0.001$ ), followed by a positive centro-parietal one peaking around 600 ms ( $t_{23} = 8.0$ ,  $p < 0.001$ ) and then at response time ( $t_{23} = 7.6$ ,  $p < 0.001$ ). We used Bayesian model selection to confirm that EEG residuals co-varied multiplicatively with the perceptual sensitivity of the subsequent decision, not additively as a bias in emotion strength, both at temporal (Bayes factor  $\approx 10^{3.4}$ ,  $p_{\text{exc}} = 0.79$ ) and centro-parietal electrodes (Bayes factor  $\approx 10^{8.9}$ ,  $p_{\text{exc}} = 0.99$ ). Critically, no difference in modulation strength was observed between THREAT+ and THREAT– combinations (temporal:  $t_{23} = -0.4$ ,  $p > 0.5$ ; centro-parietal:  $t_{23} = 0.1$ ,  $p > 0.5$ ). To determine whether this absence of significant difference is due to a genuine absence of effect (rather than a lack of statistical sensitivity), we computed Bayes factors under the same parametric assumptions as conventional statistics (see [Methods](#)). We obtained Bayes factors lower than  $10^{-4}$  at temporal and centro-parietal electrodes, indicative of no increase in ‘read-out’ weights for THREAT+ conditions. This null effect suggests that the observed enhancement in perceptual and neural sensitivity to these threat-signaling combinations of gaze and emotion is not triggered indirectly by an increase in top-down attention in these conditions.

### **Early neural encoding of threat-signaling emotions in motor preparation**

We reasoned that threat could impact not only the neural representation of the displayed emotion in visual and associative cortices, but also the preparation of the upcoming response in effector-selective structures (Conty et al., 2012). To measure response-preparatory signals in the neural data, we computed spectral power in the mu and beta frequency bands (8-32 Hz) (Donner et al., 2009; de Lange et al., 2013). The contrast between left-handed and right-handed responses at response time identified lateral central electrodes, associated with focal sources in motor cortex (Fig. 5a). Subtracting contralateral from ipsilateral signals relative to the hand assigned to the ‘fear’ response (counterbalanced across participants) provided a motor lateralization index whose sign predicts significantly the upcoming choice (anger or fear) from 360 ms before response onset (paired t-test,  $t_{23} = 4.6$ ,  $p < 0.001$ ; Fig. 5b).

We applied the previous neural encoding approach by regressing this motor lateralization index against the *signed* emotion strength (from 0 for a neutral expression, to  $\pm 7$  for an intense anger/fear expression) on a trial-by-trial basis. Parameter estimates of the regression slope diverged significantly from zero from 400 ms after stimulus onset (t-test against zero,  $t_{23} = 5.1$ ,  $p < 0.001$ ) and at response time ( $t_{23} = 5.2$ ,  $p < 0.001$ ) – reflecting stronger response preparation to stronger (i.e., more diagnostic) emotions. Computing regression slopes separately for THREAT+ and THREAT- combinations revealed that THREAT+ combinations produced a stronger encoding of emotion strength in motor preparation late at response onset ( $t_{23} = 2.9$ ,  $p < 0.01$ ), but also early around 200 ms following face presentation ( $t_{23} = 3.2$ ,  $p < 0.01$ ). While THREAT- combinations of gaze and emotion were not associated with significant neural encoding in motor preparation until 440 ms following face presentation ( $t_{23} < 0.8$ ,  $p > 0.4$ ), THREAT+ combinations resulted in significant neural encoding between 100 and 320 ms ( $t_{23} = 3.2$ ,  $p < 0.01$ ; Fig. 5c).

To determine whether this early neural encoding of threat-signaling emotions in motor preparation influences the speed of subsequent responses, we recomputed and compared regression parameters estimated separately for fast and slow responses to THREAT+ combinations, on the basis of a median split of response times informed by emotion strength. This comparison revealed a single, gradual neural encoding of emotion strength in motor preparation preceding fast, but not slow responses, arising as early as 150 ms (at a threshold  $p$ -value of 0.05) following the



presentation of the face (difference in encoding onset between fast and slow responses, jackknifed(Kiesel et al., 2008)  $t_{23} = 5.2$ ,  $p < 0.001$ ; Fig. 6a). This effect indicates that the early neural encoding of THREAT+ combinations in motor preparation is characteristic of efficient (fast) responses. We verified that this latency shift in neural encoding was selective of motor preparation signals, by performing the same comparison on the neural encoding of emotion strength at centro-parietal electrodes. This contrast revealed only a difference in peak amplitude, not onset latency, between fast and slow responses (peak amplitude:  $t_{23} = 5.1$ ,  $p < 0.001$ ; onset latency: jackknifed  $t_{23} = -1.3$ ,  $p > 0.2$ ; Fig. 6b).

Finally, we performed neural-choice correlations analyses to assess whether the early neural encoding of threat-signaling emotions in motor preparation influences not only the speed, but also the content (anger or fear) of subsequent responses. Across conditions, the neural ‘mediation’ analysis described above revealed that stimulus-independent fluctuations in motor lateralization index co-vary as an additive choice bias in the upcoming response from 400 ms following face presentation ( $t_{23} = 2.9$ ,  $p < 0.01$ ). In contrast to fluctuations in temporal and centro-parietal activity, the impact of variability in motor lateralization on emotion categorization was better described as an additive choice *bias* rather than a change in perceptual *sensitivity* (Bayes factor  $\approx 10^{36.4}$ ,  $p_{\text{exc}} = 0.98$ ) – consistent with its hypothesized role as a motor representation of the decision variable(Donner et al., 2009; de Lange et al., 2013). No difference in modulation strength was observed between THREAT+ and THREAT- combinations ( $t_{23} < 1.6$ ,  $p > 0.1$ ; Fig. 5d). Critically, even when considering THREAT+ combinations alone, residual variability in motor lateralization measured between 100 and 320 ms (where the neural encoding of threat-signaling emotions was significant) did not bias significantly the upcoming choice ( $t_{23} < 1.4$ ,  $p > 0.17$ ). This null effect was supported by Bayesian model selection which identified a genuine absence of neural-choice correlation as the most likely account of the data (Bayes factor  $\approx 10^{2.3}$ ,  $p_{\text{exc}} = 0.96$ ). This finding indicates that the early neural encoding of threat-signaling emotions in motor preparation occurs earlier than the formation of the upcoming choice.

## Anxiety-dependent neural encoding of threat-signaling emotions

In the general population, anxiety has been classically associated with an oversensitivity to threat signals in social conditions. Here, we assessed whether the enhanced neural processing of threat-signaling emotions in temporal and motor regions co-varied with the level of anxiety in our participants. For this purpose, we measured anxiety at the beginning of the experiment using the Spielberger State-Trait Anxiety Inventory (STAI). We analyzed the effect of anxiety on the behavioral and neural data in two complementary ways: 1. by splitting the participants in two equally-sized groups based on their measured anxiety, and 2. by correlating neural encoding parameters estimated at the level of individual participants with their measured anxiety. Surprisingly, we found no effect of 'anxiety on overall measures of performance ( $t_{11} < 0.05$ ,  $p > 0.9$ ), nor on the difference between THREAT+ and THREAT- combinations of gaze and emotion ( $F_{1,22} < 0.4$ ,  $p > 0.5$ ).

Nevertheless, the absence of effect of anxiety at the behavioral level was accompanied by a compensatory double dissociation in the neural data. Indeed, state anxiety influenced significantly the neural encoding of emotion strength at temporal electrodes between 200 and 400 ms following face presentation (median split, interaction:  $F_{1,22} = 10.8$ ,  $p < 0.01$ ; Fig. 7a): high-anxious observers showed no difference in neural encoding between THREAT+ and THREAT- combinations (THREAT+ :  $t_{11} = -6.5$ ,  $p < 0.001$ ; THREAT- :  $t_{11} = -6.2$ ,  $p < 0.001$ , difference:  $t_{22} = 1.1$ ,  $p = 0.28$ ), whereas low-anxious observers encoded exclusively THREAT+ combinations in the same neural activity (THREAT+ :  $t_{11} = -6.9$ ,  $p < 0.001$  ; THREAT- :  $t_{11} = -0.63$ ,  $p = 0.53$  ; difference:  $t_{22} = -3.8$ ,  $p < 0.01$ ). A parametric assessment of the relationship between state anxiety and the difference in neural encoding between THREAT+ and THREAT- combinations proved to be significant ( $r = 0.58$ , d.f. = 22,  $p < 0.01$ ; Fig. 7a). In other words, high anxiety was associated with a significant and indifferent neural encoding of negative emotions, whether threat-signaling or not, in ventral face-selective regions.

Interestingly, at the early time window (200 ms) where only THREAT+ combinations were encoded in motor signals, a reverse pattern was observed: only high anxious individuals showed a significant encoding at this latency (interaction between-subject state anxiety and gaze pairing  $F_{1,22} = 4$ ,  $p = 0.05$ ; Fig. 7b). The

more the individuals were anxious, the more they encoded observer-relevant threat signals in motor systems (correlation between parameter estimates for threat + conditions and state anxiety  $r = 0.52$ , d.f. = 22,  $p < 0.01$ ). Moreover, the encoding of THREAT+ combinations correlated with high anxious behavioral sensitivity to THREAT+ ( $r = 0.66$ , d.f. = 10,  $p = 0.01$ ; [Fig. 7b](#)), whereas it did not for low anxious individuals ( $r = -0.42$ , d.f. = 10,  $p > 0.16$ , significant difference between correlation coefficients  $p < 0.01$ ). To sum up, while high anxious individuals process all threat signals equivalently in face selective regions, they selectively encode threat signals that are relevant to them in motor specific systems, and this encoding reflects their behavioral sensitivity to threat-signaling emotions.

## Discussion

Accurate decoding of emotions in others, especially negative ones, conveys adaptive advantages in social environments. Although typical social interactions do not require an explicit categorization of the emotion expressed by others, a precise understanding of the neural mechanisms involved in emotion recognition provides important information regarding how the human brain processes socially meaningful signals. And while past work has uncovered the neural correlates of perceptual decisions (Gold and Shadlen, 2007; Heekeren et al., 2008), only few studies have addressed the issue of how such decisions are formed on the basis of socially relevant stimuli such as facial displays of emotion. As in most perceptual categorization tasks, we manipulated the ambiguity of sensory evidence – here, using controlled morphs between angry or fearful expressions and neutral ones. But owing to the social nature of our stimuli, we could simultaneously and implicitly manipulate the contextual significance of the displayed emotion in terms of implied threat for the observer, using gaze direction, and apply a model-guided approach to characterize the neural prioritization of threat-signaling information in electrical brain signals.

Gaze direction, which acts as a contextual cue in our emotion categorization task, differs from contextual cues found in perceptual decision-making studies which are typically provided hundreds of milliseconds before the decision-relevant stimulus (Rahnev et al., 2011; Kok et al., 2012; Wyart et al., 2012b; de Lange et al., 2013). Here, as in many social situations, contextual cues can co-occur with the decision-relevant stimulus – a property which strongly constrains their impact on stimulus processing. Moreover, the meaning of contextual cues (e.g., attention or expectation cues) used in perceptual decision-making studies is usually instructed explicitly, and thus processed explicitly by the participants during task execution (Kok et al., 2012; Wyart et al., 2012b). Here, by contrast, gaze direction is irrelevant for the emotion categorization task, and thus does not need to be processed explicitly. Despite these two differences with other contextual cues, we show that gaze direction tunes the neural processing of emotion information from 200 ms following stimulus onset until response in sensory, associative and motor circuits of the human brain.

Previous observations of increased subjective ratings and improved recognition of angry expressions paired with a direct gaze and fearful expressions paired

with an averted gaze have been interpreted in terms of a contextual evaluation of the displayed emotion during its processing (Adams and Kleck, 2003; Sander et al., 2007; Adams et al., 2012). In particular, ‘appraisal’ theories (Sander et al., 2007) emphasize that an angry expression paired with a direct gaze can be interpreted as behaviorally ‘relevant’ to the observer as being the target of a verbal or physical assault, whereas a fearful expression looking aside from the observer might signal a source of danger in the immediate vicinity of the observer. However, the mechanisms which instantiate the proposed contextual evaluation of emotions as a function of their implied threat for the observer have remained unclear. Gaze direction could either bias the perceived emotion towards its most relevant (threat-signaling) interpretation – i.e., anger when paired with direct gaze, or fear when paired with averted gaze, or increase the sensitivity to the most relevant emotion. The present study answers directly this issue by showing, both behaviorally (by comparing quantitative fits of the two effects to the behavioral data) and neurally (by regressing brain signals against emotion strength), that the improved recognition accuracy for threat-signaling emotions corresponds to a selective neural enhancement of perceptual sensitivity to these combinations of gaze and emotion.

Emotion information modulated EEG signals at centro-parietal electrodes from 500 ms following face presentation until response execution, a finding in accordance with the ‘supramodal’ signature of perceptual integration reported in previous studies (O’Connell et al., 2012; Wyart et al., 2012a). This centro-parietal positivity has been proposed to encode a ‘domain-general’ decision variable, as it varies with the strength of sensory evidence for both visual and auditory decisions, independently from the associated response (O’Connell et al., 2012). Here, the same centro-parietal positivity was found to increase with the emotion strength of facial expressions – which indexes the decision variable in our emotion categorization task. Importantly, the strength of this relationship was enhanced for threat-signaling emotions. This improved neural representation of threatening combinations of gaze and emotion cannot be explained by increased attentional or surprise responses, since the centro-parietal ‘P3’ potential, previously reported to vary as a function of attentional resources (Johnson, 1988) and surprise (Mars et al., 2008), was not increased in response to threat-signaling emotions. Moreover, we could also rule out the possibility

that this enhanced neural encoding is triggered indirectly by an increase in selective attention, which should have been associated with an improved 'decoding' of participants' decisions from their underlying neural signals(Nienborg and Cumming, 2009, 2010; Wyart et al., 2015). We therefore hypothesize that the enhanced neural processing of threat-signaling emotions proceeds in an attention-independent, bottom-up fashion.

Earlier contextual modulations of emotion processing were also observed in ventral face-selective areas from 170 ms following face presentation. While these findings contradict a 'two-stage' view according to which emotion and gaze information would be processed independently during the first hundreds of milliseconds(Vuilleumier and Pourtois, 2007) before being integrated as a function of their significance to the observer(Klucharev and Sams, 2004), they are in agreement with recent findings(Conty et al., 2012; El Zein et al., 2015) of early interactions between emotion and gaze information on N170 and P200 components. At these early latencies, only threat-signaling emotions were encoded by face-selective neural signals, reflecting a faster processing of emotions signaling an immediate threat to the observer as a function of their associated gaze.

More strikingly, gaze direction also modulated the encoding of emotional expressions in effector-selective regions, in parallel with the effects observed in ventral face-selective areas: only threat-signaling emotions were encoded in response preparation signals overlying human motor cortex at 200 ms following face presentation. Recent work sheds light on the adaptive function of this early representation of threat signals in motor cortex. Disrupting this motor representation using TMS impairs the facial recognition of negative (i.e., potentially threatening) emotions, not positive ones(Balconi and Bortolotti, 2012, 2013). Moreover, the perception of natural scenes engages the motor cortex at very early latencies only when the emotional valence of the scene is negative(Borgomaneri et al., 2014). Taken together, these findings support a strong connection between emotion and motor circuits(Grèzes et al., 2014) enabling the brain to react swiftly and efficiently to threat signals(Ohman and Mineka, 2001; Frijda, 2009). Our findings build on these earlier observations by showing that the brain encodes parametrically the strength of threat signals in motor cortex in parallel to their representation in face-selective, sensory regions.

Finally, our data reveal a clear functional dissociation between face- and effector-selective regions as a function of individual anxiety. The enhanced sensitivity to threat-signaling emotions in face-selective temporal cortex is driven by low-anxious observers, whereas the early enhancement measured in motor cortex is only found in high-anxious observers. The observation that high-anxious individuals encode all negative emotions as equally (and strongly) salient in face-selective regions is consistent with earlier reports of an ‘hyper-vigilance’ to potentially threatening signals in these individuals (Bishop, 2007; Cisler and Koster, 2010), and with their tendency to interpret ambiguous stimuli as threatening (Beck et al., 1985) – both associated with amygdala hyperactivity (Bishop, 2007; Etkin and Wager, 2007). Nevertheless, our findings reveal that high-anxious individuals are capable of encoding threat signals in a selective fashion in motor cortex. Consistent with the idea of a compensatory mechanism, the distinct neural enhancements of temporal and motor activity found in low- and high-anxious individuals lead to similar behavioral improvements in terms of perceptual sensitivity to threat signals. Together, this pattern of findings suggests that anxiety increases the relative contribution of the motor pathway during the processing of negative social signals, in accordance with the adaptive function of anxiety in detecting efficiently and reacting swiftly to threats in the environment (Bateson et al., 2011).

By applying theoretical models of decision-making to socially-relevant stimuli, we were able to characterize the neural and computational mechanisms underlying the integration and interpretation of facial cues in the implicit context of threat. Evolutionary pressure might have shaped the human brain to prioritize threat signals in parallel in sensory and motor systems (Darwin et al., 1998; LeDoux, 2012). Such prioritization – found to proceed in a fast, selective, yet attention-independent fashion – could increase perceptual sensitivity to other features of the sensory environment (Phelps et al., 2006) to enable rapid and adaptive responses in complex, multidimensional situations of danger.

## Methods

### Subjects

Twenty four healthy subjects (12 females; mean age,  $22.7 \pm 0.7$  years) participated in the EEG experiment. All participants were right-handed, with a normal vision and had no neurological or psychiatric history. They provided written informed consent according to institutional guidelines of the local research ethics committee (Declaration of Helsinki) and were paid for their participation.

### Stimuli

Stimuli consisted of 36 identities (18 females) adapted from the RadBoud Faces Database (Langner et al., 2010) that varied in emotion (neutral, angry or fearful expressions) and gaze direction (direct toward the participant or averted  $45^\circ$  to the left or right). Using Adobe Photoshop CS5.1 (Adobe Systems, San Jose CA), faces were modified to remove any visible hair, resized and repositioned so that eyes, nose and mouth appeared within the same circumference. All images were converted to grey-scale and cropped into a 280 x 406 pixel oval centered within a 628 x 429 pixel black rectangle.

To vary the intensity of emotional expressions, faces were morphed from neutral to angry expressions and from neutral to fearful expression using FantaMorph (Abrosoft <http://www.fantamorph.com/>). At first, we created 7 levels of morphs from neutral to angry expressions and from neutral to fearful expressions (separately for direct and averted gaze stimuli) using a simple linear morphing transformation. This resulted in 30 conditions for each identity: 7 levels of morphs \* 2 emotions \* 2 gaze directions = 28 and 2 neutral stimuli with direct and averted gaze. We then calibrated the morphing between angry and fearful expressions by performing an intensity rating pre-test of the emotional expressions and adjusting the morphs based on the results. 19 subjects (9 females, mean age,  $24.7 \pm 0.9$  years) were presented with the facial expressions for 250 ms and rated the emotional intensity perceived on a continuous scale from “not at all intense” to “very intense” using a mouse device (with a maximum of 3 seconds to respond). We adjusted for differences between emotions by linearizing the mean curves of judged intensities and creating corresponding



morphs, that were validated on 10 new subjects (4 females, mean age  $24.1 \pm 1.9$ ). To summarize, the stimuli comprise of 36 identities with an Averted gaze condition and a Direct gaze condition, each with 7 levels of Anger and 7 levels of Fear equalized in perceived emotional intensities and a neutral condition, resulting in a total of 1080 items (see [Fig.1a](#) for examples of stimuli).

## **Experimental procedure**

Using the Psychophysics-3 Toolbox (Brainard, 1997; Pelli, 1997), stimuli were projected on a black screen. Each trial was initiated with a white oval delimiting the faces that was kept during all the trial. The white oval appeared for approximately 500 ms, followed by a white fixation point presented at the level of the eyes for approximately 1000 ms (to keep the fixation to the upcoming faces natural and avoid eye movements from the center of the oval to eye regions), then the stimuli appeared for 250 ms. Participants' task was to decide whether the faces expressed Anger or Fear by pressing one of the two buttons localized on two external devices held in their right and left hands, with their right or left index correspondingly ([Fig.1b](#)). An Anger/Fear mapping was used (e.g Anger: Left hand, Fear: Right hand) kept constant for each subject, counterbalanced over all subjects. All stimuli were presented once, resulting in a total of 1080 trials. The experiment was divided in 9 experimental blocks, each consisting of 120 trials, balanced in the number of emotions, directions of gaze, gender and levels of morphs. After each block, the percentage of correct responses was shown to the participants to keep them motivated.

## **Behavioral data analyses**

Repeated-measures ANOVA was performed on the percentage of correct responses and average reaction times, with gaze direction (direct/averted), emotion (anger/fear), and intensity (7 levels of morphs) as within-subjects factors.

## **Model selection**

We performed model-guided analyses of the behavioral data to characterize the observed increase in recognition accuracy for THREAT+ combinations of gaze and emotion. We used Bayesian model selection based on the model evidence (estimated by

a 10-fold cross-validation estimation of model log-likelihood, which penalizes implicitly for model complexity without relying on particular approximations such as the Bayesian Information Criterion or the Akaike Information Criterion). We applied both fixed-effects and random-effects statistics previously described in the literature. The fixed-effects comparison assumes all participants to have used the same underlying model to generate their behavior, such that the overall model evidence for a given model is proportional to the product of model evidence for the model for all participants. Based on this model evidence, we compared different models by computing their Bayes factor as the ratio of model evidence of the compared model (Jeffreys, 1961; Kass and Raftery, 1995). The random-effects comparison is more conservative in allowing different participants to use different models to generate their behavior, and aims at inferring the distribution over models that participants draw from (Penny et al., 2010). For this comparison, we computed support for the winning model by the exceedance probability ( $p_{exc}$ ), which is the probability that participants were more likely to choose this model to generate behavior over any alternative model.

We started with the simplest model (model 0) that could account for subject's decisions using a psychometric model such as:

$$P(\text{Anger}) = \Phi(w \cdot \mathbf{x} + b)$$

where  $P(\text{Anger})$  corresponds to the probability of judging the face as angry,  $\Phi$  to the cumulative normal function,  $w$  to the perceptual sensitivity to the display emotion,  $\mathbf{x}$  to the evidence in favor of Anger or Fear in each trial (emotion strength, from -7 for an intense expression of fear to +7 for an intense expression of anger), and  $b$  to an additive, stimulus-independent bias toward one of the two responses/emotions.

We compared this 'null' model which did not allow for contextual influences of gaze direction on the decision process, with two additional models which instantiate two different mechanisms which could account for the observed increase in recognition accuracy for THREAT+ combinations of gaze and emotion. A first possibility (model 1) would be that gaze direction biases emotion recognition in favor of the interpretation signaling higher threat (anger for a direct gaze, fear for an averted gaze). Alternatively (model 2), gaze direction might selectively increase sensitivity to emotions signal-

ing higher threat in this context (modeled by a different sensitivity to emotions in THREAT+ vs. THREAT – conditions).

### **EEG acquisition and pre-processing**

An EEG cap of 63 sintered Ag/AgCl ring electrodes (Easycap) was used to record EEG activity. EEG activity was recorded at a sampling rate of 1000 Hz using a BRAINAMP amplifier (Brain Products, BRAINAMP MR PLUS) and low pass filtered online at 250 Hz. The reference channel was placed on their nose and a forehead ground was used. Impedances were kept under a threshold of 10 k $\Omega$ .

The raw EEG data was recalculated to average reference, down-sampled to 500 Hz, low-pass filtered at 32 Hz, and epoched from 1 s before to 4 s after the face stimulus onset using EEGLAB(Delorme and Makeig, 2004). First, EEG epoched data was visually inspected to remove muscle artifacts and to identify noisy electrodes that were interpolated to the average of adjacent electrodes. Second, independent component analysis (ICA) that excluded interpolated electrodes was performed on the epoched data and ICA components capturing eye blink artifacts were manually rejected. A last, visual inspection was done on the resulting single epochs to exclude any remaining trials with artifacts. After trial rejections, an average of 999  $\pm$ 10 trials per subject remained.

Time frequency analysis was performed using the Fieldtrip toolbox for MATLAB(Oostenveld et al., 2011). We were particularly interested in motor mu-bands (8-32 Hz) and thus estimated the spectral power of mu-beta band EEG oscillations using ‘multitapering’ time frequency transform (Slepian tapers, frequency range 8-32 Hz, five cycles, three tapers per window). The purpose of this multitapering approach is to obtain more precise power estimates by smoothing across frequencies. Note that this time–frequency transform uses a constant number of cycles per window across frequencies, hence a time window whose duration decreases inversely with increasing frequency.

### **EEG analysis**

#### **Time frequency: motor lateralization measures**

As the suppression of mu-beta activity in the hemisphere contralateral to the hand used for response is a marker of motor preparation to response (Donner et al., 2009; de Lange et al., 2013), spectral power from 8 to 32 Hz were calculated at each electrode and time point for all subjects. Then for each subject, to obtain the lateralization measures, the spectral power from 8 to 32 Hz for the trials where the subjects responded with their right hand was subtracted from that of the trials where the subjects responded with their left hand. After averaging on all subjects, electrodes where the motor lateralization was maximal from 200 ms before to response time were selected: 'P3','CP3','C3' for the left hemisphere and 'P4','CP4','C4' for the right hemisphere. Motor lateralization specific to 'anger' or 'fear' responses was obtained by taking into account the Anger/Fear mapping used and subtracting 'Anger' hand spectral activity to 'Fear' hand spectral activity (the average on 'P3','CP3','C3' minus the average on 'P4','CP4','C4' if participants responded 'Anger' with the left hand and vis versa if they responded 'Anger' with the right hand).

### **Regression analysis: Encoding of the emotional information**

In our emotion categorization task, evidence strength corresponds to the intensity of the displayed emotion. On the basis of recent studies (Wyart et al., 2012b, 2015), we therefore performed single-trial regressions of EEG signals against this variable. A general linear regression model (GLM) was used where emotion strength (from 0 for a neutral/emotionless expression to 7 for an intense fear/anger expression) was introduced as a trial-per-trial predictor of broadband EEG signals at each time point after stimulus onset (from 200 ms before to 1 s after stimulus onset), at each electrode. The corresponding parameter estimates of the regression, reported in arbitrary units, were measured per participant, and then averaged across participants to produced group-level averages. The time course of the parameter estimates describes the neural 'encoding' of the relevant (emotion) information provided by the presented facial expression. Electrodes and time points where the parameter estimates of the regression were maximal were selected to further compare between the conditions of interest: Anger Direct and Fear Averted vs Fear Direct and Anger Averted.

Similar general linear regressions were also performed on lateralized mu-beta activity. Once more, the intensity of the emotional expression was entered as a regressor

to predict the trial-per-trial motor lateralization activity (calculated as described above) for each time point after stimulus onset. The only important difference is that owing to the ‘signed’ nature of the motor lateralization index (positive for a contra-lateralized activity), we expressed the intensity of the emotional expression as signed by the displayed emotion, from -7 for an intense expression of fear to +7 for an intense expression of anger.

### Neural-choice correlation analyses

We determined whether residual fluctuations in single-trial EEG signals unexplained by variations in emotion strength (measured by the previous neural regressions against emotion strength) modulated the recognition of the subsequent emotion. This approach is reminiscent of ‘choice probability’ measures applied in electrophysiology to measure correlations between neural activity and choice behavior (Britten et al., 1996; Shadlen et al., 1996; Parker and Newsome, 1998) – by estimating how much fluctuations in recorded neural signals are ‘read out’ by the subsequent decision (Wyart et al., 2012a, 2015). The advantage of measuring neural-choice correlations within the framework of our computational model is that we could not only establish *whether*, but also *how* neural fluctuations influenced the subsequent behavior – either additively as a stimulus-independent bias, or multiplicatively as a change in perceptual sensitivity.

In practice, we estimated the parameters  $b_{\text{mod}}$  and  $w_{\text{mod}}$  of these neural modulation terms at each time point following face presentation via an EEG-informed regression of choice for which the neural residuals  $\varepsilon$  from the regression against emotion strength were entered either alone (additive influence, parameter  $b_{\text{mod}}$ , model 1) or as their interaction with the strength of the displayed emotion (multiplicative influence, parameter  $w_{\text{mod}}$ , model 2) as an additional predictor of the subsequent categorical choice, as follows:

$$(1) p(\text{anger}) = \Phi(w \cdot x + b + w_{\text{mod}} \cdot \varepsilon \cdot x)$$

$$(2) p(\text{anger}) = \Phi(w \cdot x + b + b_{\text{mod}})$$

We applied Bayesian model selection to compare between these two possible modulations of the decision process by neural fluctuations using both fixed-effects and random-effects statistical procedures described above.

### **Source reconstruction analysis**

Source analysis was performed using Brainstorm (Tadel et al., 2011). A source model consisting of 15002 current dipoles was used to calculate Kernel inversion matrices for each subject based on all the trials of the subject. Dipole orientations were constrained to the cortical mantle of a generic brain model taken from the standard Montreal Neurological Institute (MNI) template brain provided in Brainstorm. Individual scalp models, recorded with a Zebris device, were used to warp this template head model to EEG sensor caps. Using the OpenMEEG BEM model (Kybic et al., 2005; Gramfort et al., 2010), the forward EEG model was computed for each subject. Individual inversion matrices (15002 vertices \* 63 electrodes) were then extracted to perform single trial regressions at the source level.

### **Threat and trustworthiness rating experiment**

20 subjects participated to the experiment (10 females, mean age =  $22.7 \pm 0.6$ ). The 36 identities used in the experiment were presented in the neutral condition only. Each identity was presented twice, once with a direct, and once with an averted gaze. Faces appeared on the screen for 2 seconds after which they disappeared and 2 continuous scales were drawn on the screen. Participants rated the identities on these scales in terms of threat and trustworthiness from “not at all” to “very much” (a text appeared at the top of the scales reminding the instructions: How much is this face threatening/trustworthy?). The order of the scales was randomized across subjects. The scales stayed on the screen until the two responses were given, however subjects were instructed to answer intuitively without spending too much time to decide.

## References

- Adams RB Jr, Franklin RG Jr, Kveraga K, Ambady N, Kleck RE, Whalen PJ, Hadjikhani N, Nelson AJ (2012) Amygdala responses to averted vs direct gaze fear vary as a function of presentation speed. *Soc Cogn Affect Neurosci* 7:568–577.
- Adams RB Jr, Kleck RE (2003) Perceived gaze direction and the processing of facial displays of emotion. *Psychol Sci* 14:644–647.
- Adams RB, Kleck RE (2005) Effects of direct and averted gaze on the perception of facially communicated emotion. *Emot Wash DC* 5:3–11.
- Aviezer H, Bentin S, Dudarev V, Hassin RR (2011) The automaticity of emotional face-context integration. *Emot Wash DC* 11:1406–1414.
- Balconi M, Bortolotti A (2012) Detection of the facial expression of emotion and self-report measures in empathic situations are influenced by sensorimotor circuit inhibition by low-frequency rTMS. *Brain Stimulat* 5:330–336.
- Balconi M, Bortolotti A (2013) The “simulation” of the facial expression of emotions in case of short and long stimulus duration. The effect of pre-motor cortex inhibition by rTMS. *Brain Cogn* 83:114–120.
- Barrett LF, Kensinger EA (2010) Context Is Routinely Encoded During Emotion Perception. *Psychol Sci* 21:595–599.
- Bateson M, Brilot B, Nettle D (2011) Anxiety: an evolutionary approach. *Can J Psychiatry Rev Can Psychiatr* 56:707–715.
- Beck AT, Emery G, Greenberg RL (1985) *Anxiety Disorders and Phobias: A Cognitive Perspective*. Basic Books.
- Bindemann M, Mike Burton A, Langton SRH (2008) How do eye gaze and facial expression interact? *Vis Cogn* 16:708–733.
- Bishop SJ (2007) Neurocognitive mechanisms of anxiety: an integrative account. *Trends Cogn Sci* 11:307–316.
- Borgomaneri S, Gazzola V, Avenanti A (2014) Temporal dynamics of motor cortex excitability during perception of natural emotional scenes. *Soc Cogn Affect Neurosci* 9:1451–1457.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Cisler JM, Koster EHW (2010) Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clin Psychol Rev* 30:203–216.
- Conty L, Dezechache G, Hugueville L, Grèzes J (2012) Early binding of gaze, gesture, and emotion: neural time course and correlates. *J Neurosci Off J Soc Neurosci* 32:4531–4539.
- Darwin C, Ekman P, Prodger P (1998) *The Expression of the Emotions in Man and Animals*. Oxford University Press.
- De Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci Off J Soc Neurosci* 33:1400–1410.

- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21.
- Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol CB* 19:1581–1585.
- El Zein M, Gamond L, Conty L, Grèzes J (2015) Selective attention effects on early integration of social signals: same timing, modulated neural sources. *NeuroImage* 106:182–188.
- Etkin A, Wager TD (2007) Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry* 164:1476–1488.
- Ewbank MP, Fox E, Calder AJ (2010) The Interaction Between Gaze and Facial Expression in the Amygdala and Extended Amygdala is Modulated by Anxiety. *Front Hum Neurosci* 4:56.
- Frijda NH (2009) Emotion Experience and its Varieties. *Emot Rev* 1:264–271.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
- Graham R, LaBar KS (2007) Garner interference reveals dependencies between emotional expression and gaze in face perception. *Emot Wash DC* 7:296–313.
- Gramfort A, Papadopoulo T, Olivi E, Clerc M (2010) OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed Eng Online* 9:45.
- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics*. John Wiley and Sons.
- Grèzes J, Valabrègue R, Gholipour B, Chevallier C (2014) A direct amygdala-motor pathway for emotional displays to influence action: A diffusion tensor imaging study. *Hum Brain Mapp* 35:5974–5983.
- Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467–479.
- Jeffreys H (1961) *Theory of probability / by Harold Jeffreys*. Oxford: Clarendon Press.
- Johnson R (1988) The amplitude of the P300 component of the event-related potential: Review and synthesis.
- Kass RE, Raftery AE (1995) Bayes Factors. *J Am Stat Assoc* 90:773–795.
- Kiesel A, Miller J, Jolicoeur P, Brisson B (2008) Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. *Psychophysiology* 45:250–274.
- Klucharev V, Sams M (2004) Interaction of gaze direction and facial expressions processing: ERP study. *Neuroreport* 15:621–625.
- Kok P, Jehee JFM, de Lange FP (2012) Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75:265–270.
- Kybic J, Clerc M, Abboud T, Faugeras O, Keriven R, Papadopoulo T (2005) A common formalism for the integral formulations of the forward EEG problem. *IEEE Trans Med Imaging* 24:12–28.
- Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A (2010) Presentation and validation of the Radboud Faces Database. *Cogn Emot* 24:1377–1388.



- LeDoux J (2012) Rethinking the Emotional Brain. *Neuron* 73:653–676.
- Macmillan NA, Creelman CD (2004) *Detection Theory: A User's Guide*. Psychology Press.
- Mars RB, Debener S, Gladwin TE, Harrison LM, Haggard P, Rothwell JC, Bestmann S (2008) Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *J Neurosci* 28:12539–12545.
- N'Diaye K, Sander D, Vuilleumier P (2009) Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emot Wash DC* 9:798–806.
- Nienborg H, Cumming B (2010) Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? *Curr Opin Neurobiol* 20:376–381.
- Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92.
- O'Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat Neurosci* 15:1729–1735.
- Ohman A, Mineka S (2001) Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev* 108:483–522.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
- Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–277.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. *PLoS Comput Biol* 6:e1000709.
- Phelps EA, Ling S, Carrasco M (2006) Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol Sci* 17:292–299.
- Rahnev D, Lau H, Lange FP de (2011) Prior Expectation Modulates the Interaction between Sensory and Prefrontal Regions in the Human Brain. *J Neurosci* 31:10741–10748.
- Righart R, de Gelder B (2008) Recognition of facial expressions is influenced by emotional scene gist. *Cogn Affect Behav Neurosci* 8:264–272.
- Sander D, Grandjean D, Kaiser S, Wehrle T, Scherer KR (2007) Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *Eur J Cogn Psychol* 19:470–480.
- Sato W, Yoshikawa S, Kochiyama T, Matsumura M (2004) The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *NeuroImage* 22:1006–1013.
- Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci Off J Soc Neurosci* 16:1486–1510.

- Tadel F, Baillet S, Moshier JC, Pantazis D, Leahy RM (2011) Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci* 2011:879716.
- Vuilleumier P, Pourtois G (2007) Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* 45:174–194.
- Wyart V, de Gardelle V, Scholl J, Summerfield C (2012a) Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron* 76:847–858.
- Wyart V, Myers NE, Summerfield C (2015) Neural mechanisms of human perceptual choice under focused and divided attention. *J Neurosci Off J Soc Neurosci* 35:3485–3498.
- Wyart V, Nobre AC, Summerfield C (2012b) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A* 109:3593–3598.

Acknowledgements: This work was supported by grants from the French National Research Agency ANR-11-EMCO-00902, ANR-11-0001-02 PSL\*, ANR-10-LABX-0087 and by INSERM. We wish to thank Laurent Hugueville for his useful technical help.

## Figure legends

**Figure 1. Stimuli and experimental procedure.** (a) Examples of morphed expressions for one identity: morphs from neutral to intense fearful/angry expressions providing evidence for one or the other emotion. Stimuli displayed either an averted or a direct gaze. THREAT+ conditions (in orange) correspond to combinations of gaze and emotion that signal higher threat for the observer as compared to THREAT- conditions (in green). (b) Following fixation, a facial expression appeared for 250 ms, after which the participant had to indicate whether the face expressed anger or fear within 2 seconds. No feedback was provided after response.

**Figure 2. Enhanced recognition accuracy and perceptual sensitivity to threat-signaling emotions.** (a) Proportion of correct responses for (from left to right) averted/anger, direct/anger, averted/fear and direct/fear. THREAT+ combinations of gaze and emotion (in orange) were associated with increased recognition accuracy. (b) Psychometric function representing the proportion of 'anger' responses as a function of the evidence for anger (proportion morph, 0 for neutral, negative towards fear, and positive towards anger) for THREAT+ (orange) and THREAT- (green) combinations of gaze and emotion. Dots and attached error bars indicate the human data (mean  $\pm$  s.e.m.). Lines and shaded error bars indicate the predictions of the best-fitting model. (c) Parameter estimate for the slope of the psychometric curve (corresponding to emotion sensitivity) for THREAT+ and THREAT- combinations. \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

**Figure 3. Enhanced neural encoding of threat-signaling emotions.** (a) Middle panel: scalp topography of neural encoding at 280 ms, corresponding to its first peak of the encoding of emotion strength averaged across conditions, and expressed as mean parameter estimates in arbitrary units (a.u.). Dots indicate electrodes of interest where neural encoding was maximal. Left and right panels: encoding time course for THREAT+ and THREAT- conditions at electrodes of interest. Shaded error bars indicate s.e.m. Thick orange and green lines indicate significance against zero at a cluster-corrected p-value of 0.05. Shaded grey areas indicate significant differences between THREAT+ and THREAT- conditions at  $p < 0.05$ . (b) Same conventions as (a) at the second neural encoding peak at 500 ms. (c) Same conventions as (a) at the third neural encoding peak at response time. (e) Estimated cortical

sources of the encoding difference between THREAT+ and THREAT- conditions at the time of significant difference between conditions at 180 ms. (f) Same as (e) at 500 ms. (g) Same as (e) at response time. FG: fusiform gyrus, pSTS: posterior superior temporal sulcus, SMG: supramarginal gyrus, ANG: angular gyrus, STG: superior temporal gyrus, MTG: middle temporal gyrus, OCG: occipital gyrus, aINS: anterior insula, IFS: inferior frontal sulcus, TP: temporal pole, OFC: orbitofrontal cortex, OP: occipital pole, TPJ: temporo parietal junction, dIPFC: dorsolateral prefrontal cortex.

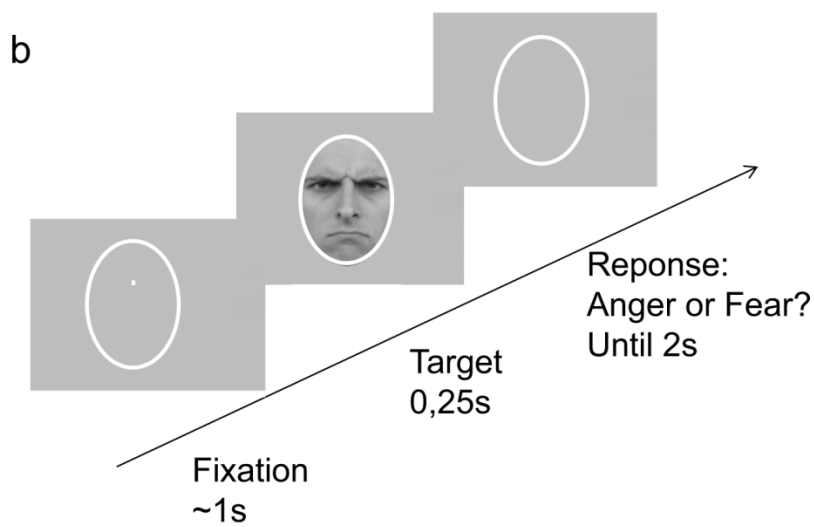
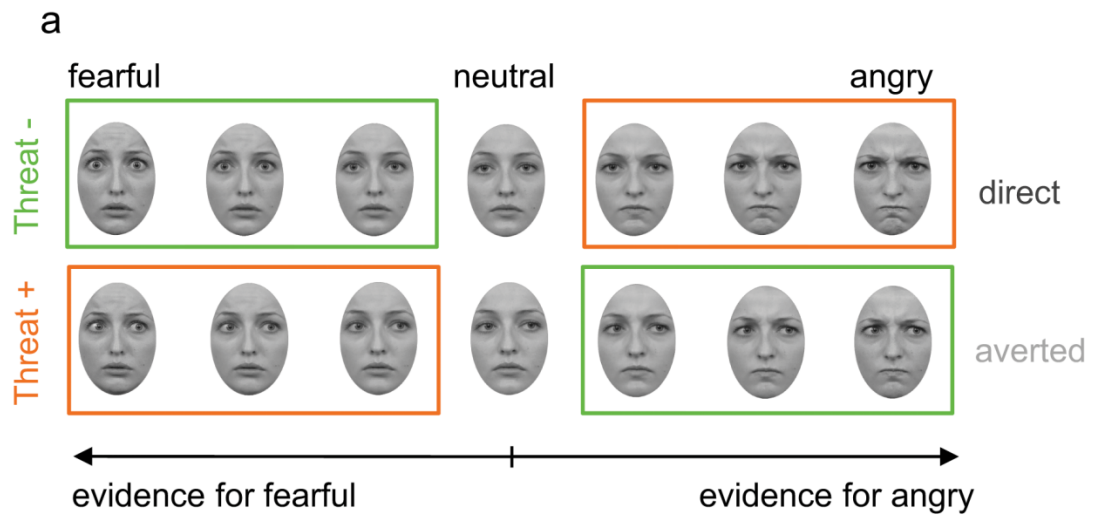
**Figure 4. Absence of threat-dependent enhancement of neural-choice correlations.** (a) Middle panel: scalp topography of neural-choice correlations, expressed as the modulation of perceptual sensitivity by EEG encoding residuals at 280 ms, same time point shown in Fig. 3a. Electrodes of interest indicated with dots are the same as in Fig.3a. Left and right panels, time course of the modulation of perceptual sensitivity by EEG encoding residuals expressed in arbitrary units (a.u.). Same conventions as in Fig. 3a. (b) Same conventions as (a) at 500 ms. (c) Same conventions as (a) at response time. The variation of the modulation strength over time is consistent with the variation of the encoding parameter estimate. No difference between THREAT+ and THREAT- is observed.

**Figure 5. Encoding of threat-signaling emotions in motor response lateralization measures.** (a) Top panel, scalp topography before response of the time frequency power in the 8-32 Hz band in the last 100 ms before response, for the trials where subjects responded with their left hand minus the trials where they responded with their right hand. Dots correspond to the selected electrodes, where the effect was maximal. Bottom panel: corresponding neural sources. (b) Time course of response lateralization (time frequency power activity from the contralateral electrodes minus ipsilateral electrodes to the hand used to respond 'fear') towards anger and fear when the choice was anger (red) or fear (blue). Shaded error bars indicate s.e.m. The shaded gray area indicates a significant difference in motor lateralization between Anger and Fear responses. (c) Encoding of emotion strength in response lateralization index for THREAT+ (orange) and THREAT- (green) conditions. Differences between conditions are observed at 200 ms after stimulus onset (stimulus-locked, upper panel) and at response time (response-locked, lower panel). Conventions are the same as in Fig.3 (d) Time course of neural-choice correlations, expressed as the modulation of additive bias by motor lateralization encoding residuals in arbitrary units (a.u.) stimulus-locked (upper panel) and response locked(lower panel). Conventions are the same as in Fig.3.

**Figure 6. Encoding of emotion strength as a function of reaction times (RT) in motor and parietal structures.** (a) Neural encoding of emotion strength for THREAT+ conditions in motor lateralization for fast and slow reaction times (RT): when RTs were fast, the encoding of emotion strength became significant at 150 ms and rose gradually until response; by contrast, when RTs were slow, the encoding of emotion strength became significant later at 540 ms. Shaded error bars indicate s.e.m. Thick dark and light grey lines indicate significance against zero at a cluster-corrected p-value of 0.05. Shaded grey bars indicate significant differences between fast and slow responses. Encoding latency is significantly different between fast and slow RTs \*\*\* $p < 0.001$  (b) Emotion strength encoding in parietal electrodes. Conventions are the same than (a). Fast responses are associated with a stronger neural encoding of emotion strength, but without any change in encoding latency.

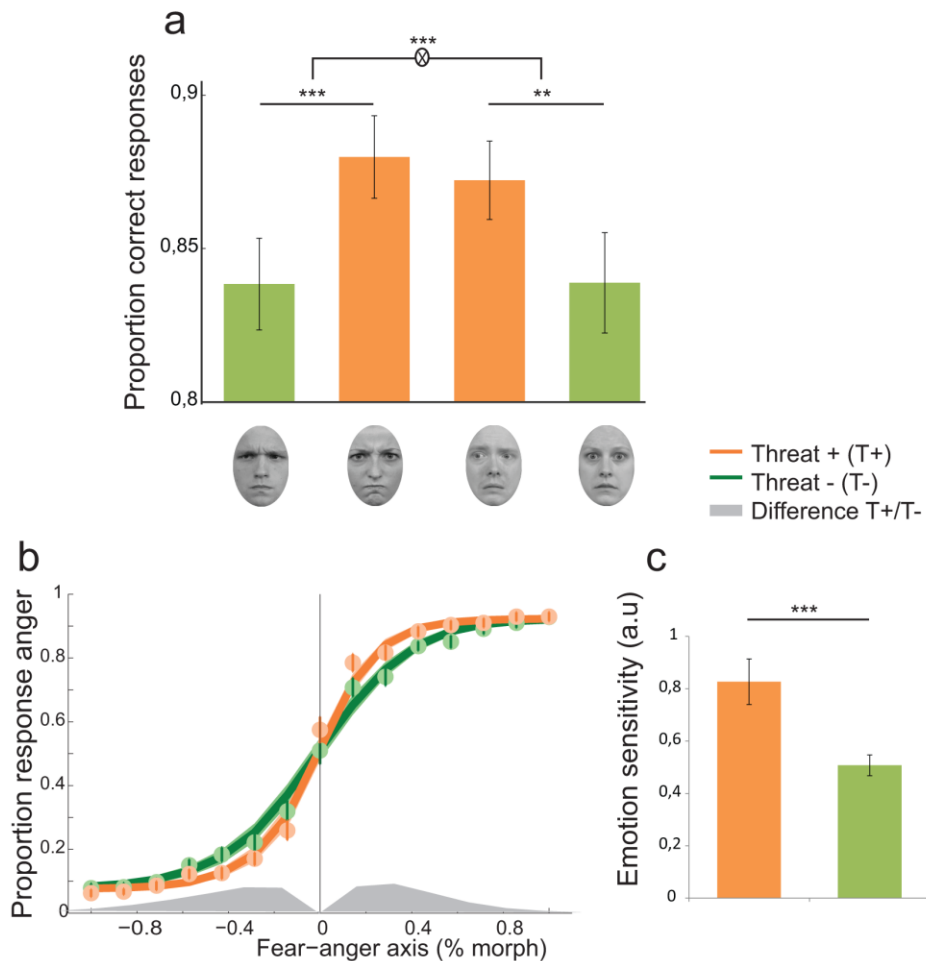
**Figure 7. Modulation of threat encoding by individual anxiety.** (a) Left panel: correlation between state anxiety and the difference of the encoding parameter estimates between THREAT+ and THREAT- conditions in temporal electrodes between 200 and 400 ms. Right panel: encoding parameter estimates in temporal electrodes split into high and low anxious individuals for both THREAT+ and THREAT- conditions between 200 and 400 ms. T+: THREAT+, T-: THREAT-. (b) Left panel: correlation between state anxiety and the encoding parameter estimates in motor lateralization signals for THREAT+ condition at 200 ms. Right panel: encoding parameter estimates in motor lateralization signals split into high and low anxious individuals for both THREAT+ and THREAT- conditions at 200 ms. \*\*\*:  $p < 0.001$ , \*:  $p < 0.05$ .

Figure 1



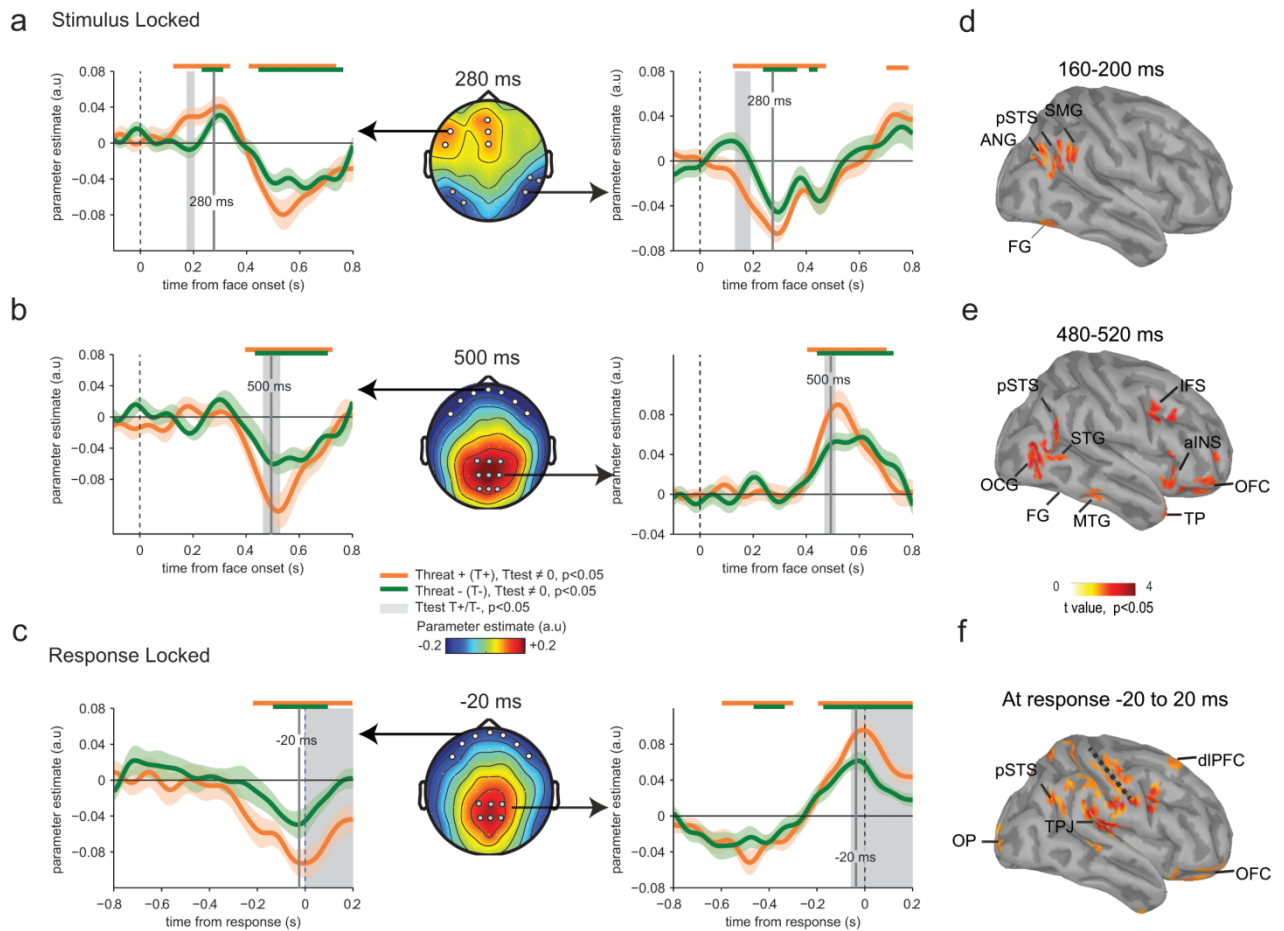
**Figure 1. Stimuli and experimental procedure.** (a) Examples of morphed expressions for one identity: morphs from neutral to intense fearful/angry expressions providing evidence for one or the other emotion. Stimuli displayed either an averted or a direct gaze. THREAT+ conditions (in orange) correspond to combinations of gaze and emotion that signal higher threat for the observer as compared to THREAT- conditions (in green). (b) Following fixation, a facial expression appeared for 250 ms, after which the participant had to indicate whether the face expressed anger or fear within 2 seconds. No feedback was provided after response.

Figure 2



**Figure 2. Enhanced recognition accuracy and perceptual sensitivity to threat-signaling emotions.** (a) Proportion of correct responses for (from left to right) averted/anger, direct/anger, averted/fear and direct/fear. THREAT+ combinations of gaze and emotion (in orange) were associated with increased recognition accuracy. (b) Psychometric function representing the proportion of 'anger' responses as a function of the evidence for anger (proportion morph, 0 for neutral, negative towards fear, and positive towards anger) for THREAT+ (orange) and THREAT- (green) combinations of gaze and emotion. Dots and attached error bars indicate the human data (mean  $\pm$  s.e.m.). Lines and shaded error bars indicate the predictions of the best-fitting model. (c) Parameter estimate for the slope of the psychometric curve (corresponding to emotion sensitivity) for THREAT+ and THREAT- combinations. \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

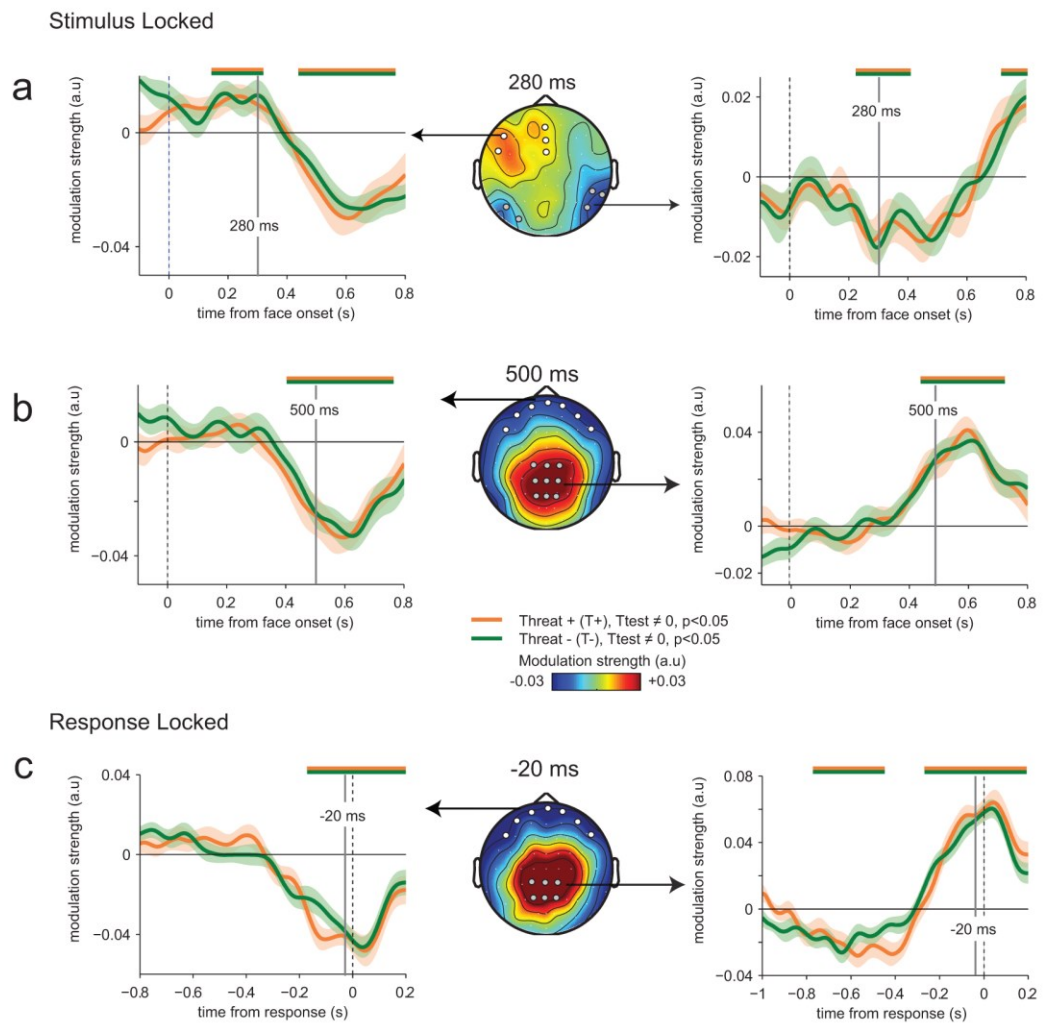
Figure 3



**Figure 3. Enhanced neural encoding of threat-signaling emotions.** (a) Middle panel: scalp topography of neural encoding at 280 ms, corresponding to its first peak of the encoding of emotion strength averaged across conditions, and expressed as mean parameter estimates in arbitrary units (a.u.). Dots indicate electrodes of interest where neural encoding was maximal. Left and right panels: encoding time course for THREAT+ and THREAT- conditions at electrodes of interest. Shaded error bars indicate s.e.m. Thick orange and green lines indicate significance against zero at a cluster-corrected  $p$ -value of 0.05. Shaded grey areas indicate significant differences between THREAT+ and THREAT- conditions at  $p < 0.05$ . (b) Same conventions as (a) at the second neural encoding peak at 500 ms. (c) Same conventions as (a) at the third neural encoding peak at response time. (e) Estimated cortical sources of the encoding difference between THREAT+ and THREAT- conditions at the time of significant difference between conditions at 180 ms. (f) Same as (e) at 500 ms. (g) Same as (e) at response time. FG: fusiform gyrus, pSTS: posterior superior temporal sulcus, SMG: supramarginal gyrus, ANG: angular gyrus, STG: superior temporal gyrus, MTG: middle temporal gyrus, OCG: occipital gyrus, alNS: anterior insula, IFS: inferior frontal sulcus, TP: temporal pole, OFC: orbitofrontal cortex, OP: occipital pole, TPJ: temporo parietal junction, dlPFC: dorsolateral prefrontal cortex.

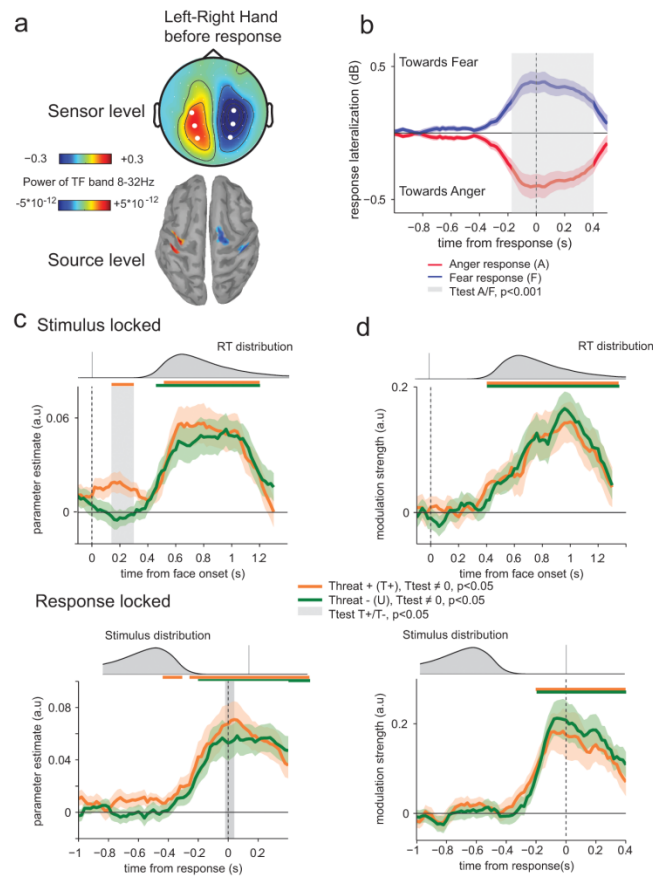


Figure 4



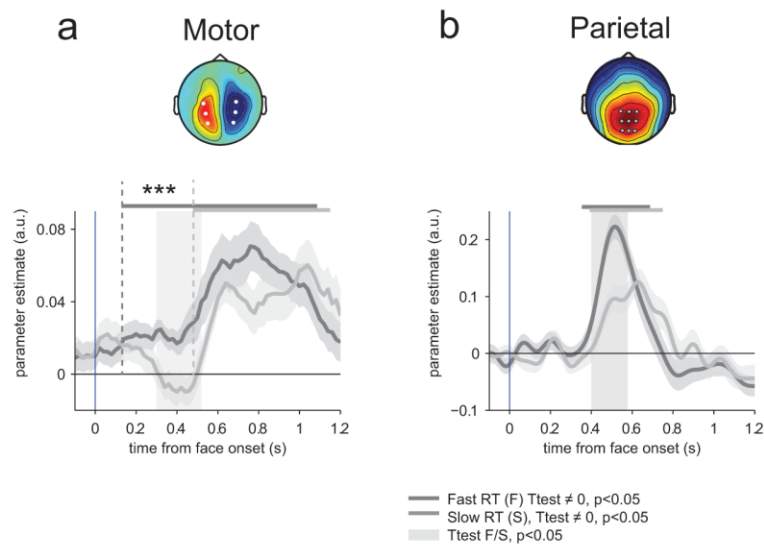
**Figure 4. Absence of threat-dependent enhancement of neural-choice correlations.** (a) Middle panel: scalp topography of neural-choice correlations, expressed as the modulation of perceptual sensitivity by EEG encoding residuals at 280 ms, same time point shown in Fig. 3a. Electrodes of interest indicated with dots are the same as in Fig. 3a. Left and right panels, time course of the modulation of perceptual sensitivity by EEG encoding residuals expressed in arbitrary units (a.u.). Same conventions as in Fig. 3a. (b) Same conventions as (a) at 500 ms. (c) Same conventions as (a) at response time. The variation of the modulation strength over time is consistent with the variation of the encoding parameter estimate. No difference between THREAT+ and THREAT- is observed.

Figure 5



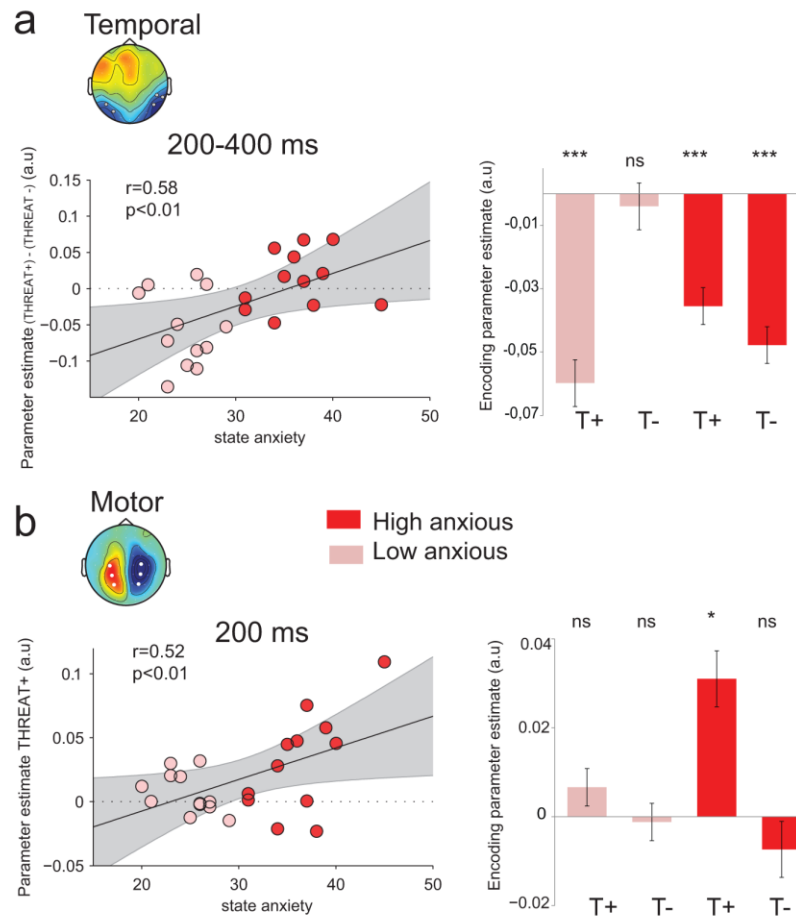
**Figure 5. Encoding of threat-signaling emotions in motor response lateralization measures.** (a) Top panel, scalp topography before response of the time frequency power in the 8-32 Hz band in the last 100 ms before response, for the trials where subjects responded with their left hand minus the trials where they responded with their right hand. Dots correspond to the selected electrodes, where the effect was maximal. Bottom panel: corresponding neural sources. (b) Time course of response lateralization (time frequency power activity from the contralateral electrodes minus ipsilateral electrodes to the hand used to respond 'fear') towards anger and fear when the choice was anger (red) or fear (blue). Shaded error bars indicate s.e.m. The shaded gray area indicates a significant difference in motor lateralization between Anger and Fear responses. (c) Encoding of emotion strength in response lateralization index for THREAT+ (orange) and THREAT- (green) conditions. Differences between conditions are observed at 200 ms after stimulus onset (stimulus-locked, upper panel) and at response time (response-locked, lower panel). Conventions are the same as in Fig.3 (d) Time course of neural-choice correlations, expressed as the modulation of additive bias by motor lateralization encoding residuals in arbitrary units (a.u.) stimulus-locked (upper panel) and response locked (lower panel). Conventions are the same as in Fig.3.

Figure 6



**Figure 6. Encoding of emotion strength as a function of reaction times (RT) in motor and parietal structures.** (a) Neural encoding of emotion strength for THREAT+ conditions in motor lateralization for fast and slow reaction times (RT): when RTs were fast, the encoding of emotion strength became significant at 150 ms and rose gradually until response; by contrast, when RTs were slow, the encoding of emotion strength became significant later at 540 ms. Shaded error bars indicate s.e.m. Thick dark and light grey lines indicate significance against zero at a cluster-corrected p-value of 0.05. Shaded grey bars indicate significant differences between fast and slow responses. Encoding latency is significantly different between fast and slow RTs  $***p < 0.001$  (b) Emotion strength encoding in parietal electrodes. Conventions are the same than (a). Fast responses are associated with a stronger neural encoding of emotion strength, but without any change in encoding latency.

Figure 7



**Figure 7. Modulation of threat encoding by individual anxiety.** (a) Left panel: correlation between state anxiety and the difference of the encoding parameter estimates between THREAT+ and THREAT- conditions in temporal electrodes between 200 and 400 ms. Right panel: encoding parameter estimates in temporal electrodes split into high and low anxious individuals for both THREAT+ and THREAT- conditions between 200 and 400 ms. T+: THREAT+, T-: THREAT-. (b) Left panel: correlation between state anxiety and the encoding parameter estimates in motor lateralization signals for THREAT+ condition at 200 ms. Right panel: encoding parameter estimates in motor lateralization signals split into high and low anxious individuals for both THREAT+ and THREAT- conditions at 200 ms. \*\*\*:  $p<0.001$ , \*:  $p<0.05$ .

## **Part 3: Gaze direction influence on emotion categorization: a few unanswered questions (follow-up behavioral studies)**

The experiment described in the previous chapter demonstrated that the perception of threatening emotions is directly influenced by the emitter's gaze direction. This was reflected in both subjects' behavior with increased performance and sensitivity to emotions relevant to the observer, and at the neural level with increased 'encoding' of these emotions in sensory and motor regions. A number of follow-up questions emerged from these findings. Do different visual characteristics of the face such as distinct spatial frequencies drive the observed effect on emotion sensitivity? Is the contextual interaction between gaze direction and emotion specific to anger and fear expressions, as hypothesized by 'appraisal' theories of emotion? Is the observed sensitivity enhancement reciprocal, in other words, does emotion similarly improve the categorization of gaze direction (direct vs. averted)? Does this effect (obtained in young healthy individuals) extend to the general population (all ages, different backgrounds?). In part 3, I will describe four behavioral studies which addressed these different questions.

### **3.1- General approach in behavioral analyses**

The same approach was used to analyze the data of the different behavioral studies and will therefore be generally described. First we started with statistical analyses on accuracy that simply consisted in running repeated-measures ANOVA with gaze direction (direct/averted), emotion (anger/fear), and intensity (levels of morphs) as within-subjects factors. Second, we performed model-guided analyses of the behavioral data to characterize effects observed on accuracy. Depending on the observed results on accuracy and on our hypotheses, we compared between different models that will be described in each section separately. We used Bayesian model selection based on the model evidence, estimated by a 10-fold cross-validation estimation of model log-likelihood, which penalizes implicitly for model complexity without making

particular (and necessarily imprecise) approximations such as the Bayesian Information Criterion or the Akaike Information Criterion. We applied both fixed-effects and random-effects statistics. The fixed-effects comparison assumes that all participants used the same underlying model to generate their behavior, such that the overall model evidence for a given model is proportional to the product of model evidence for the model for all participants. Based on this model evidence, we compared different models by computing their Bayes factor as the ratio of model evidence of the compared model (Jeffreys, 1961; Kass and Raftery, 1995). The random-effects comparison is more conservative in allowing different participants to use different models to generate their behavior, and aims at inferring the distribution over models that participants draw from (Penny et al., 2010). For this comparison, we computed support for the winning model by the exceedance probability ( $p_{exc}$ ), which is the probability that participants were more likely to choose this model to generate behavior over any alternative model.

### **3.2- Question 1: Characterization of the effect: contribution of high and low spatial frequencies during Fear/Angry categorization (Experiment 3)**

The results of Experiment 2 provided evidence for the influence of gaze direction on the sensory encoding of emotional expressions. In this experiment, we questioned whether different spatial frequencies of the emotional image contribute in dissociable manners to these mechanisms. Indeed, high and low spatial frequencies of facial expressions seem to be processed differently in the human brain. The modulation of spatial frequency conveyed by the faces reveals that different frequency bands contribute distinctively in the recognition of emotions (Smith and Schyns, 2009). Although low spatial frequencies in fearful stimuli increase the activity of emotion selective brain regions (Vuilleumier et al., 2003) and brain waves (Vlamings et al., 2009), reverse-correlation techniques during emotion recognition (Smith and Schyns, 2009) have revealed that signals of threat such as fear or anger are conveyed by high spatial frequency bands. Therefore it remains unclear whether signals of threat are transmitted through high or low spatial frequency bands and whether these two fre-

quency bands influence differently the categorization of threatening emotions. By separating high and low spatial frequencies, we sought to determine whether one of these frequency bands or both of them contribute more strongly to the observed sensitivity enhancement to threat-signaling emotions (Experiment 2). We hypothesized that increases in emotion sensitivity should be provided by high spatial frequencies (HSF), whereas low spatial frequencies (LSF) could rather change the decision criterion because of the coarseness of facial features provided by LSF.

## **Methods**

### *Participants*

2 groups of 20 participants were included in the experiment (Group 1: 10 females; mean age  $21.1 \pm 0.2$  years, Group 2: 10 females, mean age  $22 \pm 0.6$  years ). All participants were right-handed, with a normal vision and had no neurological or psychiatric history. They provided written informed consent according to institutional guidelines of the local research ethics committee (Declaration of Helsinki) and were paid for their participation. Group 1 performed the task on HSF stimuli while Group 2 performed the task on LSF stimuli.

### *Stimuli*

Stimuli consisted of the 1080 images of facial expressions used in Experiment 2 (after calibration for perceived emotional intensity), filtered into high and low spatial frequencies (Figure 23). Butterworth filters were applied to intact broad spatial frequency images (BSF) using ImageJ (Schneider et al., 2012), with a high pass  $> 24$  cycles/images for high spatial frequencies (HSF) ( $=32$ ), and a low pass  $< 6$  cycles/images for low spatial frequencies (LSF) ( $=4$ ) (Vuilleumier et al., 2003).

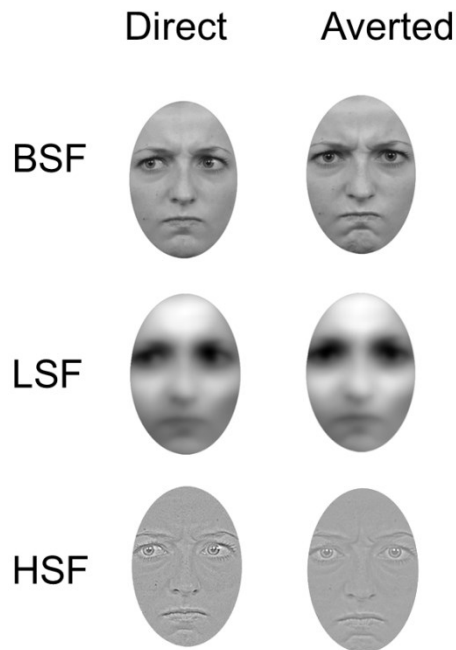


Figure 23 Example of stimuli used in Experiment 3. Left, averted gaze, Right, direct gaze. BSF=Broad Spatial Frequency, HSF=High Spatial Frequency, LSF=Low Spatial Frequency

### *Experimental procedure*

The experimental procedure is similar to the one described in Experiment 2 on a trial to trial basis. Using the Psychophysics-3 Toolbox for MATLAB (Brainard, 1997; Pelli, 1997), stimuli were projected on a black screen. Each trial was initiated with a white oval delimiting the faces that was kept during all the trial. The white oval appeared for approximately 500 ms, followed by a white fixation point presented at the level of the eyes for approximately 1000 ms (to keep the fixation to the upcoming faces natural and avoid eye movements from the center of the oval to eye regions), then the stimuli appeared for 250 ms. Participants' task was to decide whether the faces expressed Anger or Fear by pressing one of the two "ctrl" keys of the keyboard using their right left indexes. An Anger/Fear mapping was used (e.g Anger: Left hand, Fear: Right hand) kept constant for each subject, counterbalanced over all subjects.

In an inter-subject design, 20 subjects only performed the task on HSF stimuli, and 20 subjects only performed the task on LSF stimuli. All stimuli were presented once, resulting in a total of 1080 trials. The experiment was divided in 9 experimental blocks, each consisting of 120 trials, balanced in the number of emotions, directions



of gaze, gender and levels of morphs. After each block, the percentage of correct responses was shown to the participants to keep them motivated.

## Results

### *Performance*

Repeated-measures ANOVA was performed on the percentage of correct responses, with gaze direction (direct/averted), emotion (anger/fear), and intensity (7 levels of morphs) as within-subjects factors, and group type (HSF group/LSF group) as between-subjects factor.

First, accuracy increased with emotion strength (main effect of emotion strength HSF group  $F_{6,114}=143.9$ ,  $p<0.001$ , LSF group  $F_{6,114}=141.6$ ,  $p<0.001$ ). Second, as expected, an emotion by gaze interaction (HSF group  $F_{1,19}=35.7$ ,  $p<0.001$ , LSF group  $F_{1,19}=9.8$ ,  $p<0.01$ ) was observed on accuracy replicating the results of Experiment 2, with increased performances to THREAT+ conditions (anger with a direct gaze, fear with an averted gaze, see Figure 24a,b). Importantly, there was no interaction with group type on accuracy (all  $p>0.5$ ). However, a main effect of group was observed with increased accuracy in the HSF group, as expected from coarser facial features for LSF (t-test between groups  $t_{38}=-3.4$ ,  $p<0.01$ ).

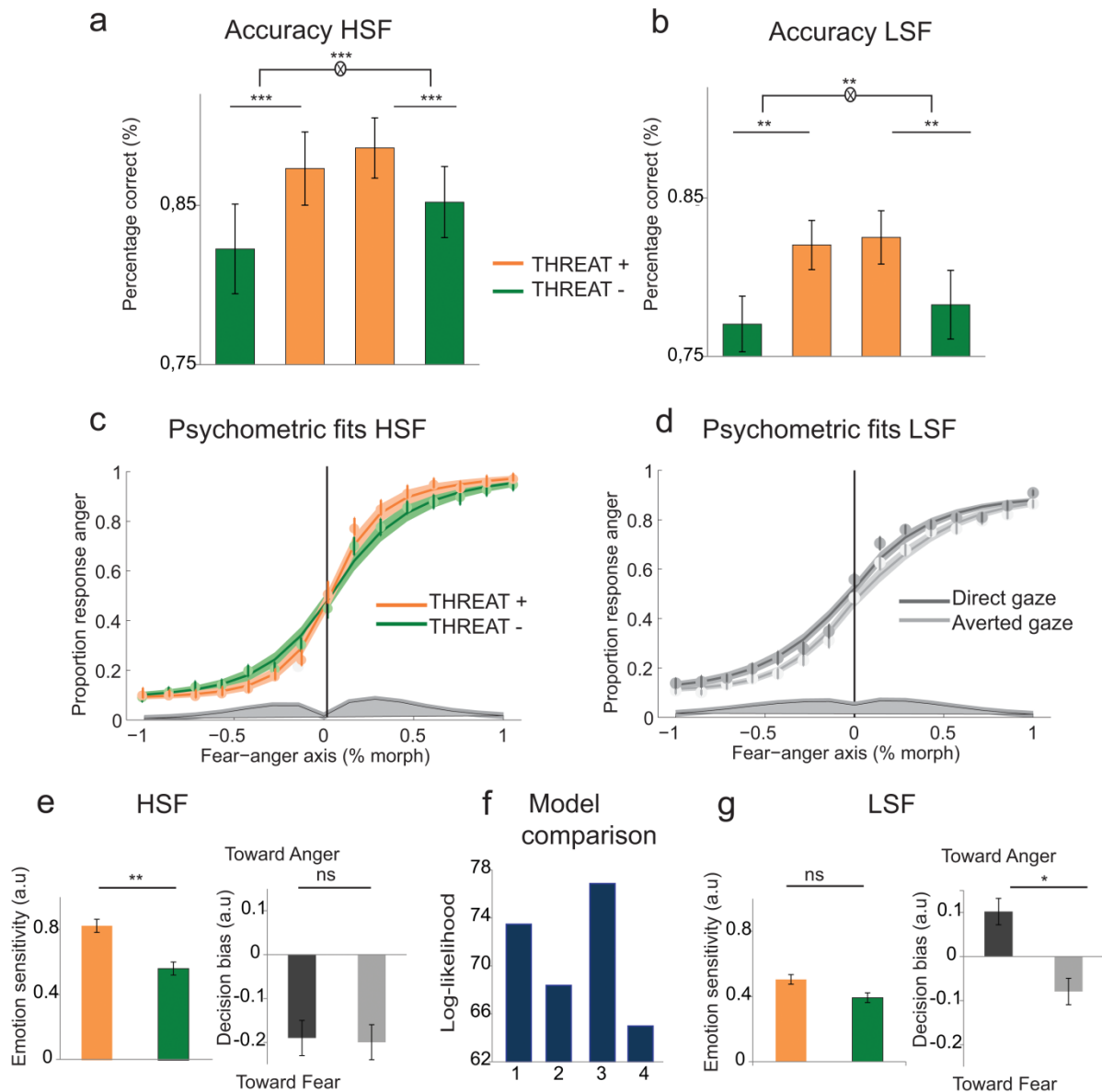


Figure 24. Results of Experiment 3. a) Accuracy for the four conditions of interest in HSF group, from left to right, anger averted, anger direct, fear averted, fear direct. b) same as a, for LSF group. c) psychometric fits of model 3 that allowed an effect on both sensitivity and bias in HSF group. Differences between the fits for direct and averted gaze are plotted in grey and show that the effect of gaze is maximal for ambiguous stimuli consistently with an effect on the sensitivity. d) same as c for LSF group. Here the different between fits for direct and averted gaze shows a difference on the neutral stimuli, consistently with an effect on the decision bias. e) left, emotion sensitivity for THREAT + (orange) and THREAT - (green) conditions; right, decision bias for direct (dark grey) and averted gaze (light grey). f) fixed-effects model comparison between 1) HSF and LSF model #2, 2) HSF and LSF model #1, 3) HSF model #1 and LSF model #2, 4) HSF model #2 and LSF model #1. When subtracting the log likelihood for model #0, the best model assigns a change in sensitivity for HSF group and a change in bias for LSF group. g) same as e for LSF group.

### Model Selection

We started with the simplest model (model #0) that could account for subject's decisions using a psychometric model such as:

$$P(\text{Anger}) = \Phi(w \cdot \mathbf{x} + b)$$

where  $P(\text{Anger})$  corresponds to the probability of judging the face as angry,  $\Phi$  to the cumulative normal function,  $w$  to the perceptual sensitivity to the display emotion,  $\mathbf{x}$  to the evidence in favor of Anger or Fear in each trial (emotion strength, from -7 for an intense expression of fear to +7 for an intense expression of anger), and  $b$  to an additive, stimulus-independent bias toward one of the two responses/emotions. Similarly to Experiment 2, we compared this 'null' model with two additional models which account for the observed increase in recognition accuracy for THREAT+ combinations of gaze and emotion. A first possibility (model 1) would be that gaze direction biases emotion recognition in favor of the interpretation signaling higher threat (anger for a direct gaze, fear for an averted gaze). A second possibility (model 2) is that gaze direction might selectively increase sensitivity to emotions signaling higher threat (modeled by a different sensitivity to emotions in THREAT+ vs. THREAT - conditions). Finally, we also considered a model (model 3) where both changes on the decision bias and the sensitivity were permitted to compare between the 2 groups of subjects.

In the HSF group, Bayesian model selection revealed that a sensitivity enhancement for THREAT+ combinations explained better the behavioral data than a criterion shift (Bayes Factor  $\approx 10^{1.5}$ ,  $p_{\text{exc}} > 0.68$ ). In opposition, in the LSF group, a criterion shift toward THREAT+ combinations explained better the behavioral data than a change in sensitivity (Bayes Factor  $\approx 10^{3.6}$ ,  $p_{\text{exc}} > 0.84$ ). To compare between the two groups, first we looked at whether effects on sensitivity and bias persisted if the model considering both changes is fitted to the data (model 3). In the HSF group, only sensitivity estimates extracted from model 3 were significantly increased for THREAT+ combinations of gaze and emotion ( $t_{19} = 2.9$ ,  $p < 0.01$ ) (Figure 24c,e). On the contrary, in the LSF group, only decision bias estimates were significantly different ( $t_{19} = 2.8$ ,  $p = 0.01$ ), with decision bias pulled toward THREAT+ combinations of gaze and emotion (Figure 24d,g). Finally, to have a measure of comparison between the two groups, we use fixed-effects analyses to contrast models that consider 1) model

1 for both HSF and LSF groups, 2) model 2 for both HSF and LSF groups, 3) model 1 for LSF group and model 2 for HSF group, and 4) model 1 for HSF group and model 2 for LSF. We used the log-likelihood of model 0 as a baseline to compare between these four possibilities. This analysis showed that the third possibility was the most likely as it maximized the cross-validated log-likelihood (Figure 24f), which further suggests that a change in sensitivity was conveyed by gaze direction in the HSF group whereas a change in decision bias was conveyed by gaze direction in the LSF group.

## **Conclusion**

The results of Experiment 3 show that the increased sensitivity to threat-signaling emotions is driven by high spatial frequencies. This result is consistent with a change in emotion sensitivity that provides a measure of the sensory precision, i.e. improved perception of visual features provided by the emotional expression. It is also compatible with the increased sensitivity observed in face selective sensory areas in Experiment 2. Threat-signaling emotions thus seem to be conveyed through high spatial frequency channels, contradicting the view that threat is primarily transmitted through low spatial frequencies (Vuilleumier et al., 2003), but in accordance with the fact that anger and fear emotions are passed on through high spatial frequencies (Smith and Schyns, 2009).

When emotion categorization had to be performed on low spatial frequencies of faces, gaze rather changed the decision bias toward threat-signaling emotions. The coarse information provided in LSF emotional faces could not change the sensitivity to threatening signals, but instead influenced the decision bias shifting decisions toward anger when gaze was direct, and choices toward fear when gaze was averted. The effect of gaze on the categorization of threatening emotions therefore seems to be robust although the mechanisms by which its influence is exerted on recognition performance varies as a function of the visual information provided.

### 3.3- Question 2: Specificity of the effect: Categorization of Happy/Angry expressions (Experiment 4)

In Experiment 2, emotions of negative valence were manipulated, and anger and fear were chosen precisely because gaze impacts their processing in an interactive fashion. Neural and behavioral correlates of this interaction were observed, and we further wanted to control whether this interaction is specific to the manipulated emotions. To do so, we replaced fearful expressions that, when compared to angry, share the valence (negative) but have an opposite effect of gaze, with happy expressions that are of opposite valence (positive vs negative) but share the same impact of gaze. This assumption is based on the *shared signal hypothesis* that suggests that happy expressions, just like angry expressions, signal approach and are therefore better recognized with a direct gaze. In opposition, fear and sadness are more 'avoidance oriented emotions' and are therefore better categorized with an averted gaze (Adams and Kleck, 2003, 2005). In Experiment 4, happy and angry faces were presented to subjects, while they performed a perceptual categorization task using the same design as in Experiment 2 (Happy or Angry?).

#### Methods

##### *Participants*

Nineteen healthy subjects (10 females; mean age,  $23.6 \pm 0.8$  years) participated in the behavioral experiment. All participants were right-handed, with a normal vision and had no neurological or psychiatric history. They provided written informed consent according to institutional guidelines of the local research ethics committee (Declaration of Helsinki) and were paid for their participation.

##### *Stimuli*

Stimuli consisted of the same 36 identities (18 females) adapted from the Radboud Faces Database (Langner et al., 2010) used in Experiment 2. They varied in emotion (neutral, angry or happy expressions) and gaze direction (direct toward the participant or averted 45° to the left or right). Using Adobe Photoshop CS5.1 (Adobe Systems, San Jose CA), faces were modified to remove any visible hair, resized and repositioned so that eyes, nose and mouth appeared within the same circumference. All images were converted to greyscale and cropped into a 280 x 406 pixel oval centered within a 628 x 429 pixel black rectangle.

To vary the intensity of emotional expressions, faces were morphed from neutral to angry expressions and from neutral to happy expression using FantaMorph (Abrosoft, <http://www.fantamorph.com/>). At first, 7 levels of morphs from neutral to angry expressions and from neutral to happy expressions (separately for direct and averted gaze stimuli) were created using a simple linear morphing transformation. This resulted in 30 conditions for each identity: 7 levels of morphs \* 2 emotions \* 2 gaze directions = 28 and 2 neutral stimuli with direct and averted gaze.

We then calibrated (linearized) the morphing between angry and happy expressions by performing an intensity rating pre-test of the emotional expressions and adjusting the morphs based on the results, as we did in Experiment 2.

To determine whether both emotions were perceived as equally intense at all the levels of the morphs, a pre-test was performed on 20 subjects (11 females, mean age, 22.7 ± 0.7 years). A main effect of emotion ( $F_{1,19}=6.8$ ,  $p<0.05$ ) on perceived intensity ratings was observed, happy stimuli being judged as more intense than angry stimuli (Figure 25). To adjust for this difference between emotions, we linearized the curve of perceived intensities of Anger as a function of the levels of the morphs. We then extracted the percentage values of morphs of angry and happy faces needed such as the perceived intensities fall into this linearized curve. New morphs were correspondingly regenerated.

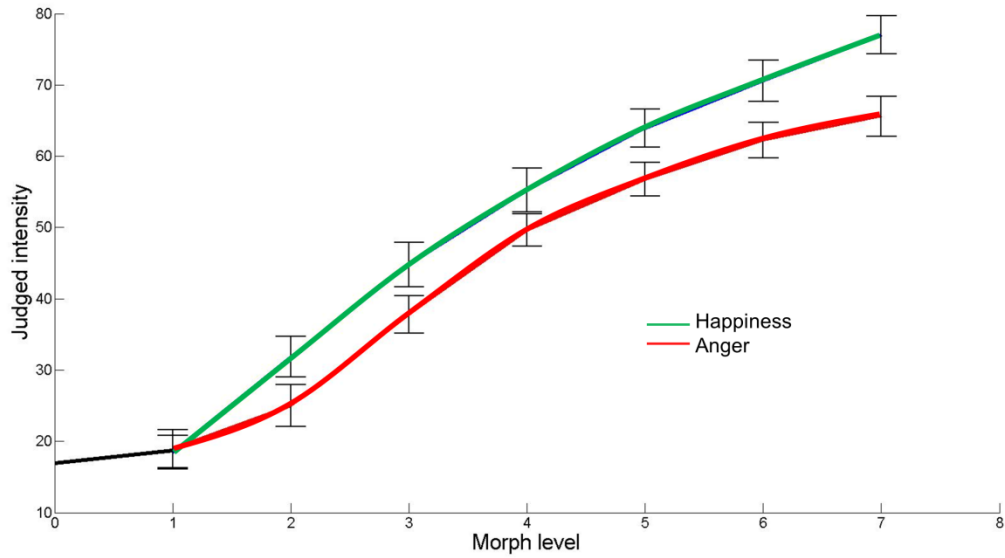


Figure 25 Pre-test results of Experiment 4. Judged intensity on happy and angry faces. The green curve represents happy stimuli and the red curve represents angry stimuli.

To summarize, the stimuli comprise of 36 identities with an averted gaze condition and a direct gaze condition, each with 7 levels of anger and 7 levels of happiness equalized in perceived emotional intensities and a neutral condition, resulting in a total of 1080 items (Figure 26).

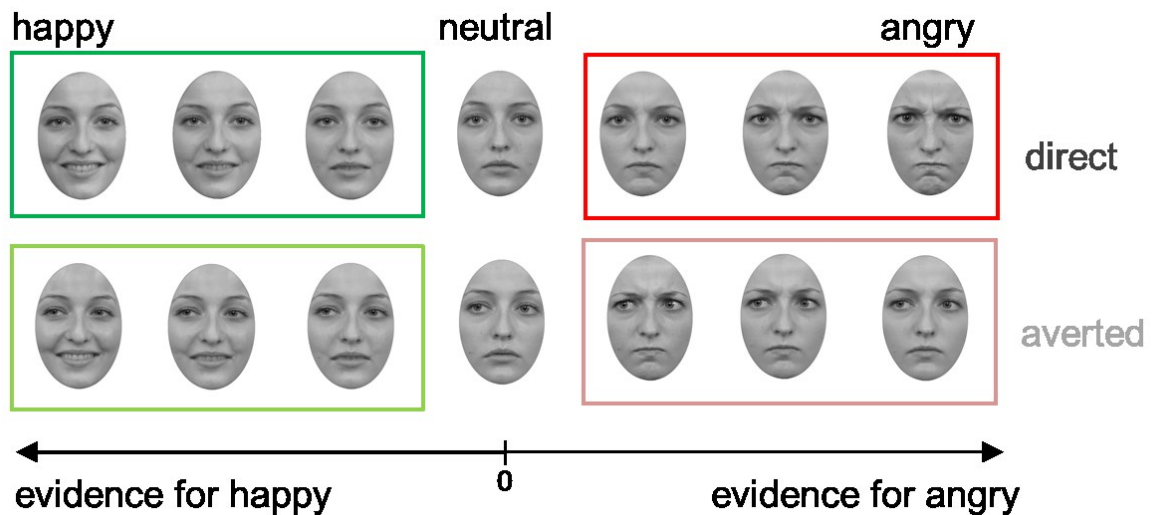


Figure 26. Examples of stimuli used in Experiment 4. Only 3 levels of morphs are shown for the example.

## Experimental procedure

Using the Psychophysics-3 Toolbox (Brainard, 1997; Pelli, 1997), stimuli were projected on a black screen. Each trial was initiated with a white oval delimiting the faces that was kept during all the trial. The white oval appeared for approximately 500 ms, followed by a white fixation point presented at the level of the eyes for approximately 1000 ms (to keep the fixation to the upcoming faces natural and avoid eye movements from the center of the oval to eye regions), then the stimuli appeared for 250 ms. Participants' task was to decide whether the faces expressed Anger or Happiness by pressing one of the two "ctrl" keys of the keyboard using their right left indexes. An Anger/Happy mapping was used (e.g Anger: Left hand, Happy: Right hand) kept constant for each subject, counterbalanced over all subjects.

All stimuli were presented once, resulting in a total of 1080 trials. The experiment was divided in 9 experimental blocks, each consisting of 120 trials, balanced in the number of emotions, directions of gaze, gender and levels of morphs. After each block, the percentage of correct responses was shown to the participants to keep them motivated.

## Results

### *Performance*

Subjects performed better for angry stimuli ( $F_{1,18}=17.5$ ,  $p<0.001$ ) as compared to happy stimuli (Figure 27a). Performances increased with the emotional intensity ( $F_{1,18}=158.4$ ,  $p<0.001$ ), and this modulation by intensity was more important for happy stimuli (Emotion by intensity interaction  $F_{6,108}=22.3$ ,  $p<0.001$ ). In addition, a triple interaction between Emotion, Gaze and intensity ( $F_{6,108}=3.2$ ,  $p<0.05$ ) was observed, and explained by gaze effects only in low intensities of happy stimuli. For the 3 first levels of happy stimuli, subjects performed better for direct gaze as compared to averted gaze ( $F_{1,18}=6.2$ ,  $p<0.05$ ).



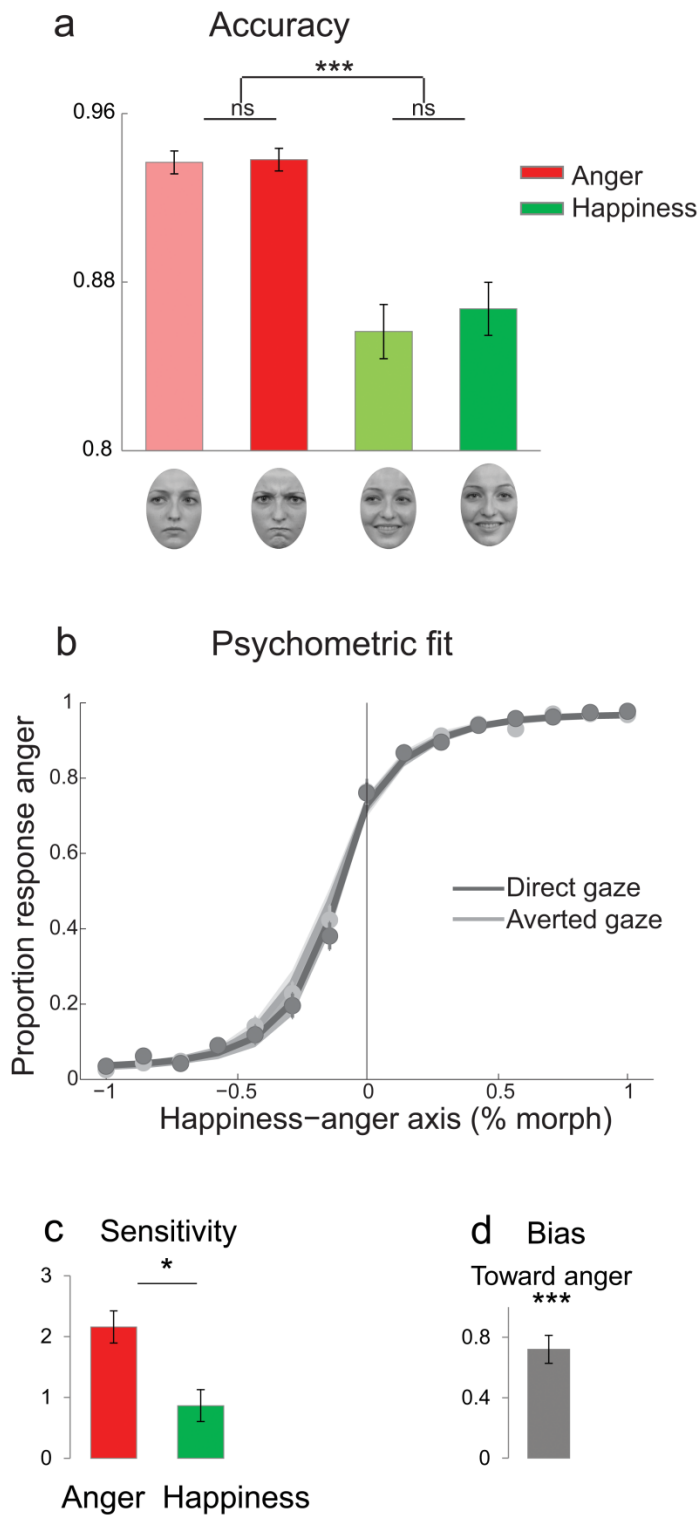


Figure 27. Results of Experiment 4. a) Accuracy for the four conditions of interest, from left to right, averted anger, direct anger, averted happiness, direct happiness. b) Psychometric fit when considering a model where all parameters are free (4 sensitivities for the four conditions and 2 biases for direct and averted gaze): no difference is observed between conditions. c) Estimated sensitivity parameters of the winning model (model 1): increased sensitivity to anger. d) Estimated bias parameter of model 1: bias toward anger. \* $p < 0.05$ ,  $p < 0.001$

### *Model selection*

We compared between numerous models describing participants' behavior:

- The 'null' model, model 0, with no effect of emotion or gaze (general sensitivity to anger and fear and general bias to direct and averted gaze).
- Model 1 allowed different emotion sensitivities between anger and happy, based on the results observed on performance that showed a strong effect of emotion.
- Model 2 allowed different sensitivities to all four conditions: anger direct/ anger averted/ happy direct/ happy averted. We considered this model based on the accuracy results that showed a small effect of gaze direction, only on low intensities of happy stimuli. This model allows differentiating between the effects of gaze on angry and happy stimuli.

Finally, based on our hypothesis about direct gaze being more relevant in the present context, we further considered models that could account for the effect of gaze:

- Model 3 allowed different sensitivities to direct and averted gaze.
- Model 4 assigned different response biases for direct and averted gaze.

Model 1 was by far the best model as its probability exceedance was above 85% when compared with all the other models (fixed-effect comparison with second best model Bayes Factor  $\approx 10^{3.6}$ ). The estimated sensitivities and bias parameters from model 1 were in accordance with the effects on performance, as subjects increased their sensitivity to anger as compared to happy, and were more biased toward angry responses (Figure 27c,d). In opposite valence context, subjects were biased towards negative valence and their sensitivity to negative stimuli was enhanced. Surprisingly, gaze direction did not affect anger categorization as in Experiment 2. However, this can be explained by the context of the experiment that included both happy and angry expressions, and therefore rendered all threatening emotions (here anger) as more relevant than non-threatening ones (happiness) independently of gaze.

## **Conclusion**

The results of this experiment suggest that when threatening emotions are presented against positive emotions, gaze direction does not play the same role as compared to when observers have to judge among negative emotions. The presence of happy faces made angry faces generally more relevant to the observer because they signal threat (whether directly relevant for the observer or not). In this setting, gaze direction facilitated the judgment of the generally less relevant emotion, happiness, when its combination with gaze was more salient and only when presented in very ambiguous displays.

Model comparisons combined with results on accuracy however show that the most marked effects only reveal both an increased sensitivity of anger expression and a response bias toward anger, highlighting the importance of threat recognition independently of gaze direction in this context.

These findings thus draw attention on the different types of settings that change emotion recognition: here the context can be viewed as a context of “judgment” (Tanaka-Matsumi et al., 1995), whereby choosing between positive and negative emotions shifted the relevant aspect to the observer to threat (anger) independently of gaze.

Finally and most importantly related to the main question addressed in this study, these findings confirm the specificity of the effect of gaze on the categorization of anger and fear emotions, as the combinations with gaze that benefited from increased sensitivities appear to apply selectively to the categorization between negative (and thus potentially threatening) emotions.

### **3.4- Question 3: Reciprocity of the effect: Emotion influence on gaze categorization (Experiment 5)**

Experiments 1, 2 and 3 reveal a strong influence of gaze on the processing of emotional expressions of threat during emotion categorization. As described in the litera-

ture, emotional displays seem to reciprocally influence gaze judgments: angry and happy (approach-oriented) facial expressions are more likely to be judged as looking directly at the observer than fearful and neutral expressions (Lobmaier et al., 2008; Ewbank et al., 2009). Averted gaze is processed faster and more accurately with a fearful face, while it is the case for direct gaze with an angry expression (Adams and Franklin, 2009). These experiments suggest that the processing of co-emitted emotion and gaze cues is interdependent. Further, Experiment 1 showed that the combination of social cues emerged independently of the task, suggesting that co-emitted social cues are combined as long as they are relevant to the observer, independently from task demands and focus of attention. Once more, these past studies did not identify the mechanisms by which emotion reciprocally biases gaze categorization. Although we expected that emotion should reciprocally influence gaze categorization, the question of whether it does so by changing the decision bias, the emotion sensitivity or both remained unanswered. The next study therefore addressed the influence of emotional displays on faces with morphed gaze directions during a gaze categorization task (gaze direct or averted?).

## **Methods**

### *Participants*

Twenty three healthy subjects (12 females; mean age,  $23.5 \pm 0.7$  years) participated in the behavioral experiment. All participants were right-handed, with a normal vision and had no neurological or psychiatric history. They provided written informed consent according to institutional guidelines of the local research ethics committee (Declaration of Helsinki) and were paid for their participation.

### *Stimuli*

Stimuli consisted of 36 identities (18 females) adapted from the Radboud Faces Database (Langner et al., 2010) that varied in emotion (neutral, angry or happy expressions) and gaze direction (direct toward the participant or averted to the left or right).

Using Adobe Photoshop CS5.1 (Adobe Systems, San Jose CA), faces were modified to remove any visible hair, resized and repositioned so that eyes, nose and mouth appeared within the same circumference. All images were converted to grey-scale and cropped into a 280 x 406 pixel oval centered within a 628 x 429 pixel black rectangle.

To vary the extent to which gaze was averted, faces were morphed from direct to averted gaze and using FantaMorph (Abrosoft <http://www.fantamorph.com/>). 8 levels of morphs from direct to averted gaze were created (for all the three emotional expressions of anger, fear and neutral) using a simple linear morphing transformation. This resulted in 27 conditions for each identity: 8 levels of averted gaze morphs \* 3 emotions = 24 and 3 direct gaze stimuli expressing either neutral, or angry or fearful expressions (Figure 28).

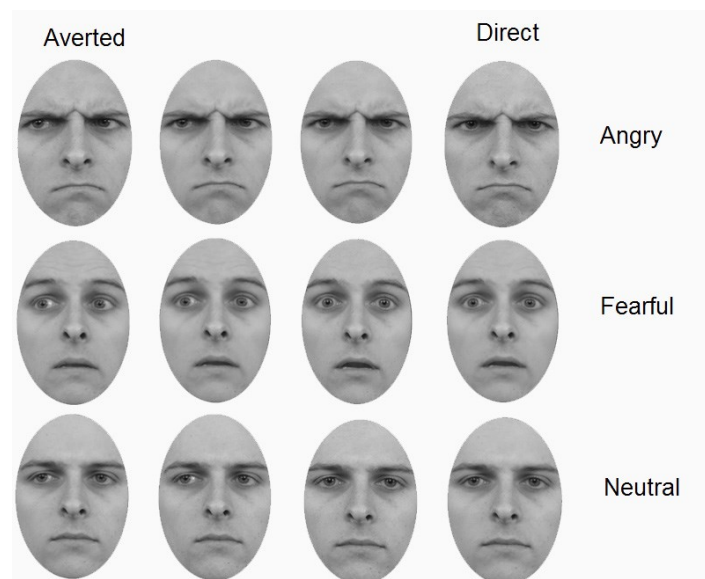


Figure 28. Examples of stimuli for Experiment 4. Gaze is morphed from direct to averted for the 3 types of emotions.

### *Experimental procedure*

Using the Psychophysics-3 Toolbox (Brainard, 1997; Pelli, 1997), stimuli were projected on a black screen. Each trial was initiated with a white oval delimiting the faces that was kept during all the trial. The white oval appeared for approximately 500 ms, followed by a white fixation point presented at the level of the eyes for approxi-

mately 1000 ms (to keep the fixation to the upcoming faces natural and avoid eye movements from the center of the oval to eye regions), then the stimuli appeared for 250 ms. Participants' task was to decide whether the faces displayed a direct or an averted gaze by pressing one of the two "ctrl" buttons of the keyboard using their right left indexes. A direct/averted mapping was used (e.g., direct: Left hand, averted: Right hand) kept constant for each subject, counterbalanced over all subjects.

All stimuli were presented once, resulting in a total of 972 trials. The experiment was divided in 9 experimental blocks, each consisting of 108 trials, balanced in the number of emotions, directions of gaze, gender and levels of morphs. After each block, the percentage of correct responses was shown to the participants to keep them motivated.

## **Results**

One subject realized that giving the 'averted' response allowed him to obtain good performance, as there are 8 times more averted gaze stimuli than direct gaze stimuli presented in each block. He therefore stopped performing the task correctly in the last 4 blocks, as he only pressed the 'averted' response at each trial. This subject was further excluded from analysis. No other subject adopted such strategy and data were thus analyzed on a total of 22 subjects.

### *Performances*

Accuracy increased ( $F_{7,147} = 224$ ,  $p < 0.001$ ) with levels of averted gaze direction. Accuracy also varied with the emotion displayed by the stimuli as it was highest for fearful stimuli (significant difference with neutral condition  $T_{21} = 6.7$  and with angry condition  $T_{21} = 13.9$ ,  $p < 0.001$ ), than for neutral stimuli, and finally the lowest for angry stimuli (significant difference between angry and neutral  $T_{21} = -11$ ,  $p < 0.001$ ). Importantly, an interaction between emotion and gaze was also observed on the percentage of correct responses ( $F_{2,21} = 42.8$ ,  $p < 0.001$ ). Performance was increased for fear as compared to anger in averted gaze condition ( $t_{21} = -14.5$ ,  $p < 0.001$ ), but the opposite applied to direct gaze condition ( $t_{21} = 4.4$ ,  $p < 0.001$ ). Neutral conditions were

at an intermediate level between anger and fear, and significantly differed from both anger ( $T_{21}=-9.8$ ,  $p<0.001$ ) and fear ( $T_{21}=8.1$ ,  $p<0.001$ ) in the averted gaze condition, and only from fear in the averted gaze condition ( $T_{21}=-3.3$ ,  $p<0.01$ ) (Figure 29a).

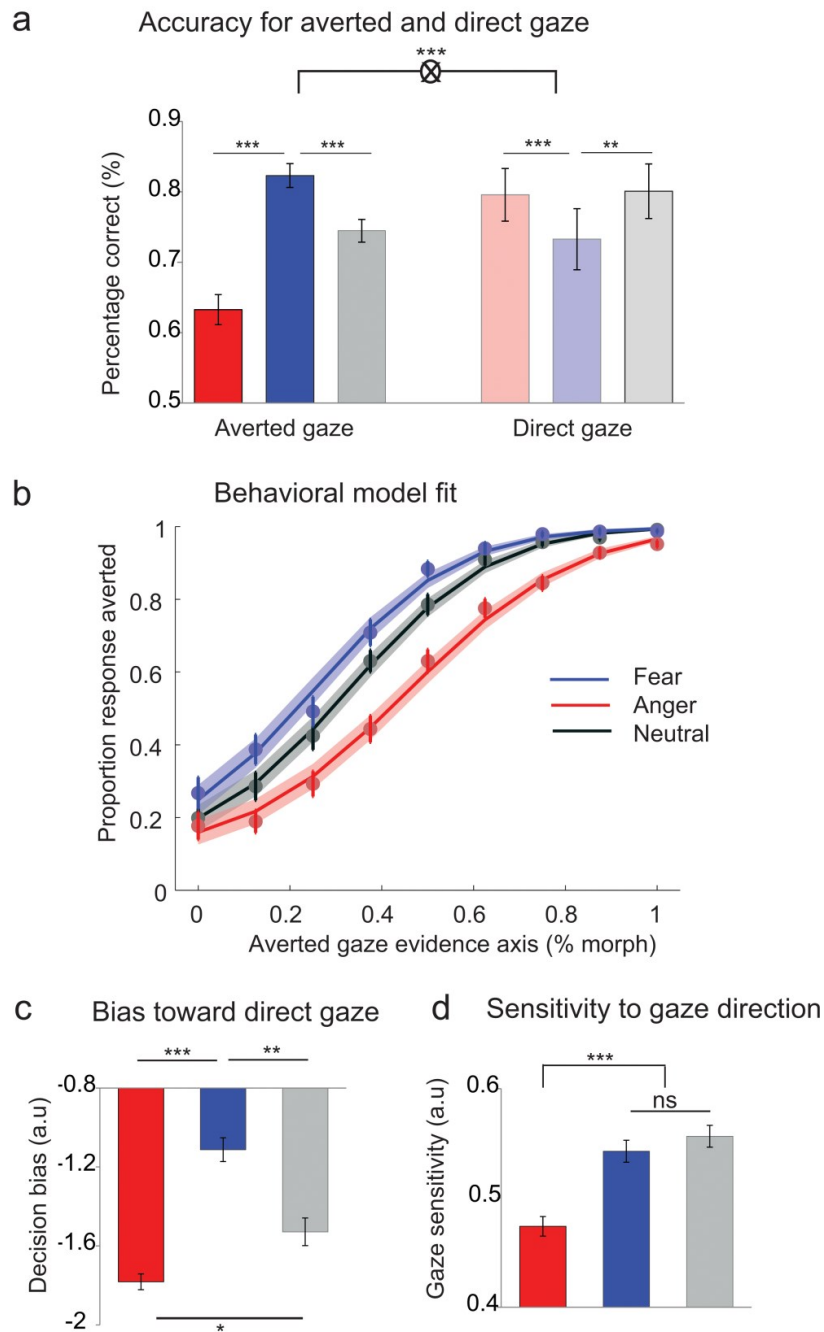


Figure 29. Results of Experiment 5. a) Accuracy for all 6 conditions, from left to right, averted gaze anger (red), fear (blue) and neutral (grey), direct gaze anger fear and neutral. b) Psychometric function representing the proportion of ‘averted’ responses as a function of the evidence for averted gaze, for anger (red), fear (blue) and neutral (black). Dots and attached error bars indicate the human data (mean  $\pm$  s.e.m.). Lines and shaded error bars indicate the predictions of the model. c) Parameter estimate of the decision bias in anger, fear and neutral conditions. d) Parameter estimate of the slope of the psychometric curve (corresponding to gaze sensitivity) for anger, fear and neutral. ns: non-significant, \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

Results on performance show the same pattern as in Experiment 2, with increased accuracy for combinations of gaze and emotion signaling higher threat (THREAT +, direct gaze with anger, averted gaze with fear), whereas accuracy for neutral seems to be at an intermediate level between both. This replicates reciprocity effects observed in the literature on morphed stimuli of gaze direction (Lobmaier et al., 2008; Adams and Franklin, 2009; Ewbank et al., 2009). Next we wanted to check whether, just like gaze direction increased the sensitivity to combinations of emotion and gaze signaling higher threat in Experiments 2 and 3, emotion increases the sensitivity to gaze direction when displaying these combinations.

### *Model selection*

We started with the simplest model (model 0) that could account for subject's decisions using a psychometric model such as:

$$P(\text{Averted}) = \Phi(w \cdot \mathbf{x} + b)$$

where  $P(\text{Averted})$  corresponds to the probability of judging the gaze as averted,  $\Phi$  to the cumulative normal function,  $w$  to the perceptual sensitivity to the display gaze,  $\mathbf{x}$  to the evidence in favor of averted gaze in each trial (from 0 for a direct gaze to +8 for the most averted gaze), and  $b$  to an additive, stimulus-independent bias toward one of the two responses/gaze directions. As effects of bias were already visible on the model-free responses of participants (Figure 29a), we chose to consider a model that consisted of 6 free parameters: 3 decision biases and 3 gaze direction sensitivity for the 3 emotions (Anger, Fear, Neutral). This allowed checking whether, in addition to the clear differences on the decision biases between the three emotions, emotion also increased gaze direction sensitivity to combinations signaling higher threat.

The parameters extracted first confirmed that there was a difference in decision biases for the 3 emotions, with differences between anger and fear ( $T_{21}=-8.2$ ,  $p<0.001$ ), anger and neutral ( $T_{21}=-2.5$ ,  $p=0.01$ ), and fear and neutral ( $T_{21}=-3.2$ ,  $p<0.01$ ) (Figure 29c). Importantly, differences were observed on sensitivity parameters between anger and the two other emotions: neutral ( $T_{21}=-4.2$ ,  $p<0.001$ ) and fear



( $T_{21}=-3.7$ ,  $p=0.001$ ). There was however no difference between the sensitivity to fear and neutral expressions (( $T_{21}=-0.4$ ,  $p>0.6$ ) (Figure 29d).

## Conclusion

By morphing gaze direction of anger, fearful and neutral emotional expressions during a gaze categorization task, this study shows that emotion improves the accuracy to observer-relevant threatening displays. While this improvement was partly driven by changes in the decision bias, it was also instantiated by changes in sensitivity to gaze direction. Here we considered the degree to which the gaze is *averted* as the perceptual evidence to the gaze direction task, and show that this sensitivity is increased in fear as compared to anger conditions. This result indirectly implies that the sensitivity to the degree to which the gaze is *direct* is increased for anger as compared to fear conditions. These findings thus demonstrate that gaze direction categorization is reciprocally influenced by emotions of anger and fear.

Contrarily to our expectation that neutral faces should lie somewhere in the middle between anger and fear emotions, the sensitivity to averted gaze direction in the neutral condition did not differ from fearful faces, but was increased as compared to angry faces. This is surprising in regards to the proposed advantage in the processing of direct gaze in neutral faces (Farroni et al., 2002; Senju and Hasegawa, 2005; Conty et al., 2007; Mareschal et al., 2013), which would rather predict an increased sensitivity to direct gaze in neutral faces (like anger). However, here gaze categorization was performed in a context where neutral faces were presented among threatening emotional faces, which possibly changed the mechanisms underlying the processing of gaze direction. In a threat context induced by the presence of anger and fear emotions, gaze direction in neutral faces seems to be processed similarly as gaze direction in fearful faces. A possible interpretation of the similarity in behavior to fearful and neutral faces is that both can signal an uncertain event to the observer when averted, potentially threat. However when displaying a direct gaze and in contrast to anger, neutral and fearful emotions do not signal a direct threat to the observer. Consistent with this suggestion, increased activity in monkey amygdala to averted as compared to direct gaze in neutral and emotional expressions is relat-

ed to greater arousal and attention when facing stimuli that signal uncertainty (Hoffman et al., 2007).

The fact that strong effects were observed on decision bias whereas it was not the case in the emotion categorization task partly stems from the experimental design. Here the extent to which gaze is averted is manipulated resulting in unequal conditions of direct and averted gaze conditions whereas there was an equal number of angry and fearful expressions in Experiment 2. Therefore, it was more likely that effects on the bias emerge due to the asymmetry which undoubtedly biases toward averted responses.

To conclude, these findings build on previous studies (Lobmaier et al., 2008; Adams and Franklin, 2009; Ewbank et al., 2009) that suggested mutual influences between emotion and gaze by showing that these effects are driven by both changes in sensitivity and decision bias favoring observer-relevant signals in the context of threat.

### **3.5- Question 4: Generalization of the effect to a representative pool of the population (Experiment 6)**

The emotion categorization task used in Experiment 2 (and tested on young healthy individuals) was conducted on a sample of population that is representative of the French population. One purpose of this study was to see whether we replicate the increased behavioral sensitivity to threat-signaling emotions in the general population. More importantly, this study was conducted among a set of other social tasks including economic games and political judgments and included questionnaires evaluating anxiety and autism scores as well as life history and backgrounds. The main goal was thus to link between individual traits and emotion categorization in the context of threat, as well as investigate whether similar profiles of people behave in selective manners in different social tasks. Here I will describe only the replication of Experiment 2 without providing results linking the different manipulations as these

analyses are ongoing at the moment. This work is conducted in collaboration with Coralie Chevallier, Nicolas Baumard, Lou Safra, Yann Algan and IPSOS.

## **Methods**

### *Participants*

Using the database from the polling organization IPSOS (<http://www.ipsos.fr>), data were collected from individuals that were representative of the French population in terms of age repartition (16 to 80 years old), socio-professional categories, and level of education. They all provided consent to participate in the study and were remunerated for their participation. 1005 individuals completed the entire set of experiments (512 females, mean age  $47.4 \pm 0.5$  years)

### *Procedure*

The emotion categorization task (Anger or Fear) was part of a set of experiments realized in these individuals. They actually performed two sessions of 22 minutes each. The first session included an anxiety questionnaire (STAI), the emotion categorization task, a task that consisted in choosing which person the participant would vote for (while varying the traits of face stimuli from trustworthy to dominant), and a questionnaire about life history of participants. The second session included 3 economic games (dictator game, trust game, public good games) and 3 questionnaires (social trust, broad autism phenotype BAPQ and social relations). Participants only performed one session of the emotion categorization task that lasted around 6-7 minutes (120 trials). Similarly as in Experiment 2, stimuli were morphed expressions from neutral to angry and from neutral to fearful expressions while gaze direction was either direct or averted. The face appeared for 250 ms and participants had to choose whether the face expressed anger or fear by pressing one of the two keys 's' or 'l' of the keyboard within 2 seconds. The position of the answers 'anger' or 'fear' were counterbalanced across subjects.

## Results

Analyses were performed on all subjects that performed the task with percentages of correct responses larger than 60 %. After exclusion, the final pool of subjects consisted of 602 individuals (325 females, mean age  $44 \pm 0.5$  years). The high rejection rate (~40%) is probably due to the nature of the task performed in this online procedure. While tasks in online procedures (including all the other tasks of the present study) usually do not require 'correct' responses (such as personality questionnaires, surveys, and preference choices), here it was crucial that participants correctly perform the task. Unfortunately, most probably because participants performing the task outside the lab environment (such as at home) might be easily distracted, a high rejection rate is not surprising and due to this quite 'difficult' task with trials that rapidly succeed one another requiring full concentration.

### *Performance*

The interaction between emotion and gaze was replicated as previously ( $F_{1,601} = 12.8$ ,  $p < 0.001$ ), with increased percentage of correct responses to direct gaze as compared to averted in anger condition ( $t_{22} = 3.8$ ,  $p < 0.001$ ) (Figure 30a). The increased accuracy for averted as compared to direct gaze in fear condition did not reach significance ( $t_{601} = 1.5$ ,  $p = 0.1$ ). Surprisingly, there was a very strong main effect of emotion, with generally increased accuracy in fear condition as compared to anger ( $F_{1,601} = 34.8$ ,  $p < 0.001$ ). As the only difference with previous studies (Experiment 2 and 3) was the few number of trials used per subject, this result suggests that participants might have an initial bias toward fear increasing their accuracy in this condition. This suggestion will be further elaborated below, in the model selection analyses.

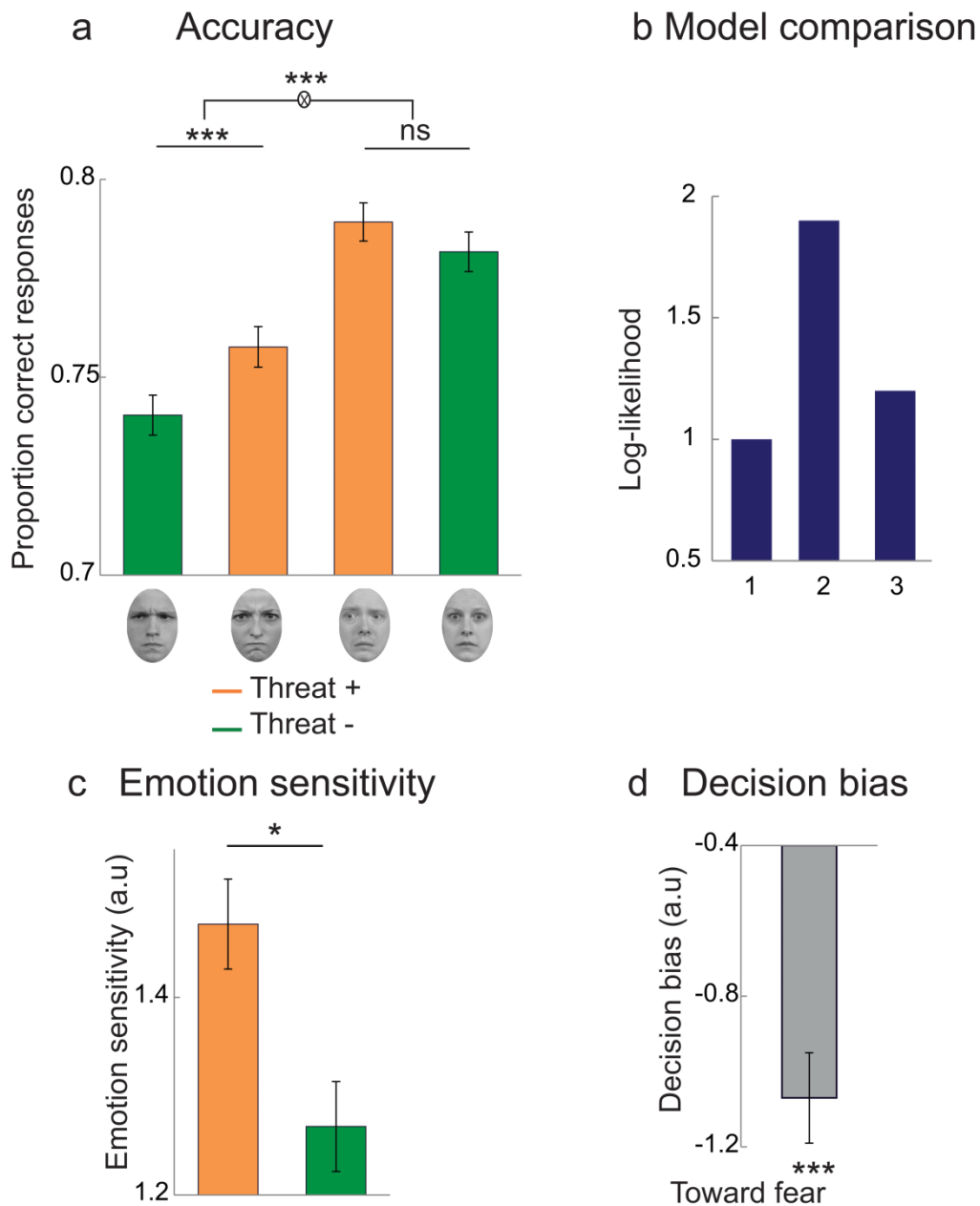


Figure 30. Results from Experiment 6. a) Accuracy in the 4 conditions, from left to right, anger averted, anger direct, fear averted and fear direct. b) Model comparison between model 1 (effect on decision bias), model 2 (effect on sensitivity) and model 3 (effect on both sensitivity and bias). c) Emotion sensitivity parameters for THREAT + and THREAT - conditions extracted from model 2. d) Decision bias parameter extracted from model 2.

### Model selection

Because of the very little amount of trials per subject, fixed-effect comparisons were used for model selection in this experiment. Similarly to Experiments 2 and 3 we compared the 'null' model which did not allow effects of gaze direction on the deci-

sion process, with three additional models that could explain the increase in recognition accuracy for THREAT+ combinations of gaze and emotion (anger with a direct gaze, fear with an averted gaze). Model 1 would be that gaze direction biases emotion recognition in favor of the interpretation signaling higher threat (anger for a direct gaze, fear for an averted gaze). Model 2 would be that gaze direction might selectively increase sensitivity to emotions signaling higher threat. Model 3 would be that both changes on the decision bias and the sensitivity explain the increased recognition of THREAT+ combinations.

When subtracting the log-likelihood of the null model to all three models, the best model was model 2 that assigned different sensitivities to THREAT+ and THREAT- (Figure 30b). However, the Bayes Factor of this difference was close to  $10^{0.38}$ , so model 2 was only approximately 2.4 times better than the other models. Nevertheless, comparison between the extracted parameters of model 2 showed that emotion sensitivity was increased for THREAT+ conditions ( $t_{601} = 2.1$ ,  $p < 0.05$ ) (Figure 30c). On the contrary, no difference between the decision bias extracted from model 1 for direct and averted gaze was evident ( $p > 0.8$ ). Besides, the decision bias extracted from model 2 was highly pulled toward fear responses ( $t_{601} = -8.5$ ,  $p < 0.001$ ) (Figure 30d). This extremely large bias can justify the increased accuracy of fearful expressions that seem to have covered the effect of gaze on the percentage of correct responses in the fear condition. Just like stated above, this bias stems from the fact that participants only performed 120 trials (in opposition to 1080 trials in Experiments 2 and 3). Further corroborating this interpretation, when analyses were performed in Experiments 2 and 3 on only the first block (120 trials), similar effects emerged on the response bias (on the total of 64 participants bias toward fear ( $t_{63} = -2.1$ ,  $p = 0.05$ )). Furthermore, when the data of 46 additional participants (25 females, mean age  $22.1 \pm 0.4$ ) that performed the same task (in a different study not reported here) is also included, the bias toward fear becomes highly significant on the first 120 trials ( $t_{109} = -3$ ,  $p = 0.003$ ), along with increased performance to fearful as compared to angry stimuli ( $t_{109} = -2.1$ ,  $p = 0.03$ ).

## Conclusion

Running the emotion categorization experiment on a large number of participants showed that the effects of gaze on recognition of threatening emotions is robust as it survived outside of the lab environment with a shortened version of the task (1/9). It importantly extends the results to the general population that varies in age (16 to 80 years old), socio-professional categories, and level of education. Certainly, model comparisons only revealed a weak difference between the model allowing an effect of gaze direction on the sensitivity and the model considering an effect on the decision bias. Nevertheless, the interaction between emotion and gaze on the performance, as well as the significant effect of gaze direction on the sensitivity but not on the decision bias, point in the same direction to enable the suggestion that gaze direction enhances the behavioral sensitivity to threat-signaling emotions.

Notably, the shortened version of the task had a drawback: an unexpected bias toward fearful responses emerged suggesting that participants have an initial bias toward fear when categorizing between angry and fearful expressions. A possibility that could explain this initial bias toward fear might be that fear has an evolutionary advantage (Ohman and Mineka, 2001; Ohman, 2005) over anger as it might signal danger of a higher level of threat as compared anger. However, this assumption is speculative, and further research is needed to understand the reason behind this initial bias. In this study, correlations between personal traits such as anxiety and the harshness of life history with the degree to which participants show this bias toward fear will be conducted as they might provide insights into the reasons that explain this effect.

More generally, further analyses of this experiment will be performed, to link the results of the emotion categorization task to the outcomes of personality questionnaires and the other social tasks conducted in this survey. As an illustration of a hypothesis that we have on these analyses, one might expect that increased sensitivity to threat-signaling emotions might be even stronger in individuals who grew up in harsh environments.

### **Conclusion part 3**

The four behavioral studies reported in this part allowed to answer a few questions raised by the results of Experiment 2, i.e. the increased sensitivity to threat-signaling emotions modulated by gaze direction during an anger-fear categorization task. First, this effect is driven by high spatial frequencies of the facial expression (Experiment 3). Second, the effect of gaze direction on emotion categorization is specific to the categorization of negative emotions (anger and fear) and does not apply to the categorization of negative (anger) against positive (happiness) emotions (Experiment 4). Third, the influence of gaze direction on emotion categorization is reciprocal in the sense that emotion similarly increases the sensitivity to gaze direction in conditions signaling higher threat to the observer (Experiment 5). Fourth and finally, gaze direction modulation of the perception of threatening emotions is proven to be robust as it was replicated on a large sample of the general population (Experiment 6). The findings of Part 3 support the increased perceptual and neural sensitivity to threat-signaling emotions and show that this sensitivity enhancement is specific, replicable, robust and independent of task demands.

Together, Experiments 1 to 6 reveal that in threat environment, gaze direction enacts as a contextual factor that increases the sensitivity to specific combinations of gaze and emotion behaviorally and neurally within 200 ms after stimulus onset in sensory and motor regions. This suggests that when threat signals are involved, evolutionary pressure shaped the brain to better perceive social displays indicating high levels of threat.

The next study manipulated perceptual experience of participants with emotion/identity associations while they did a similar fear/anger categorization task on morphed emotional expressions. Would this contextual manipulation, now not related to threat, but rather to day-to-day personal experience, similarly influence emotion categorization by changing the sensitivity to emotional information? Or alternatively, would it only change the decision bias as from an evolutionary point of view, the stakes are less high as compared to recognizing threat signals? One can speculate that while signals related to survival require a better perceptual recognition for fast appropriate reactions, associations related to daily interactions rather need a flexible



adaptation as a function of social contexts thereby only biasing perceptual choices but not changing their core encoding.

## Part 4: Expectation about the emotions of individual identities as a contextual factor (Experiment 7)

This part will be presented as a draft of a manuscript that is in preparation:

***A general mechanism for emotion categorization biases shaped by associative identity-specific expectations.***

# A general mechanism for emotion categorization biases shaped by associative identity-specific expectations

Marwa El Zein <sup>1,2,\*</sup>, Valentin Wyart <sup>1,‡</sup>, and Julie Grèzes <sup>1,‡</sup>

<sup>1</sup> Laboratoire de Neurosciences Cognitives, Inserm unit 960, Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL Research University, Paris, France

<sup>2</sup> Université Pierre et Marie Curie – Paris 6, Paris, France

\* to whom correspondence should be addressed ([marwa.el.zein@ens.fr](mailto:marwa.el.zein@ens.fr))

‡ shared senior authorship

## Abstract

Expectations in perceptual decision-making influence neural activity and behavior by facilitating the processing and recognition of expected stimuli. During social interactions, most of these expectations are based on past experience. In particular, expectations about others' characters and moods build upon previous encounters, but the mechanisms underlying their acquisition and expression in human behavior remain unclear. Here we show that experimentally *induced* associations between emotion and identity create expectations which influence emotion categorization: identities which were initially associated with a particular emotion were later more likely to be categorized as expressing that emotion. Interestingly, while this effect was apparent in categorization judgments but not reported explicitly by participants in a subsequent recall test, *incidental* expectations emerged simultaneously based on idiosyncratic (false) associations. Computational modeling demonstrated that both *induced* and *incidental* expectations shift participants' decision criterion about facial expressions of emotion. These effects were reflected in peri-decision pupillary responses which increased when participants responded against their biases about consolidated emotion-identity associations. Pupil dilation showed similar increases when responses were made against participants' *intrinsic* biases – i.e., their tendency to respond anger or fear independently of identity. Pre-decision neural activity was modulated accordingly, in the form of increased centro-parietal EEG signals when subsequent responses were made against *intrinsic* biases. Together, these findings indicate that past experience – in the form of *induced* and *incidental* emotion-identity associations as well as *intrinsic* biases, shape the perception of others' emotions. Pupil-linked neuromodulatory systems express these multiple expectations through increased responses to effortful decisions against existing associations and biases.

## Introduction

Our daily life is punctuated by the many perceptual decisions that we make on various sources of sensory information. These decisions are rarely made without biases which stem from past experience and depend on the context in which they occur. Behavioral and neural evidence has shown that prior information biases decision-making by facilitating choices congruent with one's expectations (Summerfield and de Lange, 2014).

In social contexts, faces represent the most relevant stimulus category, and the proper identification of facial features (as a function of identity or expression) is thus crucial for rich social interactions. This facial identification process is known to be influenced by several contextual factors such as biases in social judgments which depend not only on perceptual determinants (e.g., the dominance or trustworthiness of a face), but also on past experience and incidental associations (Todorov et al., 2015). As an illustration of perceptual-related influences, masculine and angry faces are perceived as more dominant (Todorov et al., 2015). Social judgements can also change independently from perceptual features as a function of initial encounters with others, experimentally induced with descriptions specific to each presented identity (Todorov et al., 2007; Schiller et al., 2009). The impact of associating specific behaviors to identities on the neural encoding of faces can be both dependent and independent of explicit memory, i.e. whether the participant recalls the link between one identity and its associated behavior (Todorov et al., 2007). Moreover, these attributions vary as a function of each person even if the same information is given about an identity, as different opinions can be formed on the same provided behaviors (Schiller et al., 2009).

These studies suggest that face encoding is influenced by personal experience, whether observers are aware or unaware of their biases, and vary considerably from an individual to another. However, these previous experiments only explicitly manipulated expectations by providing verbal descriptions about identities on neutral faces. Furthermore, they were directly interested in the neural signatures of these expectations, i.e. how identities associated with a certain attitude are encoded in the brain, but did not investigate whether they modify participants' behavior. Nevertheless, evidence from perceptual tasks about faces shows that previously encountered

identities spontaneously modify judgments about facial identities (Apps and Tsakiris, 2013; Liberman et al., 2014). They do so by changing the processing of previously encountered identities in face selective structures, the fusiform gyrus and the superior temporal sulcus (Apps and Tsakiris, 2013), possibly fostering visual stability through decreased neural computations to these identities (Liberman et al., 2014). While these studies highlight the fact that face perception is influenced by preceding context, they do not manipulate changeable features of facial identities such as emotional expressions.

Here we were interested in whether implicit associations between an emotion and an identity prompt expectations that influence subsequent emotion categorization and the cognitive mechanisms by which they do so. We implicitly manipulated the emotional characteristics of identities by presenting half of them expressing one emotion or another (anger or fear) at the beginning of experimental blocks and tested whether it influenced subsequent emotion categorization. At the end of the experiment, to assess how much participants were aware of the manipulation and whether their own incidental associations biased their decisions throughout the task, we asked them to judge the emotion of all the neutral identities they encountered.

Importantly, a parametric control over emotion strength, provided by the use of 'morphed' facial expressions, afforded fitting decision theoretical models to arbitrate between competing accounts of how expectations about facial emotions influence their categorization. Classical decision theories inspired by signal detection theory indeed distinguish between two types of mechanisms by which prior information could impact emotion categorization (Green and Swets, 1966; Macmillan and Creelman, 2004). Expectation about the emotion of an identity could bias its categorization toward the expected association, thereby additively combining identity and emotion information in accordance with 'Bayesian' accounts of decision-making (Griffiths and Tenenbaum, 2006; Oaksford and Chater, 2007). Alternatively, expectation about the emotion of an identity could increase the perceptual sensitivity to the facial features diagnostic of the emotion congruent with one's expectations, thereby enhancing the quality of processing of the sensory attributes relevant to the emotion categorization task (El Zein et al., under review).

Finally, we also recorded pupillary responses that have recently been related to individual biases both during social categorization of faces from similar or dissimilar race (Wu et al., 2012), and nonsocial perceptual discrimination of visual contrast embedded in noise tasks (Gee et al., 2014). Along these lines, we expected to find modulations of pupillary responses by individual biases, notably through increased responses to decisions against participants' inclinations (Gee et al., 2014).

## Results

Each block started with a short 'induction' period where identities were split randomly into two halves, presented in a pseudo-random order, each half expressing consistently either anger or fear, in order to induce emotion/identity associations (figure 1a). This manipulation assigned two groups of identities randomized across participants: identities that belonged to the Anger Group, and identities that belonged to the Fear Group. In the next, 'test' period, all identities expressed an equal number of angry and fearful expressions, and all analyses were performed on these trials only. In both periods, participants had to categorize the facial expression that appeared on the screen as expressing either anger or fear (figure 1b), with no explicit warning about the transition between the two induction and test periods. As gaze direction of anger and fear expressions is known to facilitate the categorization of combinations signaling higher threat to the observer (El Zein et al., under review), we further manipulated gaze direction in an implicit, blocked design to assess whether it interacts with the effects of emotion/identity associations on emotion categorization.

### **Induced manipulation of emotion/identity associations biases upcoming decisions**

The implicit manipulation of identity group successfully created expectations that influenced emotion categorization, reflected in an enhanced accuracy for angry stimuli belonging to the Anger group and fearful stimuli belonging to the Fear group (emotion by induced identity group interaction  $F_{1,30} = 7.71$ ,  $P < 0.01$ ) as compared to the other conditions (figure 2a). A main effect of emotion strength ( $F_{1,30} = 166.6$ ,  $P < 0.001$ ) confirmed that accuracy decreased with emotion ambiguity. Moreover, we replicated our previous study in regards to the effect of gaze, as accuracy was increased for emotions signaling higher threat - anger with a direct gaze and fear with an averted gaze - as compared to the 2 other conditions (emotion by gaze interaction  $F_{1,30} = 8.5$ ,  $P < 0.01$ ). No other main effect or interaction was significant (all  $F < 2$ ; all  $P > 0.1$ ) suggesting that gaze did not interact with induced emotion-identity associations.



To better characterize the effect of identity group on emotion categorization, we compared between different decision-making models (inspired by signal detection theory) that instantiate the potential mechanisms by which induced identity group can influence emotion recognition. We reasoned that induced associations between identity and emotions could either bias the responses toward congruent emotions (model 1), or increase the sensitivity of congruent emotions (model 2), or both (model 3). Using Bayesian model selection combined with a cross-validation procedure to estimate the log-evidence of the different models, we compared these models with the 'null' model (model 0) that allowed no effect of induced identity group on emotion recognition, and found that only model 1 fitted better the data than model 0 (model 1: exceedance probability  $p_{exc} = 0.90$ ; model 2:  $p_{exc} = 0.41$ ; model 3:  $p_{exc} = 0.52$ ). A direct comparison between the tested models revealed that model 1 was the most likely model (fixed-effects Bayes factor against second best model  $\approx 10^4$ , random-effects all  $p_{exc} > 0.83$ ). Associations between emotion and identity thus influenced the decision bias, by shifting the response toward angry responses when stimuli belonged to an angry group and inversely for fear (model 1, difference between predicted bias for group Anger and bias for group Fear  $T_{30} = 2.4$ ,  $P < 0.05$ ) (figure 2c).

The post-test performed by participants at the end of the experiment which consisted in deciding whether the identities expressed more anger or fear during the experiment (figure 1c) allowed assessing whether the observed bias stems from subjects' knowledge of the manipulation. No difference between the mean rating of responses for group Anger and group Fear was observed ( $T_{30} = 0.56$ ,  $P = 0.57$ ), and the receiver operating curve (ROC) calculated from rating responses was not significantly different from chance ( $= 0.504 \pm 0.003$ , t-test against chance level 0.5:  $T_{30} < 0.3$ ,  $p > 0.8$ ), suggesting that participants were at chance level in identifying the associations experimentally induced.

### **Bias toward incidental, explicitly reported emotion/identity associations**

We next wondered whether the identities rated explicitly as belonging to fear or anger groups in the subsequent post-test could have influenced individual decisions throughout the experiment. In other words, do people spontaneously associate identities with certain emotions? We therefore considered a new split between identity

groups, not the one induced by the experimenter (referred to as the ‘induced’ identity group), but the one formed through random associations of subjects and expressed in the post-test: the ‘incidental’ identity group. For each subject, the incidental anger and fear groups consisted of the identities they rated as expressing anger or fear (ratings transformed to binary decisions). Remarkably, a very strong emotion by incidental identity group interaction emerged ( $F_{1,30} = 58.2, P < 0.001$ ), with increased accuracy throughout the experiment for angry stimuli that were considered as angry in the post-test and fearful stimuli that were considered as fearful in the post-test (figure 2b). Model comparisons similar to the one described above showed that models 1, 2 and 3 fitted better the data than model 0 (model 1: exceedance probability  $p_{exc} = 1$ ; model 2:  $p_{exc} = 0.99$ ; model 3:  $p_{exc} = 0.99$ ). A direct comparison between the tested models revealed that model 1 was the most likely model (fixed-effects Bayes factor against second best model  $\approx 10^{11}$ , random-effects all  $p_{exc} > 0.93$ ). Response bias was pulled toward anger for the angry group and pulled toward fear for the fear group (difference between biases  $T_{30} = 9.2, P < 0.001$ ) (figure 2e).

We therefore decided to integrate both bias effects in one model that would allow an induced and an incidental effect on the response bias (figure 2d). Importantly, even though the incidental bias was very strong ( $T_{30} = 8.2, P < 0.001$ ), the induced bias effect remained significant ( $T_{30} = 2.1, P < 0.05$ ).

### **Incidental biases depend on both stimulus-related and observer-related biases**

The previously described findings raise the following question: is the incidental effect driven exogenously by the characteristics of the stimulus (Todorov et al., 2015) or endogenously by the participants? The between-participant consistency in the post-test rating of the different identities offers an unbiased method to address this question: if all subjects unanimously categorized each identity, then it would imply that the incidental effect is not linked to an internal bias, but to the facial characteristics of the identities (in terms of dominance, trustworthiness, etc.). To address that question, we checked whether identities were assigned to fear and anger randomly or unanimously by performing t-tests against zero on the ratings for each identity. Less than half of the identities ( $n = 14$ ) were significantly most often assigned to the same emotion (ratings significant against zero across subjects  $P < 0.05$ , 7 toward fear and 7 toward

anger, 8 males, 6 females) whereas 18 identities were not significantly assigned to anger and fear. Importantly, the incidental effect on response bias was as strong for consistently and inconsistently assigned identities as there was no difference between the bias for group Anger and group Fear for identities unanimously assigned ( $T_{30} = 5.4, P < 0.001$ ) and those randomly assigned ( $T_{30} = 4, P < 0.001$ ) (difference between the 2 effects  $T_{30} = 1.3, P > 0.18$ ).

People thus appear to create incidental associations when encountering new identities based on a combination of the perceptual features of the identity and their individual biases which likely arise from their own experience. These last results however raised a new question: what explains the fact that 14 specific identities were rated more consistently than the 18 other identities judged randomly/inconsistently? One hypothesis is that specific traits of these identities were more pronounced in favor to a specific judgment of fear or anger (Todorov et al., 2015). To test this assumption, 20 new subjects performed trait ratings in terms of threat, trustworthiness, and dominance on the neutral identities (see methods). First, as expected (Todorov et al., 2015), threat and dominance ratings were highly correlated ( $r = 0.85, p < 0.001$ ), and both threat ( $r = -0.79, p < 0.001$ ) and dominance ratings ( $r = -0.5, p < 0.01$ ) inversely correlated with trust. Second, the consistency across subjects in emotion ratings expressed as t-values of the performed t-tests for each identity were correlated with the mean subjective trait ratings of each identity: threat ratings ( $r = -0.57, P < 0.001$ ), trust ratings ( $r = 0.42, P = 0.01$ ), and dominance ratings ( $r = -0.5, P < 0.01$ ) significantly correlated with the consistency across subjects in emotion ratings, suggesting that identities that were highly rated in terms of social traits were those who led to unanimous categorizations of fear and anger (figure 2f, example for threat rating). Threat ( $r = -0.8, P < 0.001$ ), dominance ( $r = -0.8, P < 0.001$ ) and trustworthiness ( $r = 0.6, P < 0.01$ ) ratings correspondingly correlated with the mean post-test rating scores for the consistent group only (same correlations for the inconsistent group all  $r < 0.2$ , all  $p > 0.3$  and difference between the coefficient correlations using fisher r-to-z transformation, all  $p < 0.05$ ). As responses from the anger side of the scales were coded as negative and those from the fear side as positive, these correlations show that the more dominant and threatening the identities are perceived, the more they are judged as angry (figure 2 g,h), while the more the identities are rated as trustworthy the more they are rated as fearful.

Altogether, these findings show that spontaneous associations of identities with an emotion observed on decision bias stem from both stimulus-related social traits provided by the facial features and observer-related biases probably relying on individual preferences generated by past experience.

### **Pupil dilation tracks effort and individual biases**

Based on previous findings (Gee et al., 2014) showing that pupil-linked neuromodulatory systems (notably the locus coeruleus norepinephrine system) respond more strongly when participants respond against their bias, we checked whether we would find signatures of identity biases in pupillary responses. Regressions of the pupil dilation signal in response to presented identities against the main effect of choices toward vs. against biases created by emotion-identity associations, were performed. They revealed that participants dilate their pupil more strongly when they respond against their bias starting 1.5 s after stimulus onset until 2.4 s ( $T_{30} = 2.6$ ,  $p = 0.01$ ) – when the comparison concerned identities that congruently belonged to the incidental identity group and the induced identity group (figure 3a,b) (difference between congruent identities and incongruent identities  $T_{30} = -3.1$ ,  $P < 0.01$ ). Although the effect was only significant when combining induced and incidental emotion/identity associations, and thus consolidated associations, these results show that pupillary responses increase when responses are made against biases, probably requiring a stronger effort in the decision process.

### **Intrinsic biases modulate pupil and EEG responses to Anger and Fear emotions**

As results on pupil dilation showed a modulation dependent on biases of participants related to identity/emotion associations to some extent, we next wondered whether intrinsic individual biases – participants' idiosyncratic tendency to perceive facial expressions as angry or fearful, produce similar modulations of pupillary responses independently of identity. Peri-decision pupillary responses indeed tracked intrinsic individual biases toward anger or fear, as pupil diameter was increased when participants responded anger as compared to fear when their bias was toward fear ( $T_{11}$

=2.4,  $P<0.05$ ) from 1.6 to 2 sec after stimulus onset, and in opposition, the pupil dilated more strongly to fear response as compared to anger response when participants' bias was toward anger ( $T_{18}=-2.6$ ,  $P<0.05$ ) from 1.5 to 2.2 sec after stimulus onset (figure 3d). The difference between the two sets of subjects became significant starting 1.2 s until 2.4 s after stimulus onset ( $T_{29}=-3.6$ ,  $P=0.001$ ) (figure 3c). Moreover, at this timing, individual biases to fear and anger correlated with the main effect of anger/fear responses on pupillary responses ( $r=-0.58$ ,  $P<0.001$ ) suggesting that this physiological effect is highly inter-individual dependent (figure 3f).

To further interpret these results, and in particular to determine whether these bias-dependent modulations of physiological responses were present during the decision process, before the response, we checked whether these same individual biases toward anger or fear responses modulated pre-decision electrical brain activity (EEG) that was recorded in a previous experiment including 24 healthy individuals who did exactly the same experimental task - categorization of anger or fear (figure 1b) - without induced emotion-identity associations (El Zein et al., under review). Participants' behavior was explained with a simple model that assigned two parameters: the sensitivity to the emotional expressions and the individual biases. As we were interested in how individual biases modify brain activity during categorization, we isolated participants that showed a bias toward anger (12 participants) and those who showed a bias toward fear (12 participants) (figure 4a). We then searched for neural signature of such biases, and specifically looked at 'P3' event-related responses because, like pupillary responses, they have been suggested to be linked to the locus coeruleus-norepinephrine system (Nieuwenhuis et al., 2005). We had previously identified two moments where positive parietal responses peaked in encoding emotion information (El Zein et al, in revision): at 500 ms after stimulus onset, and 100 ms before response time. We therefore isolated activity in parietal electrodes at these time points and examined whether they vary with individual biases. Similarly to pupillary responses, the main effect of emotion on EEG responses (anger/fear) correlated with individual biases across the 24 participants both at 100 ms before response ( $r=-0.5$ ,  $p<0.01$ ) (figure 4c), and at 500 ms after stimulus onset ( $r=-0.46$ ,  $p<0.05$ ), showing that the more individuals are biased toward anger, the more centro-parietal EEG activity is increased preceding fear vs anger responses (and the opposite for individuals biased toward fear). Splitting the curves representing the

time-course of the EEG response main effect into participants biased toward anger and those biased toward fear shows a dissociation between the two groups from 200 ms before response until response time ( $T_{29}=3.1$ ,  $P<0.01$ ) (figure 4c). This is consistent with the fact that responding against one's bias is more costly, thereby increasing neural processing.

Together, pupillometry and EEG results converge to indicate that individual biases modulate brain responses by increasing responses to effortful decisions against biases. As both parietal responses and pupillary responses have been related to the LC-NE neuromodulatory system, these findings suggest a possible involvement of this system in modulating brain responses as a function of participants' biases – induced, incidental and intrinsic.

## Discussion

Social attributions are based on a variety of biases that come from individual experiences and contexts. Previous research had shown that social attributions not only depend on perceptual features within faces of others (Zebrowitz and Montepare, 2005, 2008; Dotsch and Todorov, 2012), but also on individual experience with others that can be manipulated experimentally (Todorov et al., 2007; Schiller et al., 2009). Here we extend these findings to the social categorization of facial expressions by showing that both induced and incidental associations of identities with specific emotions biased participants' decisions toward bias-congruent interpretations. These biases as well as identity-independent tendencies to judge faces as expressing a certain emotion were reflected in pupillary responses that were increased to effortful decisions against individual biases. Further, EEG recordings in another identical emotion categorization task supported an increased neural engagement to responses against identity-independent intrinsic biases.

The implicit manipulation of identity association with a certain emotion in only a few trials at the beginning of experimental blocks successfully biased participants' decisions by shifting their choices toward congruent emotions thereby increasing recognition accuracy in these conditions. Noticeably, these associations between identities and emotions are possible among a large number of different identities (32) in a very short period of time, thus suggesting the presence of a rather large 'reservoir' in the brain for the learning of such associations. The existence of such a 'reservoir', implying that the brain has a strong memory dedicated to these associations, is most probably due to their important social significance in daily interactions. As facial expressions of emotions encountered on a day-to-day basis are most often ambiguous, prior expectations, here related to the emotion of individual identities, offer a way to disambiguate their perceptual judgement. Moreover, participants failed to explicitly notice the manipulation, consistently with the fact that social attributions can be constructed through experience implicitly and rapidly (Todorov et al., 2015).

Interestingly, incidental associations between identities and emotions assessed in post-test ratings strongly biased decisions throughout the experiment. Although this result can appear surprising at first sight, it actually might reflect personal experience of each participant, and substantiates the variability of idiosyncratic judgments

among individuals (Schiller et al., 2009). It is very plausible that both the induced and incidental associations stem from the same source: personal experience with identities, i.e. in which expression they were previously encountered. In day-to-day life, the mood and social behaviors of people in each individual's surrounding can vary remarkably. If a particular friend is often in a bad mood, it becomes very likely that he/she will be judged as angry or upset on any given day. While in the present study we experimentally manipulated the encounter between identities and observers, incidental associations possibly emanate from resemblance between the identities presented and people from observers' surroundings. Indeed, individuals easily transfer characteristics of identities to others that are similarly looking (Kraus and Chen, 2010; Verosky and Todorov, 2010).

Although this interpretation applies to identities that were idiosyncratically judged across participants, the categorization of those identities who were unanimously assigned to anger or fear emotions may rather derive from their perceptual facial characteristics. Identities that were judged to be more dominant and threatening while displaying a neutral expression were also those who were consistently assigned to anger emotion, in accordance with the findings that angry faces are perceived as more dominant (Montepare and Dobish., 2003; Hess et al., 2007). Despite the fact that dominance and threat are probably only characteristics among others that drive the judgment of a face as angry, these associations provide a new example of how social attributions can be based on perceptual facial elements (Zebrowitz and Montepare, 2008).

Our findings demonstrate that the identity of facial expressions was processed even if not required for the emotion categorization task, supporting an automatic holistic processing of faces. Although an integrated processing of a face as a whole seems more intuitive in the context of social interactions, traditional views suggested independent processing of emotion and identity (Bruce and Young, 1986; Haxby et al., 2000). These views are however questioned (Calder and Young, 2005) as growing evidence support an inter-dependence of these two facial features that seem to reciprocally influence each other (e.g Kaufmann and Schweinberger, 2004; Fox and Barton, 2007; Chen et al., 2011; Van den Stock and de Gelder, 2014), suggesting that either of the two is processed even if not relevant to the ongoing task. Here we



corroborate the holistic processing of faces by showing that identity was implicitly processed when the only task-relevant information to participants was the emotion expressed by the face.

Pupillary responses were influenced by biases linked to identity/emotion associations. They were increased when participants categorized an identity as expressing the emotion that did not correspond to the biased association. However, this increased pupil dilation to responses against bias applied only to consolidated associations, when the effort to overcome the bias was maximal: when the identities were associated with an emotion both through induced experience and incidental associations. This result builds on the findings of increased pupil dilation when individuals responded against their preferred response during a perceptual detection task (Gee et al., 2014), by showing that responses against biases associated to the stimulus, not only to individual response preferences, similarly influence pupil dilation. Increased pupil response to identities that required responses against expectations can indicate increased arousal and effort that have been since long and repeatedly associated with greater pupillary responses (Hess and Polt, 1964; Hess, 1965; Kahneman et al., 1967; Porter et al., 2007).

Individual response biases toward an emotion or the other were also reflected in pupillary responses replicating results of the perceptual detection task mentioned above (Gee et al., 2014). Indeed, similarly to how increases in pupil dilation to 'yes' responses increased with the degree to which individuals had a conservative bias, pupil increase to anger against fear increased with participants' bias toward fear. This modulation extends the findings of the previous study that used a perceptual detection task (yes or no) to two alternative perceptual tasks. The observed effect seems to reflect a general mechanism to overcome individual biases, and should a priori apply to any other two-forced choice task, rather than being specific to emotion categorization. Furthermore, brain activity in parietal regions (P3-like components) showed a similar pattern during an identical emotion categorization task. Activity increased at 500 ms after stimulus onset and 100 ms before response when choices were against individual biases toward anger or fear. As parietal 'P3' EEG potential is associated with attention (Johnson, 1988) and surprise (Mars et al., 2008), this result

might be interpreted as increased attention and/or surprise to stimuli triggering effortful responses against biases.

In addition, both 'P3' EEG potential and pupil responses linked to cognitive processing have been suggested to stem from the neuromodulatory coeruleus-norepinephrine system (Nieuwenhuis et al., 2005). Our findings therefore suggest a possible involvement of neuromodulatory systems in conflict processing during categorization. If neuromodulators boost the gain to relevant information (Aston-Jones and Cohen, 2005), it is possible to infer that when individuals have a bias toward a specific choice, a stronger boost will be needed for them to overcome their bias and choose the option against their inclination (Gee et al., 2014).

Two types of biases were observed in the present experiment. The bias toward anger or fear choices is entirely linked to the response and can be illustrated by the general tendency of a negative person to dislike people, therefore being biased to judge new encounters as unfriendly. The bias of associating an identity with a certain emotion is linked to the stimulus and can be depicted by a negative judgment of a specific person that the judge dislikes (or that reminds him/her of a person he/she dislikes). Both effects coming from response-linked individual biases and those based on stimulus-linked associations between identities and emotions triggered similar behavioral and pupillary responses even though the latter is more social specific while the former can apply to any type of bias. This possibly suggests that both effects rely on a similar general mechanism portraying how individuals overcome their bias during decision-making. In the framework of the drift-diffusion model, which suggests that decisions are made through the accumulation of evidence until a threshold is reached (Ratcliff, 1978), this effect would manifest in lower individual thresholds when decisions correspond to the preferred option or when they are congruent with one's expectations.

In contrast to the contextual effects on individual thresholds/biases reported above, when a threat-related contextual factor was manipulated in a similar categorization task, the sensitivity to emotions was increased to threat-signaling emotions suggesting a change in the quality of the evidence accumulation process rather than in the decision threshold (El Zein et al., in revision). From an evolutionary perspective, recognizing a signal of threat requires to better perceive this signal for survival

purposes, however recognizing the mood of others is biased by personal experiences but does not necessitate an increased precision. During social interactions, judgments about others certainly require some flexibility as they can vary remarkably as a function of context (e.g., work environments or different countries). The fact that expectations about identities only change decision biases in a response-selective, but not in a less flexible stimulus-selective fashion, might represent a strategy for individuals to adapt in changing environments.

To conclude, we demonstrate that expectations about emotions expressed by identities change their recognition/perception, building on previous findings of contextual influences on social perceptual decisions. Their expression in neural processing and behavior can be described by a general mechanism in which neural engagement increases to effortful decisions against individual biases. Further research should extend the results to other emotions, especially those that we encounter on a daily basis such as happiness and disgust. Understanding how the human brain deals with social decisions influenced by past experience can provide insights into how biased choices are formed in social contexts, such as political choices.

## Methods

### Subjects

Thirty eight healthy subjects (19 females; mean age,  $22.3 \pm 0.5$  years) participated in the experiment. All participants were right-handed, with a normal vision and had no neurological or psychiatric history. They provided written informed consent according to institutional guidelines of the local research ethics committee (Declaration of Helsinki) and were paid for their participation. Seven subjects (4 females) were excluded from analysis, one because of low performances on the task, and the six others because of loss of pupillometry responses due to excessive blinks or movements (more than 40% of the experiment).

### Stimuli

Stimuli consisted of 32 identities (16 females) adapted from the RadBoud Faces Database (Langner et al., 2010) that varied in emotion (neutral, angry or fearful expressions) and gaze direction (direct toward the participant or averted  $45^\circ$  to the left or right). The intensity of emotional expressions varied with 7 levels of Anger and 7 levels of Fear equalized in perceived emotional intensities and a neutral condition (see [figure 1.a](#) for examples of stimuli). A complete description of the stimuli is provided in El Zein et al, under review.

### Experimental procedure

The task was also similar to the previous study (El Zein et al. under review) and consisted in categorizing the faces as fearful or angry. Using the Psychophysics-3 Toolbox (Brainard, 1997; Pelli, 1997), faces appeared for 250 ms on a black screen, after which subjects had to give their response by pressing one of the two 'control' buttons localized on the keyboard with their right or left index (a maximum of 2 seconds to respond before the next trial).

To manipulate the expectation of participants about identity/emotion associations, the 48 first trials of each bloc consisted of biased trials: half of the identities expressed only anger (identities belonging to Anger group), and the other half ex-

pressed only fear (identities belonging to Fear group). The assignment of identities to Anger and Fear groups varied from a participant to another, with the constraint that all identities were shown as expressing anger or fear across all subjects. To maximize the chance of inducing the associations, only the highest level of emotion strength (7) was presented in these trials, and each stimulus was repeated 3 times. Importantly, the task during these blocs was unchanged, and subjects had no explicit information about the content of the trials, they were only informed that the task might seem a little easier at the beginning of the blocs. In the rest of the bloc (112 trials), the highest level that was presented in the association blocs was excluded, and levels 1 to 6 from the morph continuum of both emotions were presented in an unbiased fashion (equal number of anger and fear for each identity) as well as the neutral stimuli. After each block, the percentage of correct responses calculated only on the unbiased trials was shown to the participants to keep them motivated. The experiment was divided in 8 experimental blocks, in half of the blocks, stimuli displayed a direct gaze and in the other half they displayed an averted gaze. Participants either started with 4 direct gaze blocs than proceeded to 4 averted gaze blocs, or the opposite and this was counterbalanced across subject. 16 identities were manipulated in the first 4 blocs, and 16 different identities were manipulated in the 4 last blocs.

At the end of the experiment, subjects performed a post-test, during which they saw the 32 neutral identities (with the corresponding gaze direction – same as they saw during the experiment), and they had to rate on a scale whether they think that each identity expressed more anger or fear during the experiment ([Figure 1c](#)). The position of anger and fear at the right and left of the scale was counterbalanced across subjects.

### **Behavioral data analyses**

Repeated-measures ANOVA was performed on the percentage of correct responses of the unbiased trials, with gaze direction (direct/averted), identity group (Anger group/Fear group), emotion (anger/fear), and intensity (6 levels of morphs) as within-subjects factors.

## Model selection

We performed model-guided analyses of the behavioural data to characterize the observed increase in recognition accuracy for congruent emotion/identity associations. We used Bayesian model selection based on the model evidence (estimated by a 10-fold cross-validation estimation of model log-likelihood, which penalizes implicitly for model complexity without relying on particular approximations such as the Bayesian Information Criterion or the Akaike Information Criterion). We applied both fixed-effects and random-effects statistics previously described in the literature. The fixed-effects comparison assumes all participants to have used the same underlying model to generate their behavior, such that the overall model evidence for a given model is proportional to the product of model evidence for the model for all participants. Based on this model evidence, we compared different models by computing their Bayes factor as the ratio of model evidence of the compared model (Jeffreys, 1961; Kass and Raftery, 1995). The random-effects comparison is more conservative in allowing different participants to use different models to generate their behavior, and aims at inferring the distribution over models that participants draw from (Penny et al., 2010). For this comparison, we computed support for the winning model by the exceedance probability ( $p_{exc}$ ), which is the probability that participants were more likely to choose this model to generate behavior over any alternative model.

We started with the simplest model (model 0) that could account for subject's decisions using a psychometric model such as:

$$P(\text{Anger}) = \Phi(w \cdot \mathbf{x} + b)$$

where  $P(\text{Anger})$  corresponds to the probability of judging the face as angry,  $\Phi$  to the cumulative normal function,  $w$  to the perceptual sensitivity to the display emotion,  $\mathbf{x}$  to the evidence in favor of Anger or Fear in each trial (emotion strength, from -6 for an intense expression of fear to +6 for an intense expression of anger), and  $b$  to an additive, stimulus-independent bias toward one of the two responses/emotions. We compared this 'null' model which did not allow for contextual influences of identity group on the decision process, with two additional models which propose two different mechanisms which could account for the observed increase in recognition accuracy for congruent combinations of identity and emotion. A first possibility (model 1)

would be that identities associated with an emotion group bias the recognition in favor of the congruent interpretation (anger for group anger, fear for group fear). Alternatively (model 2), expectations about group identity might selectively increase sensitivity to congruent emotions. Finally (model 3), effects could possibly involve both a change in the sensitivity to and response bias toward congruent emotion/identity associations. These models were used for both effects related to the experimentally induced identity group and the individual intrinsic group. After isolating winning models in each case, we used a fourth model that integrated both effects that change response bias.

### **Pupillometry preprocessing and analyses**

Pupil responses were recorded monocularly using an Eyelink 1000 eye tracking system, at a sample rate of 500 Hz. Subjects' head was positioned on a chinrest to avoid movements, and calibration was performed at the beginning of the experiment and every 2 experimental blocks (4 times in total).

Data was resampled offline to 50 Hz. Blinks were rejected by using cubic spline interpolation based on the 200 ms before the onset of the blink and 200 ms after the offset of the blink. High frequency artifacts were also removed using cubic spline interpolation. Afterwards, slow pupil fluctuations were removed using a 10 second sliding window. After removing the trial by trial baseline (-500 ms to the face onset) and standardizing (z-scoring) the activity across trials, the pupil signal was entered in a general linear model. We regressed the pupil data against main effects of responses with and against biases. Parameter estimates of this regression were extracted and averaged across participants.

### **Identity rating experiment**

20 subjects participated to the experiment (10 females, mean age=22.7±0.6). The 32 identities used in the experiment were presented in the neutral condition only. Each identity was presented twice, once with a direct, and once with an averted gaze. Faces appeared on the screen for 2 seconds after which they disappeared and 3 continuous scales were drawn on the screen. We asked participants to rate the iden-

tities in terms of threat, trustworthiness, and dominance on these 3 scales from 'not at all' to 'very much' (a text appeared at the top of the scales reminding the instructions: How much is this face threatening/trustworthy/dominant? accordingly with the scale). The order of the scales was randomized across subjects. The scales stayed on the screen until the three responses were given, however subjects were instructed to answer intuitively without spending too much time to decide.

### **EEG experiment**

24 participants had participated in a previous experiment and performed an emotion categorization task on morphed emotional expressions of Anger and Fear. (For details of the experimental procedures, EEG pre-processing and analyses, see El Zein et al. under review). We concentrated on P3 like component as it has been proposed to be tightly linked to neuromodulatory responses of the locus coeruleus system (Nieuwenhuis et al., 2005) also related to pupillometry responses (Aston-Jones and Cohen, 2005). We aimed to check whether we would find the same signatures of responses against and toward individual biases toward anger or fear. To do so, we regressed EEG signals in the isolated parietal electrodes stimulus and response locked against main effect of responses against and toward bias.



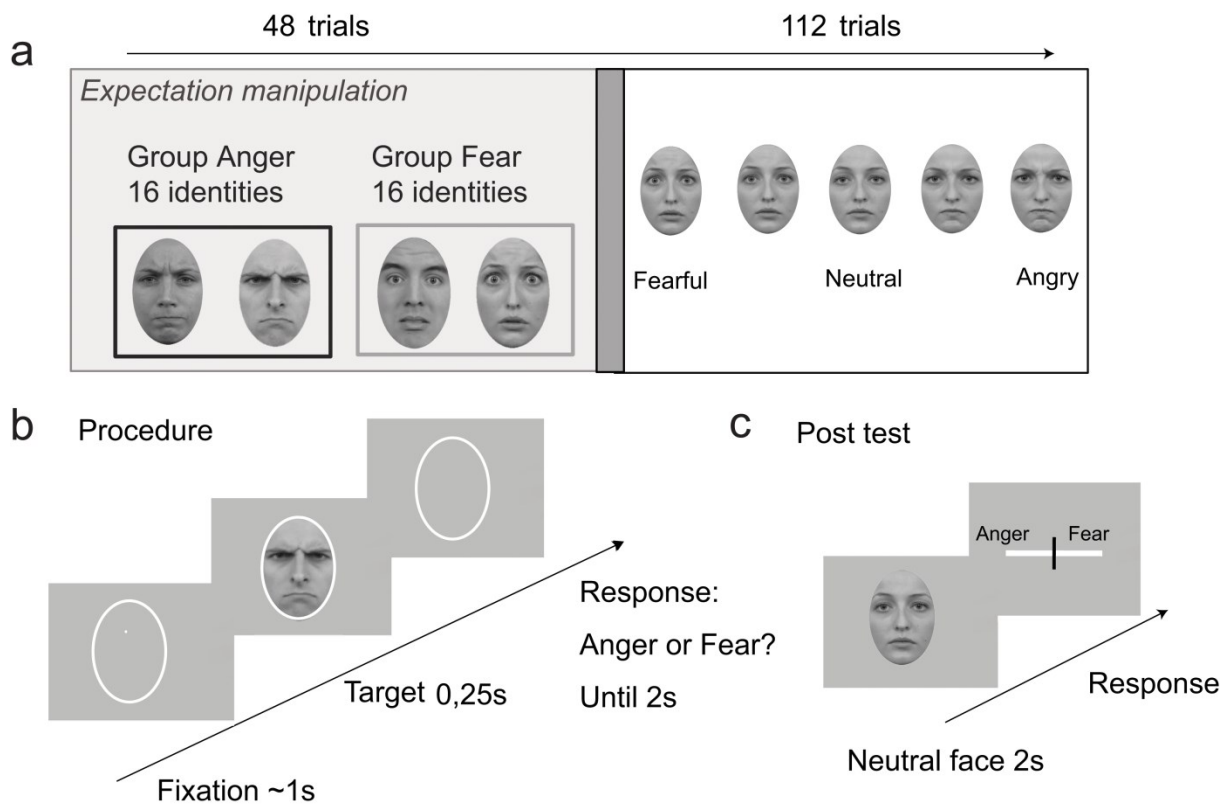
## References

- Apps MAJ, Tsakiris M (2013) Predictive codes of familiarity and context during the perceptual learning of facial identities. *Nat Commun* 4:2698.
- Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* 28:403–450.
- Bruce V, Young A (1986) Understanding face recognition. *Br J Psychol Lond Engl* 1953 77 ( Pt 3):305–327.
- Calder AJ, Young AW (2005) Understanding the recognition of facial identity and facial expression. *Nat Rev Neurosci* 6:641–651.
- Chen W, Lander K, Liu CH (2011) Matching faces with emotional expressions. *Front Psychol* 2:206.
- Dotsch R, Todorov A (2012) Reverse Correlating Social Face Perception. *Soc Psychol Personal Sci* 3:562–571.
- Fox CJ, Barton JJS (2007) What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Res* 1127:80–89.
- Gee JW de, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc Natl Acad Sci* 111:E618–E625.
- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics*. John Wiley and Sons.
- Haxby null, Hoffman null, Gobbini null (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233.
- Hess EH (1965) Attitude and pupil size. *Sci Am* Available at: <http://psycnet.apa.org/psycinfo/1965-09555-001> [Accessed August 3, 2015].
- Hess EH, Polt JM (1964) Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science* 143:1190–1192.
- Hess U, Adams RB, Kleck RE (2007) Looking at You or Looking Elsewhere: The Influence of Head Orientation on the Signal Value of Emotional Facial Expressions. *Motiv Emot* 31:137–144.
- Johnson R (1988) The amplitude of the P300 component of the event-related potential: Review and synthesis.
- Kahneman D, Beatty J, Pollack I (1967) Perceptual Deficit during a Mental Task. *Science* 157:218–219.
- Kaufmann JM, Schweinberger SR (2004) Expression influences the recognition of familiar faces. *Perception* 33:399–408.
- Kraus MW, Chen S (2010) Facial-feature resemblance elicits the transference effect. *Psychol Sci* 21:518–522.
- Liberman A, Fischer J, Whitney D (2014) Serial dependence in the perception of faces. *Curr Biol CB* 24:2569–2574.
- Macmillan NA, Creelman CD (2004) *Detection Theory: A User's Guide*. Psychology Press.

- Mars RB, Debener S, Gladwin TE, Harrison LM, Haggard P, Rothwell JC, Bestmann S (2008) Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *J Neurosci* 28:12539–12545.
- Montepare JM, Dobish H (n.d.) The Contribution of Emotion Perceptions and Their Overgeneralizations to Trait Impressions. *J Nonverbal Behav* 27:237–254.
- Nieuwenhuis S, Aston-Jones G, Cohen JD (2005) Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol Bull* 131:510–532.
- Porter G, Troscianko T, Gilchrist ID (2007) Effort during visual search and counting: insights from pupillometry. *Q J Exp Psychol* 2006 60:211–229.
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59–108.
- Schiller D, Freeman JB, Mitchell JP, Uleman JS, Phelps EA (2009) A neural mechanism of first impressions. *Nat Neurosci* 12:508–514.
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756.
- Todorov A, Gobbini MI, Evans KK, Haxby JV (2007) Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia* 45:163–173.
- Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P (2015) Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu Rev Psychol* 66:519–545.
- Van den Stock J, de Gelder B (2014) Face identity matching is influenced by emotions conveyed by face and body. *Front Hum Neurosci* 8:53.
- Verosky SC, Todorov A (2010) Generalization of affective learning about faces to perceptually similar faces. *Psychol Sci* 21:779–785.
- Wu EXW, Laeng B, Magnussen S (2012) Through the eyes of the own-race bias: eye-tracking and pupillometry during face recognition. *Soc Neurosci* 7:202–216.
- Zebrowitz LA, Montepare JM (2005) Psychology. Appearance DOES matter. *Science* 308:1565–1566.
- Zebrowitz LA, Montepare JM (2008) Social Psychological Face Perception: Why Appearance Matters. *Soc Personal Psychol Compass* 2:1497.

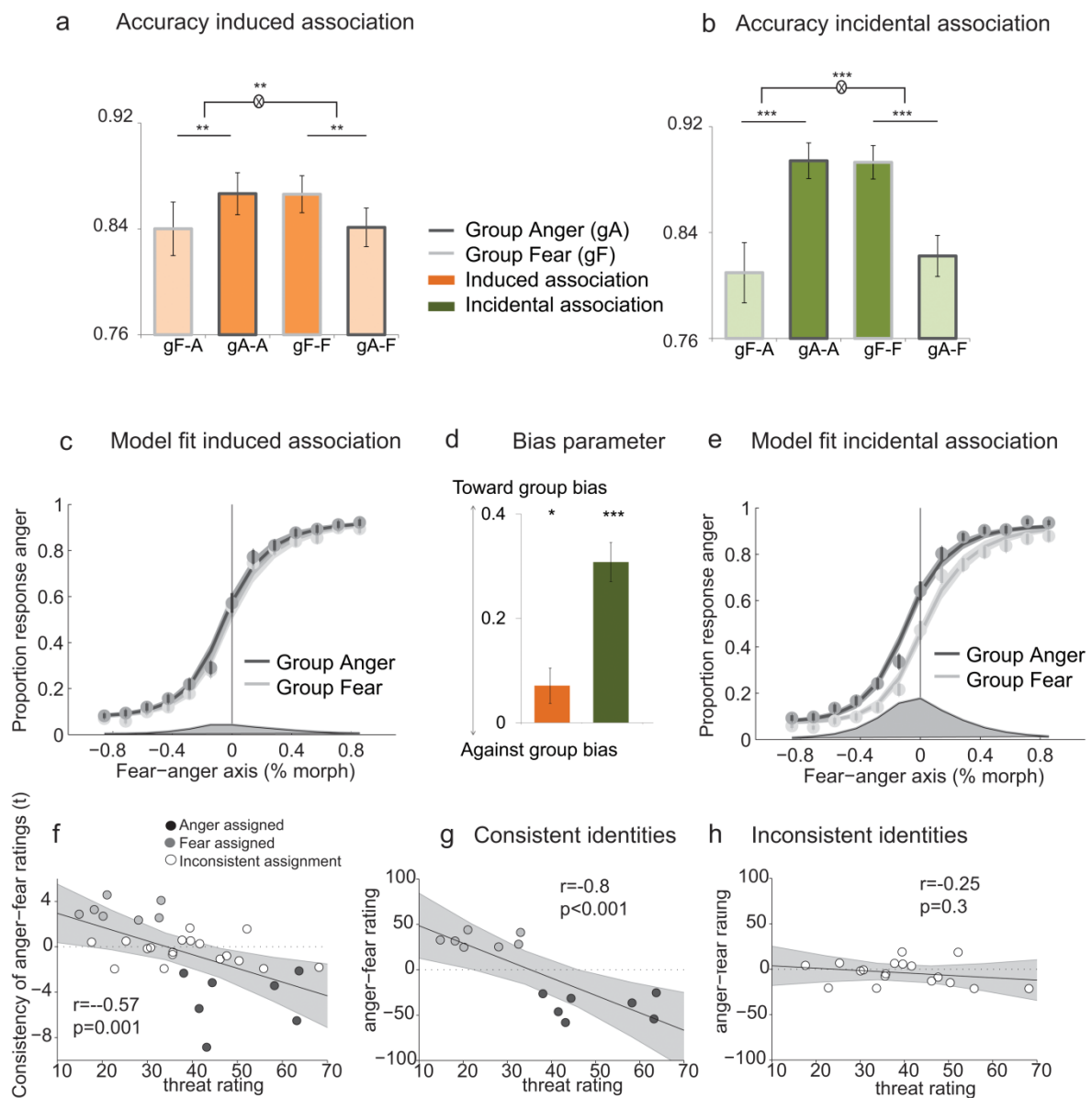
## Figures

Figure 1



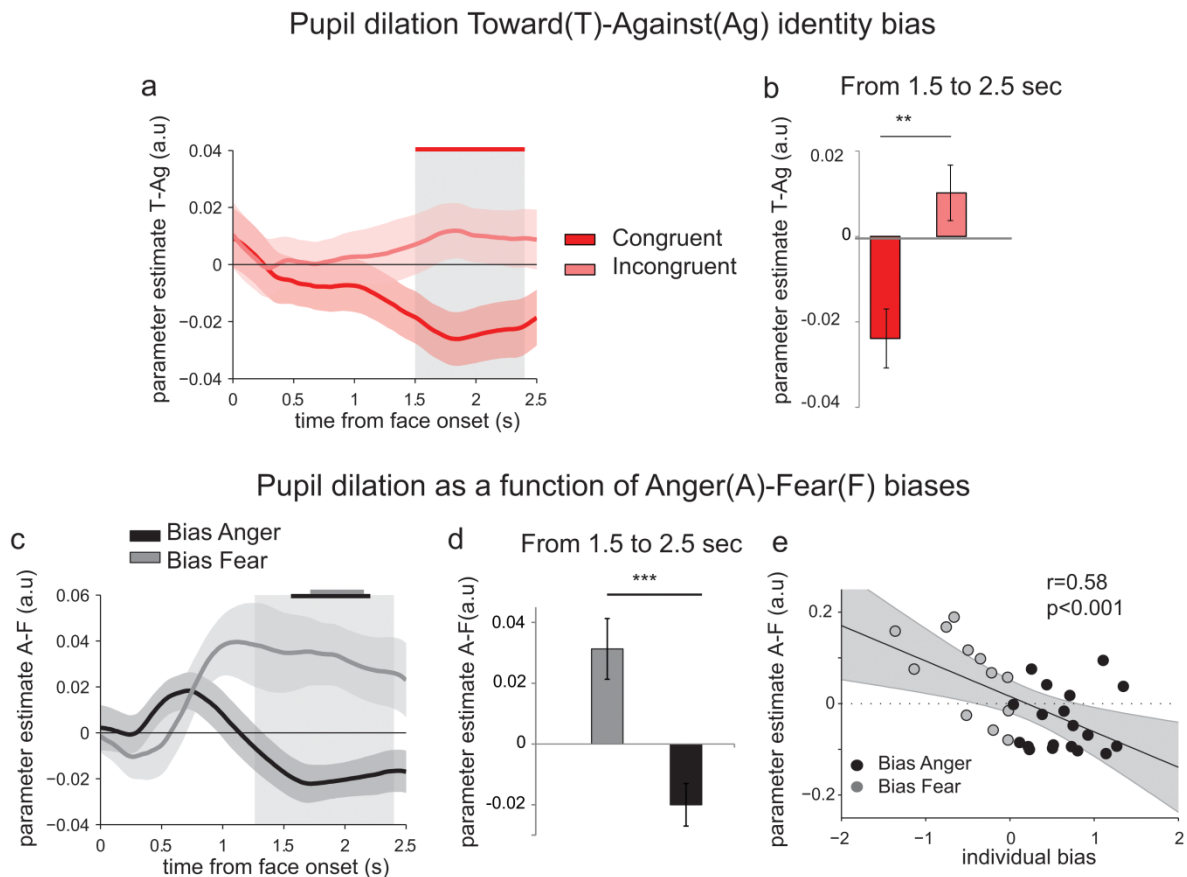
**Figure 1. Stimuli and experimental procedure.** (a) Left panel, the first 48 trials of each block consisted of identities expressing either anger or fear at the most intense level of emotion, in the purpose of creating expectations about the emotions of individual identities. Right panel, an equal number of morphed anger and fear expressions were presented in the other 112 trials of the block, examples are shown for one identity: morphs from neutral to intense fearful/angry expressions providing evidence for one or the other emotion. (b) Following fixation, a facial expression appeared for 250 ms, after which the participant had to indicate whether the face expressed anger or fear within 2 seconds. No feedback was provided after response. (c) At the end of the experiment, participants performed a post-test in which they rated on a scale whether they think each neutral identity (that appeared for 2 sec) expressed more anger or fear during the experiment.

Figure 2



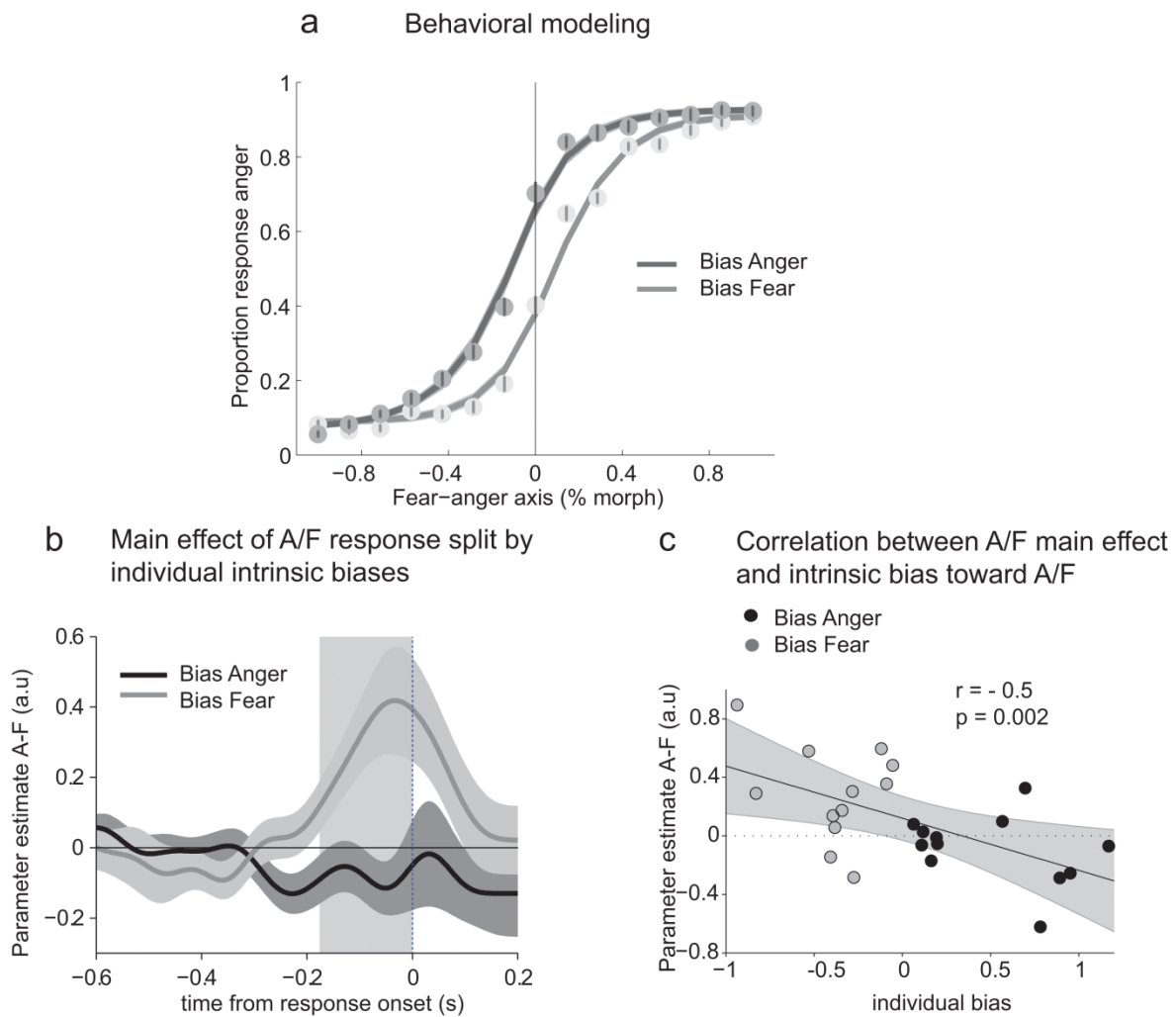
**Figure 2. Enhanced recognition accuracy and bias toward congruent displays of identity and emotions created through induced and incidental association.** (a) Proportion of correct responses for (from left to right) group Fear/anger, group Anger/anger, group Fear/fear and group Anger/fear. Congruent combinations of group and emotion created through induced association (in dark) benefited from an increased recognition accuracy. (b) Same as (a) for congruent combination of group and emotion created through incidental association. (c) Psychometric function representing the proportion of ‘anger’ responses as a function of the evidence for anger (proportion morph, 0 for neutral, negative towards fear, and positive towards anger) for Group Anger (dark grey) and Group Fear (light grey) created through induced association. Dots and attached error bars indicate the human data (mean  $\pm$  s.e.m.). Lines and shaded error bars indicate the predictions of the best-fitting model. (d) Estimated decisions bias for congruent minus incongruent identity-emotion associations for the model integrating both induced (orange) and incidental (green) effects on decision bias (e) same as (c) for incidental association. (f) Correlation between the consistency of anger-fear ratings, expressed in t-values extracted from the t-test against zero of the ratings for each identity across participants, with threat ratings for each identity. (g) Correlation between anger-fear ratings and threat ratings for identities consistently assigned. (h) Correlation between anger-fear ratings and threat rating for identities inconsistently assigned. \*:  $p < 0.05$ , \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

Figure 3



**Figure 3. Increased pupil responses for choices against individual biases.** (a) Time course of the parameter estimate of the main effect of pupil dilation responses when choice was toward vs against identity biases, for congruent (dark red) and incongruent (light red) induced and incidental associations between identities and emotions. Shaded error bars indicate s.e.m. Thick lines indicate significance against zero at a p-value of 0.05. Shaded areas indicate significant differences between congruent and incongruent conditions at  $p < 0.05$ . The negative deflection for congruent conditions indicates that pupil dilates more strongly when responses are against identity biases. (b) Parameter estimate of the main effect of pupil dilation responses when choice was toward vs against identity biases for congruent and incongruent conditions from 1.5 to 2.5 seconds after stimulus onset. (c) Time course of the parameter estimate of the main effect of pupil dilation responses when choice was toward vs against intrinsic biases (Anger vs Fear response), for individuals biased toward anger (dark grey) and fear (light grey). Same conventions as in (a). (d) Parameter estimate of the main effect of pupil dilation responses when choice was toward vs against intrinsic biases (Anger vs. Fear) for individuals biased toward fear or anger from 1.5 to 2.5 seconds after stimulus onset. (e) Correlation between anger/fear main effect on pupil dilation between 1.5 and 2.5 seconds and individual intrinsic biases toward anger or fear. \*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ .

Figure 4



**Figure 4. Increased EEG parietal responses for choices against intrinsic biases.** (a) Psychometric function representing the proportion of ‘anger’ responses as a function of the evidence for anger (proportion morph, 0 for neutral, negative towards fear, and positive towards anger) for participants biased toward anger (dark grey) and those biased toward fear (light grey). Conventions are the same as in Figure 2 (c). (b) Time course locked to the response time of the parameter estimate of the main effect of parietal EEG responses when choice was toward vs against intrinsic biases (Anger vs Fear response), for individuals biased toward anger (dark grey) and fear (light grey). Same conventions as in Figure 3 (a). (c) Correlation between anger/fear main effect on EEG parietal responses at 100 ms before response and individual intrinsic biases toward anger or fear.

## General discussion

The work conducted in this Ph.D., including 7 experiments in healthy humans, aimed at characterizing contextual influences during social perceptual decisions. Experiments 1 through 6 modulated within-emitter contextual factors including gaze direction, posture, emotion and spatial frequencies of the face. The goal of the observer was manipulated in Experiment 1 and Experiment 5 by changing the focus of attention of observers with tasks requiring to process either emotion or gaze cues. Experiments 2 and 6 manipulated elements characteristic of the observer including anxiety, age and social backgrounds. Finally, Experiment 7 manipulated observer's experience in the environment, by changing individual encounters with identities that expressed more or less one emotion over the other.

These contextual manipulations led to the first conclusion that multiple mechanisms underlie the processing of social displays, in this case emotional expressions (Figure 31). Social cues including emotion, gaze and gesture are represented simultaneously and interactively in the brain in sensory, motor and emotion-related regions. Within-emitter contextual factors thus interact and boost the bottom-up processing of combined signals relevant to the observer (e.g., signalling threat). Importantly, this interaction was found to be independent of top-down attention. Selective attention rather modulates the relative contribution of brain regions to this interaction relevant to the task at hand. Individual anxiety similarly influences the neural sources of emotion and gaze interaction by selectively increasing the involvement of action preparation regions during threat perception. Finally, prior experience enacts as a distinct top-down factor biasing responses toward expected interpretations and increasing neural activity to effortful decisions towards non-default interpretations.

The use of a model-based framework to analyse the behavioural and neural data recorded during this Ph.D. provided a compelling tool to characterize and clarify the mechanisms underlying contextual influences on social perceptual decisions. It allowed dissociating the effects of co-emitted social cues in threat contexts with top-down influences of prior experience. Behavioral modelling also provided a way to answer follow-up questions concerning the enhanced processing of threat signals,

by showing its specificity, reciprocity and generalizability. Importantly, model-based neural analysis further clarified the precise mechanisms by which gaze direction tunes emotion processing and allowed to dissociate the functional roles of sensory and motor response enhancements to threat in relation to anxiety.

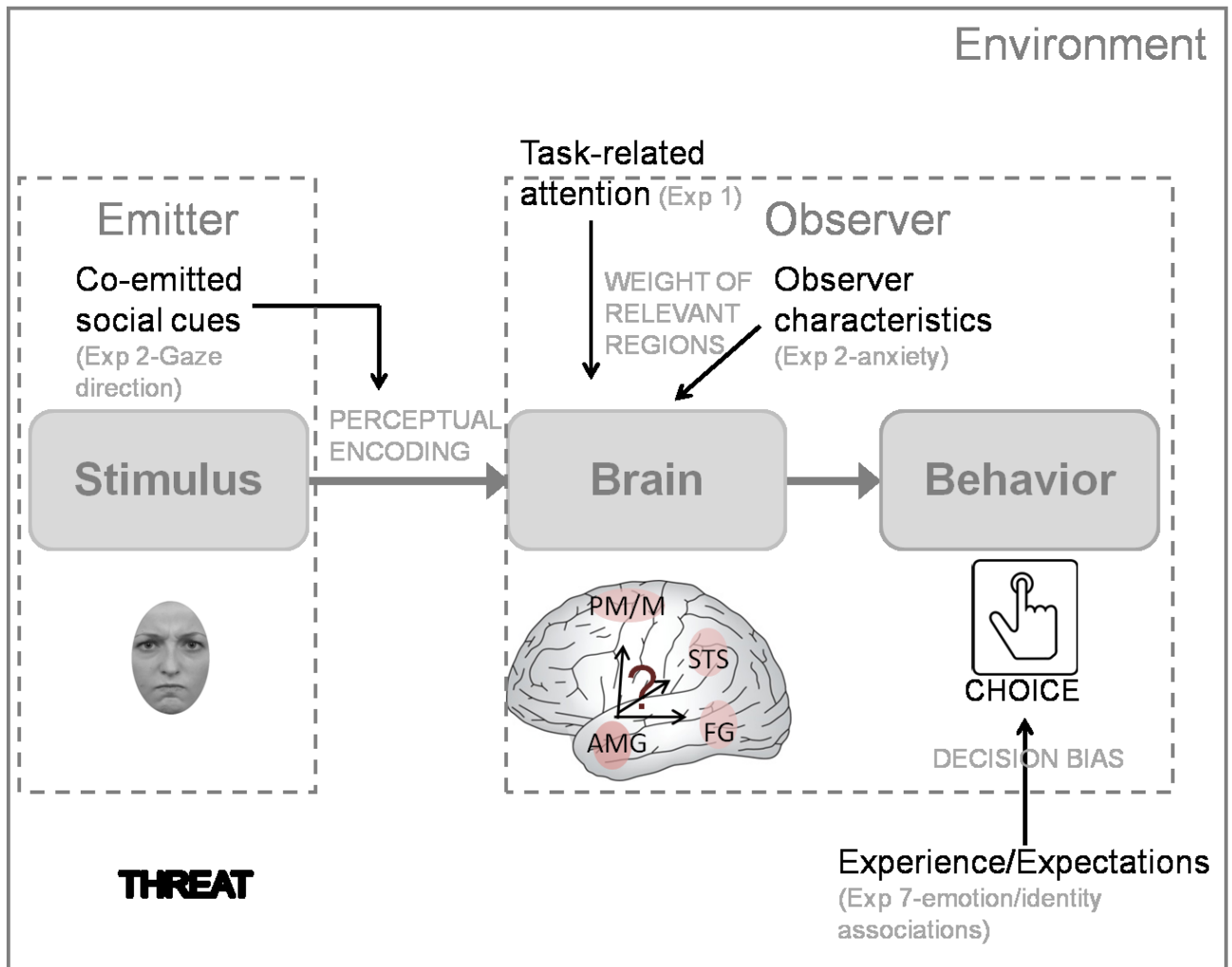


Figure 31. Summary of the main findings. Gaze direction, an inherent cue to facial expression, influenced emotion recognition by increasing the perceptual encoding of observer-relevant signals in threat contexts. Task-related attention did not modify the timing of the representation of observer-relevant signals, but changed the relative contribution of feature-relevant regions during stimulus processing. Individual anxiety similarly changed the brain regions where observer-relevant threat signals were encoded. Finally, expectations about emotion/identity associations biased the decision criterion toward congruent associations. Both sensory face-selective regions (STS: Superior temporal sulcus, FG: Fusiform gyrus) and motor related regions (PM: Premotor, M: motor) were involved in the processing of contextually threatening emotions. Whether the amygdala (AMG) drives these effects remains an open question.



## The evolutionary advantage of threat detection

The work conducted here builds on the extensive research that has been conducted on threat detection by characterizing the way social threats are represented in the human brain. We show that social cues are combined as a function of how much they signal threat to the observer. Here, this signal was provided by angry facial expressions gazing directly at the observer who becomes the target of the threat, and fearful expressions gazing away from the observer signaling a potential danger in the environment. The representation of these emotions signaling threat is independent of which social cue is task-relevant, and thus bypasses the attentional demands of the task at hand. It occurs in sensory, motor and emotion-related regions of the human brain within 170 ms after stimulus onset. This fast, selective, attention-independent encoding of social threats suggests that evolutionary pressure adapted to specifically optimize biologically relevant stimuli (Machens et al., 2005). Emotions are proposed to be expressed in facial features for adaptive reasons (Darwin, 1872), in the purpose of driving behaviors for survival (LeDoux, 2012, Panksepp, 1998, 2005). While our findings substantiate these suggestions, they also show that the functional significance of emotions is dependent on other social cues such as gaze direction. This dependence highlights the importance of considering emotions in context (Hess and Thibault, 2009) rather than as discrete entities that originate from 'hardwired' processes (Ekman, 1972, Ekman and Friesen 1978).

The use of model-based behavioral and neural analysis in Experiment 2 demonstrated for the first time to our knowledge that threat increases the *sensitivity* to emotional expressions, allowing a finer and more precise perception of sensory feature combinations signaling high threat to the perceiver. Previous studies had proposed that gaze direction tunes the processing of emotional expressions by changing behavior and neural engagement as a function of threat (Adams and Kleck, 2003, 2005; Sato et al., 2004, 2010; Sander et al., 2007; Akechi et al., 2010; Adams et al., 2012). However, this is the first evidence showing that it does so by increasing the quality of sensory processing of threat-signaling emotions in a bottom-up, attention-independent fashion. Indeed, the P300 evoked potential, that usually varies with attentional resources (Johnson, 1988), was not increased to threat-signaling

emotions. Moreover, 'mediation' analyses relating brain signals to behavior and stimulus information showed that choice predictive signals were not increased to these threat-signaling emotions, further suggesting that the increase of sensitivity to these emotions was not sustained by greater top-down attention to threat-signaling emotions (Nienborg and Cumming, 2009, 2010; Wyart et al., 2015).

Additional behavioral studies showed that the increased sensitivity to threat-signaling emotions is sustained by high spatial frequencies, stressing the role of the contribution of refined visual features to the increase in the precision of encoding. They also showed that this effect is not specific to emotion processing, but rather extends to the processing of morphed gaze directions. Emotion and gaze indeed reciprocally influence each other in the eye of threat, in accordance with previous studies (Lobmaier et al., 2008; Adams and Franklin, 2009; Ewbank et al., 2009). Here we further show that this influence changes the sensitivity to threat signals, whether the evidence to gaze direction or to emotion is manipulated. Finally, gaze direction enacted as a contextual factor modulating emotion processing specifically when anger and fear emotions were used. When categorization involved happy and angry expressions, the sensitivity to anger was increased as compared to happiness independently of gaze. This result suggests that in the context of opposite valence categorization, negative emotions (here anger) become generally more relevant to the observer, again highlighting the enhancement of the sensitivity to threat-signaling emotions. It is consistent with studies showing that negative emotions, particularly threatening ones (anger), capture attention in visual search paradigms that include positive and negative emotions (Eastwood et al., 2001; Ohman et al., 2001; Lo and Cheng, 2015).

To conclude, the behavior of participants is highly dependent on threat with increased recognition of threat-related signals sustained by greater sensitivity to these signals. Importantly, the rapid sensitization of neural responses to threat signals sheds light on the speed of threat-dependent modulations of neural processing. Finally, the results taken together stress the role of the 'perception' stage that is already modulated by the functional significance of the element to perceive (McArthur and Baron, 1983; Compton, 2003; Zebrowitz and Montepare, 2006), especially in the context of threat:

*'...invariant properties whose detection serves important adaptive actions, such as the transformational invariant fighting or the affordance danger, may be so strongly manifested in the stimulus information'* (McArthur and Baron, 1983)

### **Sensory and motor regions in emotion processing: functional roles?**

Our findings demonstrate the involvement of both sensory and motor-related regions in the processing of emotional expressions modulated by contextual gaze direction at early latencies, within 200 ms after stimulus onset. The quality of representation of emotions signaling threat to the observer was indeed enhanced in these regions. What are the functional roles of these two sets of regions in relation to the processing of contextually threatening emotions?

### **Common spatiotemporal characteristics of sensory activations during the processing of faces and observer-relevant signals from faces.**

The fusiform gyrus (FG) and superior temporal sulcus (STS) were responsible of combining gaze and emotion cues at 170 ms in both Experiments 1 and 2. These brain regions (Kanwisher et al., 1997; Puce et al., 1998; George et al., 1999), as well as the evoked potential N170 (George et al., 2005), have been repeatedly associated with face processing. Moreover, studies investigating the source of this N170 reported either the FG (Itier and Taylor, 2002; Rossion et al., 2003; Watanabe et al., 2003) or the STS (Batty and Taylor, 2003; Henson et al., 2003; Itier and Taylor, 2004; Itier et al., 2006; Watanabe et al., 2006) as responsible of N170 generation. Together, these findings thus suggest that social cues are combined into a single percept as a function of their relevance to the observer in threat contexts at the same spatiotemporal characteristics of general face processing, i.e., 170 ms in temporal regions involving STS and FG. This suggests that contextual within-emitter cues influence emotion processing at the basic level of facial sensory representations, modifying the very first step of the formation of social perceptual decisions, i.e., the perceptual 'encoding' of the facial expression.

However, according to the distributed human neural system for face perception, the FG processes invariant aspects of the face while the STS encodes variant features such as gaze and emotion (Haxby et al., 2000). Contrasting this clear dissociation, here we show that both regions can be involved in the processing of variant aspects in the face and their combination. Nevertheless, it remains unclear whether both STS and FG are acting in concert to process contextual facial expressions (Baseler et al., 2014), whether the FG receives feedback from the STS (Nguyen et al., 2014), or whether both these regions are gated by activity in other relevant structures to the processing of significant facial expressions to the observer, such as the AMG.

### **Selective involvement of motor-related regions in the representation of threat-relevant social displays**

Experiments 1 and 2 provide evidence of the contribution of motor-related regions to the representation of threatening emotions. In Experiment 1, post-central and parietal regions from the dorsal pathway are involved in the integration of emotion, gaze and gesture as a threatening signal within 200 ms after stimulus onset. In Experiment 2, at the same timing, motor-related signals selectively co-vary with the emotion strength of facial expressions signaling threat to the observer. Moreover, this early neural encoding of threat-signaling emotions in motor preparation signals occurs earlier than the formation of the upcoming choice, and thus depicts the representation of threat signals in motor cortex related to the sensory attributes of facial expressions rather than the choice. These results build on the few previous studies that did report motor activation in relation to emotion processing. Indeed, premotor or motor structures are involved in the perception of threat transmitted through the combination of emotion and gaze (Hadjikhani et al., 2008; Conty et al., 2012), the perception of threatening emotions of anger and fear conveyed by body expressions (de Gelder et al., 2004; Grèzes et al., 2007; Pichon et al., 2009, 2012), the perception of facial expressions of anger and fear (Balconi and Bortolotti, 2012, 2013), as well as the perception of negative natural emotional scenes (Borgomaneri et al., 2014). Altogether, these results converge by showing that the early motor involve-

ment during social processing is specific to threat perception, possibly for action preparation related to survival in threatening contexts (Ohman and Mineka, 2001; Frijda, 2009).

In threat settings, contextual within-emitter cues influence emotion processing in effector-specific regions, highlighting the role of motor engagement in the processing of contextually threatening emotions in parallel with face-selective sensory regions, rather than following these sensory regions only for response selection. This early motor activation, which does not predict choice at this latency, thus argues against 'sequential' views of decision-making that predict a cascade of sequential activations from the sensory feature-selective regions to associative regions and finally to effector-selective (motor) regions.

### **Modulation of the relative contribution of the dorsal pathway by task and individual-related relevance to the observer.**

The involvement of motor-related structures in the processing of the threat displays reported in Experiments 1 and 2 was modulated by task demands and individual characteristics, respectively.

#### **1. Task-related modulation**

In Experiment 1, the neural sources of the integration of emotion, gaze and gesture varied as a function of whether participants attended emotion or direction cues. While the involvement of STS and FG was consistent across tasks, emotion-related areas were more importantly involved during emotion categorization (is the person expressing anger or neutral emotions?), while postcentral and parietal regions from the dorsal pathway were rather engaged during direction of attention categorization (is the person looking away or at you?). Similarly, the premotor cortex was engaged in the combination of gaze, emotion and gesture as a function of threat when the task requested to attend directional cues (Conty et al., 2012). Although gaze direction is primarily informative as to whether the person is looking at or away from the observer, pointing gesture also contributes as a cue to determine the direction of

attention of a person. Both gaze processing (Haxby et al., 2000) and gesture processing (Grèzes & Decety, 2001 ; Rizzolatti, Fogassi, & Gallese, 2001) involve regions from the dorsal pathway including parietal and/or motor structures. Therefore, the contribution of these regions involved in directional cues processing was enhanced when attention was directed toward them.

Nevertheless, one should keep in mind that these regions were detected when analyses were ran on the study independently of task demands, thus not precluding their involvement even when attention is directed toward emotion processing. Rather, it is only the weight of their involvement that is modulated by attentional demands. Furthermore, Experiment 2 clearly shows that motor-related regions are involved in the encoding of contextual emotion processing even when attention is directed toward emotion cues. However, the stimuli as well as the analyses were very different not allowing a direct comparison between the 2 experiments. Finally, the difference in analyses that did not allow previous studies to directly measure the sensitivity to emotions might explain the lack of reports of motor involvement in the interaction between emotion and gaze cues in threat contexts (N'Diaye et al., 2009; Sato et al., 2010; Adams et al., 2012).

## **2. Anxiety-related modulation**

Experiment 2 builds on the results of Experiment 1 by showing that the involvement of ventral and dorsal pathways within 200 ms during the processing of threat signals stems from an increased sensitivity to these signals. Interestingly, a double dissociation between the effects in temporal ventral structures and dorsal motor structures as a function of anxiety was observed. While low-anxious individuals selectively encoded observer-relevant threat signals in temporal regions, high-anxious individuals did so in the motor cortex. Thus the involvement of motor related regions was modulated by individual anxiety of participants. The more participants were anxious, the more they engaged their motor cortex in the representation of threat related signals at 200 ms after stimulus onset. The normal anxiety response is suggested to represent an adaptive response that aims to prepare individuals to deal with threats (Bateson et al., 2011). In this vein, the increased involvement of the motor cortex with anxiety is

in accordance with the adaptive role of anxiety in reacting rapidly to threats. Previous studies suggested that high-anxious individuals do not properly integrate contextually relevant information but rather show a bias toward threatening stimuli (Bar-Haim et al., 2007) reflected in over-activation of the AMG (Bishop et al., 2004; Bishop, 2007). Here, the results imply that elevated anxiety within healthy individuals does not impair the processing of contextual threat, but rather shifts the representation of threat to regions consistent with the role of anxiety, the motor cortex, for the preparation to react to social threats. To further test this claim, it would be interesting to run an experiment where participants have to take action-related rather than perceptual decisions, such as approaching an individual in naturalistic environments. If the early involvement of the motor cortex in the representation of social threats truly has the role of action preparation when facing social threats, then action-related decisions (e.g., avoid threat) should accordingly be facilitated in high-anxious individuals.

### **Selective representation of threat signals: is the amygdala the starting point?**

Given the importance of the AMG in threat processing (Ohman, 2005), its contribution to the perception of ambiguous anger and fear expressions (Graham et al., 2007) and its demonstrated involvement in gaze and emotion interaction as a function of threat (Adams et al., 2003, 2012; Sato et al., 2004, 2010; N'Diaye et al., 2009), one might expect that it contributes to the observed increased sensitization to threat signals. Although in Experiment 1 the source reconstruction analyses suggested its possible involvement among ventral emotion-related structures, EEG scalp recordings lack from a precise localisation of signals in deep structures such as the AMG. The methodology we use, EEG, is thus not convenient to address whether the AMG contributes or even drives the processing of threat signals in this context. Evidence from other studies shows that the AMG is involved in threat processing even earlier than the latencies we observed, within 100 ms after stimulus onset (Luo et al., 2007; Hung et al., 2010; Liu and Ioannides, 2010). We can thus speculate that the AMG might be engaged in the earlier integration of social cues, and responsible of the enhanced sensitization to biologically relevant signals in sen-

sory and motor regions thanks to its direct projections to the ventral visual pathway including the FG and STS (Hadj-Bouziane et al., 2012), and to premotor structures in the dorsal pathway (Avendaño, 1983; Grèzes et al., 2014).

If this is true, then it would suggest that the AMG should dissociate threat-signaling emotions (anger with a direct gaze and fear with an averted gaze) at early latencies independently of anxiety. Indeed, both low and high anxious integrated emotion and gaze cues as a function of threat however in different structures, so if this integration originates in the AMG, than both high and low anxious should dissociate threat-signaling emotion in the AMG. However, results in the literature concerning the role of the AMG in threat processing in relation to anxiety are contradictory. On one hand, evidence points toward an impaired integration of contextual cues when processing threatening information associated with AMG hyperactivity in high anxious individuals (Bishop et al., 2004; Bishop, 2007; Etkin and Wager, 2007). On the other hand, one study showed that greater AMG activation to direct as compared to averted gaze anger increased with individual anxiety (Ewbank et al., 2010), suggesting that AMG integrates contextually threatening information in high anxious individuals.

Two elements can reconcile the findings. First, the influence of context might differ from a study to the other because of the nature of the task used. For example, the hyperactivity of AMG to threat signals is observed during a task that requires judging either houses or faces and reports that fearful faces are processed even when irrelevant to the task in high anxious individuals (Bishop et al., 2004). This result does not preclude that the AMG can integrate observer-relevant threat signals within fearful stimuli, as the AMG might have shown increased activity to fear with an averted gaze as compared to the presented direct gaze in that same experiment. Second, brain activity reported in fMRI studies does not provide the precise timing of the involvement of the AMG, so the discrepancy of results might simply be due to the fact that fMRI signals are not reflecting what is going on in the AMG at 100 ms. However, a recent MEG study that implemented source reconstruction algorithms capable of reaching activity in the AMG (Dumas et al., 2013) during the processing of fearful and neutral faces with direct and averted gaze, contradicts this suggestion at first sight. Indeed, this study did not reveal early interactions between gaze and



emotion cues but rather after 200 ms following stimulus onset. Nevertheless, there was no possible measure of the sensitivity to emotional expressions in this study where the faces were passively viewed, thus not excluding that the AMG might increase the sensitivity to threat-related emotions at early latencies.

To conclude, the role of the AMG as a generator of the increased sensitivity to biologically relevant stimuli (Sander et al., 2003) has still to be proven, through single neuron recordings in monkeys or implanted patients, or through MEG and EEG recordings thanks to the development of source reconstruction algorithms that are able to reach deep subcortical structures (Dumas et al., 2013).

## **Anger and Fear emotions**

The experiments conducted in this Ph.D. focused on two specific emotions: anger and fear. These emotions share an important characteristic: they can both signal threat to an observer depending on the context.

### **Do anger and fear emotions share similarities?**

When emotional faces including happiness, disgust, anger and fear are presented for short durations using backward masking, the categorization between anger and fear expressions specifically was disrupted (Sweeny et al., 2013). This could suggest that the fact that they both represent negative emotions related to threat makes them difficult to tear apart if presented very shortly. However, early stages of the perception of emotional expressions dissociate anger/disgust emotions from fear/surprise emotions, thereby not grouping anger and fear together (Jack et al., 2014). The authors interpret these groupements as originating from the fact that anger and disgust both signal stationary danger, in contrast to fear and surprise that rather signal a fast approaching danger.

### **The role of gaze direction in the dissociation of anger and fear emotions**

Facial information is integrated starting from the eyes region downward in the face as reflected in the N170 face-selective component (Schyns et al., 2007), suggesting that gaze information is necessarily processed during emotion processing. Moreover, diagnostic information in the face for the recognition of fear and anger emotions comprises the eyes (Schyns et al., 2009), and the recognition of morphed emotional expressions of anger and fear primarily relies on the eye region (Wegrzyn et al., 2015). These findings stress the importance of taking into account information in the eye, notably gaze direction, when investigating the processing of anger and fear emotion.

When gaze direction, an inseparable feature from facial expression, is taken into account, anger and fear emotions are functionally dissociated. Indeed, while anger associated with a direct gaze signals a direct danger/threat to the observer, only fear with an averted gaze signals a potential danger in the near environment (Sander et al., 2007). Danger might be considered as 'stationary' (Jack et al., 2014) in the case of direct anger as the source of threat is known to the observer (it is the emitter of the angry expression himself) while danger might be considered as 'fast approaching' (Jack et al., 2014) in the case of averted fear as it is in the periphery and unknown to the observer at the moment when he/she processes the emotion. Nevertheless, both direct anger and averted fear signal threat to the observer, directly coming from the emitter for direct anger, and coming from the environment for averted fear. The results from Experiment 2 clearly demonstrated that these two displays benefit from increased sensory processing, and exclusively involve motor related regions at their early processing. They highlight the existence of an early stage of processing, where the behavioral and biological relevance of social displays is inherent to the emotion recognition process. They also pinpoint the importance of considering contextual cues in the emotion recognition process, as they can notably change the functional meaning of emotional signals.

### **Anger and fear: what is the source of the signaled threat?**

As just stated above, the work presented in this dissertation supports the groupement of averted fear and direct anger in the same category because of their

shared communication of threat/danger to the observer. However, an important difference remains concerning these two displays of emotions. While direct anger is itself the source of the threat to the observer, averted fear only indirectly signals some potential danger in the periphery that is unknown at the moment when the observer processes the fearful expression. One might thus expect that not only averted fear should be processed efficiently, but also information in the periphery as it emanates potentially threatening information.

Accordingly, studies have shown that fearful stimuli, even with a direct gaze, increase attention to the surrounding stimuli (Phelps et al., 2006; Taylor and Whalen, 2014), and that the gaze of fearful faces and not happy faces is followed more consistently when participants are looking for a threatening target (Kuhn and Tipples, 2011). Gaze following of fearful expressions is also modulated by anxiety traits of observers as high anxious as opposed to low anxious individuals follow the gaze direction of a fearful face more than a neutral face during a perceptual detection task (Mathews et al., 2003). When compared with anger, only fearful expressions (here again, even with a direct gaze) enhance the memory of neutral words presented in the periphery, while on the contrary angry facial expressions are themselves better memorized than fearful expressions (Davis et al., 2011). The fact that even direct fear, and not only averted fear, orients attention to the periphery shows that the function of fearful expressions in signaling information in the environment extends to the perception of fear emotion in general, although in real life situations, it is only if averted that fear signals a potential danger. However, all these effects should at least be enhanced to averted fear (looking at the target to be memorized or judged).

Importantly, concerning the dissociation between anger and fear, the findings stated above draw attention on a clear dissimilarity between these emotions. On one hand, anger represents on its own an important signal that has to be efficiently processed for a proper reaction when facing threat. On the contrary, fear has to be processed efficiently along with information in the periphery possibly containing threatening information that has to be acted upon. Attention orienting to the periphery (Taylor and Whalen, 2014) and greater sensitivity to elements in the periphery (Phelps et al., 2006) following fearful faces have been reported. However, while the former study did not measure sensitivity to the elements in the periphery, the latter

did not contrast anger and fearful stimuli. Moreover, both of these studies did not contrast averted and direct gaze direction. It would therefore be interesting to test anger and fear emotions with direct and averted gaze in context with other elements in the periphery. This would allow building on the results of Experiment 2 that demonstrate that both direct anger and averted fear are processed with increased sensitivity, by possibly showing that it is only in the presence of fear, notably averted fear, that sensitivity to the other elements in the periphery will also be increased.

## Contextual effects related to congruence vs. significance to the observer

### **Different neural signatures of contextual effects related to congruence vs. significance to the observer?**

All experiments 1 to 6 manipulated among other elements contextual *within-emitter* cues, emotion and gaze direction, which interacted as a function of their significance to the observer. In these social displays however, there was no notion of ‘congruence’, i.e. an angry expressions looking away from an observer is not ‘incongruent’ but simply not ‘significant’ to the observer, who is not the source of the threat signaled by anger.

On the contrary, manipulations related to a change in the environment, involve a ‘congruency’ factor. For example, changing the emotional scene in which an emotional expression is presented (Righart and de Gelder, 2008b) creates ‘congruent’ (facial expression and scene congruent such as a happy face in a pleasant environment) and ‘incongruent’ (facial expression and scene incongruent such as a disgusted face in pleasant environment) conditions. Similarly, manipulating *within-emitter* body information in a way that it matches or not facial expressions (Meeren et al., 2005; Aviezer et al., 2011) creates congruent (e.g., a withdrawal posture with a fearful face) and incongruent conditions (a withdrawal posture with an angry face). These studies converge by showing that the recognition of a facial expression is facilitated by a congruent body posture or emotional scene. At the neural level, the results are however inconsistent: while the early visual P100 component was shown to be increased to incongruent stimuli induced by body-emotion associations but not

the face-selective N170 (Meeren et al., 2005), the opposite pattern was observed to scene-emotion associations (Righart and de Gelder, 2006, 2008b). In the latter study, only the N170, and not the P100, was modulated by congruency, by showing increased activity to congruent, not incongruent scene-emotion displays.

However, these studies focused on specific ERP components, the P100 and the N170. They do not report results on the P300, on which increased activity to incongruent displays, that should be similar to ‘surprise’ signals (Mars et al., 2008), would have been expected. If this was the case, it would clearly help dissociating such congruency effects from effects related to the significance of the stimulus to the observer. The effect related to significance to the observer induced by the manipulation of the gaze direction of emotional expression indeed did not show modulations of the P300 as a function of gaze (see Part 2, Experiment 2). In any case, additional neuroimaging experiments manipulating both congruency and observer-relevant information should be conducted to affirmatively state that these types of contextual elements, both facilitating behavioral performance, are different at the level of the brain.

Nevertheless, at the behavioral level, it is possible to dissociate increased performance related to a change in sensitivity – i.e., better discriminability of emotions, from a change related to a shift in individual decision bias – i.e., participants idiosyncratic choices. If contextual effects related to congruency differ from those related to the significance of the social signal to the observer (a change in the decision bias vs. a change in the sensitivity to the emotion), it would provide a clue into the mechanisms of these two types of effects.

### **Identity-specific expectations bias responses rather than change the sensitivity to congruent emotion-identity associations**

Experiment 7 addressed in part the issue raised above: do contextual elements related to congruence similarly influence emotion categorization than those related to the relevance to the observer? In this experiment, we altered participants’ emotional experience with individual identities (by adding initial blocks where half of the identi-

ties express one emotion or the other) to test whether and how this would affect subsequent emotion categorization. This manipulation created congruent conditions, i.e. identities expressing the emotion they were associated with in the biased blocks, and incongruent conditions, i.e. identities expressing the opposite emotion to the one they were associated with. Emotional experience with individual identities biased participants' behaviors toward congruent responses, without affecting the sensitivity to congruent displays of identity and emotion. Past experience with identities thus distinctively influenced emotion categorization in comparison with effects of gaze direction.

This difference possibly stems from the functional role of these contextual factors during emotion perception. Facial expressions encountered in everyday life are most often ambiguous. Thus, an additive prior influence rather than sensitivity enhancements to help disambiguate the perceived emotion on a daily basis, might stem from the associated biological cost for the organism. Indeed, sensitivity enhancements, that require enhancing the signal-to-noise ratio of neural processing, are probably more costly than additive biases. Accordingly, as combinations of gaze direction and negative emotions result in stimuli signaling threat to the observer, evolutionary pressure might have shaped the brain to prioritize these signals by increasing the perceptual sensitivity to these stimuli, despite the associated neural cost. In contrast, learning the general mood of a relative or a friend might be important for social interactions, but is not crucial in terms of survival. Furthermore, it can be advantageous to keep a flexible behavior when it comes to judging individuals based on experience, as associations based on experience are prone to many variations such as a specific country (different social codes) or work vs. holiday environments. Therefore, changing the decision bias, and not the more rigid perceptual encoding of stimuli when constructing interpretations about a person, could represent a strategy to adapt in changing environments.

These interpretations apply to the specific manipulation we used in Experiment 7. Certainly, it is different from previously reported experiments that varied congruency simultaneously to the presentation of the emotional expression (De Gelder and Vroomen, 2000; Meeren et al., 2005; Righart and de Gelder, 2006, 2008a, 2008b; Van den Stock et al., 2007), as here congruency effects are only

based on previously experienced associations between emotions and identities. However, one can speculate that model-based analyses of previously reported effects (for example by morphing facial expressions embedded in natural scenes) could reveal an influence of contextual elements related to congruency on the decision bias rather than the emotion sensitivity. This would contrast top-down effects triggered by expectation/congruency cues with bottom-up effects triggered by the combination of cues significant to the observer. Alternatively, this distinction could be less sharp than just stated, as contextual influences might function on a case-by-case basis depending on the precise manipulated factor. Taking again as an example faces in natural contexts, it is plausible that the sensitivity to negative emotions increases in negative contexts, because of the importance of detecting danger in unsafe environments. As an illustration, it is more crucial to perceive a threatening person when walking in a dangerous neighborhood at night, as compared to when walking in a safe neighborhood during the day. This would be consistent with the conclusion drawn from Experiments 1 to 6 suggesting that sensitivity to social signals should be increased whenever observer-relevant threat is involved.

### **No interaction between the effects of gaze direction and emotional experience with identities on emotion categorization**

In Experiment 7, gaze direction was also modulated based on our results in the 6 other experiments. Indeed, we thought it could interact with the expectation manipulation about the emotion of individual identities. More precisely, we expected that at least anger displaying a direct gaze would bias decisions more strongly than averted anger and fear in general. This assumption was based on the fact that angry facial expressions are better memorized than fearful expressions because they directly represent a signal of threat to an observer (Davis et al., 2011). However, no interaction was observed between gaze direction and effects of expectations, showing that they exert independent effects on emotion categorization. Importantly, expectations effects about the emotions of identities should be generalized to all emotions rather than being specific to anger and fear emotions (contrarily to gaze direction effects). A

useful follow-up study of Experiment 7 should thus consist of a similar design using day-to-day encountered emotions such as happiness, anger and disgust emotions.

## **Decision-making in social contexts**

Two elements seem to be pivotal to reach a comprehensive understanding of decision-making in humans. First, as it is indisputable that day to day decisions are influenced by contextual elements, taking into account these elements in the study of decision-making is essential. Second, considering decision-making in social contexts is a key step to the general understanding of human decision-making as most human choices are made in social environments. The work conducted in this Ph.D. attempted to characterize the mechanisms of decision-making in social contexts, by considering different contextual factors, and applying decision-theoretical models to understand human choice. In non-social perceptual decision-making studies, prior information that consists of either expectations or attentional cues is typically provided explicitly before the decision-relevant sensory information and is processed consciously by the participant. On the contrary, social stimuli being multi-dimensional and comprising of different complex features, these features can enact as contextual cues that co-occur with (rather than being provided before) the to-be-judged stimulus. Moreover, these contextual cues can implicitly (without being informative to the task, and thus a priori not consciously processed by the participant) influence decisions.

The manipulation of attention and expectations as contextual factors in non-social decision-making show that both types of manipulations influence decisions in sensory (Esterman and Yantis, 2010; Kok et al., 2012), motor (de Lange et al., 2013) and more domain-general (Rahnev et al., 2011a; Wyart et al., 2015) regions in the human brain. Here we similarly show that gaze direction, a simultaneous contextual cue to facial expression, increases the sensitivity to threat-signaling emotions in sensory, motor and associative domain-general brain regions. These findings are convergent with the idea that decisions form through a multi-level distributed mechanism (Cisek, 2012), encompassing sensory, associative and motor regions. Furthermore,



sensory and motor regions were recruited in tandem during perceptual decisions about emotional expressions in line with evidence suggesting that action-related regions contribute early on to the representation of the decision signal (Hernández et al., 2002; Gold and Shadlen, 2003; Romo et al., 2004; Heinen et al., 2006).

Despite the common networks involved in social and non-social decision-making, contextual gaze cue strikingly influenced the processing of emotional expressions within 200 ms after stimulus and contextual cue onset in sensory and motor pathways. Certainly, it is not possible to directly compare the effects of contextual manipulations in non-social decisions with social decisions in the classical paradigms used in the field. Indeed, as stated above, expectation cues for example can only be provided before the stimulus while gaze direction is presented together (simultaneously) with the target, i.e., the emotional expression. When expectation cues are provided before the stimulus, they influence motor preparation signals before the stimulus onset, however few hundreds of milliseconds after the cue onset (de Lange et al., 2013). The coincidence of the contextual cue (gaze direction) with the target (emotional expression) allowed demonstrating the context-specific involvement of these motor preparation signals at the same timing of the sensory encoding of stimulus information. Furthermore, the social aspect of the stimuli we used allowed to show a functional dissociation of sensory and motor regions as a function of individual characteristics, here anxiety. While high-anxious individuals encode negative emotions indifferently in sensory face-selective regions, their motor cortex activates at early latencies only in response to threat-signaling emotions. The early and concurrent involvement of motor and sensory regions draws important constraints on traditional theories of decision-making by showing parallel and functionally dissociable, rather than successive and co-dependent, responses to social threats in sensory and motor pathways.

In contrast with the contextual manipulation related to threat discussed above, the contextual factor tested in Experiment 7 was related to personal experience with identities, and was thus not presented simultaneously with the target stimulus. Although implicit in contrast to expectation cues in the nonsocial domain that are usually explicitly processed and inform the decision (e.g. Rahnev et al., 2011a; de Lange et al., 2013), this manipulation is not entirely specific to the social domain (such as

gaze direction in a facial expression). Indeed, it is possible to create associations between non-social elements such as shapes and colors for example, and maybe obtain similar effects of biased categorization. Nevertheless, this manipulation demonstrated that associations between identities and emotions are possible in a very short period of time, among 32 different identities. These rapid associations might be due to their social significance in daily interactions. However, a control study demonstrating that the same effects can be observed with 32 different shapes and 2 colors for example, would contradict the social nature of the results observed in this task. Furthermore, different types of biases were evidenced in this study: experimentally induced biases (those driven by our manipulation of adding blocks where individual identities were associated with specific emotions), incidental associations (idiosyncratic associations between identities and emotions), and intrinsic biases toward anger or fear responses independently identity. All three biases seem to operate through a general mechanism where neural engagement is increased to effortful decisions against biases. This conclusion is drawn based on the observed increased pupillary response to choices against biases, that replicates the findings of a study in which the decision was made during a perceptual detection task, thereby stressing the rather general aspect of this mechanism. Moreover, 'P3' related EEG responses showed a similar pattern, again extending the greater signal to decisions against biases to a brain wave associated with domain-general mechanisms (O'Connell et al., 2012), in line with the idea that expected events trigger diminished brain activity (Summerfield et al., 2008).

While the perceptual sensitivity to contextually threatening emotions was markedly increased within 200 ms in the brain and modulated by individual anxiety distinguishing threat signals from other stimuli, prior expectation about emotion/identity association acted as a top-down factor only influencing decisions at later stages, in accordance with Bayesian accounts of decision-making (Griffiths and Tenenbaum, 2006; Oaksford and Chater, 2007). Importantly, this dissociation sheds light on the different mechanisms by which different contextual factors influence decisions within social perceptual decisions.

## Conclusion

The work conducted in this Ph.D. reveals distinctive cognitive mechanisms by which contextual factors guided perceptual decisions on emotional expressions.

First, our findings highlight how decisions on multi-dimensional social stimuli are influenced through increased fast bottom-up processing of displays relevant to the observer. Gaze direction indeed tunes the processing of emotional expressions of anger and fear by increasing the sensitivity to signals of threat, direct anger and averted fear, within 200 ms in sensory and motor pathways. The enhanced sensitivity to direct anger and averted fear is mainly transmitted through high spatial frequencies, stressing the role of refined facial features in the better perception of threat signals. Moreover, the influence of gaze direction on emotion categorization is specific to anger and fear emotions, as choosing between happiness and anger shifted the relevant aspect to the observer to threat (anger) by generally increasing the sensitivity to anger independently of gaze. The enhanced sensitivity to threat signals, and the timing of their representation, is unchanged during emotion categorization and gaze categorization, demonstrating its independence from task related attentional demands.

Second, individual characteristics of participants influence the early (200 ms) processing of contextually threatening emotions of anger and fear: while high-anxious individuals selectively show enhanced sensitivity to threat-signaling emotions in motor cortex, the enhancement observed in face-selective temporal cortex is only driven by low-anxious individuals. This result dissociates the functions of sensory and motor systems in relation to anxiety, by pinpointing the adaptive role of anxiety in reacting rapidly to social threats, thereby recruiting the motor cortex in the early encoding of threat signals. Complementing the results described above, it sheds light on the fact that the behavioral and biological relevance of the stimuli are inherent to the emotion recognition process at an early stage of processing.

Third, when expectations about the emotions of individual identities were manipulated, they distinctively influenced emotion categorization by biasing decisions toward associated options and not by increasing the sensitivity to these options. Along with incidental associations between identities and emotions and intrinsic bi-

ases toward anger or fear responses, these induced associations reflected a general mechanism for emotion categorization biases whereby increased neural activations are required to drive unexpected decisions against existing biases. The change in the decision criterion (i.e., a bias) as a function of expectations about the emotion of identities, rather than a change of the sensitivity, could stem from the fact that such associations are related to daily interactions highly dependent on context and thus requiring a flexible adaptation. In everyday life, where emotional expressions are most often ambiguous, it can be useful to decrease the neural cost in disambiguating the perceived emotion through an additive influence on the decision bias.

Understanding how the brain forms perceptual decisions in social environments is crucial, but certainly only represents one piece of the puzzle to reach an integrative view of social functioning and the neural basis of social cognition. The work conducted here is an important step toward the understanding of at least that piece of the puzzle. By taking advantage of a model-based framework, we revealed the underlying computations of brain functioning when deciding upon socially meaningful stimuli, facial expressions (see summary Figure 31). The framework used here opens the door to the prospect of identifying what precisely is altered in psychiatric conditions associated with modified social behaviors, such as autistic individuals that are characterized by significant deficits related to social cognition. Along these lines, using the same stimuli and framework described here, ongoing work in the team investigates whether autistic individuals are capable of integrating social cues similarly as controls during social perceptual decisions. Furthermore, also using the same framework, other work in progress in the team explores how these social perceptual decisions are related to social inclusion of individuals, their social preferences, and their life history.

## References

- Adams RB, Franklin RG (2009) Influence of emotional expression on the processing of gaze direction. *Motiv Emot* 33:106–112.
- Adams RB, Gordon HL, Baird AA, Ambady N, Kleck RE (2003) Effects of gaze on amygdala sensitivity to anger and fear faces. *Science* 300:1536.
- Adams RB Jr, Franklin RG Jr, Kveraga K, Ambady N, Kleck RE, Whalen PJ, Hadjikhani N, Nelson AJ (2012) Amygdala responses to averted vs direct gaze fear vary as a function of presentation speed. *Soc Cogn Affect Neurosci* 7:568–577.
- Adams RB Jr, Kleck RE (2003) Perceived gaze direction and the processing of facial displays of emotion. *Psychol Sci* 14:644–647.
- Adams RB, Kleck RE (2005) Effects of direct and averted gaze on the perception of facially communicated emotion. *Emot Wash DC* 5:3–11.
- Adolphs R, Gosselin F, Buchanan TW, Tranel D, Schyns P, Damasio AR (2005) A mechanism for impaired fear recognition after amygdala damage. *Nature* 433:68–72.
- Adolphs R, Tranel D, Damasio H, Damasio A (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372:669–672.
- Akechi H, Senju A, Kikuchi Y, Tojo Y, Osanai H, Hasegawa T (2010) The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: an ERP study. *Neuropsychologia* 48:2841–2851.
- Allison null, Puce null, McCarthy null (2000) Social perception from visual cues: role of the STS region. *Trends Cogn Sci* 4:267–278.
- Apps MAJ, Tsakiris M (2013) Predictive codes of familiarity and context during the perceptual learning of facial identities. *Nat Commun* 4:2698.
- Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu Rev Neurosci* 28:403–450.
- Atkinson AP, Heberlein AS, Adolphs R (2007) Spared ability to recognise fear from static and moving whole-body cues following bilateral amygdala damage. *Neuropsychologia* 45:2772–2782.
- Avendaño C (1983) A fine structural study of the cells that proliferate in the partially denervated dentate gyrus of the rat. *Anat Embryol (Berl)* 166:317–332.
- Aviezer H, Bentin S, Dudarev V, Hassin RR (2011) The automaticity of emotional face-context integration. *Emot Wash DC* 11:1406–1414.
- Balconi M, Bortolotti A (2012) Detection of the facial expression of emotion and self-report measures in empathic situations are influenced by sensorimotor circuit inhibition by low-frequency rTMS. *Brain Stimulat* 5:330–336.
- Balconi M, Bortolotti A (2013) The “simulation” of the facial expression of emotions in case of short and long stimulus duration. The effect of pre-motor cortex inhibition by rTMS. *Brain Cogn* 83:114–120.
- Bankó EM, Gál V, Körtvélyes J, Kovács G, Vidnyánszky Z (2011) Dissociating the effect of noise on sensory processing and overall decision difficulty. *J Neurosci Off J Soc Neurosci* 31:2663–2674.

- Bar-Haim Y, Lamy D, Pergamin L, Bakermans-Kranenburg MJ, van IJzendoorn MH (2007) Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychol Bull* 133:1–24.
- Bar M (2004) Visual objects in context. *Nat Rev Neurosci* 5:617–629.
- Barrett LF, Lindquist KA, Bliss-Moreau E, Duncan S, Gendron M, Mize J, Brennan L (2007) Of Mice and Men: Natural Kinds of Emotions in the Mammalian Brain? A Response to Panksepp and Izard. *Perspect Psychol Sci J Assoc Psychol Sci* 2:297–312.
- Baseler HA, Harris RJ, Young AW, Andrews TJ (2014) Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cereb Cortex N Y N* 1991 24:737–744.
- Bateson M, Brilot B, Nettle D (2011) Anxiety: an evolutionary approach. *Can J Psychiatry Rev Can Psychiatr* 56:707–715.
- Batty M, Taylor MJ (2003) Early processing of the six basic facial emotional expressions. *Brain Res Cogn Brain Res* 17:613–620.
- Bishop SJ (2007) Neurocognitive mechanisms of anxiety: an integrative account. *Trends Cogn Sci* 11:307–316.
- Bishop SJ, Duncan J, Lawrence AD (2004) State anxiety modulation of the amygdala response to unattended threat-related stimuli. *J Neurosci Off J Soc Neurosci* 24:10364–10368.
- Borgomaneri S, Gazzola V, Avenanti A (2014) Temporal dynamics of motor cortex excitability during perception of natural emotional scenes. *Soc Cogn Affect Neurosci* 9:1451–1457.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Breiter HC, Etcoff NL, Whalen PJ, Kennedy WA, Rauch SL, Buckner RL, Strauss MM, Hyman SE, Rosen BR (1996) Response and habituation of the human amygdala during visual processing of facial expression. *Neuron* 17:875–887.
- Brosch T, Bar-David E, Phelps EA (2013) Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychol Sci* 24:160–166.
- Canli T, Amin Z (2002) Neuroimaging of emotion and personality: scientific evidence and ethical considerations. *Brain Cogn* 50:414–431.
- Carrasco M (2006) Covert attention increases contrast sensitivity: Psychophysical, neurophysiological and neuroimaging studies. *Prog Brain Res* 154:33–70.
- Cheadle S, Wyart V, Tsetsos K, Myers N, de Gardelle V, Herculano-Houzel S, Summerfield C (2014) Adaptive gain control during human perceptual choice. *Neuron* 81:1429–1441.
- Cisek P (2012) Making decisions through a distributed consensus. *Curr Opin Neurobiol* 22:927–936.
- Cisek P, Kalaska JF (2010) Neural Mechanisms for Interacting with a World Full of Action Choices. *Annu Rev Neurosci* 33:269–298.
- Compton RJ (2003) The interface between emotion and attention: a review of evidence from psychology and neuroscience. *Behav Cogn Neurosci Rev* 2:115–129.
- Conty L, Dezechache G, Hugueville L, Grèzes J (2012) Early binding of gaze, gesture, and emotion: neural time course and correlates. *J Neurosci Off J Soc Neurosci* 32:4531–4539.

- Conty L, N'Diaye K, Tijus C, George N (2007) When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia* 45:3024–3037.
- Cuthbert BN, Schupp HT, Bradley MM, Birbaumer N, Lang PJ (2000) Brain potentials in affective picture processing: covariation with autonomic arousal and affective report. *Biol Psychol* 52:95–111.
- Darwin, C. R. (1872). *The expression of the emotions in man and animals*. London: John Murray. 1st edition
- Davis FC, Johnstone T, Mazzulla EC, Oler JA, Whalen PJ (2010) Regional response differences across the human amygdaloid complex during social conditioning. *Cereb Cortex N Y N* 1991 20:612–621.
- Davis FC, Somerville LH, Ruberry EJ, Berry ABL, Shin LM, Whalen PJ (2011) A tale of two negatives: Differential memory modulation by threat-related facial expressions. *Emot Wash DC* 11:647–655.
- de Gelder B, Snyder J, Greve D, Gerard G, Hadjikhani N (2004) Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A* 101:16701–16706.
- De Gelder B, Vroomen J (2000) The perception of emotions by ear and by eye. *Cogn Emot* 14:289–311.
- de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci Off J Soc Neurosci* 33:1400–1410.
- Dolan RJ, Morris JS, Gelder B de (2001) Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci* 98:10006–10010.
- Donner TH, Siegel M, Fries P, Engel AK (2009) Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol CB* 19:1581–1585.
- Dotsch R, Todorov A (2012) Reverse Correlating Social Face Perception. *Soc Psychol Personal Sci* 3:562–571.
- Dubal S, Foucher A, Jouvent R, Nadel J (2011) Human Brain Spots Emotion in Non Humanoid Robots. *Soc Cogn Affect Neurosci* 6:90–97.
- Dumas T, Dubal S, Attal Y, Chupin M, Jouvent R, Morel S, George N (2013) MEG Evidence for Dynamic Amygdala Modulations by Gaze and Facial Emotions. *PLoS One* 8:e74145.
- Eastwood JD, Smilek D, Merikle PM (2001) Differential attentional guidance by unattended faces expressing positive and negative emotion. *Percept Psychophys* 63:1004–1013.
- Eimer M, Holmes A (2002) An ERP study on the time course of emotional face processing. *Neuroreport* 13:427–431.
- Einhäuser W, Stout J, Koch C, Carter O (2008) Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc Natl Acad Sci* 105:1704–1709.
- Ekman P. (1972). Universals and Cultural Differences in Facial Expressions of Emotions. In Cole, J. (Ed.), *Nebraska Symposium on Motivation* (pp. 207-282). Lincoln, NB: University of Nebraska Press.
- Ekman P (1999). Basic Emotions. In T. Dalgleish and M. Power (Eds.). *Handbook of Cognition and Emotion*. Sussex, U.K.: John Wiley & Sons, Ltd.

- Ekman P, Friesen, W, and Hagar, J.C., eds. (1978). *Facial Action Coding System Investigators Guide* (Salt Lake City: Research Nexus).
- Esterman M, Yantis S (2010) Perceptual expectation evokes category-selective cortical activity. *Cereb Cortex N Y N 1991* 20:1245–1253.
- Etkin A, Wager TD (2007) Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry* 164:1476–1488.
- Ewbank MP, Fox E, Calder AJ (2010) The Interaction Between Gaze and Facial Expression in the Amygdala and Extended Amygdala is Modulated by Anxiety. *Front Hum Neurosci* 4:56.
- Ewbank MP, Jennings C, Calder AJ (2009) Why are you angry with me? Facial expressions of threat influence perception of gaze direction. *J Vis* 9:16.1–7.
- Farroni T, Csibra G, Simion F, Johnson MH (2002) Eye contact detection in humans from birth. *Proc Natl Acad Sci U S A* 99:9602–9605.
- Fett AKJ, Shergill SS, Krabbendam L (2015) Social neuroscience in psychiatry: unravelling the neural mechanisms of social dysfunction. *Psychol Med* 45:1145–1165.
- Freedman DJ, Riesenhuber M, Poggio T, Miller EK (2003) A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci Off J Soc Neurosci* 23:5235–5246.
- Frijda, N. (1986). *The emotions*. Cambridge, England: Cambridge University Press.
- Frijda NH (2009) Emotion Experience and its Varieties. *Emot Rev* 1:264–271.
- Fusar-Poli P, Placentino A, Carletti F, Landi P, Allen P, Surguladze S, Benedetti F, Abbamonte M, Gasparotti R, Barale F, Perez J, McGuire P, Politi P (2009) Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J Psychiatry Neurosci JPN* 34:418–432.
- Gee JW de, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc Natl Acad Sci* 111:E618–E625.
- George N (2013) The Facial Expression of Emotions. In: *The Cambridge Handbook of Human Affective Neuroscience*. Cambridge University Press. Available at: <http://dx.doi.org/10.1017/CBO9780511843716.011>.
- George N, Dolan RJ, Fink GR, Baylis GC, Russell C, Driver J (1999) Contrast polarity and face recognition in the human fusiform gyrus. *Nat Neurosci* 2:574–580.
- George N, Jemel B, Fiori N, Chaby L, Renault B (2005) Electrophysiological correlates of facial decision: insights from upright and upside-down Mooney-face perception. *Brain Res Cogn Brain Res* 24:663–673.
- Gold JI, Shadlen MN (2003) The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J Neurosci Off J Soc Neurosci* 23:632–651.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
- Gosselin F, Schyns PG (2001) Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Res* 41:2261–2271.
- Graham R, Devinsky O, LaBar KS (2007) Quantifying deficits in the perception of fear and anger in morphed facial expressions after bilateral amygdala damage. *Neuropsychologia* 45:42–54.



- Green DM, Swets JA (1966) *Signal Detection Theory and Psychophysics*. John Wiley and Sons.
- Grèzes J, Decety J (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum Brain Mapp* 12:1–19.
- Grèzes J, Dezechache G, Eskenazi T (2015) Limbic to Motor Interactions during Social Perception. In: *Brain Mapping* (Toga AW, ed), pp 1027–1030. Waltham: Academic Press. Available at: <http://www.sciencedirect.com/science/article/pii/B9780123970251001287> [Accessed June 24, 2015].
- Grèzes J, Pichon S, de Gelder B (2007) Perceiving fear in dynamic body expressions. *NeuroImage* 35:959–967.
- Grèzes J, Valabrègue R, Gholipour B, Chevallier C (2014) A direct amygdala-motor pathway for emotional displays to influence action: A diffusion tensor imaging study. *Hum Brain Mapp* 35:5974–5983.
- Griffiths TL, Tenenbaum JB (2006) Optimal predictions in everyday cognition. *Psychol Sci* 17:767–773.
- Grill-Spector K, Knouf N, Kanwisher N (2004) The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 7:555–562.
- Hadj-Bouziane F, Liu N, Bell AH, Gothard KM, Luh W-M, Tootell RBH, Murray EA, Ungerleider LG (2012) Amygdala lesions disrupt modulation of functional MRI activity evoked by facial expression in the monkey inferior temporal cortex. *Proc Natl Acad Sci* 109:E3640–E3648.
- Hadjikhani N, Hoge R, Snyder J, de Gelder B (2008) Pointing with the eyes: the role of gaze in communicating danger. *Brain Cogn* 68:1–8.
- Hajcak G, Moser JS, Simons RF (2006) Attending to affect: appraisal strategies modulate the electrocortical response to arousing pictures. *Emot Wash DC* 6:517–522.
- Harris RJ, Young AW, Andrews TJ (2012) Morphing between expressions dissociates continuous from categorical representations of facial expression in the human brain. *Proc Natl Acad Sci U S A* 109:21164–21169.
- Haxby, Hoffman, Gobbini (2000) The distributed human neural system for face perception. *Trends Cogn Sci* 4:223–233.
- Hebart MN, Donner TH, Haynes J-D (2012) Human visual and parietal cortex encode visual choices independent of motor plans. *NeuroImage* 63:1393–1403.
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.
- Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG (2006) Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proc Natl Acad Sci U S A* 103:10023–10028.
- Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. *Nat Rev Neurosci* 9:467–479.
- Heinen SJ, Rowland J, Lee B-T, Wade AR (2006) An oculomotor decision process revealed by functional magnetic resonance imaging. *J Neurosci Off J Soc Neurosci* 26:13515–13522.

- Henson RN, Goshen-Gottstein Y, Ganel T, Otten LJ, Quayle A, Rugg MD (2003) Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cereb Cortex N Y N* 1991 13:793–805.
- Hernández A, Zainos A, Romo R (2002) Temporal evolution of a decision-making process in medial premotor cortex. *Neuron* 33:959–972.
- Hess U, Adams RB, Kleck RE (2007) Looking at You or Looking Elsewhere: The Influence of Head Orientation on the Signal Value of Emotional Facial Expressions. *Motiv Emot* 31:137–144.
- Hess U, Thibault P (2009) Darwin and emotion expression. *Am Psychol* 64:120–128.
- Hoehl S, Striano T (2008) Neural processing of eye gaze and threat-related emotional facial expressions in infancy. *Child Dev* 79:1752–1760.
- Hoffman KL, Gothard KM, Schmid MC, Logothetis NK (2007) Facial-expression and gaze-selective responses in the monkey amygdala. *Curr Biol CB* 17:766–772.
- Holmes A, Kiss M, Eimer M (2006) Attention modulates the processing of emotional expression triggered by foveal faces. *Neurosci Lett* 394:48–52.
- H S Bashinski VRB (1980) Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Percept Amp Psychophys* 28:241–248.
- Hsu S-M (2015) The neural mechanism underlying the effects of preceding contexts on current categorization decisions. *Neuropsychologia* 66:39–47.
- Hung Y, Smith ML, Bayle DJ, Mills T, Cheyne D, Taylor MJ (2010) Unattended emotional faces elicit early lateralized amygdala–frontal and fusiform activations. *NeuroImage* 50:727–733.
- Iidaka T, Saito DN, Komeda H, Mano Y, Kanayama N, Osumi T, Ozaki N, Sadato N (2010) Transient neural activation in human amygdala involved in aversive conditioning of face and voice. *J Cogn Neurosci* 22:2074–2085.
- Itier RJ, Latinus M, Taylor MJ (2006) Face, eye and object early processing: what is the face specificity? *NeuroImage* 29:667–676.
- Itier RJ, Taylor MJ (2002) Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *NeuroImage* 15:353–372.
- Itier RJ, Taylor MJ (2004) Source analysis of the N170 to faces and objects. *Neuroreport* 15:1261–1265.
- Jack RE, Caldara R, Schyns PG (2012a) Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *J Exp Psychol Gen* 141:19–25.
- Jack RE, Garrod OGB, Schyns PG (2014) Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Curr Biol CB* 24:187–192.
- Jack RE, Garrod OGB, Yu H, Caldara R, Schyns PG (2012b) Facial expressions of emotion are not culturally universal. *Proc Natl Acad Sci U S A* 109:7241–7244.
- Jeffreys H (1961) *Theory of probability* / by Harold Jeffreys. Oxford: Clarendon Press.
- Johnson R (1988) The amplitude of the P300 component of the event-related potential: Review and synthesis.

- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci Off J Soc Neurosci* 17:4302–4311.
- Kass RE, Raftery AE (1995) Bayes Factors. *J Am Stat Assoc* 90:773–795.
- Kelly SP, O’Connell RG (2013) Internal and External Influences on the Rate of Sensory Evidence Accumulation in the Human Brain. *J Neurosci* 33:19434–19441.
- Kelly SP, O’Connell RG (2015) The neural processes underlying perceptual decision making in humans: Recent progress and future directions. *J Physiol Paris* 109:27–37.
- Klucharev V, Sams M (2004) Interaction of gaze direction and facial expressions processing: ERP study. *Neuroreport* 15:621–625.
- Kok P, Brouwer GJ, van Gerven MAJ, de Lange FP (2013) Prior expectations bias sensory representations in visual cortex. *J Neurosci Off J Soc Neurosci* 33:16275–16284.
- Kok P, Jehee JFM, de Lange FP (2012) Less is more: expectation sharpens representations in the primary visual cortex. *Neuron* 75:265–270.
- Kosslyn SM, Cacioppo JT, Davidson RJ, Hugdahl K, Lovallo WR, Spiegel D, Rose R (2002) Bridging psychology and biology. The analysis of individuals in groups. *Am Psychol* 57:341–351.
- Krolak-Salmon P, Fischer C, Vighetto A, Mauguière F (2001) Processing of facial emotional expression: spatio-temporal data as assessed by scalp event-related potentials. *Eur J Neurosci* 13:987–994.
- Kuhn G, Tipples J (2011) Increased gaze following for fearful faces. It depends on what you’re looking for! *Psychon Bull Rev* 18:89–95.
- Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A (2010) Presentation and validation of the Radboud Faces Database. *Cogn Emot* 24:1377–1388.
- LeDoux J (2012) Rethinking the Emotional Brain. *Neuron* 73:653–676.
- Lieberman A, Fischer J, Whitney D (2014) Serial dependence in the perception of faces. *Curr Biol CB* 24:2569–2574.
- Lindquist KA, Wager TD, Kober H, Bliss-Moreau E, Barrett LF (2012) The brain basis of emotion: a meta-analytic review. *Behav Brain Sci* 35:121–143.
- Liu L, Ioannides AA (2010) Emotion separation is completed early and it depends on visual field presentation. *PLoS One* 5:e9790.
- Liu T, Pleskac TJ (2011) Neural correlates of evidence accumulation in a perceptual decision task. *J Neurophysiol* 106:2383–2398.
- Liu Y, Huang H, McGinnis-Deweese M, Keil A, Ding M (2012) Neural substrate of the late positive potential in emotional processing. *J Neurosci Off J Soc Neurosci* 32:14563–14572.
- Lobmaier JS, Tiddeman BP, Perrett DI (2008) Emotional expression modulates perceived gaze direction. *Emot Wash DC* 8:573–577.
- Lo LY, Cheng MY (2015) A quick eye to anger: An investigation of a differential effect of facial features in detecting angry and happy expressions. *Int J Psychol J Int Psychol*.
- Luo Q, Holroyd T, Jones M, Hendler T, Blair J (2007) Neural dynamics for facial threat processing as revealed by gamma band synchronization using MEG. *NeuroImage* 34:839–847.

- Machens CK, Gollisch T, Kolesnikova O, Herz AVM (2005) Testing the efficiency of sensory coding with optimal stimulus ensembles. *Neuron* 47:447–456.
- Mareschal I, Calder AJ, Clifford CWG (2013) Humans Have an Expectation That Gaze Is Directed Toward Them. *Curr Biol* 23:717–721.
- Mars RB, Debener S, Gladwin TE, Harrison LM, Haggard P, Rothwell JC, Bestmann S (2008) Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *J Neurosci* 28:12539–12545.
- Mathews A, Fox E, Yiend J, Calder A (2003) The face of fear: Effects of eye gaze and emotion on visual attention. *Vis Cogn* 10:823–835.
- McAdams CJ, Maunsell JH (1999) Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron* 23:765–773.
- McArthur LZ, Baron RM (1983) Toward an ecological theory of social perception. *Psychol Rev* 90:215–238.
- Meeren HKM, van Heijnsbergen CCRJ, de Gelder B (2005) Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A* 102:16518–16523.
- Mobbs D, Weiskopf N, Lau HC, Featherstone E, Dolan RJ, Frith CD (2006) The Kuleshov Effect: the influence of contextual framing on emotional attributions. *Soc Cogn Affect Neurosci* 1:95–106.
- Morin EL, Hadj-Bouziane F, Stokes M, Ungerleider LG, Bell AH (2014) Hierarchical Encoding of Social Cues in Primate Inferior Temporal Cortex. *Cereb Cortex N Y N* 1991.
- Müller VI, Habel U, Derntl B, Schneider F, Zilles K, Turetsky BI, Eickhoff SB (2011) Incongruence effects in crossmodal emotional integration. *NeuroImage* 54:2257–2266.
- Mumenthaler C, Sander D (2012) Social appraisal influences recognition of emotions. *J Pers Soc Psychol* 102:1118–1135.
- N'Diaye K, Sander D, Vuilleumier P (2009) Self-relevance processing in the human amygdala: gaze direction, facial expression, and emotion intensity. *Emot Wash DC* 9:798–806.
- Nesse RM (1990) Evolutionary explanations of emotions. *Hum Nat* 1:261–289.
- Nguyen VT, Breakspear M, Cunnington R (2014) Fusing concurrent EEG–fMRI with dynamic causal modeling: Application to effective connectivity during face perception. *NeuroImage* 102, Part 1:60–70.
- Nienborg H, Cumming B (2010) Correlations between the activity of sensory neurons and behavior: how much do they tell us about a neuron's causality? *Curr Opin Neurobiol* 20:376–381.
- Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92.
- Nieuwenhuis S, Aston-Jones G, Cohen JD (2005) Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychol Bull* 131:510–532.
- Oaksford M, Chater N (2007) *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. OUP Oxford.
- O'Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat Neurosci* 15:1729–1735.

- Ofan RH, Rubin N, Amodio DM (2011) Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing. *J Cogn Neurosci* 23:3153–3161.
- Ohman A (2005) The role of the amygdala in human fear: automatic detection of threat. *Psychoneuroendocrinology* 30:953–958.
- Ohman A, Lundqvist D, Esteves F (2001) The face in the crowd revisited: a threat advantage with schematic stimuli. *J Pers Soc Psychol* 80:381–396.
- Ohman A, Mineka S (2001) Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychol Rev* 108:483–522.
- Ortony A, Turner TJ (1990) What's basic about basic emotions? *Psychol Rev* 97:315–331.
- Palermo R, Rhodes G (2007) Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia* 45:75–92.
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.
- Panksepp J (2005) Affective consciousness: Core emotional feelings in animals and humans. *Conscious Cogn* 14:30–80.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Pelphrey KA, Sasson NJ, Reznick JS, Paul G, Goldman BD, Piven J (2002) Visual scanning of faces in autism. *J Autism Dev Disord* 32:249–261.
- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. *PLoS Comput Biol* 6:e1000709.
- Pessoa L, Padmala S (2005) Quantitative prediction of perceptual decisions during near-threshold fear detection. *Proc Natl Acad Sci U S A* 102:5612–5617.
- Pessoa L, Padmala S (2007) Decoding near-threshold perception of fear from distributed single-trial brain activation. *Cereb Cortex N Y N 1991* 17:691–701.
- Phelps EA, Ling S, Carrasco M (2006) Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychol Sci* 17:292–299.
- Phelps EA, O'Connor KJ, Cunningham WA, Funayama ES, Gatenby JC, Gore JC, Banaji MR (2000) Performance on indirect measures of race evaluation predicts amygdala activation. *J Cogn Neurosci* 12:729–738.
- Philiastides MG, Ratcliff R, Sajda P (2006) Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J Neurosci Off J Soc Neurosci* 26:8965–8975.
- Philiastides MG, Sajda P (2007) EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J Neurosci Off J Soc Neurosci* 27:13082–13091.
- Pichon S, de Gelder B, Grèzes J (2009) Two different faces of threat. Comparing the neural systems for recognizing fear and anger in dynamic body expressions. *NeuroImage* 47:1873–1883.
- Pichon S, de Gelder B, Grèzes J (2012) Threat prompts defensive brain responses independently of attentional control. *Cereb Cortex N Y N 1991* 22:274–285.

- Ploran EJ, Tremel JJ, Nelson SM, Wheeler ME (2011) High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex. *Cereb Cortex N Y N 1991* 21:2650–2662.
- Porter G, Troscianko T, Gilchrist ID (2007) Effort during visual search and counting: insights from pupillometry. *Q J Exp Psychol* 2006 60:211–229.
- Pourtois G, Grandjean D, Sander D, Vuilleumier P (2004) Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cereb Cortex N Y N 1991* 14:619–633.
- Preuschoff K, Hart BM 't, Einhauser W (2011) Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Decis Neurosci* 5:115.
- Privitera CM, Renninger LW, Carney T, Klein S, Aguilar M (2010) Pupil dilation during visual target detection. *J Vis* 10:3.
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci Off J Soc Neurosci* 18:2188–2199.
- Rahnev D, Lau H, Lange FP de (2011a) Prior Expectation Modulates the Interaction between Sensory and Prefrontal Regions in the Human Brain. *J Neurosci* 31:10741–10748.
- Rahnev D, Maniscalco B, Graves T, Huang E, de Lange FP, Lau H (2011b) Attention induces conservative subjective biases in visual perception. *Nat Neurosci* 14:1513–1515.
- Ratcliff R (1978) A theory of memory retrieval. *Psychol Rev* 85:59–108.
- Ratcliff R, Philiastides MG, Sajda P (2009) Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc Natl Acad Sci U S A* 106:6539–6544.
- Ratner KG, Dotsch R, Wigboldus DHJ, van Knippenberg A, Amodio DM (2014) Visualizing minimal ingroup and outgroup faces: implications for impressions, attitudes, and behavior. *J Pers Soc Psychol* 106:897–911.
- Rellecke J, Sommer W, Schacht A (2012) Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biol Psychol* 90:23–32.
- Reynolds JH, Pasternak T, Desimone R (2000) Attention increases sensitivity of V4 neurons. *Neuron* 26:703–714.
- Rhodes G, Addison B, Jeffery L, Ewbank M, Calder AJ (2012) Facial Expressions of Threat Influence Perceived Gaze Direction in 8 Year-Olds. *PLoS ONE* 7:e49317.
- Righart R, de Gelder B (2006) Context influences early perceptual analysis of faces--an electrophysiological study. *Cereb Cortex N Y N 1991* 16:1249–1257.
- Righart R, de Gelder B (2008a) Recognition of facial expressions is influenced by emotional scene gist. *Cogn Affect Behav Neurosci* 8:264–272.
- Righart R, de Gelder B (2008b) Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Soc Cogn Affect Neurosci* 3:270–278.
- Rizzolatti G, Fogassi L, Gallese V (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2:661–670.
- Romo R, Hernández A, Zainos A (2004) Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron* 41:165–173.

- Rossion B, Joyce CA, Cottrell GW, Tarr MJ (2003) Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage* 20:1609–1624.
- Rotshtein P, Henson RNA, Treves A, Driver J, Dolan RJ (2005) Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat Neurosci* 8:107–113.
- Russell JA, Fehr B (1987) Relativity in the perception of emotion in facial expressions. *J Exp Psychol Gen* 116:223–237.
- Said CP, Moore CD, Norman KA, Haxby JV, Todorov A (2010) Graded representations of emotional expressions in the left superior temporal sulcus. *Front Syst Neurosci* 4:6.
- Sander D, Grafman J, Zalla T (2003) The human amygdala: an evolved system for relevance detection. *Rev Neurosci* 14:303–316.
- Sander D, Grandjean D, Kaiser S, Wehrle T, Scherer KR (2007) Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *Eur J Cogn Psychol* 19:470–480.
- Sato W, Kochiyama T, Uono S, Yoshikawa S (2010) Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *NeuroImage* 50:1658–1665.
- Sato W, Yoshikawa S, Kochiyama T, Matsumura M (2004) The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *NeuroImage* 22:1006–1013.
- Scheller E, Büchel C, Gamer M (2012) Diagnostic features of emotional expressions are processed preferentially. *PloS One* 7:e41792.
- Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–317). Hillsdale, NJ: Erlbaum.
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9:671–675.
- Schupp HT, Ohman A, Junghöfer M, Weike AI, Stockburger J, Hamm AO (2004) The facilitated processing of threatening faces: an ERP analysis. *Emot Wash DC* 4:189–200.
- Schyns PG, Gosselin F, Smith ML (2009) Information processing algorithms in the brain. *Trends Cogn Sci* 13:20–26.
- Schyns PG, Petro LS, Smith ML (2007) Dynamics of visual information integration in the brain for categorizing facial expressions. *Curr Biol CB* 17:1580–1585.
- Senju A, Csibra G (2008) Gaze Following in Human Infants Depends on Communicative Signals. *Curr Biol* 18:668–671.
- Senju A, Hasegawa T (2005) Direct gaze captures visuospatial attention. *Vis Cogn* 12:127–144.
- Shadlen MN, Newsome WT (2001) Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey. *J Neurophysiol* 86:1916–1936.
- Sherman A, Sweeny TD, Grabowecy M, Suzuki S (2012) Laughter exaggerates happy and sad faces depending on visual context. *Psychon Bull Rev* 19:163–169.
- Siegel M, Buschman TJ, Miller EK (2015) Cortical information flow during flexible sensorimotor decisions. *Science* 348:1352–1355.

- Smith FW, Schyns PG (2009) Smile through your fear and sadness: transmitting and identifying facial expression signals over a range of viewing distances. *Psychol Sci* 20:1202–1208.
- Smith ML, Cottrell GW, Gosselin F, Schyns PG (2005) Transmitting and decoding facial expressions. *Psychol Sci* 16:184–189.
- Smith ML, Fries P, Gosselin F, Goebel R, Schyns PG (2009) Inverse Mapping the Neuronal Substrates of Face Categorizations. *Cereb Cortex*:bhn257.
- Spezio ML, Huang P-YS, Castelli F, Adolphs R (2007) Amygdala damage impairs eye contact during conversations with real people. *J Neurosci Off J Soc Neurosci* 27:3994–3997.
- Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press
- Sprengelmeyer R, Jentzsch I (2006) Event related potentials and the perception of intensity in facial expressions. *Neuropsychologia* 44:2899–2906.
- Summerfield C, Trittschuh EH, Monti JM, Mesulam M-M, Egner T (2008) Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 11:1004–1006.
- Sweeny TD, Suzuki S, Grabowecy M, Paller KA (2013) Detecting and categorizing fleeting emotions in faces. *Emot Wash DC* 13:76–91.
- Tanaka-Matsumi J, Attivissimo D, Nelson S, D'Urso T (1995) Context effects on the judgment of basic emotions in the face. *Motiv Emot* 19:139–155.
- Taylor JM, Whalen PJ (2014) Fearful, but not angry, expressions diffuse attention to peripheral targets in an attentional blink paradigm. *Emot Wash DC* 14:462–468.
- Thielscher A, Pessoa L (2007) Neural correlates of perceptual choice and decision making during fear-disgust discrimination. *J Neurosci Off J Soc Neurosci* 27:2908–2917.
- Todorov A, Gobbini MI, Evans KK, Haxby JV (2007) Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia* 45:163–173.
- Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P (2015) Social attributions from faces: determinants, consequences, accuracy, and functional significance. *Annu Rev Psychol* 66:519–545.
- Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C (2009) The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res* 168:242–249.
- Van Bavel JJ, Packer DJ, Cunningham WA (2008) The neural substrates of in-group bias. *Psychol Sci* 19:1131.
- Van Bavel JJ, Packer DJ, Cunningham WA (2011) Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *J Cogn Neurosci* 23:3343–3354.
- Van den Stock J, Righart R, de Gelder B (2007) Body expressions influence recognition of emotions in the face and voice. *Emot Wash DC* 7:487–494.
- Van den Stock J, Vandenbulcke M, Sinke CBA, Goebel R, de Gelder B (2014) How affective information from faces and scenes interacts in the brain. *Soc Cogn Affect Neurosci* 9:1481–1488.
- Van Strien JW, De Sonnevill LMJ, Franken IHA (2010) The late positive potential and explicit versus implicit processing of facial valence. *Neuroreport* 21:656–661.



- Vlamings PHJM, Goffaux V, Kemner C (2009) Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *J Vis* 9:12.1–13.
- Vuilleumier P, Armony JL, Driver J, Dolan RJ (2003) Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat Neurosci* 6:624–631.
- Vuilleumier P, Pourtois G (2007) Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia* 45:174–194.
- Wager TD, Kang J, Johnson TD, Nichols TE, Satpute AB, Barrett LF (2015) A Bayesian model of category-specific emotional brain responses. *PLoS Comput Biol* 11:e1004066.
- Watanabe S, Kakigi R, Miki K, Puce A (2006) Human MT/V5 activity on viewing eye gaze changes in others: A magnetoencephalographic study. *Brain Res* 1092:152–160.
- Watanabe S, Kakigi R, Puce A (2003) The spatiotemporal dynamics of the face inversion effect: A magneto- and electro-encephalographic study. *Neuroscience* 116:879–895.
- Wegrzyn M, Bruckhaus I, Kissler J (2015) Categorical Perception of Fear and Anger Expressions in Whole, Masked and Composite Faces. *PloS One* 10:e0134790.
- Wieser MJ, Brosch T (2012) Faces in context: a review and systematization of contextual influences on affective face processing. *Front Psychol* 3:471.
- Winston JS, O'Doherty J, Dolan RJ (2003) Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage* 20:84–97.
- Wyart V, de Gardelle V, Scholl J, Summerfield C (2012a) Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron* 76:847–858.
- Wyart V, Myers NE, Summerfield C (2015) Neural mechanisms of human perceptual choice under focused and divided attention. *J Neurosci Off J Soc Neurosci* 35:3485–3498.
- Wyart V, Nobre AC, Summerfield C (2012b) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A* 109:3593–3598.
- Yang DY-J, Rosenblau G, Keifer C, Pelphrey KA (2015) An integrative neural model of social perception, action observation, and theory of mind. *Neurosci Biobehav Rev* 51:263–275.
- Yarbus A L. (1967) *Eye Movements and Vision*. New York: Plenum Press.
- Zebrowitz LA, Montepare JM (2008) Social Psychological Face Perception: Why Appearance Matters. *Soc Personal Psychol Compass* 2:1497.
- Zebrowitz, L.A., & Montepare, J.M. (2006). The ecological approach to person perception: Evolutionary roots and contemporary offshoots. In M. Schaller, J.A. Simpson, & D.T. Kenrick (Eds.) *Evolution and Social Psychology*, (pp. 81-113). New York: Psychology Press