# In the face of consciousness:

## How emotion, orientation, and gaze modulate face perception


Renzo Lanfranco

Department of Psychology

School of Philosophy, Psychology, and Language Sciences


Thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh

2020

# In the face of consciousness:

## How emotion, orientation, and gaze modulate face perception

# Declaration

I declare that this thesis presented for the degree of Doctor of Philosophy (Ph.D.) has been composed solely by myself and that it has not been submitted in whole or in part for any other degree or professional qualification. Except where states otherwise by reference or acknowledgement, the work presented is entirely my own.

Renzo C lo Lanfranco guevara

If the doors of perception were cleansed,

everything would appear to man as it is: Infinite.

– William Blake, *The Marriage of Heaven and Hell*

# ABSTRACT

Human faces convey essential information for social behaviour, such as information about others' mental states and intentions. Crucially, many studies have claimed that several facial features such as configural facial information, emotional expressions, and gaze direction modulate how faces gain access to perceptual awareness. However, the procedures employed in said studies suffer from multiple methodological issues and limitations.

In a series of experiments, I tested whether configural facial features, emotional expressions, and gaze direction modulate how faces gain access to awareness. To achieve this, I used stringent procedures that allow measurement of perceptual sensitivity and decision criterion to the location and identity of faces. I used these measures to assess how long it takes faces to reach awareness as they overcome Continuous Flash Suppression – an interocular suppression technique that can render images invisible for several seconds. Using classical and Bayesian analyses, I found that configural face processing (which occurs for upright, but not inverted faces) promotes faces' access to awareness. Similarly, faces making eye contact gain access to awareness faster than faces looking away. Contrary to past claims, however, I found that faces expressing negative emotional expressions (anger or fear) do not enter awareness faster than neutral expressions.

In another series of experiments, I measured the minimal exposure durations required for configural facial processing, emotion processing, metacognition, and conscious access. To this end, I used a newly developed LCD tachistoscope that can present images with sub-millisecond precision and examined both behavioural (psychophysical) and neural (electroencephalography) markers of processing. I found that configural face processing promotes faces' access to awareness by showing that upright faces require shorter exposure durations than inverted faces to be seen. Crucially, only around four milliseconds of exposure were required to find this advantage. Fearful expressions, however, do not gain access to awareness faster than neutral expressions. Evidence from neural markers expanded this by showing that the exposure duration required for

configural facial processing is the same as that required for faces to reach conscious access. Finally, around six milliseconds of exposure were required for emotion processing.

Together, these findings shed light on the factors that affect access of faces to awareness: configural facial information and gaze direction can modulate faces' access to perceptual awareness; and such modulation is due to perceptual sensitivity rather than decision criterion. Furthermore, the perceptual processing of faces follows a hierarchical pattern: configural information precedes and facilitates access to awareness, and emotion processing follows awareness.

# LAY SUMMARY

Faces convey a wealth of information that is crucial for social interaction. By looking at someone's face we can learn about how they feel, what they want, and who they are. Scientists have studied how face perception occurs and how the brain became so specialised at processing faces so efficiently. But do all aspects of face perception occur consciously? Or are there aspects that do not require it to occur? Using different approaches, many scientists have claimed that a great number of facial features can be processed unconsciously, including their emotional expressions, gaze direction, and their whole configural organisation.

Most studies exploring visual perception and awareness use methods that allow them to show people images while suppressing those images from observers' awareness. These are collectively known as 'masking' techniques. They create the conditions to investigate the way that many aspects of visual processing can occur in the absence of awareness. These studies, however, are rather controversial because the methodologies they used have yet to be refined and their results do not always replicate.

In this thesis, I present fourteen experiments that study face perception and explore whether face configuration, gaze direction, and emotional expression modulate faces' access to awareness, and whether face configuration, perceptual awareness, and emotion processing occur in sequence or all at once. To do this, I employed two novel procedures, in which I always controlled how long people were exposed to images. In the first of these, I used a masking technique and examined whether a previous claim – that faces making eye contact broke through this suppression into awareness faster than faces looking away – holds up even when controlling how long participants look at the image. My findings confirmed this older claim, suggesting that processing of gaze occurs prior to – and facilitates access to – perceptual awareness.

Having this method validated, I then examined whether another previous claim about emotion holds up – that faces expressing negative emotions broke through suppression into awareness faster than faces expressing positive or no emotion. My findings cast doubt on these claims, suggesting that emotional expression cannot be processed in the absence of awareness.

Are facial features processed in a sequence of steps or all at once? The second novel procedure I used is a newly developed display system that is capable of extremely brief presentations. This allowed me to measure the minimal exposure duration required to perceive various aspects of faces, both upright and inverted. I found that it took longer for people to identify faces' emotional expressions than to integrate their isolated parts, and that this integration took longer than to discriminate intact faces from scrambled faces. This demonstrates that faces are processed in a sequence of steps.

However, it could be the case that these aspects of faces are processed simultaneously in the brain but transmitted to other areas at different speeds, resulting in the sequence just described. To explore this possibility, I measured the minimal exposure durations required for neural markers to engage with face processing, emotion processing, and perceptual awareness. My findings corroborate that face perception occurs in a sequence of steps and also suggest that the ability to process a face as an integrated configuration arises alongside perceptual awareness.

This work demonstrates that while some facial features may be processed in the absence of awareness, other features cannot, thus casting doubts on past claims that implied that perceptual awareness was not required to process a face. Furthermore, this work shows that face perception occurs in a sequence of steps, and that the ability to process a face as an integrated configuration arises alongside awareness, thus suggesting that perceptual awareness and visual integration may be related.

# ACKNOWLEDGEMENTS

I would first like to thank my supervisors David Carmel and Hugh Rabagliati for their vital and meticulous guidance, and for encouraging me to develop my own ideas when they made sense, and for advising me against it when they did not.

I would like to thank Axel Cleeremans for his kindness and generosity during my research stay in his lab at the Université Libre de Bruxelles in Belgium; for welcoming me to develop a series of experiments using his very cool LCD tachistoscope, what has become no doubt the most exciting part of my PhD. I would also like to thank Dalila for teaching me the essentials of the tachistoscope, and Santiago, Laurène, and Arnaud for giving me a hand whenever I needed one while at ULB.

I want to thank my dear friend Andrés Canales-Johnson for his support and friendship, and especially for his crucial suggestions and guidance on EEG analysis. I would also like to thank Tristán Bekinschtein for receiving me in their lab at the University of Cambridge every time I was in town, and for inviting me to present my findings.

I would like to thank Eugenio Rodríguez, Tomás Ossandón (Catholic University), and David Huepe (Adolfo Ibáñez University) for inviting me to present my research.

I would also like to thank my friends and office mates from the University of Edinburgh: Federica, Verena, Carlos, Thalia, Alistair, Emma, Qingyuan, Esperanza, Anna, Riccardo, Bill, Aurélien, Melissa, Jo, and Ilaria, among others.

I would also like to thank Alisdair Tullo, the department's programmer, whose support on coding was extremely helpful during the first year of my PhD.

I want to thank everyone who supported me years ago when I decided to leave clinical psychology for cognitive science, especially my former boss Eduardo Correa.

Last but not least, I would like to thank my family; my parents Aldo Lanfranco and Liliana Guevara, and my brother Giorgio. Without my family, I would not be here. The three of you are in everything I have achieved or will achieve. I love you.

Thanks a million!

This work was supported by a 4-year CONICYT PhD studentship, a Postgraduate Study Visit Scheme Grant awarded by the British Psychological Society, a Principal's Go Abroad Fund awarded by the University of Edinburgh, and multiple Research Support Grants awarded by the University of Edinburgh School of Philosophy, Psychology, and Language Sciences.

# Table of Contents

# LIST OF ABBREVIATIONS

| | |
|---|---|
| 2AFC | Two-Alternative Forced Choice |
| 2IFC | Two-Interval Forced Choice |
| ANOVA | Analysis of Variance |
| ASD | Autistic Spectrum Disorder |
| AU | Arbitrary Units |
| AUC | Area Under a Curve |
| BF | Bayes Factor |
| b-CFS | Breaking Continuous Flash Suppression |
| CFS | Continuous Flash Suppression |
| EEG | Electroencephalography |
| EMG | Electromyography |
| ERP | Event-related potential |
| EPN | Early Posterior Negativity |
| FFA | Face Fusiform Area |
| FIE | Face-Inversion Effect |
| fMRI | functional magnetic resonance imaging |
| GWT | Global Workspace Theory |
| ICA | Independent-Component Analysis |
| ITI | Intertrial Interval |
| KDEF | Karolinska Directed Emotional Faces |
| LFP | Local Field Potential |

| | |
|---|---|
| LP | Late Positivity |
| LPP | Late Positive Potential |
| M | Mean |
| MEG | Magnetoencephalography |
| NCC | Neural Correlate of Consciousness |
| PAS | Perceptual Awareness Scale |
| RaFD | Radboud Faces Database |
| RT | Response Time |
| SCR | Skin Conductance Response |
| SSVEP | Steady-State Visually Evoked Potential |
| SD | Standard Deviation |
| SDT | Signal Detection Theory |
| SEM | Standard Error of the Mean |
| STS | Superior Temporal Sulcus |
| TD | Typically Developed |
| VAN | Visual Awareness Negativity |
| VPP | Vertex Positive Potential |
| wPLI | weighted Phase Lag Index |
| wSMI | weighted Symbolic Mutual Information |

# Chapter 1

## 1 GENERAL INTRODUCTION

Human faces convey a wealth of crucial information that we use to guide our social behaviour, such as information about others' mental states and intentions (Grill-Spector et al., 2017; Jack & Schyns, 2015; Little et al., 2011). They are remarkably effective at capturing attention (Fox, 2002; Langton et al., 2008; Wilson & MacLeod, 2003), especially when expressing emotional states. For example, fearful and angry expressions are detected faster than neutral and happy expressions (Fox et al., 2000; Hansen & Hansen, 1988; Krysko & Rutherford, 2009). Another facial feature that is effective at capturing attention is gaze. For example, faces making eye contact draw attention toward the face whereas averted gaze draws attention to the gaze's direction (Dupierrix et al., 2014; Farroni et al., 2002; Hood et al., 1998; Senju & Hasegawa, 2005). As described below, face processing can also be affected by certain psychiatric disorders. For instance, depression enhances salience of sad expressions (Lazarov et al., 2018), anxiety enhances salience of fearful and angry expressions (Bishop et al., 2007; Mogg et al., 2007), and Autism affects face processing by inducing avoidance of eye contact (Senju & Johnson, 2009b). Therefore, understanding how facial information is processed and integrated can shed light on the cognitive mechanisms of perception and emotion both in healthy and psychopathological conditions. A myriad of studies has explored how facial information is processed in the visual system and what mechanisms are involved, including what aspects of a face are prioritised when gaining access to perceptual awareness, one of the main topics of this thesis. Do emotional expressions reach awareness faster than non-emotional ones? Are faces prioritised because of their socially relevant configural organisation? Do faces making eye contact reach awareness faster than faces looking away? Are the neural mechanisms of face processing engaged before faces gain access to awareness? As I discuss below, multiple claims have been made about how facial information gains access to awareness. However, some of those findings are inconsistent, and some have failed to replicate, among other issues addressed below.

Consciousness (or awareness[1]) has been defined in many different ways, but it ultimately refers to the ability to have a subjective experience of our surroundings and inner thoughts. Even though its nature is often described as private and qualitative – and therefore subjective – consciousness researchers agree on a key distinction for its study: On the one hand, consciousness can be thought of in terms of *state*, defined by the ability to be aware and able to respond to external stimuli (arousal level), as when we are awake (Hobson, 2007); or in pathological states, such as vegetative state, coma, and minimally conscious state, in which consciousness is thought to be partially or totally absent. Additionally, the states of consciousness normally exhibit distinct neurophysiological correlates (Bekinschtein et al., 2009; Boly et al., 2008; Cruse et al., 2011; Goupil & Bekinschtein, 2012; Monti et al., 2010; Noreika et al., 2019; Stevner et al., 2019). On the other hand, it is often thought of in terms of *contents*, defined as the experience associated with perceiving, feeling, thinking, and acting, often described as qualitative and irreducible by philosophers (Searle, 1992; Seth et al., 2008). In fact, this private and qualitative nature of subjective experience has led many thinkers of the past, like Gómez Pereira (1554) and René Descartes (1641), to support the idea that consciousness – as an apparent immaterial phenomenon – is substantially distinct from the brain, an ontological stance called mind-body dualism. This philosophical debate is probably far from resolved.

Because of its subjective nature, consciousness was not considered a valid object of scientific study until just a few decades ago. Consciousness, as a phenomenon, did not seem necessary to explain the mechanisms behind cognition and behaviour either – what does consciousness add to an explanation of how the brain works? Behaviourists such as John Broadus Watson and Burrhus Frederic Skinner, for instance, argued that private experiences – perhaps real, perhaps illusory – were irrelevant to understanding behaviour (Lashley, 1923; Pratt, 1922; Skinner, 1974; Watson, 1913). Even cognitive psychologists, like George Miller (1962), once argued that consciousness, as a concept, obscured more precise descriptions and explanations of cognition (Hilgard, 1980). This negative attitude towards consciousness as an object of scientific study may have also been partially driven

[1] In this thesis, I use the concepts 'consciousness' and 'awareness' interchangeably, referring to the ability to become aware (e.g. of face images). However, these concepts are rather nuanced in the literature: for example, 'consciousness' may refer to physiological states of consciousness whereas 'awareness' may refer to subjective experience as in visual perception, interoception, and hallucination.

by esoteric claims found in psychoanalytic theories (e.g. Freud, 1915, 1927; Lacan, 1970), which despite being circular, fallacious, and oftentimes empirically untestable – hence pseudoscientific (Bunge, 1991; Cioffi, 1985, 1998; Kuhn, 1977; Popper, 1963) – gained popularity throughout the century.

But how can we scientifically address a question about the contents of consciousness, like how faces gain access to it? William James, by integrating first-person (phenomenology) and third-person approaches (empirical study), famously distinguished consciousness from attention and thought, defending it as a valid phenomenon for scientific study a century ago (James, 1890). However, it was not until the end of the twentieth century that consciousness gained full acceptance in science (Block et al., 2014; Dehaene et al., 2006). A philosophical distinction that was crucial for this, made by David Chalmers, is that consciousness involves two kinds of problems: a hard problem and a set of easy problems. The hard problem refers to why and how we have subjective qualitative experiences (or qualia), and why physical brain processes are accompanied by experience. The easy problems refer to the functions of consciousness and their neural correlates (Chalmers, 1997). While the hard problem can be qualified as a modern summary of the philosophical enquiries of the past, and therefore might escape the scope of what science today can empirically address, the easy problems are empirically addressable. We can ask people to provide subjective reports about their experience in its simplest forms – e.g. whether they saw a light flickering or not – and measure their perceptual discrimination together with their neural correlates in relation to a condition where they were not subjectively aware of the stimulus of interest.

Of course, subjective reports are limited. While their subjective nature makes them essential for the development of other disciplines like phenomenology (Husserl, 1913; Merleau-Ponty, 1945), subjective reports alone (like in introspection) are not accepted as a valid method for scientific research – because of people's biases and introspection's limited reach, we simply cannot learn in a rigorous fashion about the brain mechanisms of consciousness by asking people about their own experience (Hixon & Swann, 1993; Nisbett & Wilson, 1977; Wilson & Brekke, 1994; Wilson & Dunn, 2004). Nevertheless, subjective reports as a source of raw data are accepted and unavoidable when studying the contents of consciousness. In fact, we need subjective reports as indices of awareness to look for objective behavioural and neurophysiological patterns

(Corallo et al., 2008; Dehaene, 2014; Fleming et al., 2010; Marti et al., 2010; McGovern & Baars, 2007; Seth et al., 2008) and to verify that a sensory stimulus is being processed while making manipulations that will lead to changes in awareness. In conclusion, subjective reports are essential to determine changes in the contents of consciousness and are usually accompanied by objective measures that can shed light on the role of awareness in cognition.

As mentioned at the beginning, faces can convey a great deal of information. Due to their richness, they are considered complex stimuli and there is a large body of research that has shown that faces go through an intricate processing hierarchy, where their configural features, emotional content, attributed intentions, personality traits, and social characteristics are extracted. But what is the processing hierarchy of facial features in their access to awareness? In the following section, I will describe how this question has been addressed in the past, what methods have been employed, and what limitations these entail. First, I will describe what methods are commonly used when addressing questions about unconscious visual processing. Secondly, I will summarise some of the findings that have been obtained with these methods when studying how faces and different facial features are perceived and gain access to conscious awareness. Finally, I will discuss the main issues and limitations that these methods involve and how they are addressed and circumvented in the experimental chapters of this thesis.

## 1.1 The study of awareness in visual perception

Visual perception is the ability to process, interpret, and become aware of visual information. Abilities such as visual detection, discrimination, and recognition, are examples of visual processing involved in visual perception. The last bit of this definition – awareness – refers to the subjective experience (or content) contained in the act of perceiving.

To answer how facial information gains access to awareness, we must first review what approaches have been taken by researchers studying visual perception and specifically awareness, including what limitations the field presents, and how we can circumvent them. As mentioned above, visual perception involves interpreting information. Those interpretations will reach awareness, but some types of visual information may enjoy priority over others when gaining access to awareness. Because simply impoverishing visual stimuli to render them faint or invisible (i.e. subjectively undetectable) will very likely eliminate any possible visual processing as a consequence, different strategies have been developed with the objective of testing what types of visual information are prioritised when entering awareness.

In this section, I will describe a series of approaches that have been employed to this end. Understanding the contributions and limitations of these approaches is essential for understanding the contribution of this thesis in the study of face processing and perceptual awareness.

## 1.2 Methods in visual processing and awareness

Multiple methods have been employed to study visual processing and awareness. These methods mainly include using brief exposures, masked stimuli presentations, and Continuous Flash Suppression (an interocular suppression technique). Both presenting stimuli for brief exposure durations and employing masking techniques to interrupt stimulus processing are widely used in the field of consciousness research. Both often involve estimating whether an objective task (e.g. stimulus detection or classification) can be performed when observers report absence of awareness of a given stimulus, thereby dissociating awareness from other visual processes (Baars, 1993; Crick & Koch, 1990, 1998; Erdelyi, 1986; Schmidt & Vorberg, 2006). Isolating awareness from other visual processes – also known as the dissociation paradigm (Augusto, 2016; Snodgrass, 2004) – provides an opportunity to investigate the dynamics of conscious and unconscious processing (Kim & Blake, 2005), including whether different stimulus features enter awareness faster than others. In this section, I provide a brief description of these methods, and thus of the logic that underlies the study of visual awareness. The limitations of these methods are addressed in section 1.4.

### 1.2.1 Brief exposure durations

Probably the simplest method there is to study how visual stimuli gain access to awareness is to present stimuli for predefined exposure durations and measure what visual features enjoy an advantage over others. By presenting stimuli for very brief exposure durations, observers can be asked to perform a detection or recognition task. For example, Intraub (1981) presented participants with sequences of unmasked pictures at rates of 114, 172, or 258 ms of exposure per picture. The target was specified by name (e.g. giraffe), by superordinate category (e.g. animal), or by negative category (e.g. the picture that is not food). Participants were instructed to press a key as soon as they saw the cued picture and to describe it briefly. They were able to detect the pictures even in the most difficult condition, when the exposure duration was the shortest – where they obtained 71% correct responses. If observers cannot consciously report stimuli presented for very brief presentation durations but can exhibit above-chance performance in an objective task (e.g. detection, recognition), this performance can be interpreted as evidence of unconscious processing. Therefore, for example, measuring neural markers during such a task can provide data about the neural correlates of consciousness. For instance, Thorpe et al. (1996) presented observers with pictures for 20 ms of exposure and asked them to perform a categorisation task (to indicate whether the picture contained an animal or not) while measuring electroencephalography (EEG). They found that the average proportion of correct responses was around 94% with just 20 ms of exposure, with an ERP frontal negativity developing roughly 150 ms after stimulus onset. Later on, Fabre-Thorpe et al. (2001) showed that both novel and highly familiar scenes could be categorised in the same 150 ms period just mentioned, and with only 20 ms of exposure, thus suggesting that the visual system does not decrease in speed when the stimuli are unknown by the observer. In a different study but following the same logic, Joubert et al. (2007) asked observers to categorise unmasked natural and manmade environments presented for 26 ms of exposure. They found that 26 ms of exposure were sufficient for participants to obtain around 96% accuracy for both categories. More recently, Liu & Tanaka (2019) presented observers with faces by using predefined exposure durations of 17, 50, 250, and 500 ms. They found that even 17 ms of exposure were sufficient to produce holistic face processing. In summary, by using predefined exposure durations to measure the minimal visual exposure that can elicit above-chance performance, conclusions can be drawn about a variety of cognitive functions, including perceptual awareness.

While this procedure is very straightforward and the results it yields are oftentimes easy to interpret, it has several limitations, which will be addressed later in section 1.4.1. along with additional studies that exemplify this approach and its limitations further.

### 1.2.2 Masking visual stimuli

Technically, it is extremely difficult for experimenters to present stimuli at briefer exposures than 16 ms, because most studies over the last few decades have presented stimuli on computer monitors, which (at least until recently) have typically had a refresh rate of around 60 Hz. Even newer monitors that have a refresh rate of more than 60 Hz, such as monitors with 100 Hz, 120 Hz, and 144 Hz cannot present stimuli at briefer exposures than 10, 8.33, and 6.94 ms, respectively. Therefore, subsequent attempts to determine the minimal exposure duration required for visual perception had to employ masking techniques to interrupt visual processing. Masking techniques allow experimenters to reduce a target stimulus' visibility – and hence observers' awareness – by presenting another image (a mask) very close in time. For example, Greene & Oliva (2009) presented observers with landscape images and asked them to perform basic-level categorisation (i.e. identifying objects in the image) and global-property classification tasks (i.e. identifying what kind of landscape was shown). Importantly, these images were suppressed from awareness by presenting a mask image right after the target stimuli were presented (backward masking; Breitmeyer & Ogmen, 2000; Kim & Blake, 2005). They found that a threshold of 75% correct responses on both tasks was achieved using exposure durations ranging from 19 to 67 ms, with high variation among participants. Shorter exposure durations were found for detection of scenes' global properties (i.e. identifying what kind of landscape is shown) than scenes' basic properties (i.e. identifying objects in the image).

As mentioned above, one key feature of masking techniques such as backward masking is that they interrupt visual processing. For example, Codispoti et al. (2009) presented observers with pleasant, neutral, and unpleasant pictures, masked and unmasked, across exposure durations ranging from 25 to 6000 ms, and measured their emotional reactivity using various techniques such as electromyography, EEG, skin conductance, and both pleasure and arousal ratings. Crucially, when masked pictures were employed, they found no evidence of emotional engagement with any measure, at any exposure duration under 80 ms, but when unmasked pictures were employed, they found evidence of emotion processing at all exposure durations. This indicates that exposure durations around 25 ms and longer are sufficiently long to reveal visual processing of scenes and faces and that, therefore, may be significantly longer than the real minimal exposure durations required for those processes to occur. However, it is still a matter of debate what specific aspects of visual processing are interrupted by masking techniques (e.g. feedforward processing, feedback processing, afterimage processing, lateral inhibition, visual cortex integration), and therefore, whether conclusions drawn from studies employing one particular masking technique can be extended to studies employing another masking technique or none. Therefore, using masking techniques may add potential confounding variables (Agaoglu et al., 2015; Breitmeyer & Öğmen, 2006; Davis & Kim, 2011). I return to this method in section 1.4.2, where I discuss critiques and limitations.

Finally, studies employing standard masking techniques may also be more vulnerable to post hoc data selection. Oftentimes, masking studies condition their analyses on a subjective report in order to test whether participants can show above-chance performance in a task while reporting no awareness of it. As shown by Vadillo et al. (2016) and Shanks (2017), performing post hoc data selection, to group trials where participants reported having no awareness of a stimulus, increases the risk of false negatives. When data from a group of participants are collected, measurement errors cancel out and the aggregate means approximate the mean true score. However, this assumption does not apply to a data subgroup selected post hoc because of regression to the mean – unaware participants may exhibit above-chance performance scores, and thus suggest unconscious cognitive processing, because of a statistical artefact.

### 1.2.3 Continuous Flash Suppression

When a different stimulus is presented to each eye at corresponding retinal locations, the observer experiences perceptual alternation between the two stimuli instead of simply experiencing two incompatibles halves or a fused image combining them together. This phenomenon is known as binocular rivalry (Blake & Logothetis, 2002). Furthermore, when one eye is presented with a constantly moving visual pattern while the other eye is presented with an often stationary and less salient image, the image is suppressed from awareness for a long period of time (usually several seconds). This phenomenon of interocular suppression has led to the development of a very important technique in the study of perceptual awareness and face processing: Continuous Flash Suppression (CFS; Tsuchiya & Koch, 2005). In CFS, one eye is flashed with a montage of different sized and coloured rectangles or circles whose luminance and locations vary randomly over time. These so-called Mondrian-like patterns are typically updated at a rate of 10 Hz and create a strong interocular suppression effect on the other eye, where the target image is introduced. Thus, CFS, unlike other masking techniques, can render stimuli invisible for up to several seconds.

There are several paradigms for the study of facial information and awareness that employ CFS (Yang et al., 2014), but the most widely used and fruitful one is Breaking Continuous Flash Suppression (Gayet et al., 2014; Stein, Hebart, et al., 2011), on which this section focuses. This paradigm is built on the assumption that stimulus categories that overcome suppression faster (thus gaining access to awareness) enjoy prioritised processing outside of awareness compared with stimuli that take longer time to overcome suppression. In the first published study using b-CFS, Jiang et al. (2007) presented either upright or inverted face images to one of the observers' eyes while their other eye was flashed with CFS masks. They asked observers to report the location of the face (left or right) as soon as they were able to see it. Upright faces elicited faster response times than inverted faces, thus suggesting that upright faces reached awareness faster due to an advantage driven by their visual configuration. This effect, known as face-inversion effect, has been replicated many times using b-CFS (Akechi et al., 2015; Gayet & Stein, 2017; Gray et al., 2013; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Hebart, et al., 2011; Stein, Senju, et al., 2011; Yang et al., 2007). It is important to note that this face-inversion effect is an instantiation of an effect found in face recognition, where upright faces are easier to recognise than inverted faces, as discussed below. Since the report by Jiang et al. (2007), b-CFS studies have adopted various tasks to estimate breakthrough times, such as stimulus detection and localisation.

The b-CFS paradigm presents one crucial advantage over other masking techniques: it provides breakthrough times, which allows researchers to estimate how long a stimulus takes to gain access to awareness. Crucially, breakthrough times can inform about two different aspects of awareness: when a stimulus gains access to awareness and whether it gains access before another stimulus of interest (or a control stimulus) does. This makes b-CFS an attractive choice when the purpose of the researcher is to test for prioritised access to awareness between different stimulus categories. Using one measure to estimate these two aspects of awareness makes b-CFS experiments immune to the methodological issues due to post hoc data selection described above. Having only one measure, as in the b-CFS procedure, prevents this problem from happening. The b-CFS procedure has other advantages, including a straightforward implementation and results that are easy to interpret. Around half of CFS studies have adopted the b-CFS task (Figure 1).

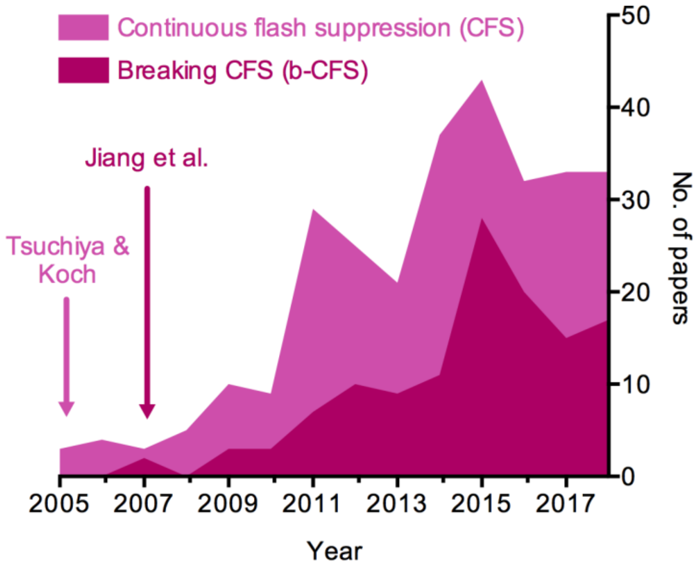



**Figure 1**. Number of publications per year using CFS and b-CFS from 2005 until August 2018. Arrows mark the year when Tsuchiya & Koch (2005) introduced the CFS procedure

and when Jiang et al. (2007) introduced the b-CFS variant. Reproduced from Stein (2019) with permission.

## 1.3 The study of awareness in face perception

Myriad studies have used the methods described above to explore how faces and their different features gain access to awareness in perception. Here, I review some of the most relevant studies that have been published on the matter, with a special focus on studies that have employed brief exposure durations, backward masking, and interocular suppression, as they are the most relevant ones for this thesis. Most of the interocular suppression studies employed the b-CFS procedure. As a consequence, most of the interocular suppression studies described below are b-CFS studies, too.

### 1.3.1 Facial configuration and the face-inversion effect

In order to perceive faces, we need to see beyond their isolated parts. We need to integrate their parts into coherent wholes. This visual integration receives the name of configural or holistic processing (Piepers & Robbins, 2012; Bruno Rossion, 2013). Faces can convey a large amount of high-level information such as emotional states, intentions, identity, gender, age, race, ethnicity, health, attractiveness, and personality traits, i.e. semantic or conceptual information contained in the integration of visual features.

One attempt to answer whether faces may enjoy prioritised access to awareness was performed by Richler et al. (2009), who explored holistic face processing using brief presentations. They presented participants with pairs of composite faces (a study face and a test face) that were made by combining the top half of a face with either its bottom half or another faces' bottom half. Both faces were presented in sequence and participants had to report whether a cued part of the test face was the same as of the study face. Because holistic processing should involve processing the face wholly and not locally, holistic processing was inferred from a congruency effect, i.e. better discrimination sensitivity in trials where uncued part in study face and cued part in test face were the same than in trials where they were different. Both faces were backward-masked and presented for different exposure durations, ranging from 17 ms to 800 ms. The researchers found that above-chance discrimination performance arose by 50 ms of exposure and suggested that holistic processing therefore emerges with very briefly presented faces.

Nowadays, the preferred method to study holistic processing of faces, however, has been b-CFS as it grants better suppression control than backward masking. CFS masking allows researchers to suppress complex visual stimuli from awareness for as long as a few seconds and provides with a simple measure of awareness access – a breakthrough time indexed by a response time, usually in the context of a detection or localisation task.

A classic and elegant strategy to study the holistic or configural processing of faces (i.e. how facial features are perceived as an integrated whole) is turning them upside down (Tanaka & Gordon, 2011). Upright faces are much easier to recognise than inverted faces. As mentioned above, this phenomenon is known as the face-inversion effect (Farah et al., 1995; Goodrich & Yonelinas, 2019; Yin, 1969). The evidence showing that this effect relies on high-level information is convergent with other, independent lines of evidence: for instance, Valentine & Bruce (1986) found that the inversion effect disrupts the recognition of faces much more than the recognition of houses. Furthermore, Searcy & Bartlett (1996) changed local elements of faces in order to make them look grotesque (e.g. blackening teeth, blurring the pupils) and then asked observers to judge their appearance in both orientations. Normal and grotesque faces were judged as different when presented upright but similar when presented inverted, a finding that was subsequently supported and extended (Leder & Bruce, 1998; Leder & Bruce, 2000). Complementing this, other studies have claimed that the fusiform face area (FFA), a key neural structure for face perception, is significantly more sensitive to upright faces than inverted faces (Yovel & Kanwisher, 2004), thus showing that such preference for upright faces – and therefore for faces with their configural characteristics intact – can be found in a very specialised neural structure. This finding was expanded by Baroni et al. (2017), who confirmed the role of the fusiform gyrus in the face-inversion effect. They presented observers with masked and unmasked images of upright and inverted faces and measured local field potentials intracranially using electrocorticographic electrodes. By using decoding analysis, they found that both ventral and lateral temporal cortices could reliably differentiate between seen and unseen faces and, crucially, that ventral electrodes located in the fusiform gyrus could reliably discriminate between upright and inverted faces.

However, it is important to note that not all configural features contained in faces are processed in the same fashion. Researchers have distinguished between first-order relational information and second-order information. First-order relational information is crucial for the successful identification of a face as a face and not as an object. It consists of the spatial relationships between facial features (e.g. eyes, nose, mouth, eyebrows). The face-inversion effect (described above) disrupts first-order information. On the other hand, second-order relational information is crucial to distinguish between different faces. It consists of the size of the spatial relationships between facial features. Whether this type of facial information can be disrupted by turning faces upside down is still a matter of debate (Civile et al., 2014; Tanaka & Farah, 1993; for a review, see: Tanaka & Simonyi, 2016). For instance, Rakover & Teucher (1997) demonstrated that isolated facial features such as foreheads, eyes, nose, mouth, and chin can exhibit inversion effects on their own, a finding that was subsequently expanded by Civile, McLaren, & McLaren (2014), who showed that the face-inversion effect can be generated by manipulating local features, too.

The fact that presenting faces upside down disrupts their holistic processing has been used in consciousness research to determine whether configural face processing makes faces gain access to awareness faster, by testing whether upright faces break through CFS masking faster than inverted faces (Axelrod et al., 2015). The first b-CFS study that addressed this question was done by Jiang, Costello, & He (2007). They presented participants with CFS-suppressed face images in an upright or inverted orientation. Face images were introduced to the left or right side of the screen and ramped up gradually from 0 to 100% contrast within a period of 1 second. Then, their contrast remained constant until the participant reported its location (left or right). Crucially, the researchers found that upright faces took, on average, about 400 ms less than inverted faces to overcome suppression, thus exhibiting a prioritised access to awareness. Because upright faces have their holistic information intact, it has been interpreted that said prioritised access is driven by holistic processing.

Many studies have replicated the face-inversion effect employing the b-CFS procedure (Akechi et al., 2015; Gayet & Stein, 2017; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Hebart, et al., 2011; Zhou et al., 2010), which appears to be particularly strong for faces in comparison to other visual stimuli. Zhou et al. (2010), for example, replicated the inversion effect for faces using a very similar b-CFS procedure. Then, they changed the face stimuli for upright and inverted house images. They did not find an inversion effect for houses, suggesting that other stimulus categories like houses do not contain the same type of high-level information and therefore may not be processed holistically, or at least not to the same extent of faces. As a consequence, houses may not be prioritised by holistic processing in their access to awareness. More recently, Kobylka et al. (2017) showed that the face-inversion effect can be found both when using localisation and categorisation b-CFS tasks. However, other reports have claimed that basic visual properties in faces, such as convexity, might explain this advantage of upright faces over inverted faces in b-CFS tasks (Moors, Wagemans, & de-Wit, 2016; Yang & Yeh, 2018a). Furthermore, other complex non-face stimuli such as human bodies have also been reported to exhibit a significant but smaller inversion effect (Stein et al., 2012). Thus, the evidence suggests that the inversion effects found with b-CFS tasks, such as the face-inversion effect, may be driven by the complexity and specificity of their configurations.

The advantage of upright faces over inverted faces has also been found in categorisation tasks. Sterzer, Jalkanen, & Rees (2009) used magnetoencephalography (MEG) to test whether participants' neural activity could discriminate between faces and objects rendered invisible using CFS. They measured: (1) target detection, by asking participants to respond whether the target was presented first or second in order; (2) object category identification, by asking them whether the target stimulus was a face or a house; and (3) the event-related potential M170 (equivalent to N170 in EEG), which indexes face processing. N170 is a negative voltage deflection found around temporooccipital areas between 140 and 200 ms after stimulus onset. This component is believed to have its neural source in the fusiform gyrus, more specifically in FFA, amongst other nearby areas (Heisz et al., 2006a; Nguyen & Cunnington, 2014; Bruno Rossion et al., 2003; Watanabe et al., 1999; Yovel et al., 2008). Participants performed at chance level in both target detection and category identification tasks, thus demonstrating that the stimuli were objectively undetectable. However, the M170 was still able to discriminate between faces and houses outside of awareness. While it was known that the M170/N170 component can discriminate between face stimuli and object stimuli categories (Bentin et al., 1996; Carmel & Bentin, 2002; George et al., 1996; Halgren et al., 2000; Liu et al., 2000;

Lu et al., 1991; Bruno Rossion & Jacques, 2012; Sams et al., 1997), these results suggest that complex high-level information such as visual stimuli categories can be processed before the stimuli enter awareness.

Conversely, there is also evidence that awareness may be required for configural face processing. For instance, Axelrod & Rees (2014) tested whether unconscious processing of a face can improve the identification of its eyes by suppressing the face using CFS but not the eye region. In their reasoning, if a face can be processed holistically, then its suppressed presentation should improve the identification of its visible eyes. To test this, they presented participants with two consecutive composite faces that contained visible or invisible eyes while the rest was suppressed from awareness. The two faces could be the same or different, whereas both pairs of eyes were always the same. Participants were asked to report whether these successive pairs of eyes were the same or different. The researchers found that the accuracy at detecting visible eyes was not affected by the invisible face in which the eyes were embedded whereas accuracy did improve when the face was visible. In a follow-up study, they tested whether a subliminal learning task with feedback could improve accuracy when using invisible faces. They tested this by adding a correct/incorrect answer indication at the end of every trial during learning sessions that had the same design as the experiments. Accuracy, however, did not improve with said feedback. Therefore, unlike visible faces, invisible faces did not influence perceptual processing of visible eyes, thus suggesting that holistic face processing either requires awareness to occur or it can occur an absence of awareness but to a limited extent.

In summary, a number of studies using masking and interocular suppression techniques, and especially b-CFS procedures, have suggested that holistic or configural features found in faces may be processed before faces gain access to awareness. Inverted faces are much more difficult to recognise than upright faces, probably due to the disruption of relational facial features that inversion entails. This face-inversion effect has been widely employed in b-CFS studies on unconscious face processing to test whether holistic facial information can facilitate access to awareness.

### 1.3.2 Emotional expressions

One question that has received great interest is whether emotional information enjoys privileged access to awareness. Most of the classic studies that addressed this question employed backward-masked facial expressions presented for brief exposure durations. For example, Esteves & Öhman (1993) explored whether observers could recognise emotional expressions in exposure durations ranging from 30 to 230 ms. They found that confident recognition of backward-masked facial expressions could be achieved with exposure durations of 100 ms. In a subsequent study, Esteves et al. (1994) presented observers with angry and happy facial expressions for exposure durations of 30 ms, both backwardly masked by a neutral face also presented for 30 ms. Importantly, using a shock unconditioned stimulus, they aversively conditioned angry expressions. By measuring skin conductance responses (SCR), they found reliable differential SCR when angry faces were conditioned but not when happy faces were conditioned, thus suggesting that angry expressions could be processed in absence of awareness. Subsequent studies demonstrated that conditioned angry faces modulate the amygdala activity differently depending on awareness: if observers are unaware of stimuli, a masked face produces enhanced neural activity in the right but not left amygdala, whereas when observers were aware, unmasked faces produce enhanced neural activity in the left amygdala instead (Morris et al., 1998), a finding that was replicated and expanded by Whalen et al. (1998).

In addition, a number of studies has provided evidence in favour of unconscious emotion processing measuring EEG event-related potentials. For instance, Balconi & Mazza (2009) presented observers with different facial expressions for either 30 ms (pre-attentive condition) or 200 ms (attentive condition) of exposure followed by a mask. In both conditions, they found a negative frontal deflection around 200 ms (N2) after stimulus onset and a positive parietal deflection around 300 ms (P3) after stimulus onset compared with the pre-attentive condition. Importantly, the amplitude of P3 was lower for both neutral and sad expressions, a difference that was found in both pre-attentive and attentive conditions. These findings were interpreted as evidence of unconscious emotion processing. Years later, Mitsudo et al. (2011) employed a similar procedure to examine ERP response to emotional expressions with and without awareness in more detail. They presented observers with masked faces (neutral or fearful) and masked objects, either shown upright or inverted in orientation, for 20, 30, and 300 ms of exposure to create subthreshold, threshold, and suprathreshold conditions, respectively. They found that the N170 component, an ERP component found around occipitotemporal areas about 170 ms after stimulus onset and normally associated with face-specific processing, was significantly smaller in response to faces than to objects in the subthreshold condition whereas this same effect but in opposite direction was found in both the threshold and suprathreshold conditions. On the other hand, P1, a component found around the visual cortex about 100 ms after stimulus onset and associated with early visual processing, was higher in amplitude for upright faces over inverted faces in all conditions. The authors interpreted these findings as evidence for unconscious face processing. Together, these studies suggest that backward masking may be able to suppress awareness while keeping aspects of visual processing intact, and that emotional content of unseen negative facial expressions is processed by engaging the amygdala differently than with seen negative facial expressions.

But what is the minimal exposure duration required for emotional discrimination of faces? Milders et al. (2008) addressed this question by presenting observers with fearful, angry, happy, and neutral backward-masked facial expressions. Observers had to discriminate between them and rate their awareness of each presented face. Signal detection analyses showed that while both sensitivity to expression and awareness ratings increased along with exposure durations, above-chance sensitivity was found at each exposure duration, ranging from 10 to 50 ms, thus indicating that exposure durations in that range are already too long to produce chance performance. I come back to this point in section 1.4.1, as it suggests a crucial limitation of this method.

Many b-CFS studies have tried to shed light on whether emotional expressions enjoy prioritised access to awareness. In fact, the second b-CFS study published addressed this question: Yang et al. (2007) reported shorter suppression times for fearful expressions than for happy and neutral expressions, suggesting an advantage of fearful expressions in gaining access to awareness. However, this effect was also found with inverted faces. Since turning faces upside down supposedly disrupts high-level information processing, it could be argued that said advantage is due to low-level features such as differences in contrast, luminance, or spatial frequency. Stein & Sterzer (2012) addressed this problem by presenting participants with schematic faces expressing angry, neutral, happy, and sad expressions. Unexpectedly, they found shorter suppression times associated with happy expressions. In a series of follow-up experiments, they demonstrated that this effect was driven by the curvature of the faces' mouths, thus suggesting that the advantage of emotional expressions in suppression times could indeed be driven by low-level features. More recently, though, Yang & Yeh, (2018b) developed an affective-priming task whereby they presented facial expressions suppressed by CFS followed by a visible emotional word. Participants were asked to judge the face's and word's emotional valence (positive or negative). They found a congruency effect – reaction times were shorter when judging the emotional valence of a word that was preceded by a congruent emotional expression. As seen here, the evidence on whether there is an advantage of emotional expressions over non-emotional expressions in the access to awareness does not converge on a single answer.

Despite the fact that findings have been inconsistent, as described above, a consistent advantage of negative emotional expressions (e.g. fear) could be interpreted as a high-level facial feature – one could expect them to rely on visual integration of different features such as open eyes, raised eyebrows, open mouth, and tense facial muscles. While some studies have indeed suggested that the advantage of emotional expressions over non-emotional expressions could be attributed to high-level information processing (Yang et al., 2007; Yang & Yeh, 2018a), most studies have suggested that such an advantage would be driven by differences in low-level features like contrast, luminance, and spatial frequency. For instance, Gray, Adams, Hedger, Newton, & Garner (2013) altered salience of faces by manipulating face orientation and luminance polarity (i.e. normal or colour-inverted), thereby affecting suppression times in a b-CFS procedure. They presented participants with fearful, happy, angry, and neutral expressions. By combining face orientation and luminance polarity, they created four stimulus categories per expression. Notably, they found that the advantage for fearful expressions over happy and neutral ones, originally reported by Yang et al. (2007), was still present even when faces were shown upside down, or colour-inverted, or both. Because turning a face upside down has demonstrated to disrupt its high-level features, and even more when inverting its colours, these findings suggest that the advantage of fearful expressions can be explained by differences in low-level features. Indeed, other studies have found that low-level features can contribute to said advantage of emotion, like differences in spatial frequency (Willenbockel et al., 2012), luminance, and contrast (Hedger et al., 2015, 2019), thus suggesting that differences in suppression times between emotional and non-emotional expressions could be explained by low-level information alone (for a meta-analysis, see Hedger, Gray, Garner, & Adams, 2016).

However, other data are still supportive of the idea that emotional information in faces is processed during CFS. In another study, researchers found that fear conditioning was associated with better performance in a localisation task of fearful expressions, in both visible and invisible conditions even when controlling for low-level visual differences (Vieira et al., 2017). More recent studies have demonstrated that regardless of what type of visual information drives this advantage for suppressed fearful faces, they draw attention resource allocation compared to other suppressed expressions. For example, researchers have reported that N170 to CFS-suppressed fearful expressions exhibits greater amplitude than to neutral expressions. More importantly for this study, they found a steady-state visual evoked potential (SSVEP) response to the Mondrian-like masks. The SSVEP response exhibited a greater decrease to suppressed fearful expressions than to suppressed neutral expressions, between 1 and 1.2 seconds after stimulus onset, suggesting that fearful expressions may call for stronger attention resource allocation even when processed unconsciously (Jiang et al., 2018). In addition, it has been shown using eye-tracking during CFS that fearful expressions attract gaze during suppression while angry faces, conversely, turn gaze away (Vetter et al., 2019). Moreover, presenting suppressed emotional expressions before a visible target expression in an emotion discrimination task was associated with shorter suppression times along with an increase in discrimination accuracy. These results were found when both stimuli (suppressed and visible) were emotionally congruent in comparison to when they were incongruent, suggesting unconscious emotion processing (Ye et al., 2014).

Irrespective of whether some emotional expressions might (or might not) gain access to awareness faster than others, there have been attempts to elucidate the neural correlates of unconscious emotion processing. The first report using CFS was published by Jiang & He (2006). They used fMRI while participants were presented with images of fearful or neutral faces, either intact or scrambled, and either visible or suppressed by CFS. Participants were asked to respond to a two-interval forced choice (2IFC) task where both intact and scrambled faces were presented. This task was employed as a sanity check – to demonstrate that observers were not aware of the faces. The researchers found strong activation for both visible and invisible faces in the bilateral amygdalae, with stronger activation to fearful over neutral expressions. In addition, activity in FFA showed stronger activation to intact faces (regardless of their expression) over scrambled faces, both when visible and invisible. On the other hand, activity in the superior temporal sulcus (STS) only showed strong activation associated with invisible fearful faces but not with invisible neutral faces. These results suggest a functional dissociation between FFA and STS in the unconscious processing of facial expressions, with the former structure more engaged in face processing and the latter more engaged in emotion processing. A follow-up study by Jiang et al. (2009) on the neural dynamics of unconscious face processing supported these claims: they used EEG event-related potentials and found that both fearful and neutral expressions evoke strong P1 and N170 responses compared to scrambled faces. While both invisible neutral and invisible fearful expressions evoked strong P1 and N170 responses compared to scrambled faces, invisible intact faces generated stronger modulations of the signal at the 140-200 ms range. In addition, invisible fearful faces elicited a significantly larger negative deflection (compared to neutral and scrambled face) starting at 220 ms, thus suggesting a temporal sequence of stimulus discrimination, first between all invisible intact faces compared to scrambled faces, and secondly, between invisible fearful faces compared to neutral and scrambled faces (Jiang et al., 2009).

Following a different approach, Stein, Seymour, et al. (2014) tested whether the unconscious processing of fearful expressions can bypass the visual cortex and reach the amygdala first through subcortical projections, possibly through the superior colliculus and the pulvinar (Tamietto & de Gelder, 2010). To do so, they created face images that differed in spatial frequency and presented them suppressed by CFS on one out of four different screen locations. Observers were asked to localise them as quickly and accurately as possible as soon as the face or any part of it became visible. Because visually responsive neurons in the subcortical route's input (superior colliculus) receive afferences mainly from magnocellular retinal ganglion cells, which are more sensitive to low-spatial-frequency information, face images with low spatial frequency should be predominantly processed by these neurons. Furthermore, because visually responsive neurons in cortical areas predominantly receive afferences from parvocellular ganglion cells, which are more sensitive to high-spatial-frequency information, face images with high spatial frequency should be mainly processed by these neurons (Livingstone & Hubel, 1988; Merigan & Maunsell, 1993). Stein, Seymour, et al. (2014) analysed differences in suppression times and found a consistent fear advantage associated with high spatial frequencies, thus suggesting that the specific suppression advantage for fearful expressions was due to high-spatial-frequency information. Interestingly, while these results contradict past findings that had suggested that unconscious emotion processing might be mediated by a specialised subcortical pathway involving the amygdala (Diano et al., 2017; Morris et al., 1998; Pessoa & Adolphs, 2010; Sato et al., 2019; Tamietto & de Gelder, 2010), they are in line with reports of unconscious fear processing in absence of amygdala (Tsuchiya et al., 2009). In conclusion, the prioritised access to awareness that fearful expressions may enjoy, as reported in many b-CFS studies, would primarily rely on visual cortical areas rather than subcortical structures. Future studies are needed to determine whether this finding is specific for CFS-suppressed expressions or generalisable to other masked or unmasked expressions.

The question of whether faces' emotional information enjoys privileged access to awareness has been explored in people who suffer from psychiatric disorders, too. People who suffer from schizophrenia, for instance, commonly present difficulties at judging emotional information in faces. Kring, Siegel, & Barrett (2014) explored whether such issues might bias faces' access to awareness. They presented participants diagnosed with schizophrenia or with schizoaffective disorder with neutral facial expressions that were preceded by CFS-suppressed smiling or scowling faces. The participants were asked to judge the expression of the neutral face. Both healthy participants and participants diagnosed with schizophrenia rated visible neutral faces as more trustworthy and warmer when preceded by a CFS-suppressed smiling face and less trustworthy and warm when preceded by a CFS-suppressed scowling face. This finding suggests that these elemental aspects of affect perception may be intact in people with schizophrenia.

People who suffer from mood disorders, including major depressive disorder (MDD) and type-1 bipolar disorder (t1BD) also exhibit abnormalities in faces' access to awareness when these communicate emotional content. People who suffer from MDD present negative cognitive biases that may influence perception (Disner et al., 2011; Vermeulen et al., 2019). In a b-CFS experiment, Sterzer et al. (2011) tested whether these MDD-related biases could affect how facial expressions gain access to awareness. They showed healthy and MDD participants different suppressed emotional expressions: neutral, fearful, happy, and sad faces. Among MDD participants, they found significantly shorter suppression times associated with sad expressions relative to neutral expressions whereas happy expressions were associated with significantly longer suppression times than neutral expressions, hence a mood-congruent information processing bias. In addition, suppression times associated with sad faces correlated with self-reported changes in depression severity after 4 weeks of treatment, thus indicating an association between depressive symptoms and unconscious emotion processing, specifically of sad and happy facial expressions.

Similarly, people who suffer from t1BD present impairments in emotion regulation that can affect their social perception. Gruber et al. (2016) explored whether invisible emotional expressions can affect the way participants judge visible neutral expressions. While they did find that t1BD participants rated neutral faces as more or less trustworthy, warm, and competent, depending on whether suppressed happy or angry faces were presented before, these results did not differ between t1BD and healthy participants, thus suggesting that emotion-related processes in t1BD may be intact at early automatic processing stages or, alternatively, that the procedure was not sufficiently sensitive to capture emotion-related effects in these patients.

Different relevant psychopathology-related personality traits modulate how facial expressions gain access to awareness as well. For example, individuals high in psychopathy typically exhibit deficits in processing fear. Sylvers et al. (2011) reported that psychopathic traits (callous-unemotional traits) in children (8-11 years of age) are associated with longer suppression times for fearful expressions whereas narcissism, impulsivity, and conduct problems are not. Additionally, psychopathic traits were associated with longer suppression times for disgusted expressions. In a similar study, Jusyte et al. (2015) explored whether psychopathic traits and aggression may affect unconscious emotion processing by presenting different emotional expressions to young antisocial violent offenders in a b-CFS task. They calculated the difference in suppression times between emotional expressions (for each expression) and their neutral counterparts – they took the resulting positive scores as index of a positive advantage in unconscious processing. Next, they found a correlation between those scores for the fearful-neutral expressions difference and psychopathic traits, suggesting that aggressive participants with higher unemotional traits exhibit a processing deficit in the unconscious detection of fear, which fits theoretical predictions in the study of antisocial behaviour. These results suggest that psychopathic traits associated to affective processing may entail changes in unconscious emotion processing.

There are other personality traits of relevance for psychopathology that have been addressed. For example, Vizueta et al. (2012) tested whether negative affectivity (NA) and dispositional fear (DF) could affect neural response to invisible fearful expressions relative to invisible neutral expressions. Both NA and DF have been related to a number of psychiatric ailments such as anxiety, depression, and general distress (Watson, 2005). The researchers presented participants with CFS-suppressed emotional and neutral expressions and measured differences in brain reactivity using fMRI. They found that only NA was a significant predictor – higher NA was associated with enhanced response in STS and right amygdala in the invisible condition, thus suggesting that amygdala reactivity may be found in the unconscious emotion processing and, by extension, may be recruited differently for individual differences related to anxiety and mood disorders.

Finally, anxiety and its relation with unconscious emotion processing has also been empirically addressed. Capitão et al. (2014) tested whether trait and state anxiety could affect suppression times of different emotional expressions in a b-CFS study. They found that trait anxiety positively correlated with how much faster fearful expressions were than happy expressions to break through CFS, thus suggesting that trait anxiety may entail biases in how facial expressions enter awareness.

In summary, multiple studies using backward masking, brief exposure durations, and b-CFS have shown that emotional expressions, in particular fearful expressions, enjoy prioritised access to awareness. However, some studies have claimed that such advantage could be well explained by differences in low-level features between face images, such as luminance, contrast, and spatial frequency, especially in the case of b-CFS studies. Moreover, a number of studies on the neural basis of unconscious emotion processing has shown that unlike what was believed in the past, this processing advantage would rely on cortical rather than subcortical pathways. Finally, researchers have also explored whether there is a link between psychiatric disorders that involve emotion processing symptoms and how facial expressions gain access to awareness. They have found that many disorders involve different patterns in breakthrough times, which can be useful to understand clinically relevant individual differences in psychiatric disorders.

### 1.3.3 Gaze direction

Eye-gaze is very informative when detecting faces and understanding other people's intentions (Bayliss et al., 2011; Edwards & Bayliss, 2019; Mundy & Newell, 2007; Scaife & Bruner, 1975). While direct-gaze faces draw attention towards themselves, faces with an averted gaze draw attention to the place they look at (Dupierrix et al., 2014; Farroni et al., 2002; Hood et al., 1998; Senju & Hasegawa, 2005), a phenomenon known as joint attention (Mundy & Newell, 2007; Scaife & Bruner, 1975). Several reports have suggested that gaze is a relevant cue for the recognition of gender, age, and emotional expression in a face (Daury, 2011; Edwards & Bayliss, 2019; Itier & Batty, 2009), especially when gaze makes eye contact, as it provides a processing advantage that makes face detection faster (Senju & Hasegawa, 2005; Senju & Johnson, 2009a); this advantage is known as the eye-contact effect.

One question that has received special attention is whether gaze direction can be processed unconsciously. For instance, Sato et al. (2007) presented observers with faces looking either at the left or right side of the screen. These backward-masked faces were employed as cues for a target subsequently presented either at the left or right side of the screen. Observers were instructed to give a detection response as soon as they became aware of the target. Crucially, response times were consistently shorter for valid (gaze looking at the same screen side of the target) than invalid gaze cues in both conditions, when the face was masked and unmasked, thus suggesting that gaze direction is processed in absence of awareness.

Most studies on this topic, however, have employed CFS to render faces invisible. For example, Stein, Senju, Peelen, & Sterzer (2011) reported shorter b-CFS suppression times for direct-gaze faces than averted-gaze faces. They concluded that these results may be due to an enhanced unconscious representation for direct gaze, probably preparing individuals for social interaction. However, they did not find an interaction between gaze direction and face orientation, i.e. the advantage of direct-gaze faces over averted-gaze faces was not disrupted by turning the face upside down, which may suggest that this effect relies on low-level information. In line with this, it has also been observed that gaze direction does not affect unconscious emotion processing, again suggesting that gaze direction might not involve high-level processing (Caruana, Inkley, et al., 2019). Chen &

Yeh (2012) addressed some of these concerns by using schematic faces without eye whites. They found shorter suppression times for direct-gaze faces, both with upright and inverted face images, and also when showing observers pairs of eyes alone. Again, this suggests that high-level information processing may not be driving the effect. Interestingly, a similar advantage has been found for faces turned towards the viewer in comparison to faces turned away, regardless of their gaze direction, indicating that a similar effect can be found with head angle alone (Gobbini et al., 2013). This effect has been successfully replicated even in conditions of high perceptual load, too (Xu et al., 2011).

Studies have also addressed the advantage of direct-gaze faces over averted-gaze faces using neural markers. Yokoyama et al. (2013) measured EEG event-related potentials during the presentation of faces with direct and averted gaze, rendered invisible using CFS. They reported that visual evoked potentials could discriminate between direct-gaze and averted-gaze faces 200, 250, and 350 ms after presentation. In addition, Madipakkam et al. (2015) presented participants with face images both visible and rendered invisible using CFS, with direct and averted gaze, while measuring fMRI signals. After presenting suppressed face image for a predefined exposure duration, they asked participants to guess the location where the target face was presented and rate their response confidence. They found greater neural response to direct-gaze than averted-gaze faces when participants reported being aware of the stimulus than when they reported being unaware. This finding was significant for activity in the FFA, STS, and intraparietal sulcus, all brain regions that have been associated with gaze direction processing (Haxby et al., 2002; Nummenmaa & Calder, 2009).

Making eye contact is commonly impaired among people diagnosed with Autism Spectrum Disorder (ASD) among other limitations in social cue processing. Akechi et al., (2014) tested whether the unconscious eye-contact effect (Stein, Senju, et al., 2011) would also be absent in adolescents with ASD. After ensuring that both typically developed (TD) adolescents and adolescents with ASD could do b-CFS tasks adequately by running a contrast discrimination task with Gabor patches, they ran a b-CFS experiment where they presented participants with both direct-gaze and averted-gaze faces. Crucially, only the control group exhibited a significant advantage of direct-gaze faces in suppression times, thus suggesting an impairment in the unconscious processing of gaze direction in ASD participants. Finally, the researchers ran a third experiment, which consisted of the same experiment but with no CFS masking. This time, both groups exhibited shorter response times for direct-gaze faces than averted-gaze faces. The authors suggest that the absence of an unconscious eye-contact effect in the ASD group might indicate that the initial and unconscious registration of eye contact is attenuated in people who suffer from ASD. More recently, Madipakkam et al. (2019) replicated and extended these findings by showing a linear relationship between the level of autistic traits in TD participants and their sensitivity to CFS-masked direct-gaze faces – the higher they scored in autistic traits, the lower their sensitivity to said faces, thus suggesting a continuum in the relationship between autistic traits and eye-contact sensitivity in a subclinical population.

In another study, Seymour et al. (2016) explored whether participants diagnosed with schizophrenia exhibited the unconscious eye-contact effect originally reported by Stein, Senju, et al. (2011), in which suppression times were shorter for faces making eye contact than for faces looking away. In a b-CFS task, Seymour et al. (2016) found this same effect both for the group diagnosed with schizophrenia and the group of healthy participants: direct-gaze faces had shorter suppression times than averted-gaze faces, suggesting that abnormalities in schizophrenia associated with social cue processing (Palmer et al., 2018), commonly described by clinicians, may take place at later processing stages, when awareness has already occurred. Interestingly, the face-inversion effect (i.e. shorter suppression times for upright over inverted faces) was replicated using the b-CFS procedure with people diagnosed with schizophrenia along with the absence of an inversion effect when using object images, thus suggesting that configural face processing before access to awareness is preserved in people who suffer from this disorder (Caruana, Stein, et al., 2019). This finding indicates that while people with schizophrenia present social cue processing problems, before and after faces enter awareness, their configural processing of faces may be preserved.

In summary, the evidence suggests that the eye-contact effect can occur before faces gain access to awareness, but that it very likely relies on processing of low-level features. Moreover, there are studies that also suggest that people diagnosed with ASD, who typically present social cue processing limitations, would not exhibit the eye-contact effect, i.e. they would not present the advantage of direct-gaze faces over averted-gaze faces in how they gain access to awareness. In schizophrenia, on the other hand, the eye-contact effect would be preserved.

### 1.3.4 Familiarity traits

Faces with familiar features also enjoy prioritised access to awareness. Geng et al. (2012), using a b-CFS procedure, found shorter suppression times for self-faces (participants' own faces) than celebrities' faces, suggesting that facial information containing self-features gains faster access to awareness. Next, they measured EEG event-related potentials during subliminal (CFS) and supraliminal (no CFS) conditions. They found enhanced N170 amplitude to self-faces in the supraliminal condition and a decrease in VPP amplitude to self-faces in the subliminal condition, thus suggesting a distinct neural modulation associated with familiar faces. Gobbini et al. (2013) expanded on these results by finding an advantage for faces of family members compared to faces of unknown people when using the b-CFS procedure. It is interesting to note, however, that while self-faces have shorter suppression times than others' faces, this effect could not be found when comparing attribute words describing the participant's characteristics to attribute words not describing the participant (Noel et al., 2017), thus suggesting that either this effect may be specific for facial information or that people are unable to read text suppressed by CFS.

Race and gender in faces have also been addressed, although research has produced contradictory findings. On the one hand, Amihai et al. (2011) reported that awareness is necessary for processing race and gender information from faces. They presented participants with suppressed faces with unequivocal masculine or feminine features in one task, and CFS-suppressed faces with unequivocal Asian or European features in another task (adaptors), and then examined whether these stimuli affected the results of subsequent gender or race classification of gender- or race-ambiguous faces, respectively. While in the visible condition the features of the unequivocal face affected the classification of an ambiguous face, leading participants to classify faces in the opposite category than that of the adaptor, this effect was not observed in the invisible condition, i.e. when the adaptor was CFS-suppressed, thus suggesting that race and gender-related information require awareness to be processed. On the other hand, Stein, End, et al. (2014) found shorter suppression times for faces matching the observer's own race or age group, in addition to larger face-inversion effects for own-race and own-age faces compared to other-race and other-age faces, suggesting that experience-based face information can be processed unconsciously. Similarly, Yuan, et al. (2017) used own-race and other-race faces as CFS-suppressed primes for an affective priming task. They found that suppressed other-race faces facilitate identification (as indexed by shorter breakthrough times) of subsequent unsuppressed negative words whereas suppressed own-race faces facilitate identification of unsuppressed positive words, suggesting that racial features in invisible faces may affect how faces gain access to awareness. Consistent with this, Yuan et al. (2019) found shorter suppression times for own-race faces than for other-race faces in a b-CFS task. However, arbitrary visual stimuli that were imbued with race-relevance through associative training were not prioritised in their access to awareness: observers made quicker and more accurate judgements for own-race match versus non-match pairings than for other-race match versus non-match pairings. The findings of this study suggest that only own-race faces, but not otherwise arbitrary visual stimuli recently learned to represent one's own race, are prioritised when accessing awareness.

Taken together, these studies suggest that familiarity in the form of own-faces and relatives' faces may enjoy a prioritised access to awareness, whereas findings of familiarity in the form of race and gender are contradictory, with some studies finding evidence in favour of own-race and own-gender faces having a prioritised access to awareness, and some finding evidence against it.

### 1.3.5 Social information in faces

People extract information from faces to evaluate social qualities such as friendliness and trustworthiness, and then use it to adjust their own social behaviour accordingly. Several studies have suggested that two major axes, trustworthiness and dominance, predominantly characterise this evaluation process (Oosterhof & Todorov, 2008; Stirrat & Perrett, 2010). Does this evaluation affect how faces access awareness?

Stewart et al. (2012) generated multiple facial expressions that covered a large range of possible combinations of trustworthiness and dominance traits. They used CFS to render these face images invisible and to test for response times in a b-CFS task. In a series of experiments, they found that faces that were either highly dominant or highly trustworthy elicited significantly longer suppression times than less dominant or trustworthy faces. Interestingly, participants who scored lower in dominance and untrustworthiness took longer to become aware of the dominant or untrustworthy faces. The researchers interpreted these results as evidence of slowed visual perception resulting from a possible passive fear response. However, a more recent study done by Stein et al. (2018) found that the results obtained by Stewart et al. (2012) could be explained by low-level visual features: They successfully replicated the dominance- and untrustworthiness-related longer suppression times, though they found the same effect when turning faces upside down and when presenting only the eye region of faces to participants.

However, a more recent study done by Abir et al. (2018) found compelling evidence of social information modulation over faces' access to awareness in a series of b-CFS studies. They replicated the face-inversion effect by presenting participants with upright faces and inverted faces and employed a reverse correlation to model the facial properties that predicted the breakthrough times obtained. They found a dimension that explained a great part of the variance, which in addition correlated with power/dominance, thus suggesting that these social traits play an important role in how fast faces break through CFS. Crucially, though, the social dimension found could still predict breakthrough times when low-level features were controlled for in the model, based on the obtained data from scrambled and inverted faces. These findings suggest that at least the social dimension power/dominance in faces can be processed in absence of awareness.

Regardless of the nature of the effects found by Stewart et al. (2012), they were successfully replicated and expanded by Getov et al. (2015). They explored individual differences in grey matter volume (measured with structural MRI) associated with suppression times. They found a dissociation: On the one hand, suppression times associated with dominance evaluation were negatively correlated with grey matter volume in right frontal operculum whereas, on the other hand, suppression times associated with untrustworthiness evaluation were negatively correlated with grey matter volume in right temporoparietal junction and bilateral fusiform gyrus, but positively correlated with grey matter volume in medial prefrontal cortex, thus suggesting that the evaluation of social features in faces depends on at least partially separable neural substrates.

Taken together, these reports suggest that social information such as power and dominance can be evaluated unconsciously, thereby affecting the time it takes for a face to break through suppression and enters awareness. The role of low-level features in this effect is controversial, nonetheless.

### 1.3.6 Attractiveness and aesthetic traits

A classic study that presented unmasked stimuli for predefined exposure durations of either 150 ms or 1000 ms found that at both durations, participants' ratings are always highly reliable regarding attractiveness of faces (Goldstein & Papageorge, 1980). However, do attractive faces gain access to awareness faster than less attractive ones? A few studies have claimed that aesthetic facial features such as attractiveness enhance faces' access to awareness. Hung et al. (2016) were the first to report this finding. In a series of b-CFS experiments, they showed that faces that had been rated higher in attractiveness were associated with shorter suppression times and higher accuracy in an orientation discrimination task using Gabor patches if an attractive suppressed face preceded it in the same location compared to if a less attractive face did. Additionally, by using a staircase procedure, they found that more attractive faces had lower visibility thresholds than less attractive ones. These results suggest that suppressed attractive faces also draw spatial attention more effectively than suppressed unattractive faces.

However, this advantage of attractive faces might, like the effects noted in previous sections, be due to low-level features. Nakamura & Kawabata (2018) successfully replicated this effect of shorter suppression times associated with attractive faces by using a b-CFS procedure, but in a series of two additional experiments, they showed that this effect could also be found when turning attractive faces upside down, i.e. when disrupting faces' holistic information. Conversely, the effect was absent when they compared intact attractive faces to scrambled attractive faces. One explanation is that the advantage effect of attractive faces is driven by low-level features such as differences in contrast and spatial frequency, given that the effect was also found with inverted faces. However, the fact that the effect disappeared when using scrambled faces (i.e. when destroying all facial information contained in face images) suggests that the effect may rely on a minimal amount of facial information to occur, perhaps just sufficient to convey that the images are face or face-like stimuli.

These studies indicate convergent evidence of an advantage of attractive faces when gaining access to awareness compared to less attractive faces. However, this effect could be due to low-level features.

### 1.3.7 Conclusions

Many studies have addressed whether different visual features can influence how faces gain access to awareness. Among them, we find configural processing such as in upright faces compared to inverted faces, emotional expressions such as fearful expressions, gaze direction such as in faces making eye contact, familiar features such as in observers' own face and their relatives', social dimensions such as in qualities like friendliness and trustworthiness, and attractiveness. They have all been studied using brief exposure durations, backward masking, b-CFS and CFS procedures, sometimes including neural measures as well. In summary, it has been claimed that faces gain access to awareness faster due to their configural properties, by comparing upright to inverted faces. Regarding their emotional expressions, it has been claimed that fearful expressions would enjoy a processing advantage in comparison to other expressions such as happy and neutral ones. Similarly, other studies have suggested that facial features such as gaze direction, familiarity, and attractiveness may be processed unconsciously, thereby affecting how fast faces gain access to awareness.

However, many studies have suggested that these advantages could be explained by low-level features such as differences in contrast and spatial frequency. Furthermore, some findings are contradictory or have not replicated. More research is needed to clarify what facial features do not need awareness to be processed and, therefore, may influence how faces gain access to awareness.

## 1.4 Main methodological critiques and limitations

All the methods and procedures described above suffer from methodological issues and limitations, which cast doubt on conclusions derived from the studies that used them. Here, I summarise these concerns.

### 1.4.1 Brief exposure durations

Using brief exposure durations to test for the minimal exposure duration required for visual perception and awareness is a straightforward approach that can directly address the main questions of this thesis, i.e. how facial information gains access to awareness and whether different facial features, such as orientation or expression, require shorter exposure durations – and thus less information accumulation – than others in order to reach awareness (suggesting they gain access to awareness faster). However, there are hardware limitations that make it extremely difficult to present visual stimuli for sufficiently brief exposure durations to assess this possibility. This is due to the low refresh rate of most computer monitors, normally between 60 and 120 Hz. For example, Liu & Tanaka (2019) presented observers with faces by using predefined exposure durations of 17, 50, 250, and 500 ms. They found that even 17 ms of exposure were sufficient to produce holistic processing. Similar findings were reported in scene categorisation when asking observers to categorise natural and manmade environments presented for 26 ms of exposure (Joubert et al., 2007). Despite claims indicating that 26 ms of exposure are required to get the gist of a scene (Joubert et al., 2007; Rousselet et al., 2005), by such exposure duration accuracy is already extremely high (> 90%), as reported in those studies, thus supporting the idea that the true minimal exposure duration required for visual perception must be significantly lower – and may be lower than computer monitors allow to reliably present. Because of monitors' low refresh rates, this approach is not appropriate to study face perception and awareness. In Chapters 4 and 5, however, I present a set of experiments that use a newly developed LCD tachistoscope that allows stimulus presentations for under 1 ms, thus circumventing the limitations just mentioned.

### 1.4.2 Masking visual stimuli

Alternatively, researchers have employed masking techniques to interrupt visual processing and impede a great variety of stimulus categories from gaining access to awareness (Lin & He, 2009). But this approach introduces a different potential confound – because we do not know either what aspects of visual processing are being interrupted by the mask, or to what extent they are being interrupted, we cannot be certain about whether different masked stimulus categories can be compared with each other. If masking techniques interrupt different aspects of visual processing besides impeding awareness, they could hence confound behavioural or neural effects associated with said stimulus categories. Furthermore, masks probably halt and replace stimuli, thus probably interacting with the masked image in various ways. In essence, the fact that it is technically extremely difficult (or impossible) to determine what backward masking specifically suppresses, could introduce unknown potential confounds.

### 1.4.3 CFS and the b-CFS procedure

CFS (and by extension b-CFS) enjoys several advantages over backward masking, such as longer suppression times. The rationale underlying b-CFS studies is quite straightforward – the time a stimulus takes to overcome suppression and thus gain access to awareness indexes unconscious processing, more specifically the transition between unconscious processing and conscious processing. Thus, by comparing suppression times between stimulus categories, we can measure which category gains access to awareness first, which reflects faster – and probably more efficient – processing.

Unfortunately, the case is not that simple. In this section, I will address the main critiques that the b-CFS paradigm has received and what could be done in order to circumvent them.

#### *1.4.3.1 The problem of disentangling detection from identification*

Most b-CFS studies ask participants to give a report as soon as the target stimulus breaks through suppression by performing a detection or localisation task. The assumption is that response times reflect breakthrough times – a faster response for detection or localisation should index prioritised unconscious processing. For this rationale to be valid, neither reporting the presence of a stimulus nor reporting its location on the screen should involve the stimulus category's identification or classification, let alone any more complex recognition processes. For example, if the task is to say whether a word is shown, or whether it was shown on the left or right, it should not need to involve reading and comprehending the word. It is therefore assumed that participants do not waste time reading the word and understanding it before responding – they simply press the relevant key as soon as they see any part of it. But is this assumption justified? Similarly, it is assumed that when asked to detect or localise a face, participants respond as soon as they see any part of it, without wasting any time on identifying whose face it is, or what emotional expression or gender it has. However, participants have control over the amount of visual information they receive since trials are self-terminated, which makes it impossible to determine whether participants' response times are the result of pure detection processes or of a combination of detection and identification processes. If identification processes influence response times, not only they could confound the results when identification performance differs between experimental conditions, but they could also bring additional identification-related post-perceptual confounders – e.g. decision criterion differences – into the equation (more on this in the next subsection). Thus, disentangling detection from identification is necessary to avoid potential confounds, especially since less information is required to detect a stimulus than to identify its nature (Kobylka et al., 2017), but we cannot assume that participants are able to suppress their identification processes just because identification is task-irrelevant.

The potentially confounding nature of identification processes in b-CFS studies raises the question of how one should interpret findings obtained with this method. For example, what is the nature of the face-inversion effect measured with b-CFS response times? Classic face perception studies have shown that turning a face upside down disrupts its recognition (Farah et al., 1995; Goodrich & Yonelinas, 2019; McKone & Yovel, 2009; Rakover & Teucher, 1997; Yin, 1969), which has been attributed to a disruption of its holistic configuration – it is much more difficult to integrate facial features when the face is inverted. If turning a face upside down disrupts its holistic processing, then why does it affect b-CFS response times when the task is to just detect the presence or location of the face? As stated before, many b-CFS studies have found shorter response times to suppressed upright faces than inverted faces (Akechi et al., 2014; Gayet & Stein, 2017; Jiang et al., 2007; Moors, Wagemans, & de-Wit, 2016; Seymour et al., 2016; Stein et al., 2012; Stein, Hebart, et al., 2011; Stein, Senju, et al., 2011; Yang et al., 2007). On the one hand, unconscious detection could be partly driven by holistic features, as suggested in the literature. On the other hand, this advantage of upright faces over inverted faces could be driven by task-irrelevant identification processes that some participants are not able to suppress, a feasible possibility given the high between-subject variability found for this effect (Gayet & Stein, 2017). Whatever may be the case, this highlights the importance of disentangling detection from identification processes.

This problem can also be found in studies exploring language. For example, Sklar et al. (2012) claimed that CFS-suppressed sentences with unusual meanings overcome suppression faster than standard sentences, suggesting that semantic processing can happen in absence of awareness. However, a series of high-powered studies performed by Rabagliati et al. (2018) could not replicated it. The response times obtained by Sklar et al. (2012) could have been confounded by identification processes such as reading, which is automatic and thus likely to occur – and influence response times – despite the words' content being task-irrelevant. Even if Rabagliati et al. (2018) had replicated their findings, it could still be that they occurred because of identification processes.

#### *1.4.3.2 The problem of post-perceptual factors*

In an ideal world, participants' b-CFS response times would purely reflect changes in perceptual sensitivity due to the stimulus overcoming suppression and thus reaching conscious awareness. Unfortunately, this may not be the case, as participants may have different decision criteria for reporting a stimulus breaking through suppression. Post-perceptual factors such as response bias (a preference to give a particular response) and decision criterion (the willingness to report a signal) are separate from perceptual sensitivity (the ability to discriminate a signal from noise). Importantly, they may confound participants' subjective reports, especially in b-CFS tasks where the main dependent variable is response time, and the amount of information collected before making a decision is controlled by the participants themselves. In the case of unconscious face processing, for example, participants could exhibit a more liberal decision criterion for reporting fearful expressions than happy expressions. Therefore, participants would need less information to report fearful expressions than happy expressions, leading to shorter response times.

Because b-CFS response times do not distinguish between perceptual sensitivity and criterion, two different approaches have been previously proposed to deal with potential criterion confounders. One approach is to use a conscious control condition that emulates all aspects of the experimental condition except the interocular suppression manipulation. Some control conditions try to achieve this by presenting the stimulus binocularly or monocularly on top of the CFS masks, thus not suppressing it from awareness. The assumption is that such a task should replicate all detection-related post-perceptual factors that could have affected b-CFS response times, such as decision criterion differences. According to this rationale, if the difference in response times to different stimulus categories is larger in the experimental condition than in the control condition, or if there is a difference in the experimental condition whereas no difference is found in the control condition, such effects could be attributed to differences in unconscious processing (Stein, Hebart, et al., 2011). For example, this is the case of the study by Jiang et al. (2007), who found significantly shorter response times to upright faces than inverted faces when the stimuli were suppressed (experimental condition). Such an advantage, however, was not found when they presented the target stimuli on top of the CFS masks (control condition).

There are, however, concerns about this approach to controlling for post-perceptual factors. Perceptual uncertainty is higher when stimuli are suppressed from awareness than when they are not, leading to wider response-time distributions and longer tails, and making stimuli easier to predict in control conditions (Stein, Hebart, et al., 2011). Therefore, as both conditions differ substantially, participants could adopt different decision criteria per condition, thereby making the control condition useless.

Another approach is to ask participants to perform a task that is orthogonal to the experimental manipulation. The assumption is that the experimental manipulation's outcome should be unaffected by differences in post-perceptual factors if these factors are unrelated to the task. For example, Gayet et al. (2016) asked participants to identify the orientation of suppressed Gabor patches. However, they were interested in the effect of the colour of the annulus surrounding the patches (for which associations had been created earlier by conditioning), making the task irrelevant to the experimental manipulation. Similarly, Salomon et al. (2013) asked participants to identify the orientation of suppressed Gabor patches presented inside a hand image, when in fact they were interested in the effect of congruency between the position of the hand image and the participants' hand. This approach has also been adopted in studies about unconscious face processing. For instance, Yang & Yeh (2018a) presented participants with different facial expressions both in upright and inverted orientations. Participants were asked to press a key as soon as any part of the face broke through suppression. Next, they were asked to report the location of the face on the screen and to rate its emotional valence. The detection and localisation tasks were probably not as orthogonal to the experimental manipulation (emotional expression) as in the previous two examples, given that participants could have guessed the purpose of the study and thereby adjusted their decision criterion to it. In fact, we cannot be certain about what specific aspects of the stimuli are relevant for each participant's decision criterion, even in the first two examples. As long as participants have control over the amount of information received, we cannot know whether their response times are confounded by criterion differences or not, let alone whether they reflect differences in perceptual sensitivity.

#### *1.4.3.3 The problem of low-level features, failed replications*

CFS paradigms, and particularly the b-CFS procedure, have been used to explore whether high-level facial features can be processed in absence of awareness. For instance, emotional expressions, identity, and attractiveness are commonly seen as high-level features since their processing allegedly involves integration of visual features and even memory recall. However, for this logic to be valid, suppressed face images should not differ in any other aspect between each other. Low-level factors such as differences in luminance and contrast (Song & Yao, 2016), spatial frequency (Yang & Blake, 2012), and retinal size (Heyman & Moors, 2014) may confound suppression times, thereby increasing false positives.

The problem of low-level features has confounded some b-CFS studies. For example, it was originally claimed that semantic relations between an object and its surrounding context could be extracted unconsciously by showing that suppressed scenes with an incongruent object (e.g. a basketball player tossing a watermelon instead of a basketball) broke through suppression faster than congruent scenes (Mudrik et al., 2011). However, studies with more stringent low-level control exploring the same effect failed to replicate it (Biderman & Mudrik, 2018; Moors, Boelens, et al., 2016). A similar case was made for a study about unconscious visual cueing, in which experimenters showed that separate fragments organised to elicit a Kanizsa triangle illusion broke through suppression faster than when presented in a disorganised way (Wang et al., 2012). A subsequent study showed this b-CFS effect to be due to a low-level confounder – the presence of collinear edges (Moors, Wagemans, van Ee, et al., 2016).

Face images are particularly vulnerable to low-level confounding factors. They naturally differ from each other (e.g. when comparing different identities and features) and, crucially, some of their high-level features may depend on low-level feature differences. For instance, emotional expressions differ in spatial frequency (Jennings et al., 2017; Kihara & Takeda, 2019; Mermillod et al., 2009; Tian et al., 2018) and importantly, certain expressions such as fearful expressions may enjoy prioritised subcortical processing thanks to their spatial frequency (Mermillod et al., 2009; Stein, Seymour, et al., 2014; Vuilleumier et al., 2003). As described above, multiple studies have shown that fearful expressions break through suppression faster than other expressions (Sterzer et al., 2011; Yang et al., 2007), but subsequent studies have indicated that differences in low-level features such as contrast could explain differences in breakthrough times (Gray et al., 2013; Hedger et al., 2015, 2019). While it may be the case that the visual system could have evolved this way, with a preference for high-contrast facial features, even involving distinct eye-movement patterns when presented with different expressions (Vetter et al., 2019), it is essential to account for potential low-level confounders to determine what factors drive the effects of interest. This may be a tricky task when studying faces – for example, Stein & Sterzer (2012) tried to replicate the advantage effect of emotional expressions over non-emotional ones using schematic faces to avoid low-level confounders. But unexpectedly, they found an advantage of happy expressions instead, which, as they demonstrated, was due to another low-level confound: a visual relation between mouth curvature and face contour.

As shown, many failed replications occurred when researchers have controlled for confounding factors, which casts doubts on the validity of the conclusions reached in b-CFS studies that did not control for potential confounds. All these failed replications cast doubts on the validity of b-CFS findings and thus call for more stringent procedures.

#### *1.4.3.4 Conclusions*

In summary, the b-CFS paradigm has received several methodological critiques, which revolve around stimuli-related and participant-related potential confounding factors. The former involves the need to control for low-level visual features that could confound participants' breakthrough times, and the latter involve task-irrelevant identification processes and post-perceptual factors that could confound participants' response times.

## 1.5 Research Questions and Hypotheses

Do upright faces gain access to awareness faster than inverted faces? In other words, does holistic processing give faces prioritised access to awareness? Do emotional facial expressions gain access to awareness faster than non-emotional facial expressions? In other words, does emotional content give faces prioritised access to awareness? As reviewed above, these questions have been addressed in several ways, by using brief exposure durations, masking techniques, and interocular suppression techniques. However, as also argued above, all those approaches suffer from important limitations that may have confounded findings.

In this thesis, I developed more stringent procedures to address whether different facial features can modulate the access of faces to awareness. I focus specifically on faces' holistic configuration, emotional expression, and gaze direction.

### 1.5.1 Developing a more stringent CFS procedure

We developed a more stringent CFS procedure that, unlike b-CFS, allowed us to measure bias-independent detection sensitivity, response bias, bias-independent identification sensitivity, and decision criterion. We achieved this by presenting participants with suppressed face images for one of seven predefined exposure durations, allowing us, rather than the participants, to control how much information they receive on each trial. Importantly, using predefined exposure durations enables us to measure decision criterion differences, which is not possible in b-CFS. Next, after face image offset, we asked participants to report, in a single response, both the side of the screen the face image was presented on, and stimulus content (either gaze direction or emotional expression, depending on the experiment). We could therefore address our main questions by describing how perceptual sensitivity to suppressed faces increases with exposure durations, and whether their orientation (Chapters 2 and 3), gaze direction (Chapter 2), and emotional expression (Chapter 3) modulate this sensitivity. Therefore, this approach presents several crucial advantages over previous methods: first, it allows us to disentangle detection from identification by measuring both during the same task. Second, by using signal detection analyses, we could estimate bias-free sensitivity and post-perceptual factors (e.g. decision criterion), separately. Third, by using the method of constant stimuli – presenting face images for predetermined exposure durations in a random order – we could estimate perceptual sensitivity while controlling the amount of visual information available for the participant.

If gaze direction modulates how faces gain access to awareness, we should find better perceptual sensitivity to suppressed direct-gaze faces than to suppressed averted-gaze faces (i.e. an eye-contact effect) at one of the predefined exposure durations, after sensitivity scores rise above zero. This question is addressed in Chapter 2. If face orientation – due to configural processing – modulates how faces gain access to awareness, we should find better perceptual sensitivity to suppressed upright faces than to suppressed inverted faces (i.e. a face-inversion effect) at one of the predefined exposure durations, after sensitivity scores rise above zero. This question is addressed in Chapters 2 and 3, and also in Chapters 4 and 5 using a different method (see below).

Do decision criteria vary across these conditions in a consistent way? If they do, we should find a consistent difference in participants' willingness to report faces' gaze direction. For example, a more liberal criterion for identifying direct gaze when faces are in an upright orientation than when they are presented upside down.

Our method also allows us to distinguish whether detection and identification vary across conditions. If they do, we should find different effects for each measure, e.g. an advantage of direct-gaze faces in identification but not in detection. If identification relies on configural facial features rather than local facial features, we should always find a face-inversion effect (i.e. better sensitivity for upright over inverted faces) in identification tasks (such as distinguishing a fearful expression from a neutral expression), even if such an effect is absent in detection.

If emotional expressions modulate how faces gain access to awareness, we should find better perceptual sensitivity to suppressed emotional expressions than to suppressed non-emotional expressions across increasing exposure durations. Furthermore, if it is true that negative expressions enjoy prioritised access to awareness than other expressions, we should find better perceptual sensitivity to suppressed angry or fearful expressions than to suppressed happy expressions.

Do decision criteria vary across these conditions in a consistent way? If they do, we should find a consistent difference in participants' willingness to report faces' emotional identity. If participants exhibit a more liberal criterion for identifying negative expressions than for positive ones, we should find a more liberal criterion associated with fearful or angry expressions than for neutral and happy expressions.

Finally, if detection and identification of facial expressions vary across conditions, we should find different effects for each measure, e.g. an advantage of emotional expressions over non-emotional expressions in identification but not in detection. These questions are addressed in Chapter 3, and also in Chapters 4 and 5 using a different method (see below).

### 1.5.2 Using an LCD tachistoscope with extremely brief presentations

We also employed the signal detection task described above with unmasked face images, by using a newly developed LCD tachistoscope with extremely brief presentations. With this new equipment, we circumvented the hardware limitations encountered with standard computer monitors when searching for the minimal exposure duration required for visual processing.

Using this new equipment, we addressed very similar questions and hypotheses to the ones posed above: do orientation and emotional expression modulate how faces gain access to awareness? Equivalent hypotheses are posed regarding perceptual detection sensitivity, response bias, identification sensitivity, and decision criterion.

We also ask an additional question: does metacognitive sensitivity to faces, measured through subjective awareness ratings, show an advantage of upright faces over inverted faces (face-inversion effect), or for emotional expressions over neutral expressions? Unlike perceptual sensitivity (first-level sensitivity), which could increase with or without awareness as exposure durations increase, metacognitive sensitivity (second-level sensitivity) requires awareness to increase. More specifically, participants could show above-chance perceptual performance while feeling that they are guessing, but in order to obtain above-chance metacognitive performance, they need to be aware of their performance. Therefore, if perceptual sensitivity arises before metacognitive sensitivity does, it may indicate unconscious processing. Rather, if perceptual and metacognitive sensitivity arise together, it may indicate that the increase in perceptual performance was accompanied by awareness. Taking this into consideration, if upright faces gain access to perceptual awareness faster than inverted faces, we should find better metacognitive sensitivity for the former than the latter as exposure durations increase. Furthermore, if emotional expressions gain access to perceptual awareness faster than non-emotional expressions, we should find better metacognitive sensitivity for the former than the latter as exposure durations increase. These questions are addressed in Chapters 4 and 5.

Importantly, by measuring all the indices just mentioned as exposure durations increase, we can test whether facial features are processed through a sequence of steps (e.g. holistic configuration before expression identification) or all at once. These questions are addressed in Chapters 4 and 5.

Finally, to test whether neural systems can process faces and their emotional expressions at shorter exposure durations than the ones required for signal detection indices, we measured EEG neural markers to explore at shorter exposure durations than those sufficiently long for the face-inversion effect and perceptual awareness to arise. If this were the case, then neural systems may be able to process faces before participants are able to report them. Similarly, we examined whether neural markers are sensitive to emotional content in faces with a shorter exposure duration than the one required for emotional expressions to be identified. If this were the case, then neural systems may be able to discriminate between emotional and non-emotional expressions before participants are able to identify them. These questions are addressed in Chapter 5.

# CHAPTER 2

# 2 EYE CONTACT FACILITATES THE UNCONSCIOUS DETECTION OF FACES DUE TO HIGHER PERCEPTUAL SENSITIVITY

## 2.1 Introduction

Facial features provide essential information about others' mental states and intentions, and are remarkably effective at capturing attention (Langton et al., 2008) even from early in infancy (Goren et al., 1975; Kwon et al., 2016). In fact, a number of reports have claimed that some facial features can even be processed unconsciously (Alpers & Gerdes, 2007; Doi & Shinohara, 2016; Gobbini, Gors, Halchenko, Rogers, et al., 2013; Jiang et al., 2007; Stein, Senju, et al., 2011; Yang et al., 2007), with the implication that faces might be special stimuli, which are given unconscious priority. However, while these claims about unconscious processing are exciting, concerns about the underlying findings have also been raised, both in terms of their replicability (Schlossmacher et al., 2017; Stein et al., 2017; Stein & Sterzer, 2012), and their interpretation. In particular, and as explained in detail below, even the findings that do replicate may not in fact reflect unconscious sensitivity to facial features, but instead could reflect differences in the biases and criteria that participants use during face processing tasks. This latter concern is particularly acute because the most popular recent method used to study unconscious face processing, the Breaking Continuous Flash Suppression technique (b-CFS) is unable to distinguish sensitivity from criterion and response bias. In this chapter, we address that issue by focusing on two specific claims about unconscious face perception – first, that upright faces reach awareness faster than inverted faces, and secondly, that faces with direct gaze reach awareness faster than faces with averted gaze. We test these claims using a new and more comprehensive method, which avoids the problems inherent in b-CFS and allows us to assess how faces from different conditions reach awareness, using measures of detection sensitivity, response bias, identification sensitivity, and decision criteria.

A rich body of studies has claimed that facial features such as gaze direction (Chen & Yeh, 2012; Stein, Senju, et al., 2011), emotional expression (Hedger et al., 2015; Yang et al., 2007), familiarity (Gobbini et al., 2013), and attractiveness (Hung et al., 2016) can be processed unconsciously. To render images invisible, these studies have employed Continuous Flash Suppression (CFS), a strong interocular suppression procedure (Tsuchiya & Koch, 2005), in which a stimulus presented to one eye is suppressed from awareness by Mondrian-like masks flashed to the other eye. In the b-CFS variant, participants are asked to provide a response as soon as the invisible stimulus breaks through suppression into awareness (Yang et al., 2014), with the assumption that stimuli which are processed with higher priority will break through into awareness faster (Gayet et al., 2014). Accordingly, previous work using this procedure has found that faces break through suppression faster if shown in an upright orientation compared to in an inverted orientation (Jiang et al., 2007), if expressing fear compared to a neutral expression (Yang et al., 2007), or if making eye contact compared to looking away (Stein et al., 2011).

While b-CFS has been widely used to provide evidence for unconscious processing, concerns about the method have been raised. For example, some prominent findings obtained in b-CFS studies have failed to replicate. Sklar et al. (2012) reported that masked sentences with unusual meanings break through suppression faster than control sentences, thus implying that semantic processing could be performed unconsciously. In a series of high-powered studies, however, Rabagliati et al. (2018) found that this result could not be replicated. Similarly, Mudrik et al. (2011) reported that complex scenes containing incongruent objects break through suppression faster than normal scenes, suggesting that awareness would not be required to extract semantic object-context relations from a scene. These results failed to replicate, though, when controlling for low-level visual confounders (Biderman & Mudrik, 2018; Moors, Boelens, et al., 2016). Similarly, Wang et al. (2012) reported that perceptually-grouped objects (Kanizsa triangles) break through suppression faster than matched-but-disorganised stimuli, but this finding failed to replicate (Moors, Wagemans, van Ee, et al., 2016). Together, these failed replications cast doubts on the validity of b-CFS findings.

But beyond the replication failures noted above, there may be an even greater problem afflicting b-CFS studies, including those that have replicated successfully: The confounding effects of decision criteria. In particular, participants may unconsciously process stimuli from different categories in a similar manner and thus possess the same perceptual sensitivity to them, but at the same time may have different criteria for reporting that they have seen members of these different categories. In b-CFS studies, suppressed stimuli do not always emerge immediately into awareness, but instead are often seen partially, which means that participants have to make a decision as to whether or not to report a partially-perceived stimulus and, importantly, their criteria for making these decisions may vary by stimulus. For instance, participants may be inclined to visually explore a certain stimulus category more exhaustively than another before deciding to commit to a response, thus leading to a more conservative criterion and thereby to a slower response. The implication of this is that differences in breakthrough times may not be due to differential sensitivity to stimulus categories – which would imply differences in unconscious processing – but rather to differential decision criteria (i.e. the willingness to report a signal).

We are not the first to note that criterion issues are a concern in b-CFS studies, and indeed some b-CFS studies have tried to control for this problem. For instance, some researchers have included a non-rivalrous control condition (where the target stimuli are shown binocularly or monocularly on top of the flashing CFS masks) with the assumption that post-perceptual effects, such as differences in decision criteria, should have similar effects on suppressed and visible stimuli (Akechi et al., 2014; Costello et al., 2009; Jiang et al., 2007; Li & Li, 2015; Madipakkam et al., 2015; Mudrik et al., 2011; Paffen et al., 2018; Stein & Sterzer, 2012; Zhou et al., 2010). The underlying reasoning is that if a non-rivalrous condition emulates all processes that are not CFS-specific but contribute to differences in response times (RTs), any larger differences between stimulus categories found in the rivalrous b-CFS condition (compared to the visible control condition) should index unconscious processing differences. However, non-rivalrous conditions do not effectively control for decision criteria. Clearly visible targets are easier to distinguish than suppressed ones (Stein, Hebart, et al., 2011), meaning there is less uncertainty about them; and the level of uncertainty is known to affect decision criteria (Charles et al., 2013) and may do so differentially for different stimulus categories. Visible conditions therefore differ in a substantive way from CFS conditions, meaning they are not valid controls.

Another proposed method for controlling for differences in decision criteria is to ask participants to perform a task that is orthogonal to the experimental manipulation. For example, rather than ask participants to report when they saw a stimulus, Gayet et al. (2016) asked participants to identify the orientation of the stimulus, a feature that was irrelevant to their experimental manipulation of stimulus colour. Similarly, Salomon et al. (2013) asked participants to identify the orientation of a stimulus presented inside a hand image, a feature irrelevant to the experimental manipulation, in which the hand image and the participant's hand were either in the same or different position. The assumption of this approach is that participants do not need to make decisions about the experimentally critical but task-irrelevant feature, and thus RTs will reflect a pure measure of processing that is unaffected by differences in decision criteria. However, such an assumption may not be justified: Participants may still perceive (and thus make decisions about) the task-irrelevant feature, and the time that participants spend collecting information on each trial may still be affected by their internal criterion for that feature irrespective of its relevance for the task. Crucially, we cannot tell what factors will affect participants' decision in any paradigm where they can freely choose how much perceptual evidence to gather (i.e. how long to look at the stimulus in a trial). To more-precisely distinguish perceptual sensitivity from decision criterion, we must use a method that does not rely on participants' willingness to commit to a response (e.g. reaction/response times), but rather on perceptual measures collected under conditions where perceptual evidence (e.g. exposure duration in a trial) is controlled by the experimenter.

Here we developed and tested a new method that does not suffer from the above problems. We used this method to test two well-established b-CFS findings that have been successfully replicated: the face-inversion effect and the eye-contact effect.

Upright faces are easier to recognise than inverted faces (Farah et al., 1995; Goodrich & Yonelinas, 2019; Jiang et al., 2007; Stein et al., 2012; Yin, 1969). Similarly, the first published b-CFS study found that upright faces overcome suppression faster than inverted faces (Jiang et al., 2007). This face-inversion effect has been repeatedly replicated with b-CFS procedures (Akechi et al., 2014; Gayet & Stein, 2017; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011) and has been interpreted as evidence of unconscious holistic face processing.

In binocular viewing conditions, eye contact draws attention towards the face whereas averted gaze draws attention towards the gaze's direction (Dupierrix et al., 2014; Farroni et al., 2002; Hood et al., 1998; Senju & Hasegawa, 2005). Stein, Senju, et al. (2011) reported that human faces with a direct gaze break through suppression faster than faces with an averted gaze, thus suggesting an unconscious processing advantage driven by eye contact. They also showed that upright faces break through faster than inverted faces, replicating the aforementioned face-inversion effect. Several other studies have supported this claim about eye contact either by directly (Akechi et al., 2014; Madipakkam et al., 2019; Seymour et al., 2016) or indirectly (Chen & Yeh, 2012; Madipakkam et al., 2015; Yokoyama et al., 2013) replicating this finding, or by expanding on its neural correlates (Madipakkam et al., 2015; Yokoyama et al., 2013).

In some of these studies, the task – to report stimulus location (on the left or right side of the screen) – was orthogonal to the stimulus category (e.g. direct/averted gaze) (Chen & Yeh, 2012; Stein, Senju, et al., 2011). However, as detailed above, participants could have exhibited shorter breakthrough times to direct-gaze faces not due to better sensitivity, but rather due to a more liberal decision criterion – requiring less evidence (and thus less time) to decide to report that they have seen a direct-gaze face.

To establish the roles of perceptual sensitivity and decision criteria in the b-CFS face-inversion effect and eye-contact effect, our new method used CFS-suppressed stimuli that were presented for a range of predefined durations. On each trial, participants saw faces that either had direct or averted gaze and were presented either upright or inverted. Following each display, participants reported the location of the stimulus and its identity (direct or averted gaze). We used signal detection analyses to establish how stimulus duration and type affect sensitivity and decision criteria. A similar stimulus-presentation approach was employed by (Stein, Hebart, et al., 2011; Experiment 3), who used four predetermined exposure durations and found that participants showed higher accuracy in reporting the location of upright versus inverted faces in all durations. Notably, however, (Stein, Hebart, et al., 2011) only measured accuracy – they did not directly assess perceptual sensitivity and decision criterion using signal-detection measures. Furthermore, they did not account for identification processes that might affect accuracy or for criterion differences in such identification processes. Here, we report two experiments: In the first, we conducted a direct replication of Experiment 2 in Stein, Senju, et al's. (2011) original b-CFS study, to verify we obtained similar results (faster RTs to direct than to averted gaze faces, and to upright than to inverted faces). The second, a pre-registered experiment[2] (Appendix A), used our new method to acquire signal-detection measures of sensitivity and criterion for both face location (left/right side of the screen) and identification (direct/averted gaze) at each of seven exposure durations, ranging from 500 ms to 5695 ms. Criterion differences may account for b-CFS findings either fully (in which case, we should find more liberal criteria for direct-gaze than averted-gaze faces, and for upright versus inverted faces, but no effects on sensitivity), partly (in which case, we should find both criterion and sensitivity differences) or not at all (in which case, we should find greater sensitivity for direct-gaze versus averted-gaze faces and for upright versus inverted faces, but no criterion differences).

## 2.2 Experiment 1

Experiment 1 was an exact replication of Stein, Senju, et al. (2011; Experiment 2), testing whether upright faces break through suppression faster than inverted faces, whether faces making eye contact break through suppression faster than averted gaze faces, and whether the factors of face orientation and gaze direction interact. We used the same Matlab scripts and stimuli as the original study but employed a larger sample (32 instead of 14 participants). The original study found a processing advantage of faces making eye contact. Additionally, upright faces broke through suppression faster than inverted faces. There was no interaction between these two effects.

[2] For Pre-registration, also see here: https://aspredicted.org/qj4wf.pdf

### 2.2.1 Method

#### 2.2.1.1 *Participants*

Thirty-two University of Edinburgh students (21 female; 4 left-handed) of mean age 23.8 ($SD_{age}$ = 4.1) provided informed consent and were paid £3 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Both experiments reported here were approved by the University of Edinburgh Psychology Research Ethics Committee.

Past b-CFS studies that have found statistically significant effects of gaze direction or face inversion on breakthrough times employed around 16 participants per experiment (e.g. Stein, Senju, et al., 2011; Jiang et al., 2007). We decided to double this number to increase power and allow counterbalancing of experimental blocks with a multiple of 8 (see Procedure section). A retrospective power analysis, conducted using G*Power 3.1.9.7 (Faul et al., 2009), to test for a difference between conditions in a repeated-measures ANOVA, with a small to medium effect size ($\eta p^2 = .04$) and alpha of .05, aiming to achieve a statistical power of 95%, determined that a sample of 19 participants would be required. If a non-sphericity correction ε of .5 were to be added – as reported in the results section, a number of tests violated this assumption – then a sample of 29 participants would be required.

#### 2.2.1.2 *Stimuli*

In both experiments, stimuli were presented on a 19-inch CRT monitor in a dimly lit room, connected to a computer running Matlab 2014a (Mathworks, Inc) using the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/cogent.php). A chin rest and mirror stereoscope were positioned 57 cm from the monitor, with a vertical divider splitting the display so each eye only saw half of the screen.

Figure 2.1a illustrates the display and stimuli. Two red frames containing binocular alignment contours (random noise pixels around the inside border of the frame) appeared side by side on the screen (squares measuring 10.6° × 10.6°, width 0.8° × 0.8°), supporting binocular alignment through the mirror stereoscope such that only a single frame was perceived. A red fixation dot (0.7° × 0.7°) was presented in the centre of each frame. High-contrast coloured Mondrian-like masks were flashed at 10 Hz to one eye while a face stimulus was presented to the other eye.

We employed the same twelve face stimuli used by Stein, Senju, et al. (2011) and others (Akechi et al., 2014; Madipakkam et al., 2015; Senju & Hasegawa, 2005; Seymour et al., 2016). In these images, the face is laterally averted either to the left or to the right, and the eyes are also averted to either the left or right, giving the impression of either averted or direct gaze, depending on whether gaze direction matched head direction. For instance, from the viewer's perspective, in the case of faces averted to the right, eyes directed to the left were classified as direct gaze and eyes directed to the right were classified as averted gaze. This was done with the intention of ruling out the potential confounding influence of greater eye symmetry present in direct-gaze faces with straight head direction (see Senju & Hasegawa (2005) for details of stimulus creation). Stimuli were cropped to oval shapes (3.3° × 4.6°), equalised for contrast and luminance and the edges were blurred into the grey background. Inverted faces were created by turning upright faces upside down (180° turn).

#### *2.2.1.3 Procedure*

Participants were instructed to focus on the fixation dot with both eyes open, trying to avoid blinking and looking elsewhere.

At the start of each trial, the red frames, binocular alignment contours, and fixation dots were presented for 1 s. The red frames and binocular alignment contours were continuously present during the experiment. In each trial, fixation dots were presented binocularly; one eye was presented the CFS mask – Mondrian-like patterns changing at 10 Hz – and a face was introduced to the other eye; its contrast ramped up linearly from 0% to 100% over 1s and then remained constant until either a response was given or 10 s had passed, at which point the face, fixation dots, and mask disappeared during a 1.5 s intertrial interval. The eye receiving the mask was the same throughout the study but varied randomly among participants.

Face stimuli were presented either to the left or to the right of the fixation dot (horizontal fixation-to-centre distance 2.7°) at a random vertical position (maximum centre-to-horizontal-midline distance 2.1°). Participants were instructed to press the left or right arrow key on the keyboard to indicate the location of the face as soon as they became aware of its presence.

The experiment consisted of 192 randomly ordered trials, which were evenly distributed over the two crossed experimental factors (gaze direction and face inversion), with the face appearing on each side of the visual field on half of the trials. A 5-minute break was given halfway through the experiment. There were no practice trials. Half of the participants viewed a version of the faces with the head averted to the left and the other half viewed a version of the faces with the head averted to the right. The full experiment took around 20 minutes to complete.

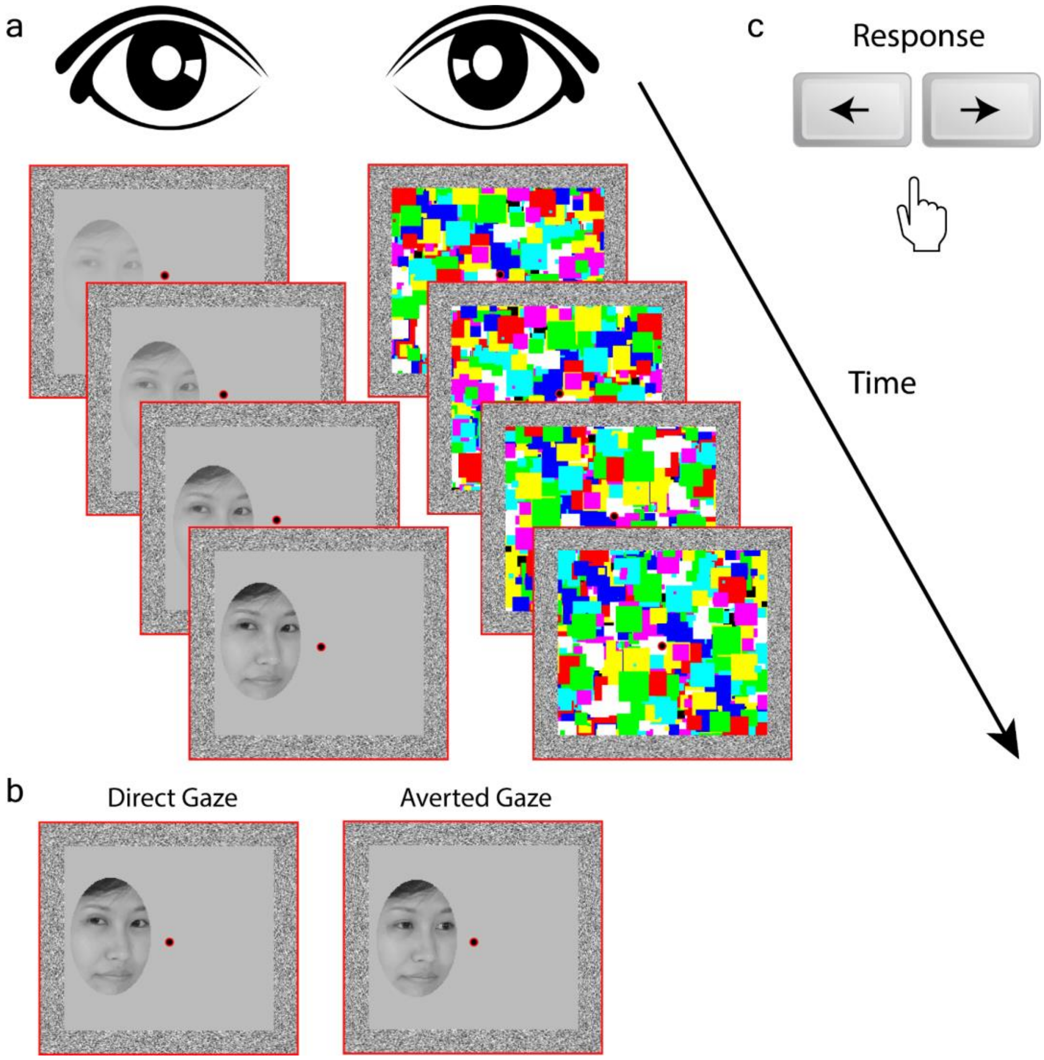



**Figure 2.1.** Schematic description of a trial in Experiment 1 (replication study). (a) The mask stimuli were shown at 100% contrast whereas the target stimulus increased in contrast linearly from 0 to 100% over 1 s. (b) Example of direct-gaze and averted-gaze faces. (c) The trial ended when the participant gave a response (left or right) or after 10 s.

### 2.2.2 Analysis and Results

We calculated mean RTs based on trials with correct responses (98.8% of all trials). Trials with no response were treated as missing data (< 5% for each participant). A preliminary mixed analysis of variance (ANOVA) on mean RTs, which included the factors of gaze direction (direct or averted) and face orientation (upright or inverted) as within-subject factors, and head direction (left or right) as a between-subjects factor, showed no main effect of head direction nor any interaction of this factor with any other factor (all relevant p-values > .1), so this factor was collapsed in further analyses.

To examine whether upright and direct-gaze faces elicit faster breakthrough reports than inverted and averted-gaze faces, as Stein, Senju, et al. (2011) found, we entered RTs into a 2 (gaze direction: direct, averted) × 2 (orientation: upright, inverted) repeated-measures ANOVA. Critically, there was a main effect of gaze direction, with faster RTs for direct-gaze faces (*M* = 3016.8 [*SD* = 962.9]) than for averted-gaze faces (*M* = 3436 [1020.3]), $(F_{(1, 31)} = 54.14, p < .001, \eta p^2 = .636)$. There was also a main effect of orientation, with faster RTs for upright faces (*M* = 2996.1 [950.5]) than for inverted faces (*M* = 3456.8 [1030.9]), $(F_{(1, 31)} = 75.72, p < .001, \eta p^2 = .710;$ Figure 2.2$)$. Finally, and similar to Stein, Senju, et al. (2011), although the simple effect of gaze was numerically larger for upright ($M_{difference}$ = 535 ms [900]) than for inverted faces ($M_{difference}$ = 303.4 ms [1008.8]), and each of these simple effects was significant $\left(t_{upright}(61.5) = -6.35, p < .001, d = -1.122;\ t_{inverted}(61.5) = -3.59, p = .004, d = -0.635\right)$, they did not differ significantly from each other, as indicated by the finding that the interaction between gaze direction and face orientation did not reach significance $(F_{(1, 31)} = 3.49, p = .071, \eta p^2 = .101)$. These results replicate all aspects of Stein, Senju, et al.'s (2011) findings.

To further examine the non-significant interaction, we calculated a Bayes factor for this result using JASP (version 0.11.1, by JASP Team (2019)), with a standard Cauchy distribution centred on zero with rate = 0.707. This provided a value of $BF_{01} = 1.343$, indicating that given the data, the null is only slightly more likely than the alternative hypothesis model. The findings are best characterized as "anecdotal evidence", and do not provide strong evidence for either the null or alternative (van Doorn et al., 2019).

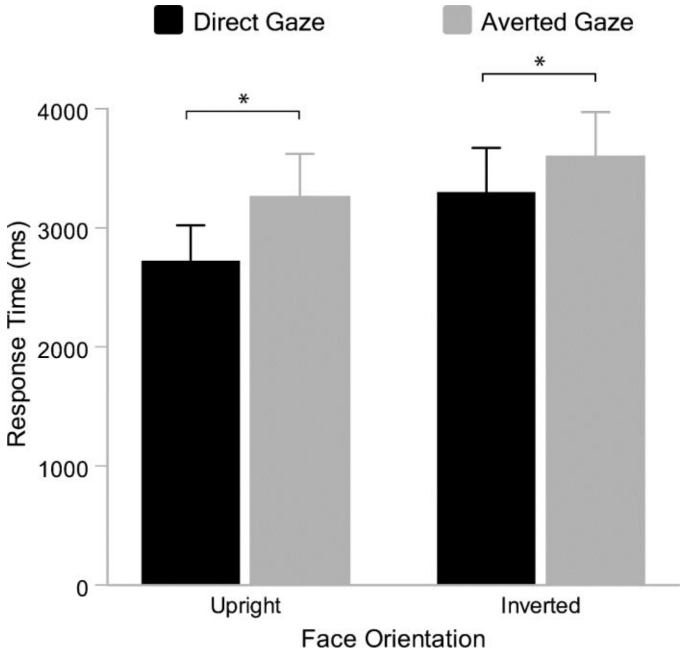



**Figure 2.2.** Results of Experiment 1. Bars indicate mean RTs for detection of CFS-masked faces. Asterisks index statistically significant differences. Error bars represent 95% confidence intervals (CI).

### 2.2.3 Discussion

This first experiment replicated Stein, Senju, et al.'s (2011) b-CFS findings: direct-gaze faces broke through CFS faster than averted-gaze faces. In addition, we found a face-inversion effect, with upright faces breaking suppression faster than inverted faces, an effect that is consistent with many other reports (Gobbini, Gors, Halchenko, Hughes, et al., 2013; Gobbini, Gors, Halchenko, Rogers, et al., 2013; Gray et al., 2013; Jiang et al., 2007; Stein et al., 2016; Stein, Hebart, et al., 2011; Stein, Senju, et al., 2011; Zhou et al., 2010). As in the original study, we did not find a significant interaction between these effects, which may have implications for the possible mechanisms behind the eye-contact effect in b-CFS; we return to this issue in the General Discussion.

Faster breakthrough of direct gaze faces has previously been interpreted as suggesting that gaze direction affects unconscious processing of faces. As discussed in the Introduction, however, such findings do not rule out the potential effects of differential criteria. Therefore, having confirmed that eye-contact and face orientation affect reported breakthrough times in b-CFS, we next examined whether these factors affect perceptual sensitivity and decision criteria when the duration of exposure to the stimulus is controlled.

## 2.3 Experiment 2

To de-confound perceptual sensitivity and decision criteria in the b-CFS paradigm, we modified the procedure so that participants were now exposed to CFS trials whose duration varied in a predetermined manner, thus controlling how much information they were exposed to on each trial. Then, after each stimulus presentation, participants judged both where on the screen the masked stimulus was shown, and what that stimulus was. We used signal detection analyses to assess sensitivity to both stimulus location and stimulus identity, as well as criteria for making these judgments. If this method reveals differences in sensitivity to stimulus categories, they cannot be attributed to differences in decision criterion. Alternatively, we might find effects on decision criteria alone, which would suggest that the orientation and eye-contact breakthrough-time effects found in Experiment 1 cannot be attributed to differential sensitivity. Finally, we could find both sensitivity and criterion differences between stimulus categories, which would suggest that both contribute to these breakthrough-time effects.

### 2.3.1 Method

#### *2.3.1.1 Participants*

Thirty-two University of Edinburgh students who had not participated in Experiment 1 provided informed consent and were paid £14 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Three participants were excluded from analysis (see the Analysis section below); the remaining 29 participants (23 female; 3 left-handed) had a mean age of 24.6 ($SD_{age}$ = 3.6).

#### 2.3.1.2 *Stimuli and Apparatus*

Face stimuli in this experiment were the same as in Experiment 1. The visual display differed slightly from that of Experiment 1: instead of red frames and a red fixation dot, binocular vergence was maintained by two vertical vergence bars (width 1°, height 8°) that appeared to the left and right of stimuli in each eye from fixation (horizontal fixation-to-bar distance 3.1°), and a black fixation cross (0.7° × 0.7°).

#### 2.3.1.3 *Procedure*

Participants were instructed to focus on the fixation cross with both eyes open and avoiding blinking during the trials.

Two textured bars were presented to each eye continuously, to maintain stable vergence. Each trial began with a fixation cross presented binocularly, between the textured bars, at the centre of each eye's visual field (Figure 2.3a). 200 ms later, the CFS mask (Mondrian-like patterns changing at 10 Hz) was presented to one eye, and a face image was introduced to the other eye, ramping up from 0% to 100% contrast over a 1-second period. On trials in which stimulus presentation was shorter than 1 s (see below), termination of presentation curtailed the change in contrast. On longer trials, face contrast remained at 100% until the end of the trial. The mask's contrast was stable across the trial. Stimuli were presented for one of seven predefined durations, spaced equally on a log scale (500; 750; 1125; 1688; 2531; 3797; 5695 ms). This range of exposure durations was determined in piloting sessions that used exposures of 300–6000 ms; importantly, it encompasses the mean breakthrough times found in Experiment 1.

After stimulus offset, a response cue consisting of four question marks was shown binocularly in the middle of the field of view, replacing the fixation cross. Participants had to provide a response within a 2 s response window: They pressed one of four keys (left control, left shift, down arrow, or up arrow) to indicate both where the face had been shown (left or right) and whether the face's gaze was directed at them or averted (Figure 2.3b). This single response thus provides measures of both detection (stimulus location), and identification (stimulus gaze). Following this response, a screen showing only the vergence bars was presented for an ITI of 1000 ms before the next trial began.

The experiment consisted of 1120 trials. Face orientation was blocked in a counterbalanced ABBABAAB BAABABBA order (70 trials/block, with A and B denoting upright and inverted faces, respectively, for half of the participants, or vice versa for the other half). Participants were given self-terminated breaks every 70 trials and a compulsory 15-minute break halfway through the experiment. Unlike Experiment 1, in Experiment 2 all participants viewed faces with the head averted to the left and faces with the head averted to the right. For each face orientation, all combinations of face side (left/right), gaze (direct/averted), head direction (left/right), and stimulus duration (seven possible durations) were presented equally often in randomised order.

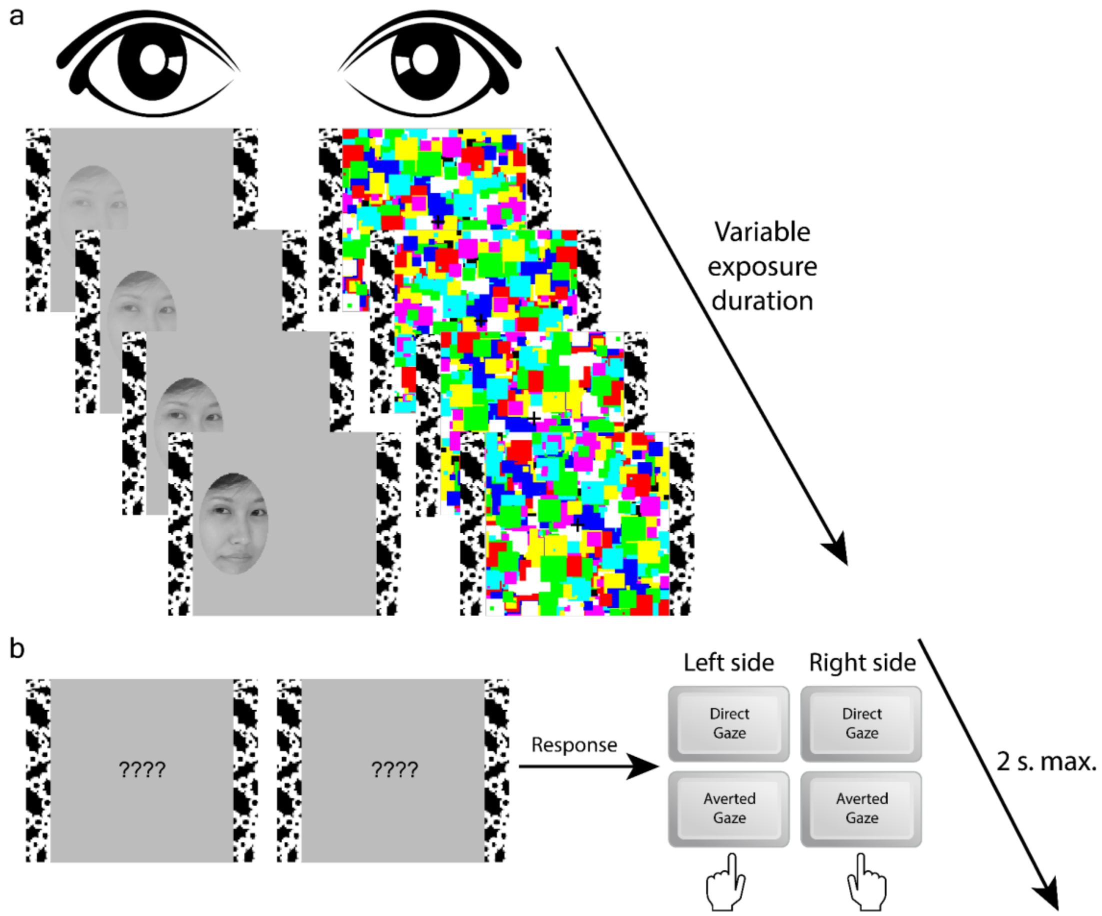



**Figure 2.3.** Schematic description of a trial in Experiment 2. (a) Stimulus presentation. Stimuli were presented for one of seven possible durations (500 – 5695 ms, equally spaced on a log scale). The contrast of the target image increased linearly from zero to 100% across the first second and then remained unchanged until the end of the trial. (b) Detection/identification response. Immediately following stimulus offset, a response cue was presented binocularly. Participants provided a single response to indicate both on which side of fixation the face had been shown and whether its gaze was direct or averted.

#### *2.3.1.4 Analysis*

We excluded data from three participants, in line with our pre-registered exclusion criteria: one failed to respond on more than 10% of trials and the other two did not show any increased accuracy as exposure duration increased (even at the longest duration, in which the mask was no longer present by the end of the trial), suggesting that they failed to attend to the task. Trials that received no response (1.7% in total) were treated as missing data (< 5% for each participant).

We used Signal Detection Theoretic (SDT) measures to assess how perceptual sensitivity and criteria to visual information changed across display durations. To determine each participant's bias-independent sensitivity to face location (left or right; henceforth referred to as location d') for each combination of duration, face orientation, and gaze direction, we employed the calculation for two-alternative forced choice (2-AFC) tasks (Macmillan & Creelman, 2004), $d'_{\text{location}} = \left(\frac{1}{\sqrt{2}}\right)(Z(Hit_{\text{location}}) - Z(FA_{\text{location}}))$, where Z(Hit) stands for the Z score associated with the probability of a Hit (defined as a trial in which a face was displayed on the right and reported on the right), and Z(FA) for that associated with the probability of a false alarm (a trial in which a face was displayed on the left but reported as being on the right). To estimate each participant's bias to respond left or right (henceforth referred to as response bias) during face localisation, we employed the calculation $C_{\text{location}} = -\left(\frac{1}{2}\right)(Z(Hit_{\text{location}}) + Z(FA_{\text{location}}))$. Positive and negative values for this measure indicate a bias toward responding "left" and "right", respectively; however, as these may cancel out across participants, we converted the results to absolute values as a measure of response bias quantity. To determine gaze direction identification sensitivity (to direct gaze; henceforth referred to as gaze identification d'), we used the calculation of d' for Yes/No detection tasks, $d'_{\text{gaze identification}} = Z(Hit_{\text{identification}}) - Z(FA_{\text{identification}})$, where a hit was defined as correctly reporting a direct-gaze face and an FA was defined as incorrectly reporting a direct-gaze face. To estimate each participant's criterion during gaze direction identification, we employed the calculation $C_{\text{gaze identification}} = -\left(\frac{1}{2}\right)(Z(Hit_{\text{identification}}) + Z(FA_{\text{identification}}))$. Lower the value of this measure indicate a more liberal criterion (i.e. the participant is more willing to report direct gaze). These measures were calculated by-condition for each participant and analysed using ANOVA as detailed below. Greenhouse-Geisser adjusted degrees of freedom were used when Mauchly's test indicated a violation of the sphericity assumption.

Both frequentist (ANOVA) and Bayesian (Bayes factors) statistical analyses were performed using Jamovi (The jamovi project, 2020) and JASP (JASP Team, 2020), and corroborated using R and SPSS. When an ANOVA indicated a significant interaction, we ran post hoc Bonferroni-corrected pairwise comparisons to look for significant effects.

Post hoc pairwise comparisons in these statistical packages use estimated marginal means based on the variance of the ANOVA model. For Bayes factor analysis, we defined the null hypothesis as no difference between conditions by using a standard Cauchy distribution centred on zero with rate of 0.707.

### 2.3.2 Results

#### *2.3.2.1 Location Sensitivity*

Individual participants' by-condition location d' scores were entered into a preliminary repeated-measures ANOVA, which included gaze direction (direct or averted), face orientation (upright or inverted), and head direction (left or right) as within-subject factors. We found no main effect of head direction $(F_{(1, 28)} = 0.123, p = .731, \eta p^2 = .005)$ nor any interaction of this factor with any other factor (all p-values > .1). Therefore, d' scores in all further analyses were collapsed across head direction conditions.

To examine how the manipulated factors affected face detection, we entered location d' scores into a 2 (gaze direction: direct, averted) × 2 (face orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. Unsurprisingly, there was a main effect of exposure duration (Figure 2.4a), whereby sensitivity increased with exposure durations $(F_{(2.58, 72.19)} = 164.411, p < .001, \eta p^2 = .854)$. Importantly, there was a main effect of gaze direction $(F_{(1, 28)} = 8.626, p = .007, \eta p^2 = .236)$, confirming that participants were more sensitive to the location of direct-gaze faces (*M* = 1.13 [1.06]) than averted-gaze faces (*M* = 1.05 [1.07]). There was also a main effect of face orientation $(F_{(1, 28)} = 12.339, p = .002, \eta p^2 = .306)$, indicating a sensitivity advantage for upright faces (*M* = 1.18 [1.1]) over inverted faces (*M* = 1.00 [1.03]). The interaction between gaze direction and exposure duration was significant $(F_{(4.56, 127.75)} = 11.67, p < .001, \eta p^2 = .294)$ but, surprisingly, Bonferroni-corrected pairwise comparisons revealed that the advantage for direct-gaze faces over averted-gaze faces only reached statistical significance at one exposure duration: 3797 ms ($t(196) = 7.61, p < .001$). The interaction between face orientation and exposure duration reached significance as well ($F_{(4.54,\ 127.25)} = 3.54, p = .007, \eta p^2 = .112$). Post hoc Bonferroni-corrected pairwise comparisons revealed that the advantage of upright faces over inverted faces was driven by significant differences at 1688 ($t(153) = 2.85, p = .007, d = 0.529$), 2531 ($t(153) = 3.75, p = .023, d = 0.696$) and at 3797 ms ($t(153) = 4.45, p < .001, d = 0.827$) of exposure. Consistent with the findings of Stein, Senju, et al., (2011) there was no interaction between gaze direction and face orientation ($F_{(1,\ 28)} = 0.093, p = .763, \eta p^2 = .003$), and no further three-way interaction with exposure duration ($F_{(4.68,\ 131)} = 0.364, p = .861, \eta p^2 = .013$).

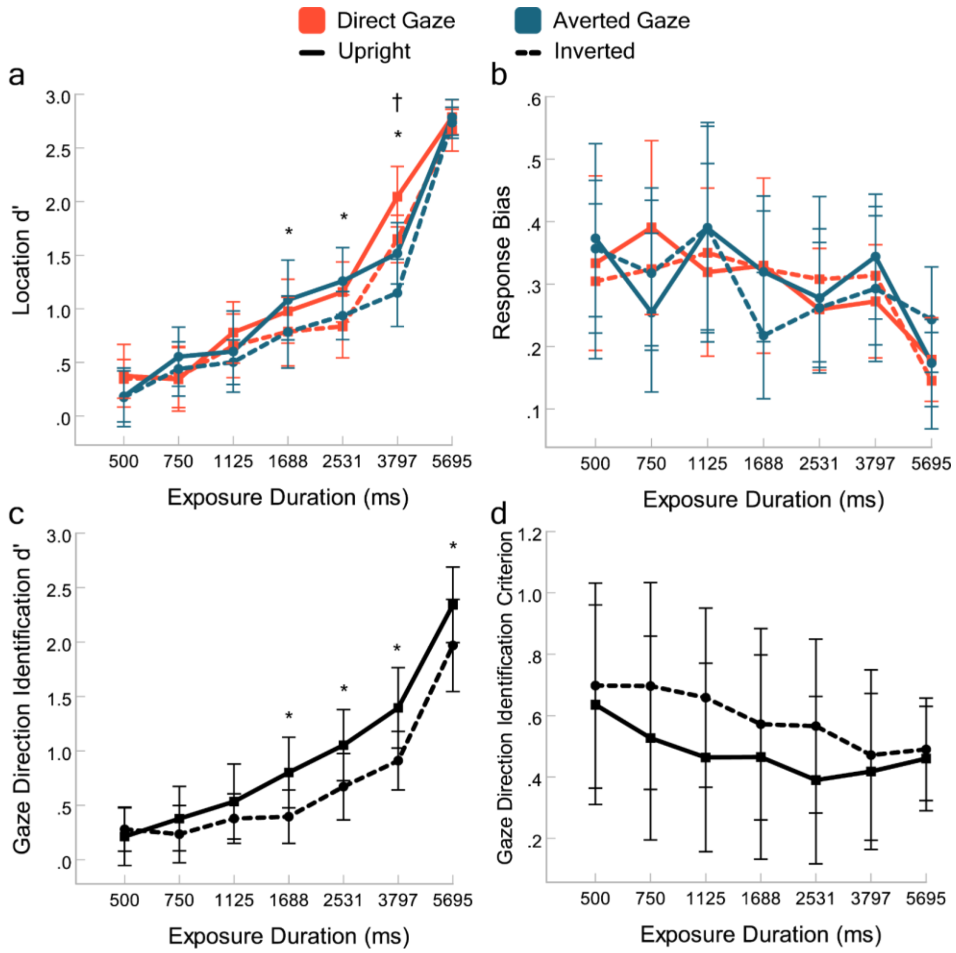



**Figure 2.4.** Results of Experiment 2. (a) Location sensitivity. Location d' increased with exposure time. A significant advantage for direct-gaze faces over averted-gaze faces is present at 3797 ms of exposure. A significant advantage for upright faces over inverted

faces is present at 1688, 2531, and 3797 ms of exposure. (b) Absolute-value response bias scores for reporting location (bias toward either left or right). The amount of bias decreased as exposure duration increased, but there was no difference in amount of response bias between gaze and orientation categories. (c) Identification sensitivity for gaze direction. Identification d' increased with exposure duration. A significant advantage in gaze direction identification for upright faces over inverted faces arises by 1688 ms of exposure. (d) Criterion scores for reporting direct gaze. Upright faces exhibit a significantly more liberal criterion (lower values) than inverted faces during gaze direction identification. Asterisks index statistically significant differences between face orientations. Daggers indicate statistically significant differences between gaze directions. Error bars represent 95% CI.

The non-significant interaction between gaze direction and face orientation is similar to that found in Experiment 1, but similarly, it does not mean that the null hypothesis is necessarily true. We calculated Bayes factors to test whether the obtained data supports the absence of an interaction here (null hypothesis model). Bayes factors indicated strong evidence in favour of the null hypothesis model ($BF_{01} = 10.069$). In other words, the results suggest that the location sensitivity advantage for direct-gaze faces did not depend on the orientation of the face.

These results, obtained with our new method, confirm the pattern of findings obtained using b-CFS in Experiment 1 and by Stein, Senju, et al. (2011): Direct-gaze faces enjoyed a detection advantage over averted-gaze faces, and upright faces enjoyed an advantage over inverted faces. The absence of an interaction between gaze direction and face orientation is also similar to that obtained in previous experiments.

#### 2.3.2.2 *Location response bias*

We examined whether participants' response bias for reporting face location varied across conditions by entering the absolute values of $C_{\text{location}}$ scores into a 2 (gaze direction: direct, averted) × 2 (face orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 2.4b). Response bias significantly decreased with exposure duration ($F_{(2.57, 71.99)} = 4.44, p = .009, \eta p^2 = .137$), indicating that as participants' ability to detect the face increased (shown by higher location d' scores) they became less likely to exhibit a systematic bias in their preference to report one side or the other. We did not find main effects of gaze direction ($F_{(1, 28)} = 0.149, p = .702, \eta p^2 = .005$), and face orientation ($F_{(1, 28)} = 0.117, p = .735, \eta p^2 = .004$), nor any interactions (all p-values > .09), suggesting that only exposure duration affected response bias. To assess whether the obtained data support the absence of an effect of gaze direction and face orientation, we estimated Bayes factors for each comparison, which indicated substantial evidence for the null hypothesis model of gaze direction ($BF_{01} = 6.899$) and anecdotal evidence for the alternative hypothesis model of face orientation ($BF_{01} = 0.449$).

#### 2.3.2.3 *Gaze Direction Identification Sensitivity*

We examined whether participants' sensitivity to identifying gaze direction varied across conditions by entering identification d' scores – taken over all trials irrespective of location response – into a 2 (face orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 2.4c). A main effect of exposure duration indicated that sensitivity to gaze direction increased with increasing duration ($F_{(2.60, 72.77)} = 64.25, p < .001, \eta p^2 = .696$). We also found a main effect of orientation, such that gaze direction identification sensitivity was significantly higher for upright faces ($M = 0.95\ [1.09]$) than for inverted faces ($M = 0.69\ [0.93]$; $F_{(1, 28)} = 34.56, p < .001, \eta p^2 = .552$). The interaction between face orientation and exposure duration also reached significance ($F_{(4.65, 130.24)} = 3.67, p = .005, \eta p^2 = .116$). Post hoc Bonferroni-corrected pairwise comparisons revealed that the advantage in favour of upright faces was significant at exposure durations of 1688 ms ($t(193) = 3.86, p = .014, d = 0.717$), 2531 ms ($t(193) = 3.63, p = .033, d = 0.675$), 3797 ms ($t(193) = 4.62, p < .001, d = 0.857$), and 5695 ms ($t(193) = 3.55, p = .044, d = 0.660$).

Thus, these results add interesting nuance to the earlier-described effect of gaze direction on location sensitivity, which was not affected by face orientation. Here, our findings indicated that orientation does affect the ability to explicitly identify gaze direction, which suggests that the identification of gaze direction may rely on high-level configural features and may be independent of how eye contact affects breakthrough from CFS. We return to this in the General Discussion.

#### *2.3.2.4 Gaze Direction Identification Criterion*

We examined whether participants' criterion for reporting direct gaze varied across conditions by entering $C_{\text{gaze identification}}$ scores into a 2 (face orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 2.4d). Lower the value of this measure indicate a more liberal criterion. There was a main effect of face orientation, indicating significantly more liberal criteria for upright (*M* = 1.05 [1.07]) than for inverted faces (*M* = 1.13 [1.06]), $(F_{(1, 28)} = 6.78, p = .015, \eta p^2 = .195)$. Despite an overall numerical trend toward more liberal criteria with increasing exposure duration, the main effect of exposure duration fell short of significance $(F_{(1.75, 48.95)} = 2.53, p = .097, \eta p^2 = .083)$. To assess whether the obtained data support the absence of an effect of exposure duration, we estimated Bayes factors, which indicated anecdotal evidence for the alternative hypothesis model $(BF_{01} = 0.445)$, i.e. the results were inconclusive on the effect of exposure duration. The interaction between both factors did not reach significance either $(F_{(4.13, 115.63)} = 1.86, p = .120, \eta p^2 = .062)$.

These results indicate a more liberal criterion for reporting 'direct gaze' for upright than for inverted faces. This criterion difference seems not to change across exposure durations. As face inversion disrupts configural processing, our findings indicate that such processing plays a role in participants' criteria for identifying gaze direction.

### 2.3.3 Discussion

In this second experiment, we measured perceptual sensitivity to gaze direction and face orientation. We found a significant advantage in location sensitivity for direct-gaze faces over averted-gaze faces. Unlike previous claims supporting this effect, our findings cannot be attributed to criterion differences between stimulus categories, and therefore, we conclude that eye contact enhances perceptual sensitivity to faces. We also found a significant advantage in location sensitivity for upright over inverted faces, demonstrating that the previously-reported face-inversion effect (Gray et al., 2013; Stein et al., 2016; Stein, Hebart, et al., 2011; Stein, Senju, et al., 2011; Zhou et al., 2010) is, like the eye-contact effect, an effect on perceptual sensitivity.

Although the eye-contact effect (greater sensitivity to the location of direct than averted gaze) did not differ between upright and inverted faces, we did find that identification of gaze-direction was better for upright than inverted faces. This finding suggests that the processes underlying detection and identification are at least partly dissociable, in line with various claims that identification is more high-level-processing dependent than detection (Gayet et al., 2016; Goodrich & Yonelinas, 2019; Leopold & Rhodes, 2010).

Interestingly, there were no effects of gaze direction on criterion for either location or identification. Participants did, however, exhibit a more liberal criterion for upright over inverted faces in gaze-direction identification, meaning they were more willing to report direct gaze for an upright face than for an inverted face. Although this result cannot explain the eye-contact effects found in b-CFS studies, it does demonstrate that decision criteria may differ across experimental conditions in studies using perceptual suppression. Generally, therefore, it is important to measure and rule out such criterion differences as they may impact and thus confound RTs in standard b-CFS studies.

These findings confirm the effects of gaze direction and face orientation on access to awareness. Crucially, and unlike b-CFS procedures, our procedure enabled us to distinguish effects on sensitivity from those on criterion.

## 2.4 General Discussion

Studies have shown that direct-gaze faces are salient and thus very difficult to ignore (Senju & Hasegawa, 2005; Senju & Johnson, 2009a). It has also been claimed that faces that make eye contact access awareness faster than faces looking away (Chen & Yeh, 2012; Stein, Senju, et al., 2011; Yokoyama et al., 2013). However, due to methodological limitations, past demonstrations of this effect could be attributed to differences in decision criterion rather than perceptual sensitivity.

To address these limitations, we developed a new method to de-confound sensitivity and criterion, providing participants with CFS trials of predetermined length, rather than relying on participants' own judgment to decide when to report that they saw a stimulus. As argued in the Introduction, RTs in the b-CFS procedure can be influenced by criterion differences as participants decide when they have accumulated enough information to make a report, and decision criteria may vary systematically between stimulus categories. By measuring sensitivity and criterion, and controlling the amount of visual information available to the participant, our method offers a more robust approach to testing unconscious processing differences.

We asked whether eye contact really makes faces reach awareness faster, or whether this effect is driven by differences in criterion. Using a traditional b-CFS method, Experiment 1 replicated the eye-contact effect originally reported by Stein, Senju, et al. (2011), and the face-inversion effect reported in several previous studies (Gray et al., 2013; Stein et al., 2016; Stein, Hebart, et al., 2011; Stein, Senju, et al., 2011; Zhou et al., 2010), but with a much larger sample, thus addressing some concerns about statistical power in psychological research (Abraham & Russell, 2008). Faces making eye contact did break through faster than faces looking elsewhere, however, as argued above, these results could still be caused by criterion differences. Therefore, we ran Experiment 2 with our new method. If response-time differences found in b-CFS studies are due to sensitivity, then greater sensitivity to direct than averted gaze should be found, even when exposure durations are the same for both types of stimulus. Indeed, the results of Experiment 2 showed an advantage for direct gaze in both location detection and gaze-direction identification, suggesting that eye contact made faces break through faster due to preferential unconscious processing. This effect was statistically significant, but not very large; nonetheless, our results are consistent with the advantage seen in b-CFS studies being due to perceptual sensitivity. In addition, participants exhibited a more liberal criterion for reporting direct gaze in the gaze direction identification task – i.e. they were more willing to report that a face had a direct gaze when it was upright.

These findings highlight the two main benefits of our method. First, the utility of using predetermined exposure durations instead of self-terminated trials to reduce the differential effects of criterion and response bias, and secondly, the importance of measuring decision criterion to test for variables other than sensitivity that could affect responses.

Interestingly, in both of our studies, the eye-contact effect was not disrupted by face inversion, both when measuring RTs (Experiment 1) and location sensitivity (Experiment 2). This consistent finding has important implications for understanding the nature of the eye-contact effect. Prior work using b-CFS RTs interpreted this result as an indication that gaze processing occurs in earlier-stage processing, prior to configural (or 'holistic') processing (Chen & Yeh, 2012; Stein, Senju, et al., 2011). Similarly, we would argue that this finding implies that gaze direction influences sensitivity to low-level visual features of the images, rather than face-specific processing. This is based on the reasoning that turning a face upside down ought to disrupt its high-level configural processing, but not the processing of low-level features (Farah et al., 1995; Goodrich & Yonelinas, 2019; Jiang et al., 2007; Stein et al., 2012; Yin, 1969). Our results may therefore indicate that face *detection* relies on low-level local features (including those that are associated with different gaze directions), and is thus not disrupted by inversion, whereas *identification* of gaze-direction depends on configural processing that is disrupted by inversion. This may be explained by the different nature of the tasks – while localisation could simply require comparing contrast differences between the two sides of the screen, gaze-direction identification requires configural face information to determine where the face is looking. Therefore, as we found in the localisation task, the eye-contact effect should be preserved regardless of the face's orientation.

Unlike prior work on b-CFS, our method also allowed us to separately measure both detection of a face and identification of that face (in terms of gaze direction). In prior b-CFS work, response-time differences have been commonly taken as an estimate of detection processes alone; however, participants in those tasks may not have been able to suppress identification processes when reporting the presence of a stimulus, and this therefore makes it difficult to confidently isolate the two processes. Asking participants to detect the location of the face and to identify its gaze direction allows us to describe more specifically how faces reach awareness and whether any condition exhibits an unconscious processing advantage. Crucially, doing so also allowed us to measure criterion differences for identification, which, as explained above, might have affected breakthrough times in b-CFS procedures. Therefore, our method presents several advantages over the b-CFS procedure and may be a better choice when testing for differences in unconscious processing of visual stimuli.

One potential limitation of our method is the fact using a single response to report two stimulus features – face location and gaze direction – may have induced greater task complexity than the single detection response required in b-CFS studies. Importantly, this additional complexity did not seem to have had a detrimental effect on participants' ability to make perceptual judgements, as indicated by the consistent main effects of exposure duration and finding both eye-contact and face-inversion effects on localisation.

A limitation of our study, however, is that we used the same face stimuli employed in previous studies that tested for the eye-contact effect (Akechi et al., 2014; Madipakkam et al., 2015, 2019; Seymour et al., 2016; Stein, Senju, et al., 2011), limiting the generalisability of our results to other face stimuli, especially if they differ from ours in low-level local features. Future studies should address this issue by including new stimulus sets.

Finally, as described above, our results suggest that the eye-contact effect may rely predominantly on low-level processing, as it does not seem to require holistic face processing to happen. These results challenge the notion that perceptual advantages related to socially-important features require perceptual integration (Ashwin et al., 2015; Palmer et al., 2018; Zaki, 2013; but see Vrancken et al., 2017). Future studies should further clarify what the nature of the eye-contact effect is: high-level or low-level. This may be especially relevant for studies with clinical populations. For example, b-CFS studies have reported that the eye-contact effect is preserved in schizophrenia (Seymour et al., 2016) but impaired in autism (Akechi et al., 2014). Determining the nature of this effect and its specific underlying neural mechanisms may help clarify the source of this impairment.

In summary, we developed a new procedure to study unconscious face processing. This procedure addresses the limitations of the most popular recent method used for this: the b-CFS procedure. Using our method, we found two effects that have also been reported with b-CFS: an advantage in detection of direct-gaze over averted-gaze faces (eye-contact effect) and of upright over inverted faces (face-inversion effect); but unlike with b-CFS, we found these effects by measuring sensitivity directly, controlling for both response bias and criterion differences. Critically, the fact that our findings confirm these previously-reported effects does not mean that all b-CFS results are reliable; on the contrary – it means that all b-CFS findings (including, and especially, the ones that have not failed replication attempts) should be submitted to rigorous methods to establish whether they are attributable to effects on perceptual sensitivity, decision criterion, or both.

# CHAPTER 3

# 3 DISENTANGLING PERCEPTUAL SENSITIVITY FROM DECISION CRITERION IN THE DETECTION OF EMOTIONAL EXPRESSIONS ACCESSING AWARENESS

## 3.1 Introduction

Facility at perceiving faces and interpreting their expressions is a critical social attribute, allowing us to quickly identify individuals, attribute mental states, and guess intentions behind actions (Grill-Spector et al., 2017; Jack & Schyns, 2015; Little et al., 2011). In this context, the processing of emotional expressions is particularly key, and a number of studies suggest that different emotions are processed in special or distinct ways. For example, angry expressions are detected faster than non-threatening expressions (Fox et al., 2000; Krysko & Rutherford, 2009) and are more quickly discriminated from happy expressions than the other way around (Hansen & Hansen, 1988). Furthermore, the presence of an angry or fearful face alone can enhance the detection of a target stimulus shown nearby (Fox, 2002; Wilson & MacLeod, 2003). A more controversial claim, however, is that the processing of emotional expressions can also happen without awareness. This claim derives from findings that faces with strong emotional expressions are more likely to break through the masking effects of different experimental suppression techniques (Alpers & Gerdes, 2007; Carlson & Reinke, 2008; Hedger et al., 2015; Sterzer et al., 2011; Yang et al., 2007).

However, the most common and effective masking procedure used to obtain this finding (Breaking Continuous Flash Suppression, or b-CFS) does not distinguish perceptual sensitivity (the ability to discriminate between a signal and noise) from decision criterion (the willingness to report a signal). In b-CFS tasks, participants are presented with masked stimuli and are asked to report them as soon as they break through into awareness, thus letting participants decide how much information to be exposed to before deciding to respond. In this chapter, we focus on the specific claim that emotional expressions reach awareness faster than neutral expressions. We test this claim by using the method we developed in Chapter 2, which addresses the problems inherent in b-CFS by assessing detection sensitivity, identification sensitivity, and decision criteria to faces presented for predetermined exposure durations.

Multiple studies have claimed that faces and their emotional expressions can be processed without awareness, based on evidence from Continuous Flash Suppression (CFS), a strong interocular suppression procedure that renders stimuli (like faces) invisible by presenting them only to one eye, while continuously flashing high-contrast Mondrian-like masks to the other eye. In the b-CFS variant, participants are asked to report when the target stimulus breaks through suppression into awareness, and the time that the masked stimuli take to overcome suppression has been taken as an index of unconscious processing, with the assumption being that faster breakthrough times indicate faster, more efficient, or higher priority unconscious processing. This technique has been used to investigate unconscious processing of various aspects of faces, including both low- and high-level facial information. Holistic face processing – processing faces as a whole – is an example of the latter, and was addressed by Jiang et al. (2007), who found that upright faces break through suppression faster than inverted faces, thus suggesting when faces' whole configuration is intact, they are prioritised in their unconscious processing. This finding has been replicated multiple times (Akechi et al., 2015; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011).

The b-CFS method has provided significant amounts of evidence for unconscious processing of another high-level feature: emotional expression. Yang et al. (2007) found that fearful expressions broke through suppression faster than neutral expressions, which suggests that emotional information can be extracted unconsciously, and this finding has been repeatedly replicated as well. For example, Yang & Yeh (2018) also found that fearful expressions broke through faster than neutral expressions, and that turning faces upside down increased breakthrough times overall. In addition, Capitão et al. (2014) tested whether anxiety can affect breakthrough times of fearful, happy and neutral expressions. While they found a main effect of expression, they only found an advantage of fearful expressions over happy expressions, unlike in the original report by Yang et al. (2007), which found an advantage of fearful expressions over neutral and happy expressions.

Nevertheless, Capitão et al. (2014) showed an effect of emotion over breakthrough times, with shorter breakthrough times for fearful expressions than happy expressions. In addition, they found a modulation of anxiety over differences in breakthrough times between fearful and happy faces, expressed by a correlation between higher trait anxiety scores and an increased speed to detect fearful compared to happy expressions. These results suggest differential unconscious emotion processing. Finally, Sterzer et al. (2011) tested whether breakthrough times to fearful, happy, sad, and neutral expressions vary between healthy participants and participants diagnosed with major depressive disorder (MDD). They measured breakthrough times of fearful, happy, sad, and neutral expressions. Then, they estimated the differences between each emotional expression and neutral expressions for each individual; they found a reduction of breakthrough times of sad expressions and an increase in breakthrough times of happy expressions, thus exhibiting a different pattern of findings than previous non-clinical b-CFS studies. Therefore, their results suggest differential unconscious emotion processing as well, due to MDD. Thus, these studies also suggest that emotional information can be extracted unconsciously. Furthermore, clinical studies such as Sterzer et al.'s (2011) and Capitão et al. (2014), which showed that both anxiety and depressive symptoms can influence suppression times, suggest not only that this effect of emotional expression is real but also that it interacts with participants' personal characteristics.

However, the interpretation of these findings is quite controversial, and not every study has replicated the original findings. For example, Stein & Sterzer (2012) found that happy expressions broke through suppression faster than neutral, angry, and sad expressions when using schematic faces. Importantly, though, their results showed that this effect was driven by a low-level confound in the mouth area. By creating positive (happy) and negative (sad) schematic faces with maximised mouth-contour congruency and incongruency, respectively, they found shorter breakthrough times for happy expressions even when the subjective impression of a face was greatly diminished by replacing the eyes with another line. These results suggest that this effect is driven by a low-level confound, which led the researchers to question whether faster detection in b-CFS tasks really involves dedicated face- or emotion-specific processes. Several possible mechanisms have been suggested to explain what drives findings such as the faster breakthrough of fearful than neutral expressions – whether this advantage relies on high-level information (e.g. emotional content), low-level information (e.g. differences in contrast or line curvature; Gray et al., 2013; Hedger et al., 2015, 2019; Stein & Sterzer, 2012), or might be a result of decision criteria rather than perceptual sensitivity, as we argued in Chapter 2.

These worries about the validity of b-CFS studies of emotional expressions mirror two broader sets of methodological concerns about the b-CFS method. The first concern is that many studies using this methodology have failed to replicate, or at least have been shown to be highly sensitive to very particular analytic decisions, mirroring broader concerns about replicability in psychological science. For example, Sklar et al. (2012) claimed that people could unconsciously understand the meanings of sentences and perform arithmetic operations; these findings either failed to replicate (Rabagliati et al., 2018) or when they did replicate were not robust to different analytic strategies (Moors & Hesselmann, 2018). Similarly, findings of unconscious perceptual grouping (Wang et al., 2012) and unconscious context-object integration (Mudrik et al., 2011) have also failed to subsequently replicate (Biderman & Mudrik, 2018; Moors, Boelens, et al., 2016; Moors, Wagemans, van Ee, et al., 2016). Thus, there are worries about the reliability of the b-CFS methodology in general.

The second concern deals with what, precisely, b-CFS measures. An assumption of b-CFS studies is that differences in breakthrough time between conditions are solely due to differences in unconscious processing (Gayet et al., 2014; Yang et al., 2014). However, differences in breakthrough times could also be caused by differences in the criteria that participants use to make decisions about the stimuli, i.e. their willingness to report a signal. Notably, breakthrough from CFS unfolds over a brief – but not immediate – time period; typically, a small part of the suppressed stimulus breaks through first, and visibility then expands to the rest of the stimulus. The amount of breakthrough that a participant requires in order to commit to reporting the stimulus may vary by condition. For example, participants may possess the same perceptual sensitivity to each emotional expression category, but may be more willing (or require the accumulation of less information) to report that they have seen a fearful expression than a neutral expression, and thus report the former faster, even if both stimuli take the same time to break through suppression.

To account for this, prior studies have used two solutions. One is to include a control condition that aims to precisely mimic the experience of a b-CFS task, without actually using CFS, such as by measuring response times to stimuli that are superimposed upon Mondrian masks (Akechi et al., 2014; Costello et al., 2009; Jiang et al., 2007; Li & Li, 2015; Madipakkam et al., 2015; Mudrik et al., 2011; Paffen et al., 2018; Stein & Sterzer, 2012; Zhou et al., 2010). But these control conditions usually differ from experimental conditions in multiple ways, e.g. visible targets are much easier to distinguish than suppressed ones (Stein, Hebart, et al., 2011), which is known to affect decision criteria (Charles et al., 2013; also see Chapter 2) and may do so differentially for different stimulus categories. As a consequence, the masked and unmasked conditions will differ in breakthrough time distributions – b-CFS RTs exhibit more spread and longer tails than their non-rivalrous counterparts (Stein, Hebart, et al., 2011), suggesting that the central assumption of this control – that conditions without suppression are similar to the experimental conditions in everything but the suppression – is not valid.

The second solution is to ask participants to perform a task that is orthogonal to the experimental manipulation, with the assumption that participants will ignore task-irrelevant features and thus their RTs will not be affected by criterion differences (Gayet et al., 2016; Salomon et al., 2013). For example, Gayet et al. (2016) asked participants to report the orientation of Gabor patches (irrelevant manipulation) whilst testing for an effect of the patches' annulus colours on suppression times. Nonetheless, with this approach participants may still integrate aspects of ignored features into their criteria, which can impact breakthrough times. It is simply not possible to determine what factors will affect participants' criteria in a task that allows them to decide for themselves how much information to collect before committing to a response.

In Chapter 2, we developed a method to circumvent these issues by specifically disentangling perceptual sensitivity from decision criterion in b-CFS tasks. This method employs Signal Detection Theoretic (SDT) indices, which allow measurement of both detection and identification sensitivity, in addition to decision criteria for each (Macmillan & Creelman, 2004). Measuring detection and identification separately allows us to test for identification processes that may contribute to performance in a method where detection and identification are not measured individually, such as in b-CFS tasks. Our method employs predetermined increasing exposure durations, allowing the experimenter to control the amount of visual information participants receive. Therefore, unlike b-CFS, this method can track changes in different types of sensitivity and criterion as the available amount of information increases, thus providing a more comprehensive view of how different stimulus categories enter awareness.

Here we use this method to explore whether emotional expressions such as fearful, happy, and angry expressions break through suppression faster than neutral expressions, with an emphasis on the claim that fearful expressions are prioritised over neutral expressions (Yang et al., 2007). While the effect of fearful expressions has been replicated several times (Alpers & Gerdes, 2007; Capitão et al., 2014; Carlson & Reinke, 2008; Sterzer et al., 2011; Yang & Yeh, 2018), its interpretation is still unclear: As mentioned above, participants could have exhibited shorter breakthrough times for fearful faces due to a more liberal criterion rather than a difference in perceptual sensitivity – they could simply need less time to commit to report a fearful expression. Moreover, the concerns detailed above about the replicability of b-CFS studies mean that it is important to replicate this finding under more stringent conditions.

Thus, our studies aimed to better-understand the degree to which emotional expressions break through CFS faster than non-emotional expressions, and the mechanisms by which such an effect occurs. In our task, participants saw faces for predetermined exposure durations, and we varied their expression (fearful, angry, happy, and neutral) and, in some of the experiments below, their orientation (upright or inverted). At the end of each trial, participants reported, with a single keypress, both the location of the face (left or right) and its expression (emotional or non-emotional). We used these responses to determine participants' sensitivity and criterion for both face location (left/right side of screen – a measure of detection) and expression category (emotional/non-emotional expression – a measure of identification). In the first two experiments, we tested whether angry (Experiment 3) and fearful expressions (Experiment 4) break through suppression faster than happy and neutral expressions. In the second pair of experiments, we verified that our method is sufficiently sensitive to detect modulatory effects on breakthrough, by replicating the face-inversion effect (FIE) we had obtained in Experiment 2 of Chapter 2. The FIE is a detection or identification advantage of upright faces over inverted faces (Farah et al., 1995; Goodrich & Yonelinas, 2019; Rakover & Teucher, 1997), that has repeatedly found with b-CFS studies (Akechi et al., 2015; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011). Having verified in Chapter 2 that our method is sensitive enough to detect this effect, we examined whether the FIE is modulated by the emotional content of fearful (Experiment 5) and happy (Experiment 6) expressions, compared to neutral faces. In these experiments, we measured detection through a two-alternative forced-choice (2AFC) task by presenting the stimuli to one of two possible screen locations. In Experiment 7, we replaced our 2AFC task with a Yes-No detection task, in order to measure detection more closely to how it was measured in original b-CFS reports on unconscious emotion processing. Finally, in Experiment 8 we addressed the concern that our predetermined durations may be sufficiently distant from each other to miss – completely or partially – the relevant window of durations at which an effect of emotion arises; to prevent this, we used a staircase procedure to estimate breakthrough thresholds for fearful and neutral expressions presented both in upright and inverted orientations.

In summary, we present a methodologically rigorous assessment of whether emotional expressions break through CFS faster, and which factor – sensitivity, criterion, or both – might be responsible for such faster breakthrough. If emotional expressions reach awareness faster than non-emotional expressions, then we should find better sensitivity for the former than the latter. But even if we found such difference in sensitivity, it would not necessarily mean that sensitivity fully accounts for b-CFS findings. Differences in criterion could influence findings too. If criterion differences account for b-CFS findings on unconscious emotion processing, then we should find more liberal criteria for fearful and angry expressions than for other expressions, i.e. participants would be more willing to report fearful and angry expressions than other expressions. Furthermore, if b-CFS findings can be fully attributed to criterion differences, we should find no sensitivity differences whereas if they can only be partly attributed to criterion differences, we should find both a more liberal criterion and a higher sensitivity for fearful expressions when compared to other expressions. In summary, these studies allowed us to determine to what extent sensitivity and criterion, measured separately, could have contributed to b-CFS findings on unconscious emotion processing.

## 3.2 Experiment 3

In this experiment, participants saw multiple blocks of faces presented for different durations, always masked by CFS. Each block contained either angry and neutral expressions, or happy and neutral expressions. If angry expressions genuinely break suppression faster, then sensitivity to angry faces should be greater than either happy or neutral expressions, probably between 2 and 3 s in average, as shown by breakthrough times in b-CFS studies (Gray et al., 2013; Yang et al., 2007).

### 3.2.1 Method

#### *3.2.1.1 Participants*

Thirty-four University of Edinburgh students provided informed consent and were paid £14 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Two participants were excluded from analysis (see Analysis section below): the remaining 32 participants (17 female; 2 left-handed) had a mean age of 23.5 ($SD_{age}$ = 3.8). All the studies reported here were approved by the Psychology Research Ethics Committee of the University of Edinburgh.

Past b-CFS studies that have found statistically significant effects of emotional expression on breakthrough times employed around 16 participants per experiment (e.g. Yang, et al., 2007). We decided to double this number to increase power and allow counterbalancing of experimental blocks with a multiple of 8 (see Procedure section). A retrospective power analysis, conducted using G*Power 3.1.9.7 (Faul et al., 2009), to test for a difference between conditions in a repeated-measures ANOVA, with a small to medium effect size ($\eta p^2 = .04$) and alpha of .05, aiming to achieve a statistical power of 95%, determined that a sample of 19 participants would be required. If a non-sphericity correction ε of .5 were to be added – as reported in the results section, a number of tests violated this assumption – then a sample of 29 participants would be required.

#### 3.2.1.2 *Stimuli and Apparatus*

In all experiments, stimuli were presented on a 19-inch CRT monitor in a dimly lit room, connected to a computer running Matlab 2014a (Mathworks, Inc) using the Cogent 2000 toolbox (http://www.vislab.ucl.ac.uk/cogent.php). A chin rest and mirror stereoscope were positioned 57 cm from the monitor, with a vertical divider splitting the display so each eye only saw half of the screen.

To maintain binocular alignment, two vertical textured vergence bars (width 1°, height 8°) appeared to the left and right of fixation in each eye (horizontal centre-to-centre distance 3.1°) such that only one pair of vergence bars was perceived. A black fixation cross (0.7° × 0.7°) was presented in the centre of each pair of vergence bars. Grey Mondrian-like masks were flashed at 10 Hz to one eye while a face stimulus was introduced to the other eye.

Stimuli were 60 human faces taken from the Karolinska Directed Emotional Faces (KDEF) database (Goeleven et al., 2008), all seen from a front angle, and classified as either angry (20), happy (20), or neutral (20) expressions. Images were cropped to show only the internal facial features and transformed to greyscale. Luminance was equated for all the resulting images using the Matlab SHINE toolbox (Willenbockel et al., 2010). Finally, background colour was replaced with uniform grey matching the screen's background colour (see Appendix B). Stimuli were then sorted into 6 different sets, 2 for each facial expression. Using the KDEF norms, we matched sets for expression identification, intensity, and arousal, as well as gender. Our stimuli present two advantages over those used in prior studies: they are more numerous (previous studies used between 4 and 8, Capitão et al., 2014; Gray et al., 2013; Yang et al., 2007; Yang & Yeh, 2018) and are matched in expression identification ($M_{\text{angry}}$ = 85%; [$SD_{\text{angry}}$ = 16.87]; $M_{\text{happy}}$ = 94.6% [6.14]; $M_{\text{neutral}}$ = 83.02% [8.36]) and intensity ($M_{\text{angry}}$ = 5.68; [0.93]; $M_{\text{happy}}$ = 5.85 [0.69]; $M_{\text{neutral}}$ = 5.15 [0.41]; not all b-CFS studies did this (e.g. Capitão et al., 2014 and Yang et al., 2007 did not).

#### 3.2.1.3 *Procedure*

Each trial began with a fixation cross, presented binocularly at the centre of each eye's visual field between two vergence bars, allowing participants to maintain stable binocular vergence. After 200 ms, the fixation cross remained superimposed on both screens, but a changing Mondrian-like mask was presented to one eye; 200 ms later a face image was introduced to the other eye either to the left or to the right side of the fixation cross (Figure 3.1a). For half of the participants, their left eye was assigned to receive the mask and their right eye to receive the stimulus. Following the procedure developed by Yang et al. (2007), the face image's contrast ramped up linearly from 0% to 25% over 1 second (s) after which its contrast remained stable. Emotional expression was blocked (70 trials/block) with block order counterbalanced across participants in an ABBABAAB BAABABBA order. Thus, participants went through blocks of angry/neutral and happy/neutral expressions. Stimuli were presented for one of seven predefined durations, spaced equally on a log scale (600; 900; 1350; 2025; 3038; 4557; 6836 ms), with an equal number of trials for each duration. This range of exposure durations was determined in piloting sessions that used exposures of 500 – 8000 ms. On trials in which stimulus presentation was shorter than 1 s, termination of presentation curtailed the change in contrast. On longer trials, face contrast remained at 100% until the end of the trial. The Mondrian-like masks consisted of fields of grey circles with randomly differing sizes, grey-levels, rotation, and position, changing at a rate of 10 Hz. After 1 s, the mask's contrast began decreasing linearly until reaching zero at 6 s (Figure 3.1b).

Participants were instructed to focus on the fixation cross with both eyes open, trying to avoid blinking and looking around. After the end of each stimulus presentation, four question marks were presented on the screen, replacing the fixation cross. Participants were instructed to report the location of the face image and whether its expression was emotional or non-emotional, by pressing one of four keys (left control and left shift to report that the face was on the left, down arrow and up arrow to report it was on the right; the top or bottom key on each side was used to denote an emotional face, with the other denoting a neutral expression. The keys used to identify emotional expression were counterbalanced across participants, Figure 3.1c). The next trial began after a response was given or, if the participant did not respond, after a 2 s response window. Following this response, a screen showing only the vergence bars was presented for an ITI of 1000 ms before the next trial began. Participants were given self-terminated breaks every 70 trials and a compulsory 15-minute break halfway through the experiment. Before beginning the experiment, participants completed 60 training trials to ensure that the stereoscope was properly calibrated and that they had understood the task, as done in some past CFS studies (Gayet et al., 2020; Schlossmacher et al., 2017; Wormwood et al., 2019; Yang et al., 2007; Zhou et al., 2010).

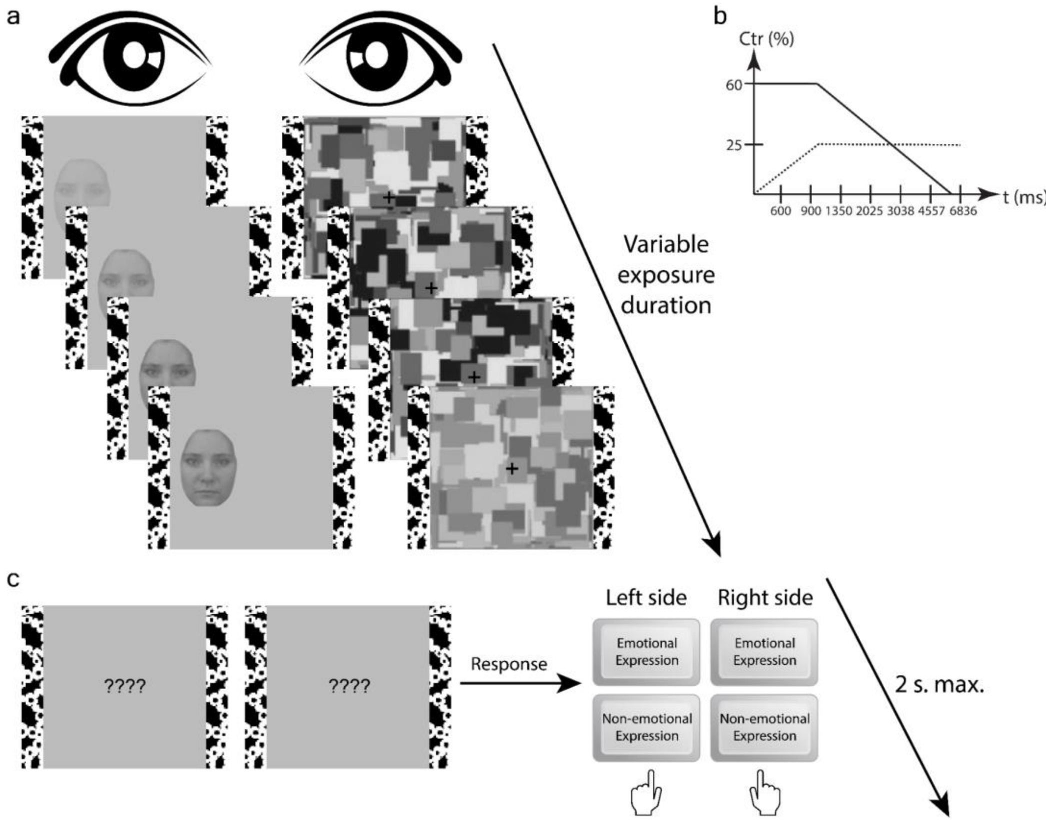



**Figure 3.1.** Schematic description of a trial in experiments 3 through 6. (a) Stimulus introduction and CFS. The duration of each display was chosen from seven equally likely exposure durations: 600, 900, 1350, 2025, 3038, 4557, or 6836 ms. (b) Changes in contrast. The contrast of the target image increased linearly until reaching 25% contrast at 1 s. Then, contrast remained unchanged until the end of the display. Mondrian-like patterns started at 60% contrast until reaching 1 s, after which contrast linearly decreased until reaching 0% at 6 s of exposure. (c) Detection/identification task. Following stimulus offset, participants were presented with a binocular response cue composed of question

marks. Participants provided a single response to indicate both on which side of the fixation the face had been shown and whether its expression was emotional or neutral.

### 3.2.1.4 Analysis

We excluded data from two participants, one who failed to respond on more than 5% of trials, and one who did not become more accurate on the task as exposure time increased (accuracy at chance level), suggesting that they failed to attend to the task. Trials that received no response were treated as missing data.

To assess how perceptual sensitivity to visual information and response criteria change across display durations, we calculated signal-detection measures.

First, we examined reports on the faces' location. To determine each participant's bias-independent sensitivity to face location (location d') for each combination of duration and emotional expression, we employed the calculation for 2AFC tasks (Macmillan & Creelman, 2004), $d'_{location} = \left(\frac{1}{\sqrt{2}}\right)\left(Z(Hit_{location}) - Z(FA_{location})\right)$, where Z(Hit) stands for the Z score associated with the probability of a Hit (defined as a trial in which a face was displayed on the right and reported on the right), and Z(FA) for the Z score associated with the probability of a false alarm (a trial in which a face was displayed on the left but reported on the right). To estimate each participant's bias to respond left or right (henceforth referred to as response bias) during face location, we employed the calculation $C_{location} = -\left(\frac{1}{2}\right)\left(Z(Hit_{location}) + Z(FA_{location})\right)$. Positive and negative values for this measure indicate a bias toward responding "left" and "right", respectively; however, as these may cancel out across participants, we converted the results to absolute values as a measure of overall amount of response bias.

Second, we analysed judgments of whether an emotional expression had been shown, defining emotional expression (angry or happy, depending on the block) as the signal, and neutral expressions as signal-absent. To determine expression identification sensitivity, we used the calculation of d' for Yes/No detection tasks, $d'_{identification} = Z(Hit_{identification}) - Z(FA_{identification})$, where a hit was defined as correctly reporting an emotional expression and an FA was defined as incorrectly reporting a neutral expression as emotional. To estimate each participant's criterion during expression identification, we employed the calculation $C_{identification} = -\left(\frac{1}{2}\right)\left(Z(Hit_{identification}) + Z(FA_{identification})\right)$, where positive/negative values indicate a greater bias to report seeing an emotional expression.

Thus, location d', response bias, expression identification d', and emotion decision criterion were estimated for each combination of emotional expression and exposure duration. Subsequently, a repeated-measures analysis of variance (ANOVA) was run on each measure to test for differences between emotion conditions as exposure duration increased. Wherever Mauchly's test indicated that the assumption of sphericity was violated, degrees of freedom were corrected using Greenhouse-Geisser estimates. Where null results were of theoretical interest, we calculated Bayes factors to evaluate the strength of the evidence for the null.

Both frequentist (ANOVA and t-tests) and Bayesian (Bayes factors) statistical analyses were performed using Jamovi (The jamovi project, 2020) and JASP (JASP Team, 2020), and corroborated using R and SPSS. When an ANOVA indicated a significant interaction, we ran post hoc Bonferroni-corrected pairwise comparisons to look for significant effects. Post hoc pairwise comparisons in these statistical packages use estimated marginal means based on the variance of the ANOVA model. For Bayes factor analysis, we defined the null hypothesis as no difference between conditions by using a standard Cauchy distribution centred on zero with rate of 0.707.

### 3.2.2 Results

#### *3.2.2.1 Location sensitivity*

We calculated mean location d' scores for angry, happy, and neutral expressions. Since we obtained two different d' scores for neutral expressions as there were twice the amount of neutral-expression trials than of angry-expression and happy-expression trials, we collapsed neutral expression trials into one condition, making three conditions in total

($M_{angry} = 1.196$ [1.027]; $M_{happy} = 1.236$ [1.053]; $M_{neutral} = 1.178$ [1.026]). To confirm that our decision to collapse neutral trials was justified, we compared whether location d' scores for neutral trials did not differ across the two block types, using a 2 (neutral-expression trial groups) × 7 (exposure durations) into a repeated-measures ANOVA. There was no effect of trial group ($F_{(1, 31)} = 0.520, p = .476, \eta p^2 = .017$). However, given that absence of a significant effect cannot be taken as evidence of a null effect, we also estimated Bayes factors. The results suggest that the data are substantially better explained under the null hypothesis model ($\mathrm{BF}_{01} = 9.317$), thus supporting our decision to collapse neutral trials.

In the main analysis, we examined whether participants' sensitivity to the location of suppressed faces varied across the emotion conditions and exposure durations by entering location d' data into a 3 (expression: angry, happy, neutral) × 7 (exposure durations) repeated-measures ANOVA. As Figure 3.2a shows, location d' increased with increasing exposure duration ($F_{(2.77, 85.74)} = 104.537, p < .001, \eta p^2 = .771$), such that participants were close to ceiling at the longest exposure durations. Our key question, however, was whether sensitivity levels would differ across the different facial expressions, either overall or in interaction with exposure duration. However, we did not find a main effect of expressions, meaning that the emotional expression of the faces did not significantly affect participants' sensitivity ($F_{(1.73, 53.78)} = 2.422, p = .105, \eta p^2 = .072$). The interaction between expression and exposure duration did not reach significance either ($F_{(7.71, 239)} = 0.831, p = .572, \eta p^2 = .026$). To quantify whether these latter null results provide evidence for the null hypothesis, we used Bayes factors. We found very strong evidence for the null hypothesis model ($BF_{01} = 46.825$), i.e. no effect of expression on location d'. Thus, this first analysis provided no evidence that faces with emotional expressions have preferential access to awareness.

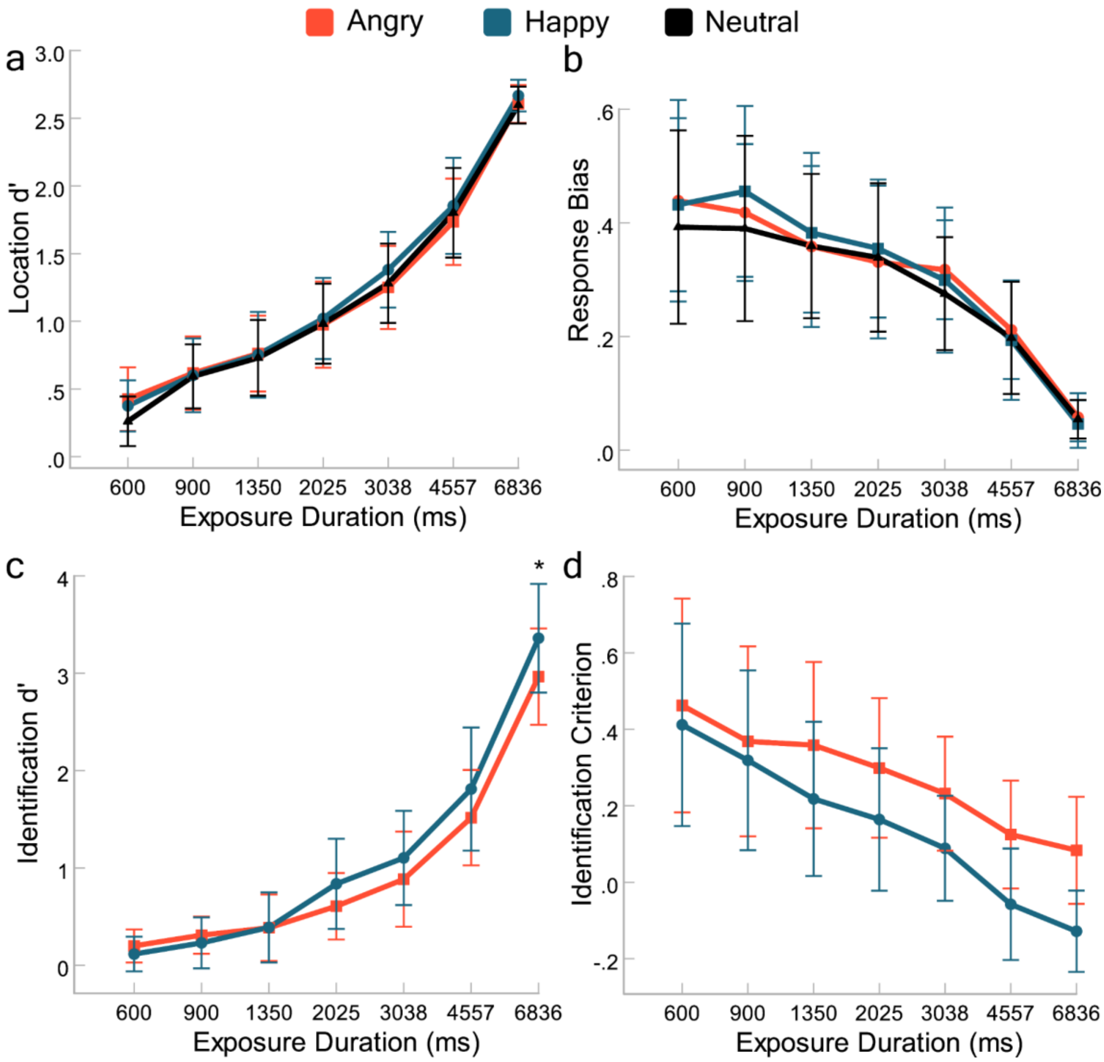



**Figure 3.2.** Results of Experiment 3. (a) Location sensitivity. Location d' increased with exposure duration, but there was no difference among expressions. (b) Response bias. Absolute-value response bias scores for reporting location decreased with increasing exposure durations, but there was no difference between expressions. (c) Expression identification sensitivity. Expression identification d' increased with exposure duration; there was higher sensitivity to happy faces than angry faces at 6836 ms of exposure. (d) Expression identification criterion. Identification criterion became more liberal with increasing exposure durations; happy faces exhibited a significantly more liberal criterion than angry faces. Asterisks index statistically significant differences ($p < .05$) between emotional expressions. Error bars represent 95% CI.

### 3.2.2.2 *Location Response Bias*

d' is a bias-independent measure of discrimination between signal and noise – it is not affected by any response bias participants might have, for instance a bias towards responding 'right'. Thus, we next examined whether the stimuli may have affected participants' biases. Specifically, we entered the absolute values of response bias scores into a 3 (expression: angry, happy, neutral) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration – response bias significantly decreased with increasing exposure durations suggesting that response bias decreases as visibility increases ($F_{(1.97, 61.17)} = 14.515, p < .001, \eta p^2 = .319$), (Figure 3.2b). However, we did not find a main effect of emotional expression, suggesting that emotional expression did not affect response bias ($F_{(1.88, 58.24)} = 1.198, p = .307, \eta p^2 = .037$). To assess whether the evidence supports this null hypothesis, we estimated Bayes factors, which indicated very strong support for the null hypothesis of expression, i.e. no effect of emotional expression on response bias ($BF_{01} = 37.885$). Similarly, we estimated Bayes factors to assess whether the evidence supports a null hypothesis of no interaction. We found extreme support for the null hypothesis of no interaction ($BF_{01} > 100$).

### 3.2.2.3 *Expression identification sensitivity*

We examined whether participants' sensitivity to identifying the emotional expression of the suppressed face varied across the two emotions assessed and exposure duration conditions. Recall that here, identification sensitivity describes the ability to distinguish emotional expressions from neutral expressions. We entered identification d' scores to a 2 (emotional expression: angry, happy) × 7 (exposure durations) repeated-measures ANOVA. As Figure 3.2c shows, identification d' scores increased with increasing exposure duration ($F_{(2.69, 83.29)} = 81.23, p < .001, \eta p^2 = .724$), such that participants were close to ceiling at the longest exposure durations. In addition, we found a main effect of expression ($F_{(1, 31)} = 4.47, p = .043, \eta p^2 = .126$): happy expressions exhibited an advantage over angry expressions ($M_{angry} = 0.981$ [1.387]; $M_{happy} = 1.121$ [1.619]). The interaction between expression and exposure duration also reached significance $(F_{(4.21,\,130.51)} = 4.37, p = .002, \eta p^2 = .124)$. To examine this interaction in more detail, we ran post hoc Bonferroni-corrected pairwise comparisons between the two expressions at each exposure duration. These tests revealed that the advantage of happy expressions over angry expressions was only significant at 6836 ms of exposure $(t(149) = -3.672,\ p = .03,\ d = -0.649)$. Thus, while emotional expressions did not affect judgments of where a face was, they did affect the ability to judge what expression the face showed: there was greater sensitivity to the emotional content of happy expressions than angry expressions at long exposure durations.

#### 3.2.2.4 *Expression decision criterion*

As discussed in the Introduction, one potential explanation for past findings that emotional expressions have preferential access to awareness, is that emotional expressions like anger might enjoy a more liberal identification criterion than positive ones. Thus, we tested how participant's identification criteria varied across expressions and exposure durations. We entered identification criterion scores into a 2 (expression: angry, happy) × 7 (exposure durations) repeated-measures ANOVA. Participants' willingness to report an emotional expression (indexed by lower criterion scores) increased with increasing exposure durations, as evidenced by a main effect of exposure duration $(F_{(1.65,\,51.03)} = 7.66, p = .002, \eta p^2 = .198)$. Crucially, participants also exhibited a significantly more liberal criterion for happy expressions than for angry expressions, as shown by the main effect of expression $(M_{angry} = 0.276\ [0.564]; M_{happy} = 0.145\ [0.549]; F_{(1,\,31)} = 19.80, p < .001, \eta p^2 = .390)$, (Figure 3.2d). The interaction did not reach significance $(F_{(3.95,\,122.49)} = 1.71, p = .153, \eta p^2 = .052)$. To assess whether the evidence supports this null effect, we estimated Bayes factors, which indicated very strong support for the null hypothesis, i.e. no significant interaction between expression and exposure duration $(\mathrm{BF}_{01} = 63.502)$.

These results demonstrate that participants exhibited a more liberal criterion for the identification of happy expressions than of angry expressions.

### 3.2.3 *Discussion*

Experiment 3 tested whether emotional expressions enjoy priority over non-emotional ones during unconscious processing, using a methodologically stricter variant on the breaking Continuous Flash Suppression paradigm. This method allowed us to control for a concern about previous studies which have reported shorter breakthrough times for emotional over non-emotional expressions – that these results could be due to criterion differences rather than perceptual sensitivity. If that were the case, with this more stringent procedure we should expect no difference between emotional and non-emotional expressions in perceptual sensitivity. On the other hand, a more liberal identification criterion should be found for emotional expressions such as angry faces. Our data supported this latter prediction: sensitivity to the location of a suppressed face did not vary based upon its emotional expression.

By asking participants to identify each face's expression, we measured their identification sensitivity. Surprisingly, we found that participants exhibited better sensitivity for happy expressions than angry expressions, i.e. the former were easier to identify from their neutral counterparts than the latter. Based on past b-CFS findings, which showed shorter breakthrough times to negative expressions over neutral and positive ones, we expected both better identification sensitivity and a more liberal criterion for angry expressions than happy ones. However, it is important to note that the identification advantage of happy expressions over angry ones only reached significance at the longest exposure duration, once CFS masks had disappeared. Hence, this effect could have been driven by inherent differences between happy and angry expressions in the KDEF stimuli set – despite having minimised these differences during stimuli selection, the happy expressions selected had higher emotion recognition than angry ones, which fits this stimuli set's norms (see: Calvo & Lundqvist, 2008; Goeleven et al., 2008). Nevertheless, the fact that we found a more liberal decision criterion for happy over angry expressions – meaning that participants were more willing to report a happy expression than an angry expression – reinforces the idea that post-perceptual factors could affect detection responses. As argued in the Introduction, b-CFS response times do not distinguish between detection and identification – participants decide for themselves when to give a response. If their response times are confounded by identification processes, then differences in decision criterion for expression identification may affect those response times, too. Thus, our data support the importance of disentangling detection from identification by showing that participants can exhibit differences in decision criteria.

Our null finding regarding location sensitivity could indicate that no type of emotional expression has preferential access to awareness, however, it is also possible that the particular emotional expressions used here – happy versus angry – may not be ideal for eliciting the intended effect. In particular, the meta-analysis by Hedger et al. (2016) found that fearful expressions presented a more robust and consistent effect in b-CFS studies searching for differential unconscious emotion processing. Therefore, we replicated this study using fearful expressions rather than angry expressions.

## 3.3 Experiment 4

In this experiment, we tested whether fearful expressions make faces break through suppression faster than happy and neutral expressions. If fearful expressions enjoy an unconscious processing advantage over happy and neutral expressions as repeatedly reported in the literature (Capitão et al., 2014; Hedger et al., 2015; Oliver et al., 2015; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018; Zhan et al., 2015), higher detection and/or identification sensitivity should be found for fearful expressions than for the other expressions. If b-CFS findings were due to criterion differences, we may find more liberal identification criterion (i.e. willingness to report an emotional expression) for fearful expressions than for neutral expressions. As in the previous experiment, the hypotheses concerning sensitivity and criterion are not mutually exclusive.

### 3.3.1 Method

#### 3.3.1.1 *Participants*

Thirty-two students of the University of Edinburgh (19 female; 3 left-handed), with a mean age of 21.4 ($SD_{age}$ = 4.2), participated for a payment of £14.

#### 3.3.1.2 *Stimuli and Procedure*

The methods were the same as in Experiment 3, except that fearful faces were employed instead of angry faces. These images were again selected from the KDEF database and matched to the happy and neutral faces on gender and expression identification while minimising differences in expression identification ($M_{fearful}$ = 71%; [7.44]; $M_{happy}$ = 96.09% [4.49]; $M_{neutral}$ = 65.1% [18.25]) and intensity ($M_{fearful}$ = 5.82; [0.61]; $M_{happy}$ = 6.13 [0.84]; $M_{neutral}$ = 4.85 [0.66]; Appendix B). They were all then equated in luminance with the Matlab SHINE toolbox, following the same processing procedure as in Experiment 3.

### 3.3.2 Results

#### 3.3.2.1 *Location sensitivity*

As in Experiment 3, we calculated mean location d' scores for each condition. We started off by collapsing neutral trials into one condition as they did not differ between each other[3].

[3] Location d' scores again were entered into a 2 (neutral-expression trial groups) × 7 (exposure durations) repeated-measures ANOVA. As expected, we found a main effect of exposure duration ($F_{(2.39, 74.19)} = 112.504, p < .001, \eta p^2 = .784$), but we did not find a main effect of trial group ($F_{(1, 31)} = 0.103, p = .750, \eta p^2 = .003$). We did not find an interaction between trial group and exposure duration either

For the main analysis, we examined whether participants' sensitivity to the location of the suppressed face varied across expression ($M_{fearful} = 1.674$ [$SD = 0.941$]; $M_{happy} = 1.673$ [0.947]; $M_{neutral} = 1.663$ [0.942]) and exposure duration conditions. To this end, we entered location d' scores into a 3 (expression: fearful, happy, neutral) × 7 (exposure durations) repeated-measures ANOVA. As seen in Figure 3.3a, location sensitivity dramatically increased with increasing exposure duration ($F_{(2.09, 64.82)} = 111.961, p < .001, \eta p^2 = .783$). Neither the effect of expression ($F_{(1.84, 57.15)} = 0.103, p = .887, \eta p^2 = .003$) nor the interaction between expression and exposure duration ($F_{(6.81, 211.26)} = 1.042, p = .402, \eta p^2 = .033$) reached significance. These results suggest emotional expressions do not affect participants' detection sensitivity. To quantify whether these latter null results provide evidence for the null hypothesis of expression (i.e. no difference between expressions), we estimated Bayes factors. We found very strong evidence in favour of the null hypothesis of expression ($BF_{01} = 58.586$), i.e. no difference among emotional expressions in location sensitivity. We also estimated Bayes factors for the null hypothesis model of no interaction between expression and exposure duration. We found extreme evidence in favour of the null hypothesis of no interaction ($BF_{01} > 100$).

Therefore, these results suggest that emotional expressions do not enjoy higher sensitivity than non-emotional ones entering awareness.

($F_{(3.84, 119.12)} = 0.048, p = .995, \eta p^2 = .002$). We also estimated Bayes factors; we defined the null hypothesis as no difference between neutral-expression conditions. Again, the results suggest the data are substantially better explained under the null hypothesis model ($BF_{01} = 9.57$). We also estimated Bayes factors for the null hypothesis of no interaction between trial group and exposure durations and found that the data are extremely better explained under it ($BF_{01} > 100$). Thus, the decision of collapsing neutral-expression trials was justified.

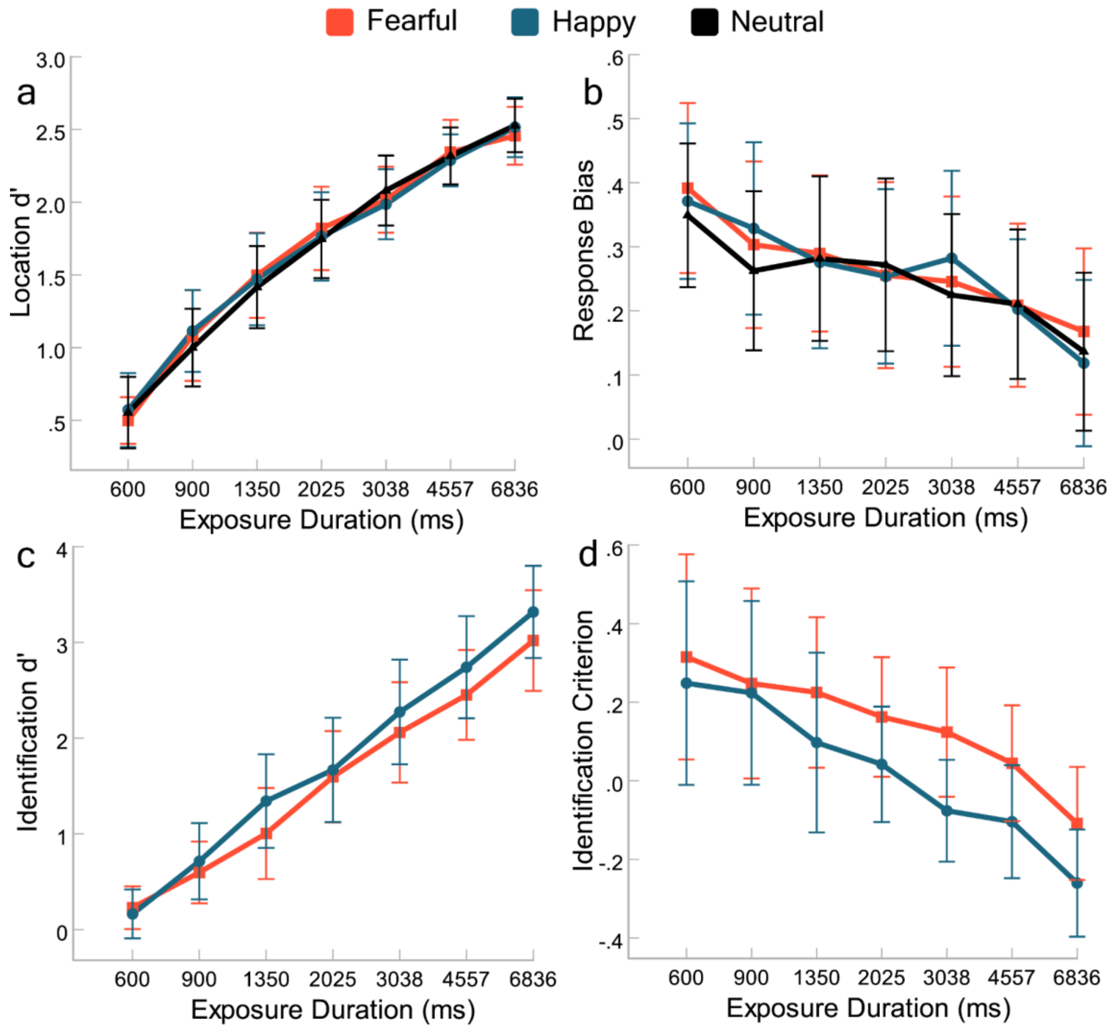



**Figure 3.3.** Results of Experiment 4. (a) Location sensitivity. Location d' increased with exposure duration, but there was no difference among expressions. (b) Response bias. Absolute-value response bias scores for reporting location decreased with increasing exposure durations, but there was no difference among expressions. (c) Expression identification sensitivity. Expression identification d' increased with exposure duration; happy faces exhibited higher sensitivity than fearful faces. (d) Expression identification criterion. Identification criterion became more liberal with increasing exposure durations; happy faces exhibited a significantly more liberal criterion than fearful faces. Error bars represent 95% CI.

### 3.3.2.2 *Location response bias*

We examined whether participants' location responses varied across conditions by entering the absolute values of response bias scores into a 3 (expression: fearful, happy, neutral) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration – response bias significantly decreased with increasing exposure durations, suggesting that response bias decreases as visibility increases ($F_{(2.84,\,88.13)} = 8.001, p < .001, \eta p^2 = .205$), (Figure 3.3b). However, we did not find a main effect of expression, suggesting that expression did not affect response bias ($F_{(1.74,\,53.90)} = 0.562, p = .550, \eta p^2 = .018$) or an interaction between expression and exposure duration ($F_{(8.19,\,253.89)} = 0.753, p = .647, \eta p^2 = .024$). Bayes factors analyses indicated very strong support for the null hypothesis of no effect of expression ($BF_{01} = 35.461$) and extreme support for the null hypothesis of no interaction between expression and exposure duration on response bias ($BF_{01} > 100$).

### 3.3.2.3 *Expression identification sensitivity*

We examined whether participants' identification sensitivity varied across the expression ($M_{fearful} = 1.566\ [1.537]; M_{happy} = 1.746\ [1.661]$) and exposure duration conditions. We entered identification d' scores into a 2 (expression: fearful, happy) × 7 (exposure durations) repeated-measures ANOVA. Similar to Experiment 3, identification d' scores dramatically increased with increasing exposure duration ($F_{(2.41,\,74.7)} = 76.89, p < .001, \eta p^2 = .713$), such that participants were close to ceiling at the longest exposure duration (Figure 3.3c). In addition, we found a main effect of expression ($F_{(1,\,31)} = 9.86, p = .004, \eta p^2 = .241$), with an advantage of happy expressions over fearful expressions. The interaction between expression and exposure duration also reached significance ($F_{(5.32,\,165.01)} = 2.34, p = .04, \eta p^2 = .07$). Inspection of Figure 3.3c suggests that identification effect was larger at longer durations, but Bonferroni-corrected pairwise comparisons did not reveal significant differences between expressions for any of the exposure durations. Therefore, emotional expressions did affect the ability to judge what expression the face showed: there was greater sensitivity to the emotional content of happy expressions than angry expressions.

#### 3.3.2.4 *Expression decision criterion*

As in Experiment 3, we tested whether participant's identification criteria varied across expressions and exposure durations. We entered expression criterion scores into a 2 (expression: fearful, happy) × 7 (exposure durations) repeated-measures ANOVA. Participants' willingness to report an emotional expression (indexed by lower criterion scores) increased with increasing exposure durations, as evidenced by a main effect of exposure duration $(F_{(1.44, 44.58)} = 7.54, p = .004, \eta p^2 = .196)$, (Figure 3.3d). We also found evidence of a more liberal identification criterion for happy expressions than for fearful expressions $(M_{fearful} = 0.276\ [0.539]; M_{happy} = 0.145\ [0.546]; F_{(1, 31)} = 16.08, p < .001, \eta p^2 = .342)$. The interaction between expression and exposure duration did not reach significance $(F_{(5.09, 157.94)} = 1.72, p = .131, \eta p^2 = .053)$. To assess whether the evidence supports this null interaction, we estimated Bayes factors, which indicated very strong support for the null hypothesis, no significant interaction between expression and exposure duration $(BF_{01} = 69.572)$.

These results suggest that participants employed a more liberal criterion during the identification of happy expressions than of fearful expressions.

### 3.3.3 *Discussion*

Experiment 4, like Experiment 3, did not find an advantage in location sensitivity for emotional expressions, this time using fearful instead of angry expressions. Importantly, the fact that neither angry nor fearful expressions gained access to awareness faster than happy or neutral expressions suggests that negative threatening faces do not enjoy priority in how they enter awareness.

However, and perhaps surprisingly, expression identification again exhibited higher sensitivity and a more liberal criterion for happy expressions than for fearful expressions. These results reinforce the idea that emotional expressions may not affect faces' breakthrough times when suppressed by CFS.

Similar to Experiment 3, Experiment 4 found better identification sensitivity for happy expressions than fearful expressions, and a more liberal decision criterion for happy expressions than fearful ones. Again, as argued in Experiment 3, these results stress the importance of disentangling detection from identification so post-perceptual factors that may affect the latter do not confound the former.

Whilst the results of experiments 3 and 4 are consistent with each other, we cannot rule out the possibility that our procedure may not be sufficiently sensitive to capture effects of emotional expression on detection (location sensitivity). In order to test whether this is the case or not, we decided to run an experiment that allowed us to simultaneously test for an effect of emotional expression while also attempting to capture a different face-related effect that has been replicated several times in the b-CFS literature: the face-inversion effect (FIE). Multiple b-CFS studies have shown that upright faces have shorter breakthrough times than inverted faces (Akechi et al., 2015; Gayet & Stein, 2017; Jiang et al., 2007; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011; Zhou et al., 2010). Thus, combining a test of emotional expression with a test of face inversion can help us to calibrate the sensitivity of our method.

In summary, our data indicate that location sensitivity increases with exposure duration, however, we were not able to replicate the emotion effect – of negative expressions (fearful expressions in this experiment) breaking through CFS faster than neutral expressions – reported by Yang et al. (2007). In other words, the probability of a face to break into awareness increased alongside exposure, but there was no differential access to awareness in favour of fearful expressions. To calibrate the sensitivity of our procedure and thus make sure it is sensitive to face processing effects, we add a face orientation manipulation to test the FIE in the next experiment.

## 3.4 Experiment 5

In this experiment, we only used fearful and neutral facial expressions, but presented both in upright and inverted orientations. If our method is indeed sensitive to the properties of these stimuli, then we expect to see inversion affecting location judgments.

### 3.4.1 Method

#### *3.4.1.1 Participants*

We recruited thirty-two University of Edinburgh students (17 female; 4 left-handed) of mean age of 23.1 ($SD_{age}$ = 3.9), who were paid £14 for their participation.

#### *3.4.1.2 Stimuli and Procedure*

The methods were the same as in Experiment 3, except that we employed fearful (instead of angry) and neutral faces, in upright and inverted orientations. We selected all the stimuli images from the KDEF database again to ensure they could be matched on the criteria described in Experiment 3 while minimising differences in expression identification ($M_{\text{fearful}}$ = 73.1%; [9.82]; $M_{\text{neutral}}$ = 73.06% [13.34]) and intensity ($M_{\text{fearful}}$ = 5.9; [0.78]; $M_{\text{neutral}}$ = 4.98 [0.54]; Appendix B).

### 3.4.2 Results

#### *3.4.2.1 Location sensitivity*

We examined whether participants' sensitivity to the location of the suppressed face varied between emotional expression ($M_{fearful} = 1.511$ [0.981]; $M_{neutral} = 1.553$ [1.011]), orientation ($M_{upright} = 1.623$ [0.981]; $M_{inverted} = 1.440$ [1.009]), and exposure duration conditions. To this end, we entered location d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure duration) repeated-measures ANOVA. As seen in Figure 3.4a, location d' scores dramatically increased with increasing exposure duration ($F_{(2.05, 63.47)} = 185.273, p < .001, \eta p^2 = .857$). As in Experiments 3 and 4, we did not find a main effect of expression ($F_{(1, 31)} = 2.03, p = .164, \eta p^2 = .061$), and a Bayes factors suggested strong evidence in favour of the null hypothesis ($BF_{01} = 10.551$). On the other hand, we did find a main effect of orientation ($F_{(1, 31)} = 50.409, p < .001, \eta p^2 = .619$), with an advantage for upright faces over inverted faces. Therefore, while we did not find an effect of expression on location sensitivity, we did find an effect of the face orientation on this sensitivity. In addition, the interaction between orientation and exposure time was significant ($F_{(4.53, 140.29)} = 3.576, p = .006, \eta p^2 = .103$). We ran post hoc Bonferroni-corrected pairwise comparisons to determine specific significant differences between upright and inverted faces per exposure duration, and found such differences at 1350 ms ($t(209) = 5.182,\ p < .001,\ d = 0.916$) and 3038 ms ($t(209) = 5.596,\ p < .001,\ d = 0.989$) of exposure. Finally, we did not find an interaction between expression and exposure duration ($F_{(4.17, 129.30)} = 0.576, p = .688, \eta p^2 = .018$), expression and orientation ($F_{(1, 31)} = 2.753, p = .107, \eta p^2 = .082$), or between these three factors ($F_{(4.77, 147.95)} = 0.590, p = .70, \eta p^2 = .019$). Bayes factors suggested extreme evidence in favour of the null interaction between expression and exposure duration ($BF_{01} > 100$) and very strong evidence in favour of the null three-way interaction ($BF_{01} = 88.297$).

These results thus replicate the FIE, demonstrating that our method is sensitive to holistic face processing, and they continue to suggest that fearful expressions do not break through suppression faster than neutral expressions.

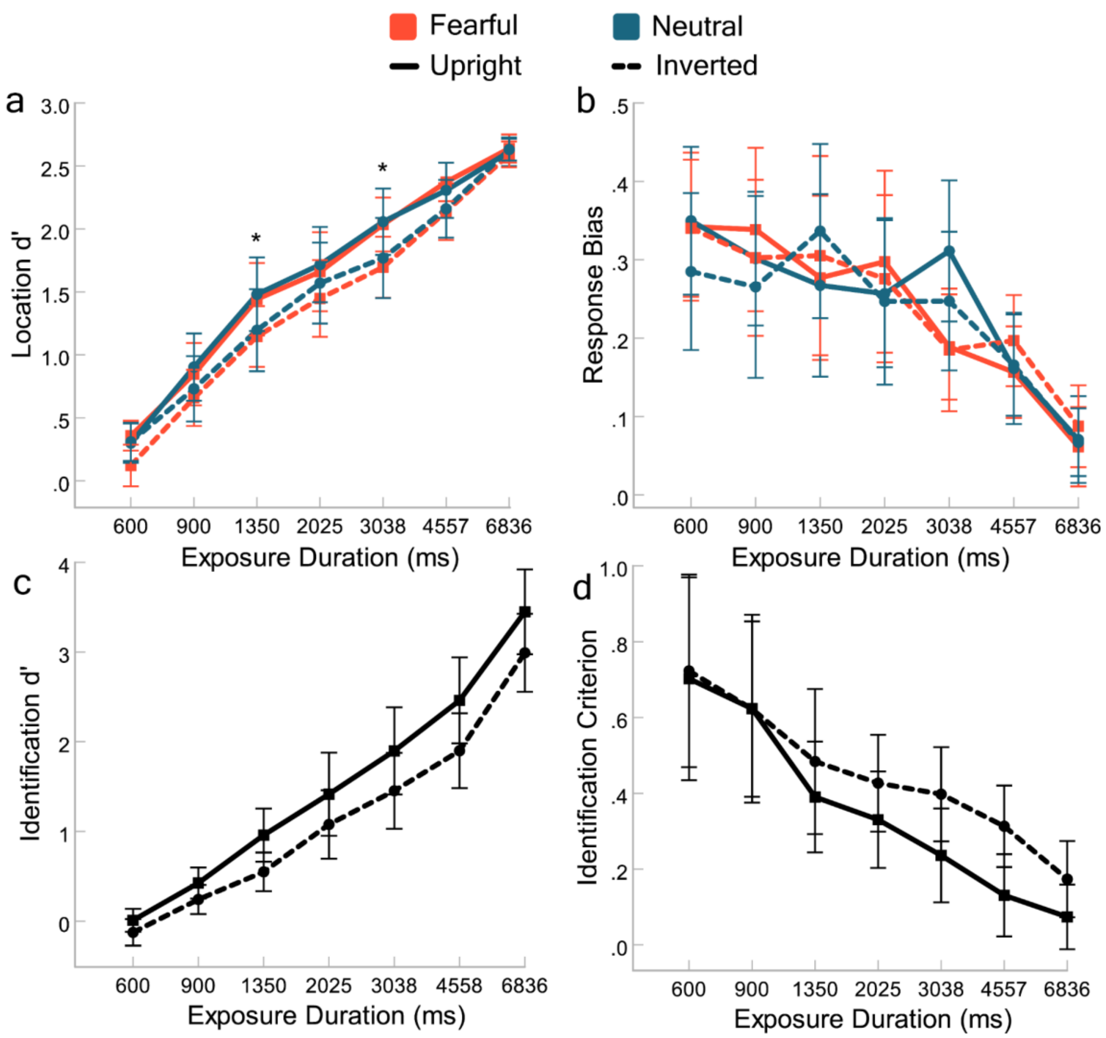



**Figure 3.4.** Results of Experiment 5. (a) Location sensitivity. Location d' increased with exposure duration. There was an advantage of upright faces over inverted faces, but there was no difference between expressions. (b) Response bias. Absolute-value response bias scores for reporting location decreased with increasing exposure duration, but there was no difference between expressions or between orientations. (c) Expression identification sensitivity. Expression identification d' increased with exposure duration; upright faces exhibited higher sensitivity than inverted faces. (d) Expression identification criterion. Identification criterion became more liberal with increasing exposure durations; upright faces exhibited a significantly more liberal criterion than inverted faces. Asterisks index

statistically significant differences (p < .05) between face orientations. Error bars represent 95% CI.

### 3.4.2.2 Location response bias

We examined whether participants' location responses varied across conditions by entering the absolute values of response bias scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration: response bias significantly decreased with increasing exposure durations, suggesting that response bias decreases as visibility increases ($F_{(2.82,\, 87.42)} = 16.202, p < .001, \eta p^2 = .343$), (Figure 3.4b). However, we found no main effect of emotional expression on response bias ($F_{(1,\, 31)} = 0.025, p = .874, \eta p^2 = .001$). We did not find a main effect of face orientation ($F_{(1,\, 31)} = 0.151, p = .7, \eta p^2 = .005$), suggesting that face orientation did not affect response bias either. To assess whether the evidence supports these null hypotheses, we estimated Bayes factors, which indicated strong support for the null hypothesis of emotion expression ($\mathrm{BF}_{01} = 13.638$) and extreme support for the null hypothesis of face orientation ($\mathrm{BF}_{01} = 12.261$) on response bias. Unexpectedly, we found a significant interaction between expression and exposure duration ($F_{(4.59,\, 142.38)} = 2.426, p = .043, \eta p^2 = .073$). However, Bonferroni-corrected pairwise comparisons did not reveal significant differences between expressions at any exposure duration. Finally, no other interaction reach significance: between expression and orientation ($F_{(1,\, 31)} = 0.654, p = .425, \eta p^2 = .021$), orientation and exposure duration ($F_{(4.75,\, 147.40)} = 1.122, p = .351, \eta p^2 = .035$), or the three-way interaction between expression, orientation, and exposure duration ($F_{(4.55,\, 141.07)} = 0.513, p = .750, \eta p^2 = .016$).

### 3.4.2.3 *Expression identification sensitivity*

We examined whether participants' identification sensitivity to the expression of the suppressed face varied across the face orientation ($M_{upright} = 1.517$ $[1.528]; M_{inverted} = 1.156$ $[1.349]$) and exposure duration conditions. We entered identification d' scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration – identification sensitivity significantly increased with increasing exposure duration ($F_{(2.54,\ 78.69)} = 88.46, p < .001, \eta p^2 = .740$), (Figure 3.4c). We also found a main effect of orientation – higher identification sensitivity for upright faces than inverted faces ($F_{(1,\ 31)} = 28.38, p < .001, \eta p^2 = .478$). Finally, we found a marginal interaction between the factors ($F_{(4.37,\ 135.53)} = 2.36, p = .051, \eta p^2 = .071$). A Bayes factor analysis indicated very strong evidence for the null hypothesis model ($BF_{01} = 39.946$). These results indicate that participants were more sensitive to the emotional expression in upright faces than in inverted faces.

### 3.4.2.4 *Expression decision criterion*

We examined whether participant's identification criteria varied across orientation ($M_{upright} = 0.355$ $[0.518]; M_{inverted} = 0.449$ $[0.502]$) and exposure duration conditions. We entered expression criteria scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 3.4d). Participants' willingness to report an emotional expression (indexed by lower criterion scores) increased with increasing exposure durations, as evidenced by a main effect of exposure duration ($F_{(1.5,\ 46.57)} = 14.34, p < .001, \eta p^2 = .316$). We also found a main effect of orientation, with a more liberal criterion for upright faces than inverted faces ($F_{(1,\ 31)} = 6.93, p = .013, \eta p^2 = .183$). The interaction between the two factors also reached significance ($F_{(5.17,\ 160.12)} = 2.30, p = .046, \eta p^2 = .069$). Although the numerical pattern suggests generally larger differences between orientations at longer exposure durations, post hoc Bonferroni-corrected pairwise comparisons did not reveal significant differences at any specific duration. These results suggest participants exhibited a more liberal identification criterion for upright faces than for inverted faces.

### 3.4.3 *Discussion*

The findings of Experiment 5 provided further evidence that there is no difference in the speed with which fearful and neutral expressions are unconsciously processed. As before, we found no effect of expression on location sensitivity, but this time we also found that face inversion affects location sensitivity, which importantly suggests that our method is sensitive to the nature of the stimuli used. The FIE was also seen in expression identification sensitivity in the direction one would expect, i.e. higher identification sensitivity for upright faces than inverted faces. Therefore, the failure to find an effect of expression could be taken as evidence of no difference in unconscious processing between fearful and neutral expressions.

One remaining possibility for why we did not find an effect of expression in experiment 5 is that we compared fearful expressions to neutral expressions, and neutral expressions may not be the best comparison as they can be more ambiguous in terms of their expression recognition. Therefore, we ran an additional experiment in which we used happy expressions instead of neutral expressions. This allowed us to test for unconscious processing comparing between fearful faces as clearly threatening expressions, and happy faces as clearly non-threatening expressions.

## 3.5 Experiment 6

We used the same methods as Experiment 5, but comparing fearful and happy facial expressions, presented both in upright and inverted orientations.

## 3.5.1 Method

### *3.5.1.1 Participants*

We recruited thirty-two University of Edinburgh students (25 female; 2 left-handed) with mean age of 20.1 ($SD_{age}$ = 2.2), who were paid £14 for their participation.

### *3.5.1.2 Stimuli and Procedure*

The methods were as in Experiment 5, except that we employed fearful and happy expressions, in upright and inverted orientations. These images were again selected from the KDEF database and matched on gender and expression identification while minimising differences in expression identification ($M_{fearful}$ = 76.38%; [5.55]; $M_{happy}$ = 93.75% [8.71]) and intensity ($M_{fearful}$ = 6.12; [0.79]; $M_{happy}$ = 5.86 [0.97]; Appendix B).

However, participants gave a slightly different response. Where previously participants judged whether an expression was emotional or not, now participants judged whether the face presented showed a fearful or happy expression. Trials with fearful expressions were considered signal-present, whereas trials with happy expressions were considered signal-absent. Therefore, a hit was defined as a trial with a fearful expression that was reported as presenting a fearful expression, whereas false alarm was defined as a trial with a happy expression that was reported as presenting a fearful expression.

### 3.5.2 Results

#### 3.5.2.1 *Location sensitivity*

We examined whether participants' sensitivity to the location of the suppressed face varied across emotional expression ($M_{fearful} = 1.697\ [1.124]; M_{happy} = 1.761\ [1.121]$), orientation ($M_{upright} = 1.758\ [1.121]; M_{inverted} = 1.700\ [1.125]$), and exposure duration conditions. We entered location d' scores into a 2 (expression: fearful, happy) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. As seen in Figure 3.5a, location d' scores increased with increasing exposure duration ($F_{(2.23,\ 69.16)} = 160.99, p < .001, \eta p^2 = .839$) as in our previous studies. But two additional results were quite unlike our previous studies. First, unlike in Experiment 5, this time the effect of orientation was only marginal ($F_{(1,\ 31)} = 3.511, p = .070, \eta p^2 = .102$). Second, unlike our three previous studies, this time we found a main effect of expression ($F_{(1,\ 31)} = 9.406, p = .004, \eta p^2 = .233$), driven by higher sensitivity to happy than fearful faces. Finally, we found a marginal interaction between emotional expression and exposure duration ($F_{(4.99,\ 154.59)} = 2.20, p = .058, \eta p^2 = .066$). To quantify whether the evidence supports the marginal effect of orientation, we estimated Bayes factors. We found substantial evidence in favour of the null hypothesis of orientation ($BF_{01} = 9.174$), as well as very strong evidence in favour of the null hypothesis model of the interaction between expression and exposure duration ($BF_{01} = 50.234$). The results from this experiment indicate that sensitivity to the location of the face was higher for happy expressions than for fearful expressions, and that upright faces only enjoyed a marginal sensitivity advantage over inverted faces. However, Bayes factor analysis showed contradictory results in relation to these effects, suggesting they might be spurious.

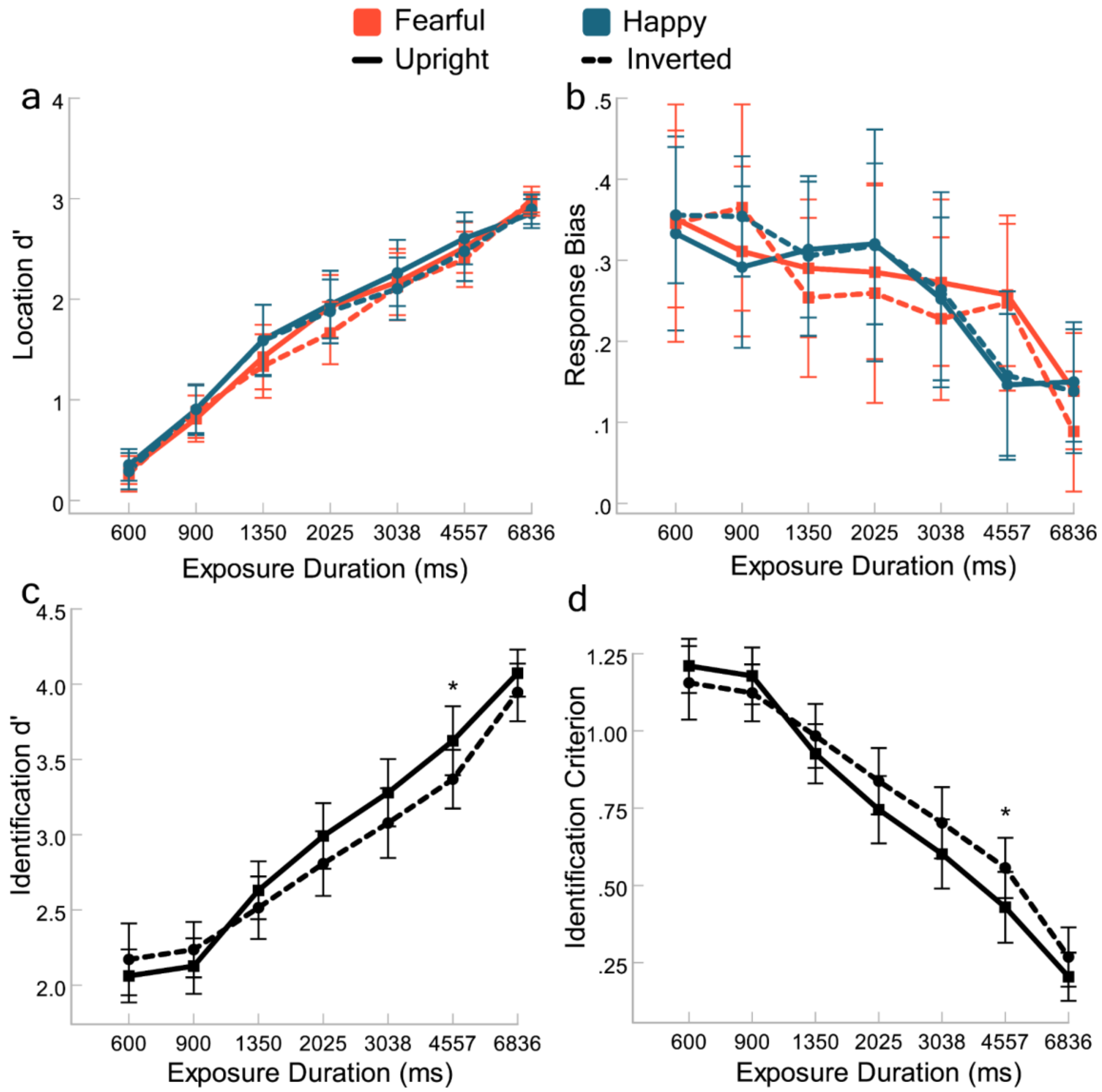



**Figure 3.5.** Results of Experiment 6. (a) Location sensitivity. Location d' increased with exposure duration. There was a marginally significant advantage of upright faces over inverted faces. There was also an advantage of happy expressions over fearful expressions (b) Response bias. Absolute-value response bias scores for reporting location decreased with increasing exposure duration, but there was no difference between expressions or between orientations. (c) Expression identification sensitivity. Expression identification d' increased with increasing exposure duration; upright faces exhibited higher sensitivity than inverted faces. (d) Expression identification criterion. Identification criterion became more liberal with increasing exposure durations; upright faces exhibited a significantly more liberal criterion than inverted faces. Asterisks index statistically significant differences ($p < .05$) between face orientations. Error bars represent 95% CI.

### 3.5.2.2 *Location response bias*

We examined whether participants' location responses varied across conditions by entering the absolute values of response bias scores into a 2 (expression: fearful, happy) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration – response bias significantly decreased with increasing exposure durations, suggesting that response bias decreases as visibility increases $(F_{(3.26, 101.13)} = 8.411, p < .001, \eta p^2 = .213)$, (Figure 3.4b). However, we did not find a main effect of expression $(F_{(1, 31)} = 0.001, p = .974, \eta p^2 = 0)$ or of orientation $(F_{(1, 31)} = 0.018, p = .895, \eta p^2 = .001)$. Similarly, we did not find an interaction between expression and exposure duration $(F_{(4.86, 150.6)} = 1.972, p = .088, \eta p^2 = .06)$, between expression and orientation $(F_{(1, 31)} = 1.314, p = .261, \eta p^2 = .041)$, or between orientation and exposure duration $(F_{(4.67, 144.85)} = 0.762, p = .571, \eta p^2 = .024)$. We did not find a three-way interaction either $(F_{(4.02, 124.48)} = 0.047, p = .996, \eta p^2 = .002)$. To assess whether the evidence supports these null effects of expression and orientation, we estimated Bayes factors, which indicated strong support for the null hypothesis of expression $(\mathrm{BF}_{01} = 13.546)$ and extreme support for the null hypothesis of orientation $(\mathrm{BF}_{01} = 13.662)$.

### 3.5.2.3 *Expression identification sensitivity*

We examined whether participants' identification sensitivity to the expression of the suppressed faces varied across the face orientation $(M_{upright} = 2.970\ [0.626]; M_{inverted} = 2.875\ [0.585])$ and exposure duration conditions. We entered identification d' scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. Expression identification sensitivity significantly increased with increasing exposure time $(F_{(2.56, 79.41)} = 131.95, p < .001, \eta p^2 = .810)$, (Figure 3.5c). We also found a main effect of orientation $(F_{(1, 31)} = 4.27, p = .047, \eta p^2 = .121)$. The interaction between both factors also reached significance $(F_{(4.84, 150.06)} = 5.52, p < .001, \eta p^2 = .151)$. Bonferroni-corrected pairwise comparisons revealed a marginally non-significant advantage for upright over inverted expressions at 4557 ms of exposure $(t(147) = 3.449, p = .067, d = 0.610)$. These results indicate that participants were more sensitive to facial expression in upright faces than in inverted faces.

#### 3.5.2.4 *Expression identification criterion*

We examined whether participant's identification criteria varied across orientation $(M_{upright} = 0.757\ [0.442]; M_{inverted} = 0.804\ [0.413])$ and exposure duration conditions. We entered identification criterion scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found that participants were more willing to report a happy face as exposure increased $(F_{(2.56, 79.38)} = 131.96, p < .001, \eta p^2 = .810)$, (Figure 3.5c). We also found a main effect of orientation, showing a more liberal criterion in expression identification for upright over inverted faces $(F_{(1, 31)} = 4.27, p = .047, \eta p^2 = .121)$. The interaction was also significant $(F_{(4.84, 150.1)} = 5.51, p < .001, \eta p^2 = .151)$. Post hoc Bonferroni-corrected pairwise comparisons revealed a marginally significant more liberal criterion for upright than inverted expressions at 4557 ms of exposure $(t(147) = -3.445,\ p = .068,\ d = -0.610)$. These results suggest that participants exhibited a more liberal identification criterion for upright faces than for inverted faces.

### 3.5.3 *Discussion*

The findings of Experiment 5 suggested that there is no difference in how fearful and neutral expressions break through CFS. However, neutral expressions may not be the best to compare fearful expressions with given that they can be more ambiguous in terms of their emotion recognition. Therefore, we ran Experiment 6 to compare fearful expressions to happy expressions, which are less ambiguous in their emotional expression. Surprisingly, we found that happy expressions broke through suppression faster than fearful expressions. This contradicts b-CFS studies where researchers found shorter breakthrough times for fearful expressions than for happy expressions, with one exception where researchers employed schematic faces (Stein & Sterzer, 2012). In addition, the FIE was marginal in the localisation task and significant in the expression identification task, which may suggest that localisation tasks relied less on holistic face processing than the identification tasks. Or rather, that localisation tasks can be performed in absence of holistic processing. Whatever the case may be, this poses the question to what extent FIEs found in b-CFS studies may have been driven by holistic face processing.

## 3.6 Interim summary

We tested whether angry (Experiment 3) or fearful (Experiment 4) expressions break through suppression faster than happy or neutral ones. However, we did not find any detection advantage for either of those emotional expressions. One possible explanation is that emotional expressions are not prioritised in unconscious emotion processing. Another one is that our method is not sensitive to the properties of face stimuli. To test this possibility, we ran Experiment 5, which tested whether fearful expressions broke through suppression faster than neutral expressions both presented in upright and inverted orientations. This orientation manipulation allowed us to test for the face-inversion effect (FIE) – a repeatedly reported detection advantage of upright faces over inverted faces. As expected, we found sensitivity to the location of upright faces to be higher than for inverted faces, but again we did not find an advantage of fearful expressions over neutral expressions. These results again may suggest that emotional expressions do not enjoy prioritised unconscious processing. However, another possibility is that comparing fearful expressions to neutral expressions is not an ideal test given that neutral expressions are more ambiguous in terms of their emotion recognition. Therefore, we ran Experiment 6 to compare fearful expressions with happy expressions. We presented them in both upright and inverted orientations to test again whether our method is sensitive to the properties of face stimuli. We found a marginally significant effect of orientation. However, as in our previous experiments, we did not find an advantage of emotional expressions over non-emotional ones.

So far, our experiments have employed a 2AFC procedure, where participants had to judge the location of the face out of two possible options. However, original b-CFS procedures employed a detection task where participants simply had to press a key as soon as they became aware of the face. Therefore, participants did not have to make a choice reaction. To make our task more similar to original b-CFS studies, we turned our procedure into a Yes-No detection task. Like in our previous experiments, we used predefined exposure durations instead of self-terminated trials. Half of the trials, however, did not present a face. By asking participants to report the presence (or absence) of a face stimulus, we calculated their sensitivity to the presence of face stimuli.

## 3.7 Experiment 7

In Experiment 7, we measured sensitivity to emotional expressions by asking participants to judge the presence or absence of a face on the screen. As in Experiment 5, we used fearful and neutral expressions presented in upright and inverted orientations. Faces were presented either on the left or on the right side of the screen as in our previous experiments, but this time they were presented on only half of the trials, and participants were simply required to respond ‘Yes’ (without indicating the side) if they thought a face was shown and ‘No’ if they thought no face was shown. Importantly, this task did not involve an expression identification report, thus bearing a closer resemblance to original b-CFS experiments that tested for unconscious emotion processing. If fearful expressions break through suppression faster than neutral expressions, we should find higher detection sensitivity for fearful expressions than for neutral expressions. If decision criterion can contribute to explaining past b-CFS results showing an advantage of fearful over neutral expressions, we should find evidence that participants are more willing to report fearful expressions as present than neutral expressions. If we found an effect on detection sensitivity, with higher sensitivity to fearful expressions than to neutral expressions, but no effect on decision criterion, we could suggest the advantage of fearful expressions over neutral expressions can be fully explained by sensitivity. Conversely, if we found an effect of decision criterion, with a more liberal criterion for fearful expressions, but no effect on sensitivity, we could suggest the advantage of fearful expressions over neutral expressions can be fully explained by decision criterion. As in the previous experiments, these possible effects are not mutually exclusive.

### 3.7.1 Method

#### *3.7.1.1 Participants*

We recruited thirty-two students of the University of Edinburgh (21 female; 2 left-handed) with a mean age of 23.2 ($SD_{age}$ = 3.1), who were paid £14 for their participation.

#### *3.7.1.2 Stimuli and Procedure*

The same stimuli from Experiment 5 were employed: fearful and neutral expressions in upright and inverted orientations. As in previous experiments, the total number of trials was 1120. However, half of the trials presented a face and other half did not present a face. In half of the face-present trials the face had a fearful expression, and in the other half it had a neutral expression. Expression was blocked as described in Experiment 5, maintaining the proportion 50/50 between face-present and face-absent trials. Participants were instructed to only report on the presence or absence of a face image (up arrow to report that the face was present, down arrow to report it was absent; the keys were counterbalanced across participants). They were informed that half of the trials would contain a face image while the other half would not.

## 3.7.2 Results

### 3.7.2.1 *Detection sensitivity*

We examined whether participants' sensitivity to the detection of the suppressed face varied across emotional expression ($M_{fearful}$ = 2.422 [1.207]; $M_{neutral}$ = 2.402 [1.204]), orientation ($M_{upright}$ = 2.460 [1.110]; $M_{inverted}$ = 2.364 [1.228]), and exposure duration conditions. We entered detection d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. Detection sensitivity scores dramatically increased with increasing exposure duration ($F_{(2.54, 78.67)} = 141.812, p < .001, \eta p^2 = .821$), (Figure 3.6a). However, we did not find a main effect of expression ($F_{(1, 31)} = 0.560, p = .460, \eta p^2 = .018$) or of orientation ($F_{(1, 31)} = 2.191, p = .149, \eta p^2 = .066$). We estimated Bayes factors to see whether the evidence supported these null effects. We found substantial evidence for the orientation null hypothesis ($BF_{01} = 4.947$) and strong evidence for the emotion null hypothesis ($BF_{01} = 13.102$). These results show that detection sensitivity only increased as a function of exposure duration.

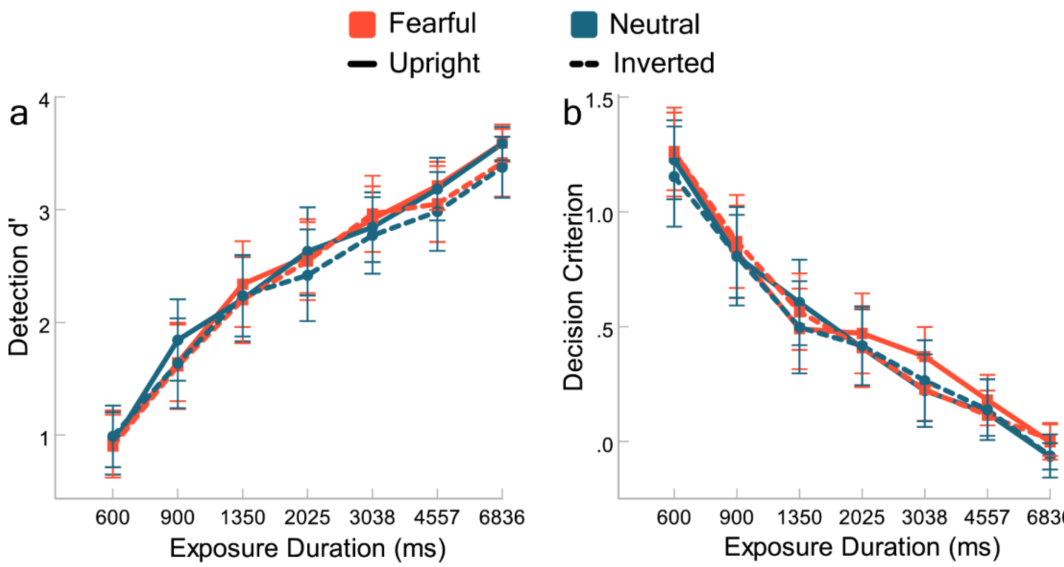



**Figure 3.6.** (a) Detection sensitivity. Sensitivity increased as a function of exposure duration only. (b) Decision criterion. Criterion decreased with increasing exposure

duration with a significantly more liberal criterion for neutral expressions. Error bars represent 95% CI.

#### 3.7.2.2 Detection decision criterion

We had predicted that if b-CFS reports showing shorter breakthrough times to fearful expressions than neutral expressions were fully or partly due to a more liberal criterion for reporting fearful expressions than neutral expressions, we should find in this experiment a more liberal criterion for fearful than neutral expressions. To test this, we examined whether participant's decision criteria varied across expression ($M_{fearful} = 0.504\ [0.585]; M_{neutral} = 0.467\ [0.603]$), face orientation ($M_{upright} = 0.495\ [0.580]; M_{inverted} = 0.477\ [0.607]$) and exposure duration conditions. We entered criterion scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found that criteria turned more liberal across increasing exposure duration ($F_{(2.37,\ 73.60)} = 115.271, p < .001, \eta p^2 = .788$), (Figure 3.6b). We also found a main effect of expression ($F_{(1,\ 31)} = 6.554, p = .016, \eta p^2 = .175$), but we did not find a main effect of orientation ($F_{(1,\ 31)} = 0.165, p = .688, \eta p^2 = .005$). While this main effect of expression may suggest that fearful expressions were associated with more conservative decision criteria than neutral expressions, the effect was small and probably mostly driven by a couple of higher scores obtained at 2025 and 3028 ms of exposure.

### 3.7.3 Discussion

In this experiment, we asked participants to judge the presence or absence of a face on the screen and assessed how emotional expressions affect sensitivity and criterion of those judgments. Neither expression nor orientation affected sensitivity, suggesting neither the manipulation of emotional expression (fearful expressions versus neutral expressions) nor the manipulation of holistic face processing (upright faces versus inverted faces) affected unconscious processing. Thus, when we used a task that was perhaps better-matched to traditional b-CFS tasks, we still found no evidence that emotional expressions broke through awareness faster.

There are different factors that could explain these unexpected results. First, the fact that we obtained rather high sensitivity scores across all exposure durations may suggest that the simpler nature of this task decreased its perceptual difficulty, thus increasing the chance of break through CFS at shorter exposures. If this were the case, it could also be expected that effects found in previous experiments would be diluted between exposure durations that are now excessively long for the task. The fact that we found a numerical trend towards better sensitivity for upright faces may support this view. However, not having found a significant effect of orientation undermines the sensitivity of our procedure in this specific experimental setting, meaning that the fact that we did not find an effect of expression could be simply due to lack of sensitivity of our procedure rather than due to a non-existent effect.

Despite not having found effects in perceptual sensitivity, we did find an effect in decision criterion, with fearful expressions associated with a more conservative criterion than neutral expressions. As argued in the Introduction, if b-CFS findings are due to a more liberal criterion for fearful expressions, we should have found a more liberal criterion for fearful expressions. These results suggest that criterion might have played a role in b-CFS studies, but perhaps not in a way that explains those previous findings.

In our final study, we addressed one final potential objection to our method: Perhaps the predetermined exposure durations used were unsuitable for measuring unconscious processing? We cannot rule out that the distances between our exposure durations may have been too large to capture sensitivity differences between stimulus categories, or that we might have missed the right range of durations. This may be particularly the case of Experiment 7, where participants only had to report whether a face was shown. Although b-CFS findings have typically shown differences between facial expressions for response times of around 2.5 to 3 seconds (Gray et al., 2013; Yang et al., 2007), this might be an overestimate of the time at which breakthrough may occur, due to the lag involved in preparing and executing a motor response. The fact that in Experiment 7 detection d' scores at the shortest exposure durations were already sufficiently high to index above-chance detection supports this interpretation. To account for this possibility, we designed a staircase procedure to estimate detection threshold for each stimulus category. If there were effects of expression and of orientation on yes/no detection that our predefined exposure durations could not capture, we may be able to find them with a staircase procedure.

## 3.8 Experiment 8

In this experiment we assessed how emotional expressions affected the presentation time that elicited threshold detection of stimuli. Participants were presented with faces in the same way as in experiments 3 through 6. But unlike in those experiments, exposure durations changed depending on the participant's performance, following a staircase procedure, as detailed below.

A similar approach was taken by Stein, Hebart, et al. (2011; Experiment 6), who tested – using a similar 2AFC task to ours – whether (neutral) upright faces evoked shorter breakthrough thresholds than (neutral) inverted faces. They tested this in two conditions: one presenting faces rendered invisible with CFS, and a control condition in which the faces were presented binocularly on top of the masks. Their results show shorter thresholds for upright (around 1200 ms) than inverted faces (around 1450 ms) in the CFS condition, but no significant threshold difference in the control condition. We developed a similar but methodologically stricter staircase procedure: ours contained both an ascending staircase starting at the shortest possible exposure duration, and a descending staircase starting at the longest duration, each of which was run twice. Thus, our procedure yielded 16 staircases: 2 expressions × 2 orientations × 2 staircase types (ascending and descending) × 2 repetitions, thereby providing more robust threshold estimates. Additionally, we employed 32 participants whereas Stein, Hebart, et al. (2011) had only 13.

### 3.8.1 Method

#### *3.8.1.1 Participants*

We first recruited twenty University of Edinburgh students. However, their data were not used given that an error was found in the experiment code that prevented staircases from replacing their exposure duration vector after reaching five reversals (see below). After fixing the issue, we recruited a new group of thirty-three students (23 female; 4 left-handed) with a mean age of 22.7 ($SD_{age}$ = 2.8) who were paid £14 for their participation.

#### *3.8.1.2 Stimuli and Procedure*

A staircase procedure with 23 possible exposure durations was employed, following a 1-up, 2-down rule, meaning that two consecutive correct responses decreased the exposure duration by one step, and one incorrect answer increased the exposure duration by one step, with the constraint that exposure durations were not allowed to fall below 50 ms or to exceed 6826 ms. The 23 steps of the initial exposure duration vector were spaced equally on a log scale from 200 to 6326, each step 1.17 times bigger than the one before. The experiment comprised 1120 trials, which were sorted into 16 different staircases, each defined by a combination of the stimuli's facial expression (fearful or neutral), orientation (upright or inverted), and staircase direction (ascending and descending), with two repetitions for each staircase. Ascending staircases started with an exposure duration of 200 ms whereas descending staircases started at 6326 ms. After reaching five reversals, the exposure duration vector was replaced by a new vector of 23 values where the value of the fifth reversal was defined as exposure duration number 12 (the median of the possible durations), and the rest of the values were defined with a step-size of 50 ms. After this change, the staircase remained unchanged for the rest of the procedure.

Participants were told that each trial would contain a face. They were instructed to report the location of the face on the screen (either left or right of fixation). They were not given details about the staircase procedure.

#### *3.8.1.3 Analysis*

The final five reversal values (durations at which the staircase changed direction) were averaged for each individual staircase. Then, the means were collapsed into four categories: fearful upright, fearful inverted, neutral upright, and neutral inverted faces. The threshold means were entered into a repeated-measures ANOVA described below.

### 3.8.2 Results

#### *3.8.2.1 Detection threshold estimates*

We examined whether the time at which suppressed faces reached detection threshold varied across expression and orientation conditions. To this end, we entered threshold means into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) repeated-measures ANOVA. We did not find a main effect of expression ($F_{(1, 32)} = 0.309, p = .582, \eta p^2 = .010$) nor a main effect of orientation ($F_{(1, 32)} = 0.153, p = .698, \eta p^2 = .005$; Figure 3.7). The interaction did not reach significance either ($F_{(1, 32)} = 0.126, p = .725, \eta p^2 = .004$). These results suggest that detection thresholds did not differ between expressions or orientations. To test whether the evidence supports these two null effects, we estimated Bayes factors. They indicated substantial evidence for the null effect of expression ($\mathrm{BF}_{01} = 5.403$) and of orientation ($\mathrm{BF}_{01} = 4.544$), and for the null model of an interaction between them ($\mathrm{BF}_{01} = 3.979$). These results indicate that neither expression nor orientation had an effect on detection threshold estimates.

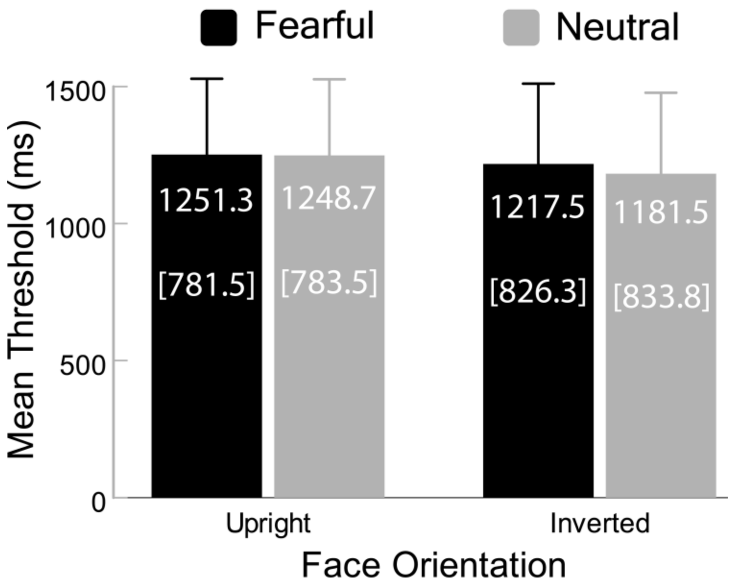



**Figure 3.7.** Mean location detection threshold estimates in ms. Error bars show 95% CI.

#### *3.8.2.2 Face-inversion effect consistency between participants*

Surprisingly, we did not find an effect of orientation in Experiment 7 and 8. It could be the case that the FIE is not very consistent among participants and therefore may be diluted in our data. To assess the FIE's consistency among participants, we estimated the area under the curve (AUC) of each psychometric function in the experiments that involved orientation manipulation. We subtracted AUC values of inverted faces from AUC values of upright faces (Figure 3.8). To test for evidence in favour of the FIE, we then ran a series of one-sample t-tests against zero. If the AUC subtraction scores are significantly above zero it would indicate evidence in favour of FIE. AUC differences were significantly above zero in Experiment 5 $(t(31) = 7.62,\ p < .001,\ d = 1.347)$, but marginally significantly above zero in Experiment 6 $(t(31) = 2.01,\ p = .053, d = 0.355)$, and not significant in Experiment 7 $(t(31) = -0.187, p = .853, d = -0.0331)$ and Experiment 8 $(t(31) = -0.391, p = .698, d = -0.068)$. Then, to see if this inconsistency was present in identification (a task that allegedly cannot be done without engaging in holistic processing), we replicated this same analysis but on AUC differences based on identification-related psychometric functions. We found that the AUC differences were significantly above zero in both: Experiment 5 $(t(31) = 5.35, p < .001, d = 0.945)$ and 6 $(t(31) = 2.37, p = .024, d = 0.419)$.

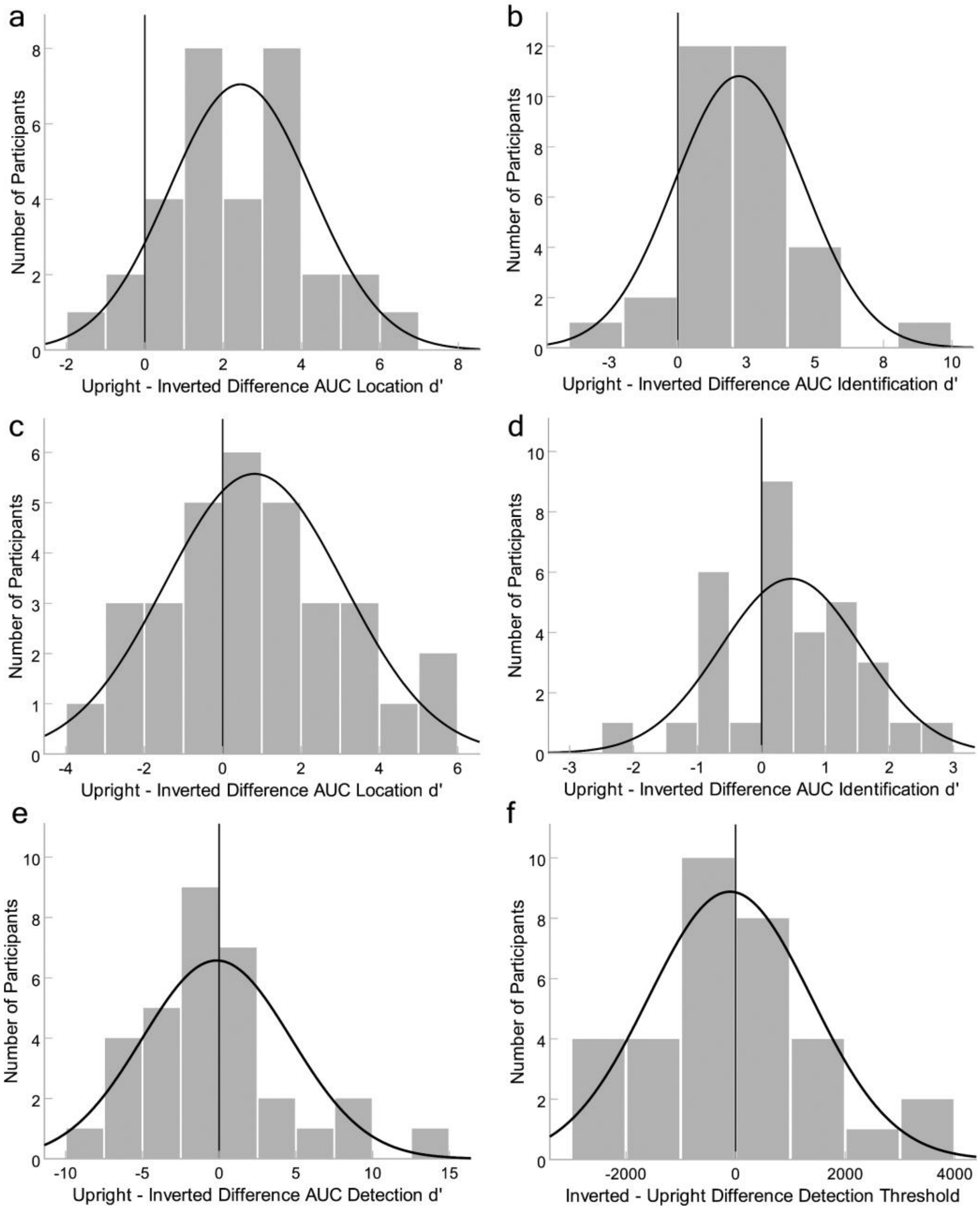



**Figure 3.8.** Face Inversion Effect (FIE) distribution across observers. Experiment 5: (a) Location sensitivity; (b) Emotion identification sensitivity. Experiment 6: (c) Location sensitivity; (d) Emotion identification sensitivity. Experiment 7: (e) Location sensitivity. Experiment 8: (f) Location detection threshold. In Experiments 5-7, these values were

determined by calculating the area under the curve (AUC) of each psychometric function. Then AUC values of inverted faces were subtracted from AUC values of upright faces. In Experiment 8 (f), however, means of detection thresholds of upright faces were subtracted from the means of inverted faces.

### 3.8.3 Discussion

We did not find differences in detection threshold estimates for either expression or orientation. These results suggest that neither expression nor identification modulate detection thresholds of suppressed faces.

It could be the case that most participants adopted a very liberal criterion and decided to report the presence of a face on the correct side based only on differences in contrast between screen sides. In such a case, these detection threshold estimates would correspond to low-level features thresholds (e.g. related to the faces' oval shape or contrast) rather than to high-level facial features (e.g. emotional expression). However, if this were the case, then why have b-CFS studies found effects of expression and orientation? Another possibility is that the way we selected and processed our stimuli decreased visual differences between faces that may have been driving said effects. For example, it has been shown that low-level features may explain shorter breakthrough times of fearful expressions over other expressions (Gray et al., 2013; Hedger et al., 2015; Stein & Sterzer, 2012).

Finally, it was unexpected not to have found an effect of orientation, which Stein, Hebart, et al. (2011) found, given that our procedure was stricter – it employed a higher number of trials, staircases, and it controlled for stimuli low-level features.

## 3.9 General discussion

Do emotional expressions enter awareness faster than non-emotional ones? Past studies using CFS to render facial expressions invisible have claimed that they do (Capitão et al., 2014; Hedger et al., 2015; Oliver et al., 2015; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018; Zhan et al., 2015). However, those studies used RTs as a measure of unconscious processing – participants had to press a key as soon as they became aware of the target stimulus. Importantly, using RTs in a b-CFS procedure as measure of unconscious processing entails methodological issues. It does not isolate detection from irrelevant identification processes. Therefore, identification-related criterion differences could affect detection RTs. For example, if participants exhibited more liberal criteria to emotional expressions than to non-emotional expressions, shorter RTs for emotional expressions could be attributed to criterion differences rather than to sensitivity differences. Here, we employed a recently developed procedure (see Chapter 2) to address this problem. First, by using predefined exposure durations we controlled the amount of visual information available per trial, unlike in the b-CFS procedure, where participants get to decide how much information they receive before committing to a response. Second, by asking them to report both the location and the emotional expression of the face presented, we could estimate sensitivity and criterion.

Our new method allowed us to understand how emotional expressions affected breakthrough from CFS in a more controlled manner, that pulled apart sensitivity to location, response bias, sensitivity to the emotional identity of the face, and criterion when judging identity. Over 6 studies, emotional expressions only affected location sensitivity once, and when it did so (Experiment 6) it showed the opposite pattern from prior reports (i.e. we found advantage for happy expressions, whereas they found advantage for fear (Capitão et al., 2014; Gray et al., 2013; Yang et al., 2007; Yang & Yeh, 2018), though see Stein & Sterzer (2012), who found an advantage of happy faces). Nonetheless, emotional expressions did affect expression identification, showing that happy expressions were better discriminated from their neutral counterparts compared with angry (Experiment 3) or fearful expressions (Experiment 4). We found a similar case regarding face orientation. Face orientation affected location sensitivity twice (significantly in Experiment 5 though only marginally in Experiment 6, out of 4 studies), but it did affect expression identification by showing fearful expressions were better discriminated from their neutral counterparts when presented in upright orientation (Experiments 5 and 6, the only ones that asked to identify the expression while varying orientation).

Our approach used a range of exposure durations that capture location sensitivity from near-chance to high sensitivity (Experiments 3-6). However, we did not find a consistent advantage of emotional expressions over neutral expressions in sensitivity. This absence of effect was also supported by Bayes factors analysis. Some more recent studies have suggested that the advantage of emotional expressions over non-emotional expressions found with the b-CFS procedure may have been due to low-level visual features like spatial frequency and contrast (Gelbard-Sagiv et al., 2016; Gray et al., 2013; Schlossmacher et al., 2017; Stein et al., 2018). In such a case, the fact that we only controlled luminance and still did not find an advantage of emotional expressions may suggest that differences in low-level features are not the only cause of previously found effects.

With the idea that an effect of emotional expression could be taking place between predetermined exposure durations, or diluted by between-subjects variability, we adapted our procedure into a staircase procedure (Experiment 8) to estimate thresholds for location detection. However, we found no difference between fearful and neutral expressions.

Nevertheless, using our procedure, we did replicate the FIE (albeit inconsistently), which has been interpreted as an index of holistic face processing (Axelrod & Rees, 2014; Farah et al., 1995; Goodrich & Yonelinas, 2019; Rakover & Teucher, 1997; Tanaka & Gordon, 2011). Arguably, the FIE relies on high-level facial information just like emotional expressions given that both may require integrating local facial features. The fact that our procedure was sufficiently sensitive to capture the FIE may suggest that it may also be sufficiently sensitive to capture an advantage of emotional over non-emotional expressions, if it exists (for a summary of these two effects, see Table 1). Interestingly, the FIE for expression identification was present in both experiments that involved an orientation manipulation and an expression identification estimation.

**Table 1.** Summary of effects on location sensitivity across experiments.

| | Type of effect | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Effect of exposure | | | Effect of expression | | | Effect of orientation | | |
| | $F$ | p | $\eta p^2$ | $F$ | p | $\eta p^2$ | $F$ | p | $\eta p^2$ |
| Exp. 3 | 104.5 | **< .001***  | .771 | 2.42 | .097 | .072 | - | - | - |
| Exp. 4 | 111.9 | **< .001*** | .783 | 0.104 | .887 | .003 | - | - | - |
| Exp. 5 | 250.8 | **< .001*** | .857 | 2.03 | .164 | .061 | 50.4 | **< .001*** | .610 |
| Exp. 6 | 160.9 | **< .001*** | .839 | 9.406 | **.004*** | .233 | 3.511 | **.07ϕ** | .102 |
| Exp. 7 | 141.8 | **< .001*** | .821 | 0.560 | .460 | .018 | 2.191 | .149 | .066 |
| Exp. 8 | - | - | - | 0.309 | .582 | .001 | 0.153 | .698 | .005 |

Note. Report of F tests, p-values, and partial eta-squared effect sizes. The single significant effect of expression indicated an advantage of happy expressions over fearful expressions. Both the significant and marginal effects of orientation indicated an advantage of upright over inverted faces.

* Significant effects    ϕ Non-significant marginal effects    - Not applicable

In Experiments 5 and 6, we found that most participants exhibited at least a slight advantage of upright faces over inverted faces, whereas in Experiments 7 and 8 such advantage seemed non-existent or at least highly heterogeneous among participants and therefore statistically meaningless (see Figure 8). The difficulty that we had in repeatedly replicating the FIE is interesting because it chimes with a recent argument by Heyman et al. (2019) that the magnitude of the FIE might have been overestimated in previous studies. One possibility is that the FIE is inconsistent among participants because it relies on high-level information in a task that allegedly could be done by only using low-level information – some participants may report mainly based on stimulus configural information whereas others may report mainly based on perceived contrast differences between screen locations. On the one hand, our results show that the FIE was always present in the expression identification task, a task that requires observers to integrate facial features holistically. On the other hand, it was inconsistent in the localisation task (Table 1). The localisation task could be answered by comparing differences in contrast between the two sides of the screen, whereas the expression identification task requires visual integration. Our results may suggest that the FIE for location sensitivity, when present, may have been driven by identification processing that participants were not able to inhibit. Crucially, if this were the case, one could ask whether b-CFS studies testing for emotional expression effects can distinguish between detection and identification. Their findings may be due to residual identification processes that participants could not inhibit.

Interestingly, we did find an effect of decision criterion in expression identification. Surprisingly, however, this effect was consistent across experiments and in the opposite direction to what we had expected – happy expressions enjoyed a more liberal criterion than fearful and angry expressions during the expression identification task. This finding demonstrates that participants may have consistent decision criteria during identification tasks, which stresses the importance of disentangling detection from identification.

In conclusion, our results suggest that emotional expressions such as angry and fearful expressions do not enjoy an advantage over neutral expressions in their path to awareness. However, while we found consistent differences in identification criteria during expression identification, these could not explain why b-CFS studies found shorter breakthrough times for fearful expressions than for neutral expressions. Furthermore, the FIE was consistent when measuring expression identification sensitivity, but it was inconsistent when measuring location sensitivity. It may be the case that the FIE for detection reported in b-CFS studies was due to identification processing that participants could not inhibit. Future studies should use more stringent procedures to address how consistent and high-level-based the FIE is, and how much of unconscious emotion processing can actually be attributed to high-level features.

# CHAPTER 4

# 4 MINIMAL REQUIRED EXPOSURE REVEALS GRADED ACCESS TO AWARENESS OF FACIAL CONFIGURATION AND EMOTIONAL EXPRESSION

## 4.1 Introduction

Faces communicate important information about others' mental states and intentions (Grill-Spector et al., 2017; Jack & Schyns, 2015; Little et al., 2011). They are remarkably effective at capturing attention (Fox, 2002; Langton et al., 2008; Phelps et al., 2006; Wilson & MacLeod, 2003), especially when communicating emotional states. For example, fearful and angry expressions are detected faster than neutral and happy expressions (Fox et al., 2000; Hansen & Hansen, 1988; Krysko & Rutherford, 2009). Face processing is also relevant to emotional disorders. For instance, depression often enhances salience of sad expressions (Burkhouse et al., 2017; Lazarov et al., 2018) whereas anxiety enhances salience of fearful and angry expressions instead (Bishop et al., 2007; Ladouceur et al., 2009), indicating that face processing can be affected by personal characteristics. A more controversial claim, however, is that faces, and in particular emotional expressions, enjoy prioritised access to awareness. In Chapter 2 and Chapter 3, we investigated whether holistic face processing and emotion processing modulate how faces gain access to awareness. As detailed below, here we employed a complementary approach to investigate how faces gain access to awareness – we measured the minimal exposures required for holistic face processing, emotion processing, and perceptual awareness, by using a newly developed LCD tachistoscope that enables extremely brief visual presentations with submillisecond precision. This tool allowed us to directly test whether the orientation (enabling holistic processing or not) and emotional content (facial expression) of a face modulate the minimal exposure required to perceive them. Notably, this approach allows us to present stimuli without any form of masking, avoiding the potential confounds that arise with stimulus-mask interactions.

Multiple claims have been made about how faces gain access to awareness. Some researchers have claimed that emotional facial expressions enjoy prioritised access to awareness in comparison to non-emotional ones. For example, Yang et al. (2007) found shorter response times to suppressed fearful expressions than to suppressed neutral expressions. This claim derives from a long list of findings using masking techniques to suppress visual information from awareness, such as Continuous Flash Suppression (CFS), a strong interocular suppression technique that renders stimuli shown to one eye invisible by flashing high-contrast Mondrian-like patterns to the other eye. In fact, as described in previous chapters, it has been claimed that faces expressing negative emotions overcome suppression – and thus enter awareness – faster than faces expressing positive emotions and no emotion (Capitão et al., 2014; Hedger et al., 2015; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018a), though this effect has failed to replicate a number of times (see Chapter 3). Similarly, it has been claimed that upright faces overcome suppression faster than inverted faces. For example, Jiang et al. (2007) found shorter response times to suppressed upright faces than to suppressed inverted faces, suggesting that upright faces gain access to awareness faster due to their configural or holistic properties, an effect that has been repeatedly replicated (Akechi et al., 2015; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011). However, the usefulness of masking and interocular suppression techniques hinges on their ability to interrupt visual processing; and because it is impossible to determine the extent of this interruption – unless relative to a control condition – said techniques cannot provide with minimal exposures required for holistic face processing or emotion processing, let alone with the minimal exposure required for perceptual awareness. Crucially, by measuring the minimal required exposures of upright and inverted faces, and of emotional and non-emotional expressions, we can directly test whether there is an inherent hierarchy in how faces are perceived. Are all facial features processed together or do they involve different bottlenecks in processing?

Several approaches have been taken to determine minimal required exposures in visual perception. The main limitation, though, is that the shortest exposure durations that have been employed seem to be sufficiently long for visual detection and identification to occur. In a classic study, Intraub (1981) presented observers with sequences of unmasked pictures at 114, 172, or 258 milliseconds (ms) of exposure per picture. The target was specified by name (e.g. giraffe), by superordinate category (e.g. animal), or by negative category (e.g. the picture that is not food). Participants were asked to press a key as soon as they saw the cued picture and to describe it briefly. They were able to detect the pictures even in the most difficult condition, when the exposure duration was the shortest. Similarly, Thorpe et al. (1996) presented observers with brief unmasked pictures and asked them to perform a categorisation task (to indicate whether the picture contained an animal or not) while measuring electroencephalography (EEG). By measuring event-related potentials (ERP) and reaction times to pictures that could contain or not an animal in a go/no-go categorisation task, they showed that successful visual categorisation was possible even at 20 ms of exposure. Later, in a follow-up study, researchers showed that even using novel scenes (i.e. never-before seen by the observers) did not decrease categorisation performance, even with 20 ms of exposure (Fabre-Thorpe et al. 2001). Similar findings were reported in scene categorisation when asking observers to categorise unmasked natural and manmade environments presented for 26 ms (Joubert et al., 2007). These results demonstrate that with an exposure duration of 20 ms the visual system receives sufficient information to discriminate elements contained in unmasked pictures. Therefore, despite claims indicating that 26 ms of exposure are required to get the gist of a scene, by such exposure duration accuracy is already extremely high (> 90%), as reported by Joubert et al. (2007) and Rousselet et al. (2005), thus implying that the true minimal exposure duration required for visual perception must be shorter. Additionally, these results obtained using unmasked pictures may raise the potential issue of afterimages – it is difficult to determine whether visual processing stopped after stimulus removal.

Technically, it has been difficult to present stimuli at briefer exposures than 16 ms, because most studies over the last few decades have presented stimuli on computer monitors, which (at least until recently) have typically had a refresh rate of around 60 Hz. Even newer monitors that have a refresh rate of more than 60 Hz, such as monitors with 100 Hz, 120 Hz, and 144 Hz cannot present stimuli at briefer exposures than 10, 8.33, and 6.94 ms, respectively. Therefore, subsequent attempts to determine the minimal exposure required for visual perception had to employ masking techniques to interrupt visual processing. But this introduces a potential confound, because it will always be unclear if differences in minimal exposure are due to differences in the time necessary to process relevant stimulus properties, or in the degree to which visual masking affected processing of those properties. For example, Greene & Oliva (2009) determined through a scene classification task that participants needed exposure durations of between 19 and 67 ms to exhibit 75% correct-response thresholds, with high between-subject variability and, importantly, with shorter exposure durations for detection of scenes' global properties (i.e. identifying the kind of landscape) than their basic properties (i.e. identifying objects in the image). However, because they used a dynamic masking paradigm to limit sensory processing following image presentation, it is impossible to determine whether these differences in threshold reflect the speed with which these properties could be identified, or the degree to which the masking interfered with processing these properties. Indeed, masking has been demonstrated to significantly affect visual processing with brief presentations: Codispoti et al. (2009), for example, presented observers with pleasant, neutral, and unpleasant pictures, masked and unmasked, across exposure durations ranging from 25 to 6000 ms, and measured their emotional reactivity using various techniques such as electromyography, EEG, skin conductance, and both pleasure and arousal ratings. Crucially, when masked images were employed, they found no evidence of emotional engagement with any measure with 25 ms of exposure, and some evidence for such processing with 80 ms of exposure (at this point, some participants could reliably rate pictures as pleasant or unpleasant); but when unmasked images were employed, they found reliable evidence of emotion processing at all exposure durations. Importantly, even when masked images gained sufficient visibility as with 80 ms, their physiological responses differed from responses found with unmasked pictures shown for 25 ms, suggesting that masking may not only suppress but also alter visual processing. Together, these studies indicate that exposure durations of 25 ms and above are sufficiently long to reveal visual processing of scenes and faces. Essentially, because it has not been determined what specific aspects of visual processing are disrupted by backward masking – e.g. image processing, afterimage processing – we simply cannot determine whether findings obtained using masked and unmasked stimuli can be compared. Therefore, even when masked stimuli are perceived, they may still not be comparable to perceived unmasked stimuli, as shown by Codispoti et al. (2009), who reported that physiological responses also differ between perceived masked and unmasked pictures.

Another approach that researchers have taken to investigating visual processing with minimal exposure is to test whether visual perception can occur in the absence (or near absence) of attention. For example, Li et al. (2002) trained a group of participants on a demanding letter discrimination task. They gradually shortened the exposure duration of the stimuli every time participants' performance improved, until they individually reached a duration at which they could no longer improve (the researchers took this as evidence that the task was exhausting attentional resources). The researchers found that letter discrimination could be performed at around 77% performance requiring between 133 ms and 240 ms of exposure. Importantly, even at the briefest possible duration for the letter task, when the researchers added peripheral images, observers were still able to reliably make discriminations about them (e.g. whether they contained an animal), indicating that these durations still enabled perception across the visual field, even when attentional resources were exhausted by the main task. Reddy et al. (2004) reported similar results using masked human faces. They asked observers to simultaneously perform a face-gender discrimination task in the screen periphery and a letter discrimination task in the centre of the screen. They found better discrimination for upright faces than inverted faces in near absence of attention (when attention was fully employed by the fixated stimuli), with exposure durations ranging between 133 and 160 ms per participant. However, these studies may have failed to deplete attention. For example, Evans & Treisman (2005) reported that classification tasks could widely vary in difficulty depending on their nature (detection or identification) and stimulus complexity (number of elements in a scene). Thus, one could question whether tasks used to deplete attention were sufficiently demanding. Cohen et al. (2011) went one step further and found that when the distracting task is sufficiently demanding, perception of natural scenes is fully impaired, thus casting doubts on the efficacy of this approach. In addition to these limitations, because these studies employed backward masking, they are subject to the same limitations described in the previous paragraph.

In summary, due to hardware limitations, previous studies examining fast visual processing had to employ exposure durations that while brief, were sufficiently long for participants to exhibit both above-chance detection and discrimination of scenes and faces, in addition to differential psychophysiological reactivity to hedonic images. Alternatively, researchers have also employed masking techniques to either assess the minimal required exposure of visual processing or to prevent conscious processing and thus assess unconscious processing. However, these two approaches suffer from the same problem: they cannot determine how much of the findings is due to processing of the masks themselves, given that we cannot determine what aspects of visual processing are being interrupted, nor to what extent they are interrupted.

Here we present a novel solution to determine the minimal required exposure of visual perception, using a newly developed LCD tachistoscope that enables sub-millisecond presentations to study visual discrimination, emotion processing, and awareness in face perception. In two experiments, participants had to discriminate the location (Experiment 9) and presentation order (Experiment 11) of a face from that of a scrambled face, in extremely brief presentations (without a mask) ranging in exposure duration from 0.8 to 6.2 ms in the former, and from 0.6 to 6 ms in the latter. To accomplish this, we used an adapted version of the procedure described in Chapters 2 and 3. This procedure allowed us to measure how perceptual sensitivity to the location of the intact face and to its emotional expression changed across this range of exposure durations. We also asked participants to rate the clarity of their visual experience, which allowed us to estimate metacognitive sensitivity, i.e. how sensitive participants' subjective experience is to objective discrimination. As explained below, metacognitive sensitivity can be a useful and reliable measure of perceptual awareness.

## 4.2 Experiment 9

In Experiment 9, we measured how perceptual sensitivity and decision criterion to the location and emotional expression of human faces changed across exposure durations. To do so, we adapted the procedure we developed in Chapter 2. We selected seven exposure durations, equally spaced on a linear scale, that encompass the range of values from 0.8 to 6.2 ms. After each stimulus presentation, participants judged both where on the screen the stimulus was presented and what the stimulus was (an emotional or neutral expression). Next, participants judged in the perceptual awareness scale (PAS) how clear their visual experience of the face was. The PAS is a 4-point scale that allows participants to rate their subjective experience by selecting one of the following alternatives: “no experience”, “vague impression”, “almost clear experience”, and “clear experience” (Ramsøy & Overgaard, 2004; Sandberg & Overgaard, 2015). Signal Detection Theory (SDT) distinguishes between type-1 indices, which assess how well an observer can discriminate between stimuli (e.g. between a signal and noise), and type-2 indices that assess how well an observer’s confidence (or awareness) ratings can discriminate between their own correct and incorrect stimulus classifications. We used type-1 signal detection analyses to assess their sensitivity to stimulus location (location d’) and expression (identification d’), as well as their criteria for making these judgments. Then, because perceptual sensitivity alone does not necessarily reflect changes in awareness – e.g. an observer could exhibit above-chance performance in a visual discrimination task without being aware of the stimuli employed (e.g. Weiskrantz et al., 1974) – we used type-2 signal detection analyses to assess how sensitive their subjective awareness was to their perceptual sensitivity (i.e. metacognitive sensitivity or meta-d’) as exposure durations increased.

### 4.2.1 Methods

#### *4.2.1.1 Participants*

Thirty-five students of the Université Libre de Bruxelles provided informed consent and were paid €10 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Three participants were excluded from the analysis (see Analysis section). The remaining 32 participants (17 female; 4 left-handed) had a mean age of 22.6 ($SD_{age}$ = 2.7; range: 18 – 26). The studies reported here were approved by the Université Libre de Bruxelles ethics committee.

Past studies that have found statistically significant effects of emotional expression on unconscious visual processing employed around 16 participants per experiment (e.g. Yang, et al., 2007). We decided to double this number to increase power and allow counterbalancing of experimental blocks with a multiple of 8 (see Procedure section). A retrospective power analysis, conducted using G*Power 3.1.9.7 (Faul et al., 2009), to test for a difference between conditions in a repeated-measures ANOVA, with a small to medium effect size ($\eta p^2 = .04$) and alpha of .05, aiming to achieve a statistical power of 95%, determined that a sample of 19 participants would be required. If a non-sphericity correction $\varepsilon$ of .5 were to be added – as reported in the results section, a number of tests violated this assumption – then a sample of 29 participants would be required.

### *4.2.1.2 Apparatus and stimuli*

We used a custom-made LCD tachistoscope (Figure 4.1), an adapted version of the design described by Sperdin et al. (2013). In both versions, two LCD screens are employed with one placed upright vertically and the other placed horizontally, aligned to the top of the other screen. Both screens are branded Philipps 223V5LHSB2 with a resolution of 1920 × 1080 pixels, with a size of 476.6 mm in width and 268.11 mm in height, giving a pixel pitch of 0.248 × 0.248 mm. Both screens are fixed in a rigid aluminium frame, which is covered both internally and externally by matte-finished Plexiglass, which in addition to providing structural support also prevents light reflection. Between the two screens, a diagonally placed semi-permeable mirror allows light to pass from the vertical screen while it reflects light emitted from the horizontal screen placed above it. Therefore, both screens are presented to the observer superimposed on each other if their backlights are simultaneously turned on, as when a target stimulus is presented with submillisecond precision. This semi-permeable mirror is a Pilkington MirroView™ 50/50 glass of 418 × 504 mm, 6 mm thickness, toughened for robustness (to avoid image deformation due to gravity-caused deflection). We can control which screen is visible to the observer at each time point with a precision of 2±1 microseconds (μs) by controlling the screens' backlights, which are powered by an independent power supply (36V, 108 Watt) with 20 Ohms rheostat in series to dim screen luminosity. The semi-permeable mirror is not perfectly 50:50, resulting in one screen being more luminous than the other. Rheostats are therefore employed in order to decrease the voltage of the more luminous screen, so it matches the screen with lower luminosity and corrects for screen disparities. Unlike the design of Sperdin et al. (2013), in our design backlights are controlled by a dedicated micro-controller (ATmega328 AVR) instead of using a parallel port. The micro-controller receives the instruction via serial communication (USB 2.0) in order to switch on the backlight for the stimulus' predetermined presentation duration. Using a dedicated micro-controller offers two advantages: it prevents hardware compatibility issues since parallel ports are becoming very uncommon; and it provides higher precision and consistency given that its only function is controlling backlight thus it is not affected by computer workload. A minor disadvantage, however, is that the absolute stimulus presentation time is affected by the time it takes for the computer to send the instructions, which is not constant and can take a few milliseconds. Nevertheless, while the onset of stimulus presentation may vary slightly, the exact stimulus presentation duration is controlled with a precision of $2 \pm 1$ µs for presentations under 16 ms and of 20 µs for stimulus presentations between 16 ms and 10 seconds. The tachistoscope included a signal output to synchronise stimulus presentation durations with other hardware (e.g. EEG signal triggers). In the experiments presented in Chapters 4 and 5, the vertical screen presented all images that did not involve extremely brief exposure durations, including fixation, placeholders, response cue, etc., and its backlight was therefore always on, whereas the horizontal screen presented the stimuli that involved extremely brief exposure durations (described below), and its backlight was therefore off, except during stimulus presentations when it came on for a very brief period; during this period, the content of both screens was visible to the observer.

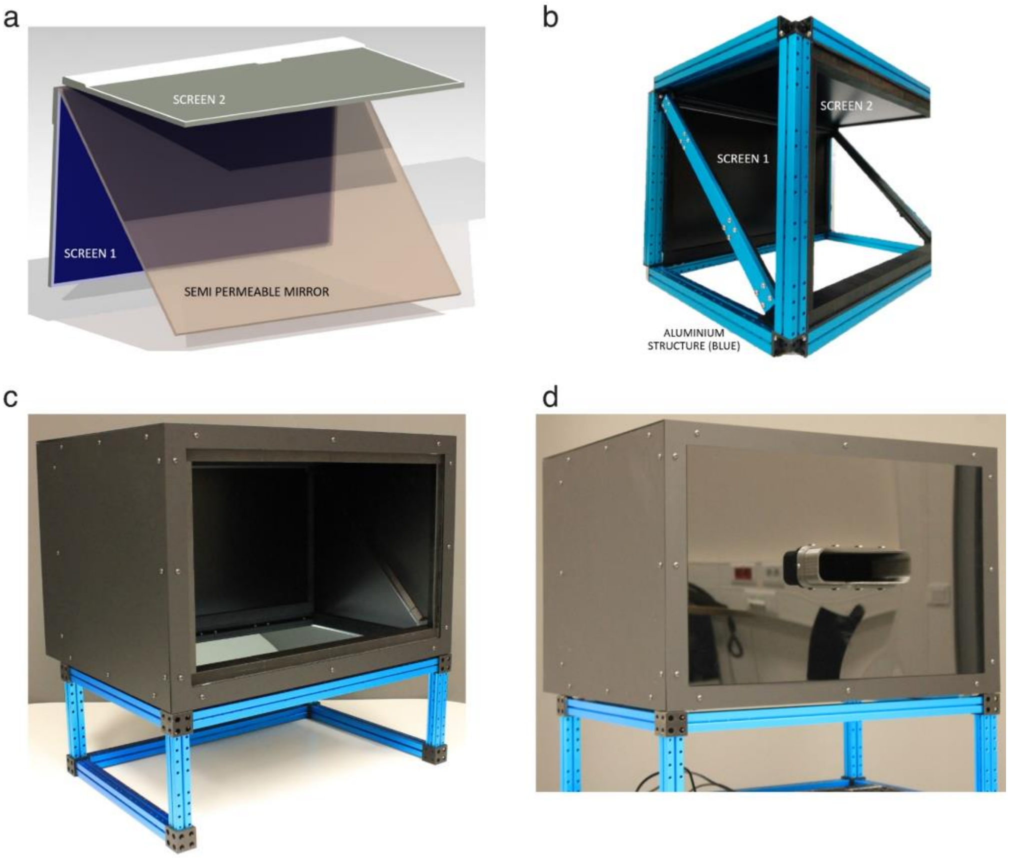



**Figure 4.1**. LCD Tachistoscope. (a) Vertical screen (screen 1), horizontal screen (screen 2), and semi-permeable mirror. (b) Both screens and the mirror assembled together using an aluminium structure. (c) The tachistoscope assembled without lid and (d) with lid, showing the eyecup through which participants look at the display.

Stimuli were 20 human faces (10 fearful and 10 neutral, the same 10 identities for both, 5 female; see Appendix C) taken from the Radboud Faces Dataset (RaFD; Langner et al., 2010), all seen from a front angle. Images were classified as either fearful or neutral expressions, and were matched between these categories, using ratings from the stimulus dataset, for agreement with the intended emotional expression ($M_{\text{fearful}}$ = 94.9% [$SD_{\text{fearful}}$ = 4.12]; $M_{\text{neutral}}$ = 97.5% [4.03]) while minimising intensity differences (by definition, fearful expressions are higher in intensity than neutral expressions: $M_{\text{fearful}}$ = 4.27 [0.2]; $M_{\text{neutral}}$ = 3.69 [0.27]). Images were cropped to remove hair and create a uniform oval shape, transformed to greyscale, and equated for luminance using the Matlab SHINE toolbox (Willenbockel et al., 2010). Scrambled images were created by selecting an oval-shaped area that encompassed all relevant facial features (eyes, nose, and mouth; see Appendix D). Then, pixels were divided into 40 square patches, which were scrambled by using the Scramble Filter in Photoshop™ CC 2019. The luminance of both intact and scrambled images was equated (mean luminance = 179.8 cd/m$^2$). Background colour was replaced with uniform grey (luminance = 220 cd/m$^2$). Images (2.56° × 3.67° in size) were presented either to the left or to the right of a fixation cross (horizontal centre-to-centre distance 2.82°), between two dots that were placed above and below each stimulus location, serving as placeholders (see Figure 4.2). The vertical distance between placeholder dots was 4.7°.

Exposure durations used in the first experiment were chosen after a pilot study in which we tested 20 participants from the same pool with different duration values spanning 0.5 to 10 ms (linearly spaced at 0.8 ms intervals, 160 trials per duration), using the procedure described below. We measured participants' sensitivity scores for location discrimination at each duration and used these scores to select 7 exposure durations (covering chance performance, d'≈0, to stable high performance, d'≈2) for use in the experiment.

All the experiments reported here took place in a dimly lit room. Participants viewed the display through a rectangular eyecup, positioned 55 cm away from the vertical screen. The tachistoscope was connected to a computer running Matlab (version 2018a) and the experiments were written using the Psychophysics Toolbox extensions (Brainard, 1997).

#### *4.2.1.3 Procedure*

Each trial began with the presentation of a black fixation cross (0.41° × 0.41°) at the centre of the screen and a pair of placeholder dots on each side, marking the locations where stimuli would be shown (Figure 4.2). Participants were instructed to focus on the fixation cross with both eyes open, avoiding blinking during the trials. 300 ms after trial onset, two images were displayed, one on each side, for one of seven equally spaced exposure durations (range 0.8 – 6.2 ms). One of the images was an intact human face with either a fearful or neutral expression, in either an upright or inverted orientation. The other image was the face's scrambled counterpart. After stimulus offset, only the placeholder dots remained on the screen for 200 ms; then, a response cue ("????") was presented at the centre of the screen, prompting participants to report the location of the intact face (left or right) and identify its expression (emotional or non-emotional) with a single keypress. The 'left Control' and 'Left Shift' keys of a standard keyboard were used for 'left' reports, and the 'up arrow' and 'down arrow' keys were used for 'right' reports. Mapping of keys to expressions was counterbalanced across participants. The response cue remained on the screen until the participant either pressed a key or 2 seconds had elapsed. Following response-cue offset and an additional 300 ms, participants were shown the question "How clear was your visual experience?" and 4 response options (Perceptual Awareness Scale, PAS; Sandberg & Overgaard, 2015), until the participant either pressed a key or 3 seconds had passed. Participants were instructed to describe their visual subjective experience of a face, irrespective of its location on the screen or expression. Mapping of keys to PAS response was counterbalanced across participants. If the participant did not give either of the required responses by the end of its response window, an 800-ms message ("You have taken too long to respond!") was displayed. Trials that did not receive both responses were not included in the analyses (no-response trials constituted < 2.5% for each of the participants included in the analyses). Following the PAS response, only a pair of placeholder dots on each side remained on the screen for a 1-second intertrial interval.

Participants performed 40 practice trials with randomly selected stimuli; these were followed by 1120 experimental trials. Face orientation was blocked (70 trials/block), with block order counterbalanced across participants in an ABBABAAB BAABABBA order (70 trials/block, with A and B denoting upright and inverted faces, respectively, for half of the participants, or vice versa for the other half). Participants were given self-terminated breaks every 70 trials and a compulsory 15-minute break after completing 560 trials.

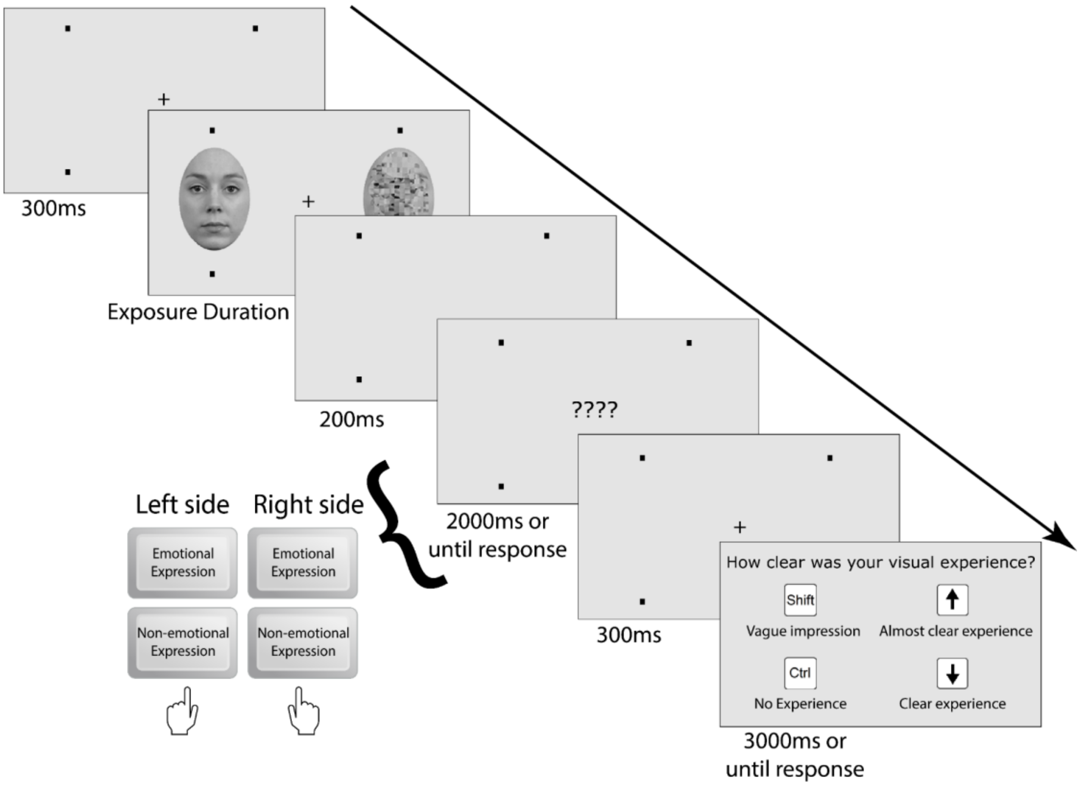



**Figure 4.2**. Schematic description of a trial in Experiment 9. Stimuli were presented for one of seven possible exposure durations (0.8 – 6.2 ms, equally spaced on a linear scale). After stimulus offset, participants judged the location of the intact face (left or right) and its expression (emotional or non-emotional) by pressing one key. Next, they judged the clarity of their visual experience (PAS).

#### *4.2.1.4 Analyses*

We excluded data from three participants: two failed to provide a response on more than 5% of trials (missing data) and one showed chance accuracy, with performance that did not increase as exposure durations increased, suggesting that the participant failed to pay attention.

We used Signal Detection Theoretic (SDT) measures to assess how perceptual sensitivity, metacognitive sensitivity, and decision criteria changed across display durations. To determine each participant's bias-independent sensitivity to face location (left or right; henceforth referred to as location d') for each combination of duration, face orientation, and emotional expression, we employed the calculation for two-alternative forced choice (2AFC) tasks for perceptual sensitivity and criterion (type-1 SDT; Macmillan & Creelman, 2004), $d'_{location} = \left(\frac{1}{\sqrt{2}}\right)(Z(Hit_{location}) - Z(FA_{location}))$, where Z(Hit) stands for the Z score associated with the probability of a Hit (defined as a trial in which a face was displayed on the right and reported on the right), and Z(FA) for that associated with the probability of a false alarm (a trial in which a face was displayed on the left but reported as being on the right). To estimate each participant's bias to respond left or right (henceforth referred to as response bias) during face location, we employed the calculation $C_{location} = -\left(\frac{1}{2}\right)(Z(Hit_{location}) + Z(FA_{location}))$. Positive and negative values for this measure indicate a bias toward responding "left" and "right", respectively; however, as these may cancel out across participants, we converted the results to absolute values as a measure of response bias quantity. To determine how sensitive each participant's awareness judgment (PAS ratings) was to their location sensitivity performance (metacognitive sensitivity; henceforth referred to as meta-d') and the bias in such judgments (metacognitive bias; henceforth referred to as meta-bias), we employed the maximum likelihood estimation procedure developed by Maniscalco & Lau (2012, 2014, 2016; see http://www.columbia.edu/~bsm2105/type2sdt/).

To determine emotional identification sensitivity (to fearful expressions versus neutral expressions; henceforth referred to as identification d'), we used the calculation of d' for Yes-No detection tasks, $d'_{identification} = Z(Hit_{identification}) - Z(FA_{identification})$, where a hit was defined as correctly reporting a fearful expression and FA was defined as incorrectly reporting a fearful expression. To estimate each participant's identification criterion during emotion expression identification, we employed the calculation $C_{identification} = -\left(\frac{1}{2}\right)(Z(Hit_{identification}) + Z(FA_{identification}))$. These measures were calculated by-condition for each participant and analysed using analysis of variance (ANOVA) as detailed below. Greenhouse-Geisser adjusted degrees of freedom were used when Mauchly's test indicated a violation of the sphericity assumption.

Both frequentist (ANOVA and t-tests) and Bayesian (Bayes factors) statistical analyses were performed using Jamovi (The jamovi project, 2020) and corroborated using JASP (JASP Team, 2020) and R. When an ANOVA indicated a significant interaction, we ran post hoc Bonferroni-corrected pairwise comparisons to look for significant effects.

Post hoc pairwise comparisons in these statistical packages use estimated marginal means based on the variance of the ANOVA model. For Bayes factor analysis, we defined the null hypothesis as no difference between conditions by using a standard Cauchy distribution centred on zero with rate of 0.707.

### 4.2.2 Results

#### *4.2.2.1 Location sensitivity*

To examine how the manipulated factors affected location discrimination, we entered location d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration, whereby sensitivity increased with increasing exposure duration ($F_{(4.90,\ 151.98)} = 15.331, p < .001, \eta p^2 = .331$). Crucially, as shown in Figure 4.3a, location d' scores went from showing no sensitivity (chance performance) at the shortest exposure duration to showing high sensitivity at the longest exposure durations, indicating that changes in exposure duration in the order of a few milliseconds had a great impact on participants' ability to discriminate an intact face from a scrambled face. Importantly, there was a main effect of face orientation ($F_{(1,\ 31)} = 34.198, p < .001, \eta p^2 = .530$), indicating a sensitivity advantage for upright faces ($M = 1.169\ [SD = 0.905]$) over inverted faces ($M = 0.963\ [0.701]$): a face-inversion effect. However, the effect of expression was only marginally significant ($F_{(1,\ 31)} = 3.633, p = .066, \eta p^2 = .105$), with a slight numerical advantage of neutral expressions ($M = 1.088\ [0.977]$) over fearful expressions ($M = 1.043\ [0.966]$). Importantly, the interaction between face orientation and exposure duration was significant ($F_{(4.90,\ 151.98)} = 15.331, p < .001, \eta p^2 = .331$). To see at what specific exposure duration upright faces enjoyed significantly better sensitivity over inverted faces, we ran post hoc pairwise comparisons, which revealed an advantage of upright faces at 4.4 ms ($t(184) = 5.987, p < .001, d = 1.058$), 5.3 ms ($t(184) = 8.164, p < .001, d = 1.443$), and 6.2 ms of exposure ($t(184) =$

$5.477, p < .001, d = 0.968$). Therefore, 4.4 ms of exposure were sufficient to exhibit a face-inversion effect. We did not find an interaction between expression and face orientation ($F_{(1, 31)} = 0.002, p = .970, \eta p^2 = 0$), or between expression and exposure duration ($F_{(5.10, 158.13)} = 0.912, p = .476, \eta p^2 = .029$), or a three-way interaction ($F_{(4.96, 153.87)} = 15.033, p = .400, \eta p^2 = .032$). Thus, expression had no direct or modulatory effect on location sensitivity.

Although we did not find any effect involving expression, absence of evidence is not necessarily evidence of absence. Therefore, we calculated Bayes factors to test whether the obtained data support the absence of an effect of expression (null hypothesis model). Bayes factors indicated strong evidence in favour of the null hypothesis model ($BF_{01} = 10.571$), suggesting that these data are 10.571 times more likely to be observed under the null hypothesis model of expression. This analysis suggests that fearful expressions are not prioritised for perceptual discrimination when compared to neutral expressions.

Finally, to determine the minimal required exposure for above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 1.7 ms for upright neutral ($M = 0.166\ [0.381];\ t(31) = 2.47,\ p = .019,\ d = 0.436$), inverted fearful ($M = 0.228\ [0.396];\ t(31) = 3.25,\ p = .003,\ d = 0.575$), and inverted neutral faces ($M = 0.268\ [0.363];\ t(31) = 4.17,\ p < .001,\ d = 0.738$), and 2.6 ms for upright fearful faces ($M = 0.627\ [0.463];\ t(31) = 7.66,\ p < .001,\ d = 1.354$). Thus, our results suggest that above-chance discrimination of an intact face stimulus from its scrambled counterpart requires around 2 ms of visual exposure.

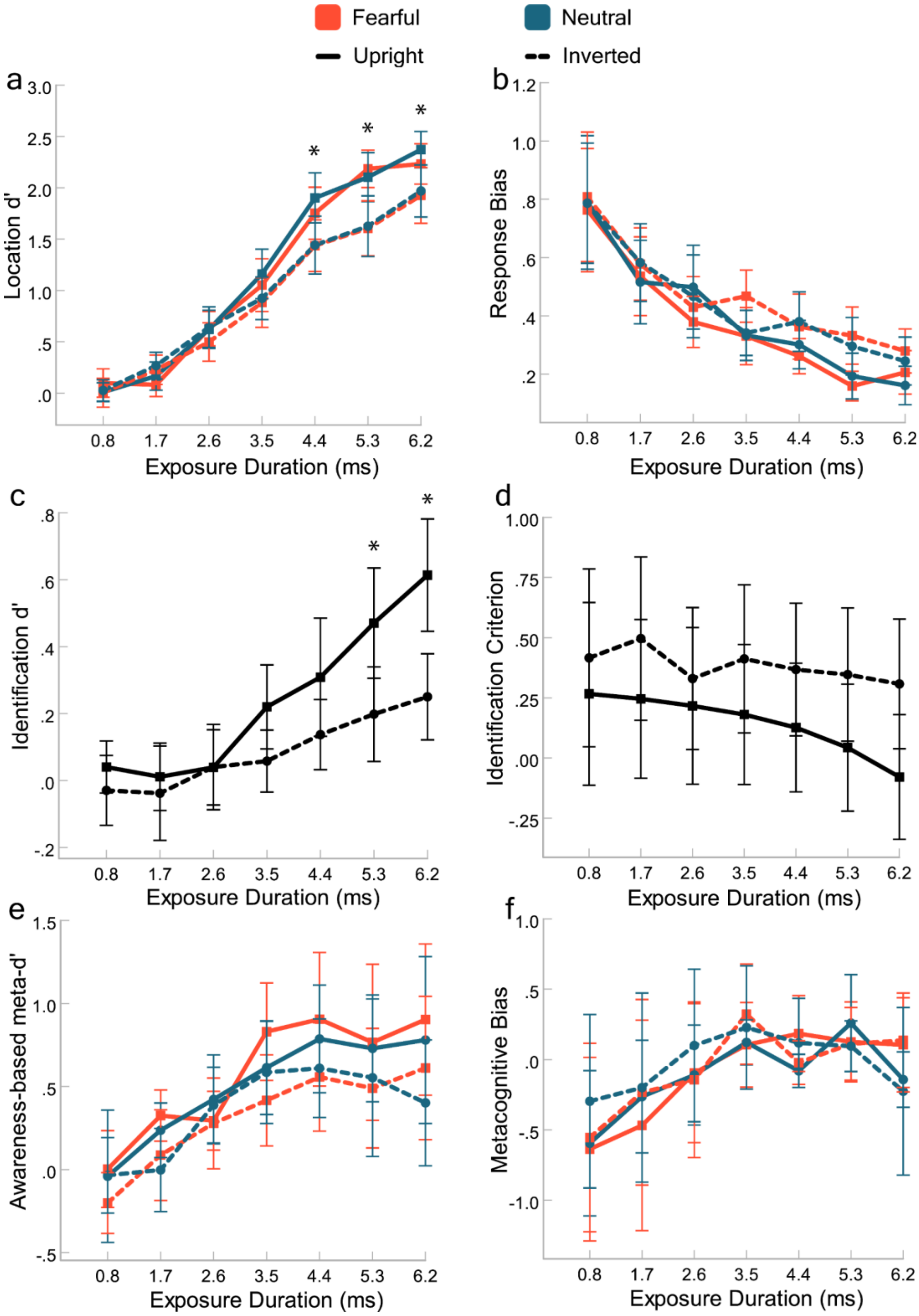



**Figure 4.3.** Results of Experiment 9. (a) Location sensitivity. Location d' increased with increasing exposure duration, departing from chance around 2 ms. A significant advantage for upright faces over inverted faces is present at 4.4, 5.3, and 6.2 ms of exposure. (b) Absolute-value response bias scores for reporting location (bias toward either left or

right). The amount of response bias decreased as exposure duration increased, with slightly (but significantly) greater bias for inverted faces over upright faces. However, there was no statistically significant difference between fearful and neutral expressions (c) Identification sensitivity for expression. Identification d' increased with exposure duration. A significant advantage in expression identification for upright faces over inverted faces arises by 5.3 ms of exposure. (d) Criterion scores for reporting expression. Although we did not find main effects of exposure duration and face orientation, we did find a significant interaction between these two factors. However, post hoc comparisons did not reveal significant differences between orientations at any exposure duration. (e) Awareness-based metacognitive sensitivity. Meta-d' increased with exposure duration. Upright faces enjoyed higher scores than inverted faces overall, suggesting awareness ratings were more sensitive to face discrimination when faces were in an upright orientation than in an inverted orientation. (f) Metacognitive bias scores for reporting subjective awareness. Lower scores denote a more liberal bias to report higher confidence. Meta-bias increased as exposure duration increased, suggesting that participants were more willing to report lower confidence than in shorter exposure durations. Asterisks index statistically significant differences between face orientations. Error bars represent 95% CI.

#### *4.2.2.2 Location response bias*

We examined whether participants' response bias for reporting face location varied across conditions by entering the absolute values of $C_{location}$ scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. Response bias significantly decreased with exposure duration ($F_{(1.73,\,53.66)} = 18.754, p < .001, \eta p^2 = .377$), indicating that as participants' ability to discriminate the face increased (shown by higher location d' scores) they became less likely to exhibit a systematic bias in their preference to report one side or the other (Figure 4.3b). We did not find a main effect of expression ($F_{(1,\,31)} = 0.026, p = .872, \eta p^2 = .001$), but we did find an effect of face orientation ($F_{(1,\,31)} = 5.718, p = .023, \eta p^2 = .156$), indicating slightly greater bias for inverted faces ($M =$

$0.454\ [0.177]$) over upright faces ($M = 0.388\ [0.210]$). To assess whether the obtained data support the absence of an effect of expression, we estimated Bayes factors, which indicated strong evidence for the null hypothesis model ($BF_{01} = 10.380$). No interaction reached significance (all p > .078).

### 4.2.2.3 *Expression identification sensitivity*

We examined whether participants' sensitivity at identifying emotional expression varied across conditions by entering identification d' scores – taken over all trials irrespective of location response – into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.3c). A main effect of exposure duration indicated that sensitivity to expression increased with increasing exposure duration ($F_{(4.22,\ 130.93)} = 12.89, p < .001, \eta p^2 = .294$). We also found a main effect of face orientation ($F_{(1,\ 31)} = 19.54, p < .001, \eta p^2 = .387$), such that expression identification d' was significantly higher for upright faces ($M = 0.243\ [0.235]$) than for inverted faces ($M = 0.088\ [0.111]$). The interaction between face orientation and exposure duration also reached significance ($F_{(4.75,\ 147.15)} = 3.12, p = .012, \eta p^2 = .091$). To determine the minimal exposure duration that elicited the advantage of upright faces over inverted faces in expression identification, we ran post hoc pairwise comparisons. They revealed a significant advantage of upright faces over inverted faces at 5.3 ms ($t(209) = 3.563$, $p = .041$, $d = 0.630$) and 6.2 ms ($t(209) = 4.762$, $p < .001$, $d = 0.842$) of exposure.

To determine the minimal required exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance identification was 3.5 ms for upright ($M = 0.220\ [0.349]$; $t(31) = 3.57$, $p = .001$, $d = 0.631$) and 4.4 ms for inverted faces ($M = 0.137\ [0.291]$; $t(31) = 2.67$, $p = .012$, $d = 0.471$).

These results indicate that sensitivity to expression identity requires about 5.3 ms of exposure to exhibit a face-inversion effect, thereby suggesting that this minimal required exposure was sufficient for the successful integration of facial features in expression identification.

#### 4.2.2.4 *Expression identification criterion*

We examined whether participants' criteria for reporting fearful expression varied across conditions by entering $C_{identification}$ scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.3d). The lower the value of this measure, the more willing a participant is to report a fearful expression (liberal criterion). We did not find a main effect of exposure duration $(F_{(1.55,\ 48.11)} = 2.05, p = .150, \eta p^2 = .062)$ or of face orientation $(F_{(1,\ 31)} = 3.11, p = .088, \eta p^2 = .091)$. However, we did find a significant interaction between these two factors $(F_{(3.51,\ 108.74)} = 3.16, p = .021, \eta p^2 = .092)$. To determine differences between face orientations per each exposure duration, we used post hoc pairwise comparisons, but no comparison of relevance reached significance. We calculated Bayes factors to test whether the data support this absence of an orientation effect. Surprisingly, Bayes factors indicated substantial evidence in favour of the alternative hypothesis model $(BF_{01} = 0.0003)$, thus suggesting that participants may have exhibited a more liberal criterion when reporting fearful expressions. We also calculated Bayes factors to test whether the data support this absence of an effect of exposure duration. Bayes factors indicated strong evidence in favour of the null hypothesis model $(BF_{01} = 15.959)$, confirming that criterion scores did not vary across exposure durations. These results suggest that identification criterion did not become more liberal across increasing exposure duration but might have been more liberal for the identification of fearful faces when the face was presented in an upright orientation than in an inverted orientation.

#### 4.2.2.5 *Awareness-based metacognitive sensitivity*

We examined whether awareness scores were sensitive to participants' location sensitivity scores by estimating meta-d', a measure of metacognitive sensitivity. To examine whether metacognitive sensitivity varied across conditions, we entered meta-d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.3e). A main effect of exposure duration indicated that meta-d' increased with increasing exposure duration ($F_{(3.69,\ 114.38)} = 12.922, p < .001, \eta p^2 = .294$). We also found a main effect of face orientation ($F_{(1,\ 31)} = 6.475, p = .016, \eta p^2 = .173$), which indicates better metacognitive sensitivity to upright faces ($M = 0.540\ [0.328]$) than inverted faces ($M = 0.339\ [0.271]$). However, we did not find a main effect of expression ($F_{(1,\ 31)} = 0.110, p = .742, \eta p^2 = .004$), suggesting that emotional expression did not affect metacognitive sensitivity. We did not find an interaction between face orientation and exposure duration ($F_{(4.11,\ 127.48)} = 0.512, p = .732, \eta p^2 = .016$), between expression and exposure duration ($F_{(4.80,\ 148.72)} = 0.588, p = .702, \eta p^2 = .019$), between face orientation and expression ($F_{(1,\ 31)} = 1.048, p = .314, \eta p^2 = .033$), and no three-way interaction either ($F_{(4.48,\ 138.77)} = 0.321, p = .882, \eta p^2 = .01$). These results suggest upright faces reach conscious access faster than inverted faces.

As described, we did not find a main effect of expression, therefore we calculated Bayes factors to test whether the obtained data support this absence of an effect. Bayes factors indicated strong evidence in favour of the null hypothesis model ($BF_{01} = 13.320$), confirming that neither facial expression is prioritised by metacognitive sensitivity.

Finally, to determine the minimal required exposure that exhibited above-chance metacognitive sensitivity (meta-d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 1.7 ms for upright fearful ($M = 0.325\ [0.427];\ t(31) = 4.31, p < .001, d = 0.762$) and upright neutral faces ($M = 0.237\ [0.456];\ t(31) = 2.94, p = .006, d = 0.519$), and 2.6 ms for inverted fearful ($M = 0.278\ [0.758];\ t(31) = 2.07, p = .047, d = 0.366$) and inverted neutral faces ($M = 0.388\ [0.632];\ t(31) = 3.47, p = .002, d = 0.613$). Thus, our results suggest that it takes around 2 ms of exposure for a face stimulus to reach above-chance metacognitive sensitivity.

#### 4.2.2.6 *Metacognitive bias*

Metacognitive bias (meta-bias) is the tendency to give high confidence ratings regardless of actual performance. In this experiment, however, we used the PAS, a more exhaustive measure of awareness than confidence ratings (Sandberg et al., 2010). The PAS allows participants to rate their visual experience using four ratings covering from "no experience" to "clear experience" reports. Participants described their visual experience of a face and, therefore, metacognitive bias describes the tendency to describe one's visual experience as clear. To examine whether metacognitive bias varied across conditions, we entered meta-bias scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.3f). A main effect of exposure duration indicated that as exposure duration increased, meta-bias became more conservative ($F_{(3.02, 93.65)} = 5.371, p = .002, \eta p^2 = .148$), i.e. participants' tendency to give high confidence ratings decreased with exposure duration. However, we did not find an effect of orientation ($F_{(1, 31)} = 0.749, p = .393, \eta p^2 = .024$) or expression ($F_{(1, 31)} = 0.003, p = .958, \eta p^2 = 0$). No interaction reached significance either: between expression and orientation ($F_{(1, 31)} = 0.072, p = .791, \eta p^2 = .002$), expression and exposure duration ($F_{(3.22, 99.88)} = 0.469, p = .718, \eta p^2 = .015$), orientation and exposure duration ($F_{(3.48, 107.84)} = 0.275, p = .870, \eta p^2 = .009$), and the three-way interaction ($F_{(3.41, 105.61)} = 0.299, p = .850, \eta p^2 = .01$).

We calculated Bayes factors to test whether the obtained data support this absence of an effect of orientation and expression. Bayes factors indicated strong evidence in favour of the null hypothesis model of orientation ($BF_{01} = 10.022$) and of the null hypothesis model of expression ($BF_{01} = 13.329$), confirming that metacognitive bias was unaffected by emotional expression and face orientation.

### 4.2.3 Discussion

This first experiment found that discrimination between intact and scrambled faces increases from no sensitivity (chance performance) to above-chance sensitivity in less than three milliseconds of exposure. Therefore, very little visual information was required for stimulus discrimination. In addition, we found a face-inversion effect – locations of upright faces were more easily discriminated than inverted faces – which required only a little over four milliseconds of exposure to emerge. This finding suggests that holistic face processing provided an advantage, starting from 4.4 ms of exposure. However, we did not find an advantage of fearful expressions over neutral expressions, suggesting that perceptual sensitivity is not enhanced for emotional content. These findings were complemented by type-2 signal detection analyses, as we also found a metacognitive sensitivity advantage for upright faces over inverted faces, suggesting that the former reach higher meta-d' scores than the latter as exposure duration increases. Similar to our type-1 SDT findings, we did not find an advantage of fearful expressions over neutral expressions on this measure either, suggesting that emotional content is not prioritised for faces' access to awareness.

Expression identification required more visual exposure than location discrimination to reach above-chance sensitivity – around four milliseconds. Expression identification was accompanied by a face-inversion effect, found (only slightly later) with 5.3 ms of exposure, suggesting that holistic processing may make a crucial contribution to identifying a facial expression as emotional or non-emotional. Indeed, fearful expressions were more easily identified when shown in upright orientation.

Importantly, these findings demonstrate that face discrimination and awareness arise together: perceptual performance does not seem to arise earlier (with between 1.7 and 2.6 ms as indicated by type-1 SDT measures) than awareness (also with between 1.7 and 2.6 ms as indicated by type-2 SDT measures); overall, we find no convincing evidence for unconscious face processing that takes place before faces gain access to awareness.

Taken together, these findings demonstrate that face processing requires a minimal exposure duration of just a few milliseconds, and suggest that different aspects of faces might be processed in a hierarchical sequence: as exposure duration increased, participants exhibited evidence first for detection of faces, then for holistic face processing, and finally for emotion processing; this was indicated by both type-1 SDT measures of perceptual performance and type-2 SDT measures of perceptual awareness. Importantly, fearful and neutral expressions gained access to perception and awareness in a similar manner as exposure duration increased, suggesting that emotional content does not modulate early visual processing of faces.

In the absence of masking, however, image processing may not have ended at stimulus offset; we cannot rule out a contribution of afterimage processing. The image stimuli presented in Experiment 9 could have left a retinal impression (afterimage) for a longer duration than the exposure durations we predefined. How much did afterimage processing contribute to perceptual sensitivity? We address this question in the following control experiment.

## 4.3 Experiment 10 (Control)

To test whether perceptual discrimination in Experiment 9 – expressed in sensitivity scores – could have been affected by afterimage processing, we developed this first control experiment. Afterimages are retinal impressions that persist after stimulus removal. They exhibit inverted lightness levels and complementary colours to the stimulus that caused it (Shimojo et al., 2001). We thus tested whether stimuli that have these properties can elicit the effects found in Experiment 9. To emulate afterimages, we inverted the stimuli's colour and presented participants with those new face images instead. To avoid creating a positive afterimage (i.e. an afterimage of the afterimage-like stimuli), we used backward masking. In addition, we only used one exposure duration of 10 ms, an exposure duration that should be sufficiently long to elicit the face-inversion effect both with location sensitivity scores and with expression identification sensitivity scores, as shown in Experiment 9. If afterimage processing made a substantial contribution to the findings reported in Experiment 9, we should find these two face-inversion effects with these new afterimage-like stimuli. If we do not, we can argue that these findings in Experiment 9 could not be attributed to afterimage processing.

### 4.3.1 Methods

#### *4.3.1.1 Participants*

All the participants of Experiment 9 did this control experiment 10 minutes after having finished Experiment 9.

#### *4.3.1.2 Stimuli*

Stimuli were the same as in Experiment 9, but with inverted colours to emulate negative afterimages, including both intact and scrambled face images (Figure 4.4; for full layout of afterimage stimuli, see Appendix E).

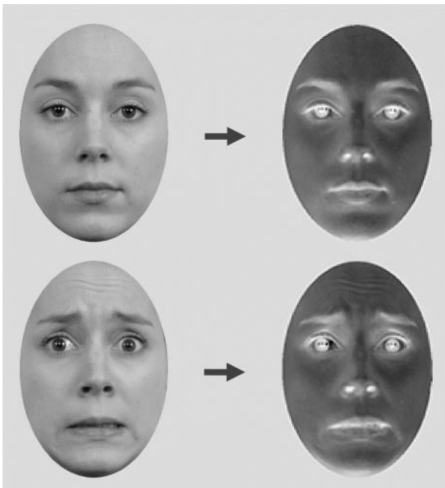


**Figure 4.4.** Examples of afterimage-like stimuli. (Left) Intact face with neutral (top) and fearful (bottom) expression. (Right) Their colour-inverted afterimage-like counterparts.

#### 4.3.1.3 *Procedure*

The procedure was similar to that of Experiment 9, but with only one exposure duration (10 ms); hence this experiment had fewer trials in total (160). In addition, backward masking was used to prevent the new stimuli from creating afterimages. To achieve this, on each trial we presented the non-colour-inverted scrambled version of the face used in that trial. The mask was displayed for 50 ms immediately after stimulus offset in both intact and scrambled stimulus locations. Then, as in Experiment 9, participants were asked to first judge the location and expression of the intact face stimulus, followed by rating their visual experience (PAS).

### 4.3.2 Results

#### 4.3.2.1 *Location sensitivity*

To examine whether face orientation and expression modulate participants' ability to discriminate the location of afterimage-like face stimuli, we entered location d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) repeated-measures ANOVA (Figure 4.5a). We did not find an effect of expression $(F_{(1,31)} = 0.038, p = .846, \eta p^2 = .001)$ or of orientation $(F_{(1,31)} = 1.926, p = .175, \eta p^2 = .058)$, indicating that neither factor modulated location d'. The interaction between these two factors did not reach significance either $(\mathrm{F}_{(1,31)} = 0.073, \mathrm{p} = .789, \eta p^2 = .002)$, even though stimuli were presented for substantially longer than the duration that enabled such discriminations for intact faces. To assess whether the obtained data support the absence of an effect of expression and of orientation, we estimated Bayes factors, which indicated substantial evidence for the null hypothesis model in the former $(BF_{01} = 5.153)$ and in the latter $(BF_{01} = 4.974)$.

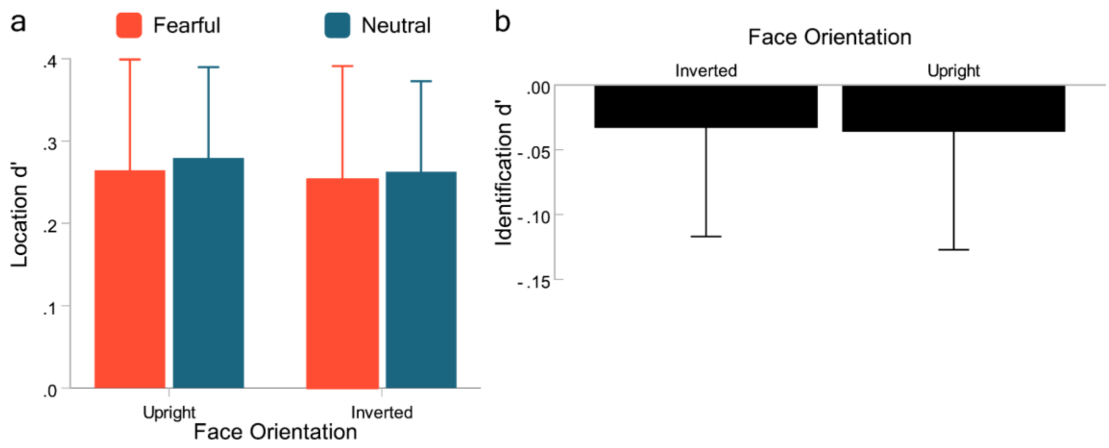



**Figure 4.5.** Results of Experiment 10 (Control). (a) Location sensitivity. Location d' scores were very low and unaffected by expression or orientation. (b) Identification sensitivity for expression. Identification d' was at chance level both for upright and inverted faces. Error bars represent 95% CI.

To examine whether participants could tell, above chance, on what side the intact afterimage-like faces were, we ran a series of uncorrected one-sample t-tests against zero. We found that upright fearful $(M = 0.264\ [0.374];\ t(31) = 4.00,\ p < .001, d = 0.707)$, upright neutral $(M = 0.280\ [0.306];\ t(31) = 5.18,\ p < .001, d = 0.915)$, inverted fearful $(M = 0.253\ [0.382];\ t(31) = 3.76,\ p < .001, d = 0.664)$, and inverted neutral faces $(M = 0.263\ [0.305];\ t(31) = 4.87,\ p < .001, d = 0.860)$ were discriminated above chance, thus indicating that participants could discriminate between intact and scrambled afterimage-like faces, though with low sensitivity.

### 4.3.2.2 *Expression identification sensitivity*

To examine whether face orientation modulates participants' ability to identify the expression of afterimage-like face stimuli, we compared identification d' scores for upright and inverted faces by using a paired-sample t-test. We did not find significant differences between face orientations in identification sensitivity $(t(31) = -0.808, p = .787, d = -0.143)$. To assess whether the obtained data support the absence of an effect of face orientation, we estimated Bayes factors, which indicated substantial evidence for the null hypothesis model $(BF_{01} = 3.920)$. These results suggest that participants are not able to identify fearful expressions and distinguish them from neutral expressions by using afterimage processing (Figure 4.5b).

### 4.3.3 Discussion

In this control experiment, we presented participants with images that resembled the afterimages that may have been generated by viewing a face. This approach allowed us to test whether afterimage processing could fully or partly explain the findings of Experiment 9. Crucially, we did not find face-inversion effects in location or identification sensitivity, indicating that holistic face processing was not engaged with afterimage-like faces. Therefore, these results suggest that afterimage processing could not explain our findings of Experiment 9. Similarly, we found extremely low expression identification sensitivity for both upright and inverted faces. These results suggest the findings of Experiment 9 cannot be attributed to afterimage processing.

## 4.4 Experiment 11

The visual task used in Experiment 9 presented participants with two images per trial, one on each side of the screen. Therefore, visual processing involved in both the location and expression identification tasks mainly relied on peripheral vision. In the retinal periphery, the concentration of photoreceptors is much lower than in the foveal region (i.e. centre of the visual field), thereby enjoying lower resolution than the latter (Baden et al., 2020). In Experiment 11, we adapted our task so visual processing depended on foveal vision, where resolution is the highest. To accomplish this, we turned the previous 2AFC task into a two-interval forced choice (2IFC) task. Intact and scrambled face images were presented at the centre of the screen one after the other. Participants had to discriminate when (first or second) the intact face image was shown. To anticipate the possibility that foveal vision would allow better perceptual performance than peripheral vision, we used shorter exposure durations than in Experiment 9 (0.6 – 6 ms, each duration 200μs-shorter than its Experiment 9 counterpart). Importantly, the design of this experiment allowed us to examine whether the effect of orientation, observed in Experiment 9, would replicate for foveal vision; and whether an effect of expression, absent in Experiment 9, might arise when images are presented foveally. Measuring perceptual sensitivity in peripheral (Experiment 9) and foveal vision (Experiment 11) allows for a more general description of how faces are visually processed and gain access to awareness.

### 4.4.1 Methods

#### *4.4.1.1 Participants*

Thirty-four students of the Université Libre de Bruxelles provided informed consent and were paid €10 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Two participants were excluded from the analysis (see Analysis section). The remaining 32 participants (19 female; 5 left-handed) had a mean age of 24.6 ($SD_{age}$ = 5.5; range: 18 – 28).

#### *4.4.1.2 Stimuli*

Stimuli were the same as in Experiment 9, but the range of presentation durations differed in that each duration in this experiment was 200 μs shorter than its Experiment 9 counterpart: 0.6, 1.5, 2.4, 3.3, 4.2, 5.1, and 6 ms.

### 4.4.1.3 *Procedure*

The procedure was similar to that of Experiment 9 (Figure 4.6), with the following differences: each trial began by presenting a fixation cross at the centre of the screen along with one pair of placeholder dots – one above and one below the fixation cross – for 400 ms. After fixation cross offset, the two placeholder dots were left on the screen for 300 ms. Then, one stimulus was displayed at the centre of the screen for a predetermined exposure duration (fixation cross offset was separated from stimulus onset by 300 ms to prevent the possibility that the visual transient of the cross's disappearance might forward-mask the stimulus). After stimulus offset, only the placeholder dots were left on the screen, for 300 ms. Then, the fixation cross was presented in the centre of the screen again for 400 ms. After fixation cross offset, only the placeholder dots were left on the screen for 300 ms. The second stimulus was displayed at the centre of the screen for the same predetermined exposure duration as the first one. Finally, after the second stimulus' offset, only the placeholder dots were left on the screen for 300 ms, after which the response cue was presented at the centre of the screen, prompting participants to report the presentation order (whether the intact face had appeared first or second) and identify its expression (emotional or non-emotional) with a single keypress using the same keys as in Experiment 9: the 'left Control' and 'Left Shift' keys were used for "first" reports, and the 'up arrow' and 'down arrow' keys were used for "second" reports. Block structure, breaks, number of trials, and practice were the same as in Experiment 9.

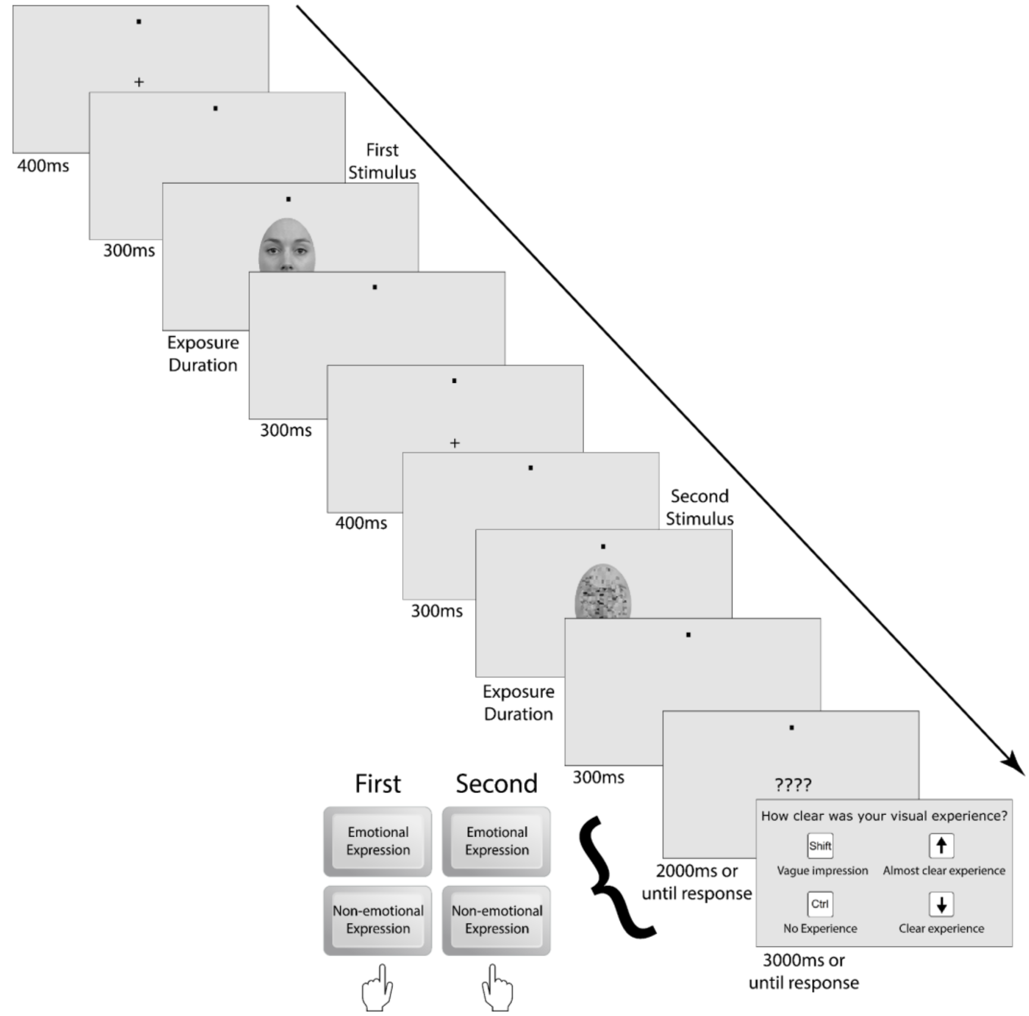



**Figure 4.6.** Schematic description of a trial in Experiment 11. Stimuli were presented for one of seven possible exposure durations (0.6 – 6 ms, equally spaced on a linear scale). After stimulus offset, participants judged the presentation order (whether the intact face was first or second) and its expression (emotional or non-emotional) by pressing one key. Next, they judged the clarity of their visual experience (PAS).

#### *4.4.1.4 Analyses*

Analyses were the same as in Experiment 9, with one exception: instead of location d' we estimated order d', i.e. sensitivity to the presentation order of the intact faces. A hit was defined as a trial in which a face was displayed first and reported first, and a false alarm was defined as a trial in which a face was displayed second but reported as shown first.

## 4.4.2 Results

### 4.4.2.1 *Order sensitivity*

To examine how the manipulated factors affected face discrimination, we entered order d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration, whereby sensitivity increased with increasing exposure duration ($F_{(2.23,\ 69.02)} = 180.787, p < .001, \eta p^2 = .854$). As in Experiment 9, order d' scores went from showing no sensitivity (chance performance) at the shortest exposure duration to showing high sensitivity at the longest exposure durations (Figure 4.7a). We also found a main effect of face orientation ($F_{(1,\ 31)} = 49.058, p < .001, \eta p^2 = .613$), indicating a sensitivity advantage for upright faces ($M = 0.970\ [0.785]$) over inverted faces ($M = 0.709\ [0.616]$), but did not find a main effect of expression ($F_{(1,\ 31)} = 0.761, p = .390, \eta p^2 = .024$), indicating no advantage of fearful expressions over neutral expressions. Importantly, the interaction between face orientation and exposure duration was significant ($F_{(4.98,\ 154.24)} = 10.713, p < .001, \eta p^2 = .257$). To see at what specific exposure durations upright faces enjoyed significantly better sensitivity over inverted faces, and thus answer our main question, we ran post hoc Bonferroni-corrected pairwise comparisons, which revealed said advantage at 3.3 ms ($t(187) = 4.737,\ p < .001,\ d = 0.837$), 4.2 ms ($t(187) = 7.515,\ p < .001,\ d = 1.328$), 5.1 ms ($t(187) = 6.694,\ p < .001,\ d = 1.183$), and 6 ms of exposure ($t(187) = 4.837,\ p < .001,\ d = 0.855$). Crucially, 3.3 ms of exposure were sufficient to elicit a face-inversion effect when perceptual discrimination relied on foveal vision. As in Experiment 9, we did not find an interaction between expression and face orientation ($F_{(1,\ 31)} = 0.504, p = .483, \eta p^2 = .016$), or between expression and exposure duration ($F_{(5.11,\ 158.36)} =$

$0.328, p = .899, \eta p^2 = .010)$, or a three-way interaction $(F_{(4.97,\,154.08)} = 1.313, p < .262, \eta p^2 = .041)$.

To determine the minimal exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 1.5 ms for upright fearful $(M = 0.197\ [0.299];\ t(31) = 3.72, p < .001, d = 0.658)$, and 2.4 ms for upright neutral faces $(M = 0.450\ [0.534];\ t(31) = 4.77, p < .001, d = 0.842)$, inverted fearful faces $(M = 0.358\ [0.364];\ t(31) = 5.57, p < .001, d = 0.984)$, and inverted fearful faces $(M = 0.367\ [0.534];\ t(31) = 3.90, p < .001, d = 0.689)$. Our results suggest that it takes around 2 ms of exposure for a face stimulus to reach above-chance discrimination from its scrambled counterpart when relying on foveal vision.

Like in Experiment 9, we did not find a main effect of expression. Therefore, we calculated Bayes factors to test whether the obtained data support the absence of this effect (null hypothesis model). Bayes factors indicated strong evidence in favour of the null hypothesis model $(BF_{01} = 11.891)$. This analysis suggests that fearful expressions are not prioritised by perceptual sensitivity over neutral expressions.

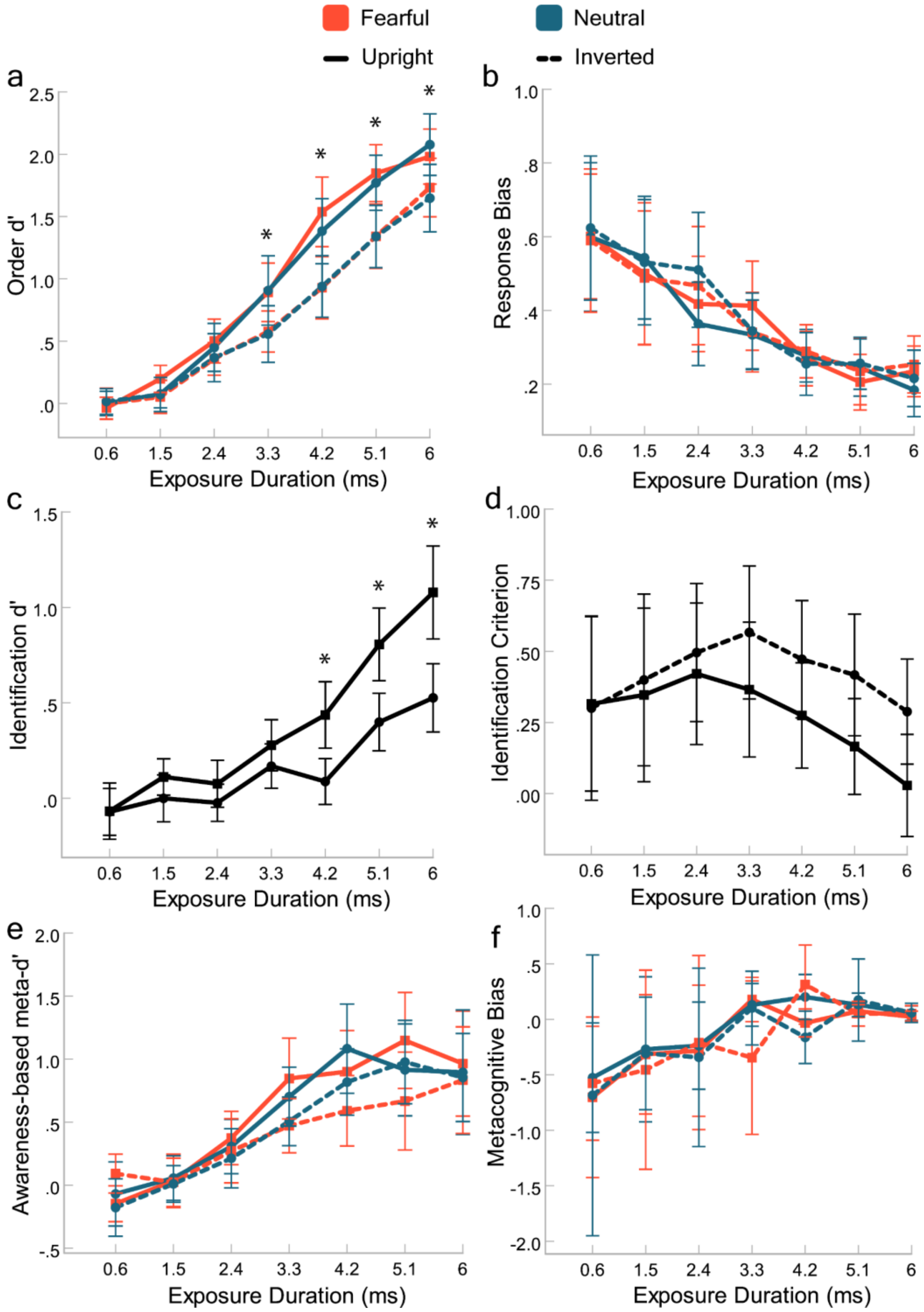



**Figure 4.7**. Results of Experiment 11. (a) Order sensitivity. Order d' increased with increasing exposure duration. The face-inversion effect arises by 3.3 ms of exposure. (b) Absolute-value response bias scores for reporting presentation order of the intact face (bias toward reporting either first or second). The amount of bias decreased as exposure duration increased, but there was no difference in amount of response bias between expressions and orientations. (c) Identification sensitivity for expression. Identification d'

increased with increasing exposure duration. A significant advantage in expression identification of upright faces over inverted faces arises by 4.2 ms of exposure. (d) Criterion scores for reporting expression. Upright faces exhibit a significantly more liberal criterion (lower values) than inverted faces during expression identification. (e) Awareness-based metacognitive sensitivity. Meta-d' increased with increasing exposure duration. Upright faces enjoyed higher scores than inverted faces overall, suggesting awareness ratings were more sensitive to presentation order discrimination when faces were in an upright orientation than in an inverted orientation. (f) Metacognitive bias scores for reporting subjective awareness. Lower scores denote a more liberal bias to report higher confidence. Meta-bias increased as exposure duration increased, suggesting that participants were more willing to exhibit lower confidence than in shorter exposure durations. Asterisks index statistically significant differences between face orientations. Error bars represent 95% CI.

#### *4.4.2.2 Order response bias*

We examined whether participants' response bias for reporting the intact face presentation order varied across conditions by entering the absolute values of $C_{order}$ scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. As in Experiment 9, response bias significantly decreased with increasing exposure duration $(F_{(1.98,\, 61.31)} = 12.581, p < .001, \eta p^2 = .289)$, indicating that as participants' ability to discriminate the face increased, they became less likely to exhibit a systematic bias in their preference to report one presentation order or the other (Figure 4.7b). We did not find a main effect of expression $(F_{(1,\, 31)} = 0.014, p = .906, \eta p^2 = 0)$, or of face orientation $(F_{(1,\, 31)} = 1.099, p = .303, \eta p^2 = .034)$. No interaction reached significance (all p > .268).

To assess whether the obtained data support the absence of an effect of expression and of orientation, we estimated Bayes factors, which indicated strong evidence for the null hypothesis model of expression $(BF_{01} = 13.746)$ and of orientation $(BF_{01} = 10.287)$.

### 4.4.2.3 *Expression identification sensitivity*

We examined whether participants' sensitivity at identifying expression varied across conditions by entering identification d' scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.7c). We found a main effect of exposure duration, indicating that sensitivity to expression increased with increasing exposure duration ($F_{(3.36,\ 104.12)} = 36.20, p < .001, \eta p^2 = .539$). We also found a main effect of face orientation ($F_{(1,\ 31)} = 27.87, p < .001, \eta p^2 = .473$), such that expression identification sensitivity was significantly higher for upright faces ($M = 0.389\ [0.418]$) than for inverted faces ($M = 0.155\ [0.227]$). The interaction between orientation and exposure duration also reached significance ($F_{(4.99,\ 154.54)} = 6.08, p < .001, \eta p^2 = .164$). To determine the minimal exposure duration that elicited this face-inversion effect in expression identification sensitivity, we ran post hoc Bonferroni-corrected pairwise comparisons. They revealed significant advantages of upright faces over inverted faces at 4.2 ms ($t(198) = 3.967, p = .009, d = 0.701$), 5.1 ms ($t(198) = 4.632,\ p < .001,\ d = 0.819$), and 6 ms ($t(198) = 6.279, p < .001, d = 1.110$) of exposure.

To determine the minimal exposure that exhibited above-chance performance (d' > 0) in expression identification, we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 1.5 ms for upright faces ($M = 0.112\ [0.262];\ t(31) = 2.41,\ p = .022,\ d = 0.427$), and 3.3 ms for inverted faces ($M = 0.169\ [0.322];\ t(31) = 2.96,\ p = .006,\ d = 0.524$).

### 4.4.2.4 *Expression identification criterion*

We examined whether participants' criterion for reporting fearful expression varied across conditions by entering $C_{identification}$ scores into a 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA. We did not find a significant main effect of exposure duration ($F_{(1.46,\ 45.36)} = 2.66, p = .096, \eta p^2 = .079$), but we did find a main effect of face orientation $(F_{(1,31)} = 8.91, p = .005, \eta p^2 = .223)$, indicating a more liberal criterion for upright faces $(M = 0.274\ [0.135])$ than inverted faces $(M = 0.420\ [0.102])$, i.e. participants were more willing to report a fearful expression for an upright face than for an inverted face (Figure 4.7d). The interaction between face orientation and exposure duration reached significance $(F_{(3.79,117.44)} = 3.26, p = .016, \eta p^2 = .095)$. To determine differences between face orientations per each exposure duration, we ran post hoc Bonferroni-corrected pairwise comparisons. However, no comparison of interest reached significance. To assess whether the obtained data support the absence of an effect of exposure duration, we estimated Bayes factors, which unexpectedly indicated anecdotal evidence for the alternative hypothesis model $(BF_{01} = 0.046)$, i.e. in favour of an effect of exposure duration. These results suggest that identification criterion was only affected by face orientation. However, the data are inconclusive regarding whether exposure duration contributed to identification criterion.

#### *4.4.2.5 Awareness-based metacognitive sensitivity*

We examined whether awareness scores were sensitive to participants' order sensitivity scores by estimating meta-d'. To examine whether metacognitive sensitivity varied across conditions, we entered meta-d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.7e). A main effect of exposure duration indicated that meta-d' increased with increasing exposure duration $(F_{(3.41,105.73)} = 29.216, p < .001, \eta p^2 = .485)$. We also found a main effect of face orientation $(F_{(1,31)} = 6.176, p = .019, \eta p^2 = .166)$, which indicates better metacognitive sensitivity to upright faces $(M = 0.573\ [0.461])$ than inverted faces $(M = 0.440\ [0.370])$. However, we did not find a main effect of expression $(F_{(1,31)} = 0, p = .977, \eta p^2 = 0)$, thus suggesting emotional expression did not affect metacognitive sensitivity. No other interaction reached significance (all $p > .298$). These results suggest that upright faces reach awareness faster than inverted faces.

As described, we did not find a main effect of expression, therefore we calculated Bayes factors to test whether the obtained data support this absence of an effect. Bayes factors indicated strong evidence in favour of the null hypothesis model ($BF_{01} = 13.343$), thus suggesting that fearful expressions are not prioritised by metacognitive sensitivity when compared to neutral expressions.

Finally, to determine the minimal exposure that exhibited above-chance metacognitive sensitivity (meta-d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 2.4 ms for upright fearful faces ($M = 0.374$ [0.587]; $t(31) = 3.60, p = .001, d = 0.637$), upright neutral faces ($M = 0.307$ [0.597]; $t(31) = 2.91, p = .007, d = 0.514$), inverted fearful faces ($M = 0.273$ [0.706]; $t(31) = 2.18, p = .037, d = 0.386$), and 3.3 ms for inverted neutral faces ($M = 0.507$ [0.535]; $t(31) = 5.36, p < .001, d = 0.948$). Thus, our results suggest that it takes around 3 ms of exposure for a face stimulus to reach above-chance metacognitive sensitivity.

#### 4.4.2.6 *Metacognitive bias*

We estimated metacognitive bias as in Experiment 9. To examine whether metacognitive bias varied across conditions, we entered meta-bias scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 7 (exposure durations) repeated-measures ANOVA (Figure 4.7f). A main effect of exposure duration indicated that metacognitive bias became less liberal with increasing exposure duration ($F_{(3.21, 99.55)} = 4.127, p = .007, \eta p^2 = .117$), thus indicating that as participants' metacognitive sensitivity increased, they became less likely to systematically report having no experience (or clear experience) of the intact stimulus shown. However, we did not find an effect of face orientation ($F_{(1, 31)} = 0.420, p = .522, \eta p^2 = .013$) or of expression ($F_{(1, 31)} = 0.139, p = .712, \eta p^2 = .004$). No interaction reached significance either (all p > .550).

We calculated Bayes factors to test whether the obtained data support this absence of an effect of orientation and expression. Bayes factors indicated strong evidence in favour of the null hypothesis model of orientation ($BF_{01} = 11.326$) and of expression

$(BF_{01} = 12.287)$, thus suggesting that neither fearful (compared to neutral) expressions nor upright (compared to inverted) faces are prioritised by metacognitive sensitivity.

### 4.4.3 Discussion

In this experiment, we tested how discrimination of face presentation order and expression identification increase across increasing exposure durations. Unlike in Experiment 9, where we measured sensitivity in a 2AFC task that relied on peripheral vision, in Experiment 11 we measured sensitivity in a 2IFC task, which relied on foveal vision, where resolution is the highest. We found equivalent results to those reported in Experiment 9, in all sensitivity and criterion measures. Interestingly, the face-inversion effect was elicited at 3.3 ms, whereas in Experiment 9 it arose at 4.4 ms, suggesting that holistic face processing might have benefited from foveal-based visual discrimination.

Taken together, findings from both Experiment 9 and Experiment 11 converge to show how stimulus discrimination, holistic face processing, and emotion identification emerge, in this order, across six milliseconds of visual exposure. Both experiments also converge on metacognitive sensitivity by showing that upright faces gain access to awareness faster than inverted faces, whereas emotional content in faces does not affect either perceptual or metacognitive sensitivity. Importantly, these findings demonstrate that face discrimination and awareness arise together: perceptual performance does not seem to arise earlier (with between 1.5 and 2.4 ms as indicated by type-1 SDT measures) than awareness (with between 2.4 and 3.3 ms as indicated by type-2 SDT measures), taking into account that the data were noisy; overall, as in Experiment 9, we find no convincing evidence for unconscious face processing that takes place before faces gain access to awareness.

## 4.5 Experiment 12 (Control)

In Experiment 10 (Control), we tested whether findings in Experiment 9 could be affected by afterimage processing. Our findings suggested that neither face orientation nor expression could be discriminated with afterimage-like stimuli presented for 10 ms (over 3 ms longer than the longest-duration stimuli in Experiment 9). However, that approach presented two limitations. First, we used backward masking to prevent those afterimage-like stimuli from creating additional afterimages. Yet backward masking could have also disrupted perceptual sensitivity, thus making the procedure less reliable. Second, we only used one exposure duration, making it impossible to determine whether exposure duration modulates the discrimination of intact afterimage-like stimuli from their scrambled afterimage-like counterparts. In this second control experiment, we addressed these issues with the 2IFC task used in Experiment 11.

We employed three exposure durations that covered the range of values used in Experiment 11: from when order sensitivity was at chance (0.6 ms), when the face-inversion effect arose (3.3 ms), and when location sensitivity was very high (6 ms). We used the same afterimage-like stimuli of Experiment 10 (Control), but this time we did not use any masking. If these stimuli exhibit a modulatory effect of face orientation on order sensitivity (a face-inversion effect), it is likely that afterimage processing may have contributed to these same effects in Experiment 11. Conversely, if these stimuli do not exhibit said effect, it would suggest that the face-inversion effect found in Experiment 11 cannot be explained by afterimage processing.

### 4.5.1 Methods

#### *4.5.1.1 Participants*

All the participants of Experiment 11 did this control experiment 10 minutes after having finished Experiment 11.

### 4.5.1.2 *Stimuli, Procedure, and Analyses*

Stimuli were the same as in Experiment 10 (Control), and procedures were the same as in Experiment 11, but with only three exposure durations, hence fewer trials in total (480). Analyses were the same as in Experiment 11 for sensitivity and criterion.

## 4.5.2 Results

### 4.5.2.1 *Order sensitivity*

To examine how the manipulated factors affected order presentation discrimination of afterimage-like faces, we entered order d' scores into a 2 (expression: fearful, neutral) × 2 (orientation: upright, inverted) × 3 (exposure durations) repeated-measures ANOVA. Importantly, a main effect of exposure duration was found ($F_{(1.70,\, 52.74)} = 114.565, p < .001, \eta p^2 = .787$) indicating that order sensitivity increased with increasing exposure duration when using afterimage-like faces (Figure 4.8a). However, we did not find an effect of expression ($F_{(1,\, 31)} = 0.011, p = .917, \eta p^2 = 0$) nor of orientation ($F_{(1,\, 31)} = 2.698, p = .111, \eta p^2 = .080$), thus suggesting neither expression nor face orientation affected sensitivity to afterimage-like faces. Importantly, this finding of a null effect of orientation suggests that the orientation effect found in Experiment 11 cannot be attributed to afterimage processing. The interaction between face orientation and exposure duration reached significance ($F_{(1.94,\, 60.18)} = 4.228, p = .020, \eta p^2 = .120$). However, post hoc Bonferroni-corrected pairwise comparisons did not reveal significant differences between upright and inverted faces at any exposure duration. No other interaction reached significance (all p > .456). To assess whether the obtained data support the absence of an effect of expression and of face orientation, we estimated Bayes factors, which indicated substantial evidence for the null hypothesis model in the former ($BF_{01} = 8.394$) and in the latter ($BF_{01} = 5.849$). These results suggest that discrimination of afterimage-like faces increases with exposure duration similarly as it did with regular images in Experiment 11. However, face orientation did not modulate discrimination this time, which may suggest that either afterimage processing does not contribute to holistic face processing or that it does but requires longer exposure durations to reveal a face-inversion effect.

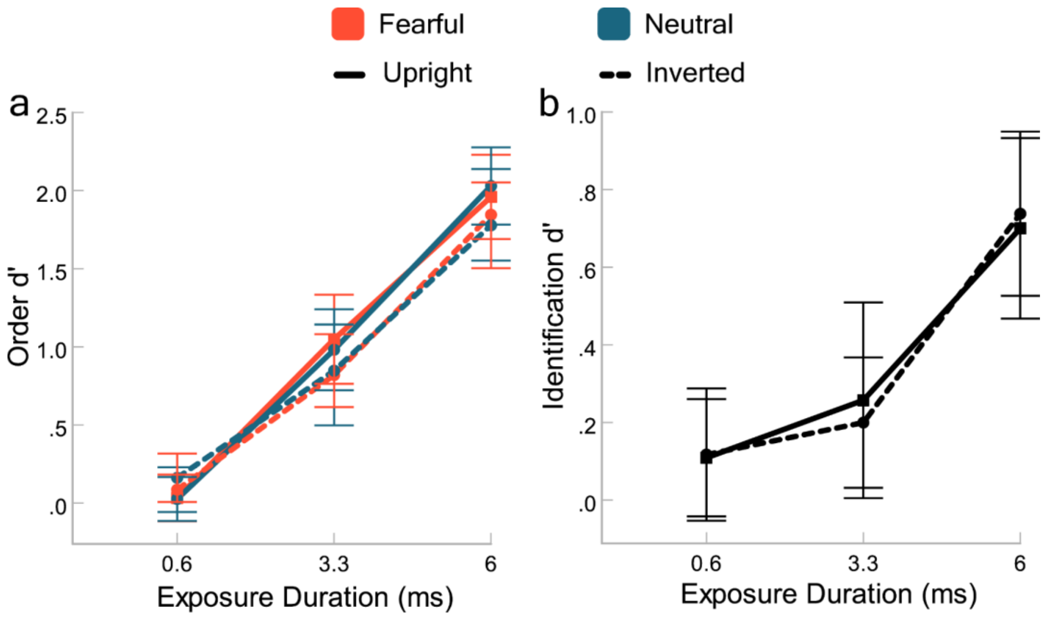



**Figure 4.8**. Results of Experiment 12 (Control). (a) Order sensitivity. Order d' increased with increasing exposure duration from chance discrimination to high discrimination but was unaffected by expression or orientation. (b) Identification sensitivity for expression. Identification d' increased with increasing exposure duration but was unaffected by orientation. Error bars represent 95% CI.

### 4.5.2.2 *Expression identification sensitivity*

To examine how the manipulated factors affected expression identification of afterimage-like faces, we compared identification d' scores for upright and inverted faces by entering the data into a 2 (orientation: upright, inverted) × 3 (exposure durations) repeated-measures ANOVA. A main effect of exposure duration was found ($F_{(1.88,\ 58.26)} = 16.521, p < .001, \eta p^2 = .348$), suggesting that expression identification sensitivity increased with increasing exposure duration (Figure 4.8b). However, we did not find an effect of face orientation $(F_{(1,\,31)} = 0.004, p = .952, \eta p^2 = 0)$, indicating that face orientation did not affect expression identification. Likewise, the interaction between these factors did not reach significance $(F_{(1.84,\,57.10)} = 0.195, p = .806, \eta p^2 = .006)$. To assess whether the obtained data support the absence of an effect of face orientation, we estimated Bayes factors, which indicated substantial evidence for the null hypothesis model $(BF_{01} = 6.648)$. These results suggest that expression identification increases with increasing exposure duration regardless of face orientation. Because turning faces upside down disrupts holistic face processing, these results might rely on discrimination of local face features preserved in afterimage-like faces rather than on holistic features.

### 4.5.3 Discussion

When a visual image is presented and then removed, retinal impressions (afterimages) can be formed. Because we could not measure, in Experiments 9 and 11, whether such afterimages were formed, to what extent they were processed, and whether they contributed to participants' sensitivity, we ran two control experiments to assess participants' sensitivity to afterimage-like images of faces and assess whether the results of Experiments 9 and 11 could be attributed – in whole or in part – to afterimage processing. In Experiment 10, we used afterimage-like faces to emulate afterimage processing by measuring participants' sensitivity to relatively long presentations of those stimuli (compared to the durations used in Experiment 9), and backward masking to prevent said afterimage-like stimuli from creating additional afterimages. We found above-chance (though weak) location sensitivity; participants were unable to identify expressions, though, and the inversion effect seen in Experiment 9 was absent. However, our procedure in Experiment 10 was different from that of Experiment 9 – our use of backward masking may have interacted with stimulus processing, and the use of a single presentation duration did not allow us to measure whether exposure duration modulated perceptual sensitivity.

In Experiment 12 (Control), we presented participants with unmasked stimuli, using three of the exposure durations used in Experiment 11. We found that participants could discriminate intact afterimage-like faces from their scrambled counterparts.

Crucially, they did not exhibit a face-inversion effect either in order discrimination or expression identification. These results suggest that in both Experiment 9 and Experiment 11 participants' sensitivity to faces might have partially relied on information coming from afterimage processing, but it is unlikely that such information contributed to holistic face processing as indexed by the face-inversion effect.

## 4.6 General Discussion

What is the minimal exposure duration required for face perception? Due to hardware limitations, past studies could not employ sufficiently brief exposure durations to address this question. Researchers had to rely on different strategies such as using masking techniques to interrupt visual processing. However, it is impossible to determine what aspects of visual processing masking techniques interrupt, or to what extent they are interrupted. In this Chapter, we circumvented the technical limitations of standard computer monitors by using a newly developed LCD tachistoscope that allows extremely brief visual presentations. This allowed us to present observers with unmasked faces using exposure durations ranging from less than one millisecond to around six milliseconds. This range of exposure values was sufficiently sensitive to capture changes in face discrimination, holistic processing, emotional expression identification, and perceptual awareness. By using type-2 signal detection analyses, we could describe changes in metacognitive sensitivity, i.e. how awareness discriminated changes in location discrimination performance across exposure durations. Thus, using extremely brief exposure durations, we obtained signal detection indices that have provided us with a convergent description of the sequence of processing steps that unfolds in the early stages of face processing.

We ran two experiments with seven linearly spaced exposure durations to track changes in stimulus discrimination and test whether upright faces enjoy an advantage over inverted faces, as shown in masking studies by others (Akechi et al., 2015; Jiang et al., 2007; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Senju, et al., 2011) and by us (see Chapters 2 and 3). In addition, we tested whether fearful expressions enjoy an advantage over neutral expressions, as claimed in other studies (Capitão et al., 2014; Hedger et al., 2015; Oliver et al., 2015; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018a; Zhan et al., 2015). We found remarkably similar minimal exposure durations for face discrimination, both when stimulus discrimination relied on peripheral vision (Experiment 9) and on foveal vision (Experiment 11). For holistic face processing, as indicated by a discrimination advantage in favour of upright over inverted faces, 4.4 ms of exposure were required when relying on peripheral vision (Experiment 9) and 3.5 ms of exposure when relying on foveal vision (Experiment 11), where resolution is the highest. These results elucidate the minimal required exposure duration for face-specific processing. Thus, we found convergent evidence for a minimal exposure duration required for holistic face processing.

Unlike with upright and inverted faces, we did not find evidence of better location (Experiment 9) or order discrimination (Experiment 11) for fearful expressions over neutral expressions. This was true for both type-1 and type-2 SDT measures, suggesting that emotional expressions are not prioritised for either perceptual processing or for access to perceptual awareness. Nonetheless, we did measure the minimal exposure duration for above-chance emotional expression identification, finding it to be between 3.5 ms (for upright faces) and 4.4 ms (for inverted faces) of exposure in Experiment 9 (peripheral vision). Similarly, we found that between 1.5 ms (for upright faces) and 3.3 ms (for inverted faces) of exposure were sufficient for above-chance identification performance in Experiment 11 (foveal vision). In addition, we found an advantage in expression identification for upright over inverted faces, i.e. a face-inversion effect in expression identification, with 4.4 ms of exposure in Experiment 9 and with 3.3 ms of exposure in Experiment 11. Therefore, the face-inversion effect for expression identification arose at an exposure of between 1 and 3 ms longer than the face-inversion effect for stimulus discrimination. These results may contradict past claims about an advantage of fearful expressions over neutral expressions in how faces gain access to perceptual awareness (Capitão et al., 2014; Oliver et al., 2015; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018a; Zhan et al., 2015). However, as previously suggested (see Chapter 3), those previous findings may have been due to low-level visual confounds (Gray et al., 2013; Hedger et al., 2015; Schlossmacher et al., 2017; Stein & Sterzer, 2012). Our stimuli, on the other hand, were matched for emotion recognition and equated in luminance and contrast. Therefore, these controls minimise the effects of low-level confounders that could provide fearful expressions with a perceptual advantage.

So far, our findings have presented a very consistent picture of the sequence of events in face processing: stimulus discrimination is followed by holistic processing, and then emotional expression identification. Our data provide evidence that the minimal durations required for perceptual processing and awareness of basic stimulus discrimination are similar. However, is this the case of other aspects of face processing and awareness such as holistic processing? In Experiments 9 and 11, we estimated metacognitive sensitivity by testing how awareness ratings predicted changes in location and order discrimination across exposure durations. While we did find evidence of better metacognitive sensitivity for upright over inverted faces overall, this advantage did not interact with exposure duration. Therefore, we cannot say whether this face-inversion effect occurred at a specific exposure duration.

As argued above, our results suggest that both face holistic processing and awareness of faces require roughly similar minimal exposure duration. This temporal convergence may suggest that the two processes are related. As described in Chapter 1, many claims have been made about what cognitive functions can do in absence of awareness. Some researchers have claimed that arithmetic operations (Sklar et al., 2012), semantic processing (Alsius & Munhall, 2013; Sklar et al., 2012; Yang & Yeh, 2011), object-background integration (Mudrik et al., 2011), visual configuration integration (Wang et al., 2012), and emotion detection (Capitão et al., 2014; Sterzer et al., 2011; Yang et al., 2007; Yang & Yeh, 2018a), do not require awareness, whereas other researchers have found that those claimed effects were possibly confounded by other factors or simply do not replicate (Gray et al., 2013; Hedger et al., 2015; Moors, Boelens, et al., 2016; Moors, Wagemans, van Ee, et al., 2016; Rabagliati et al., 2018). Our results suggest that awareness may be necessary for face holistic processing and emotion processing. In addition, the fact that both holistic processing and awareness required roughly similar minimal exposure durations might suggest that a function of awareness is information integration.

Our new method presents a crucial advantage over past methods in the study of complex visual stimuli. By using minimal exposure durations with sub-millisecond precision, we could reveal steps and bottlenecks in visual processing of faces without the problems inherent in other brief-presentation methods. However, employing unmasked visual stimuli presents one important limitation: participants could have produced negative afterimages, thereby confounding their performance. To see to what extent the effects seen in Experiment 9 and Experiment 11 could be attributed to afterimage processing, we ran two control experiments to estimate whether participants' sensitivity could have benefited from afterimage processing. In Experiment 10 (Control), we presented participants with masked colour-inverted stimuli to mimic the effects of afterimages and did not find a face-inversion effect, thus ruling out a contribution of afterimage processing on holistic face processing. Because backward masking could have disrupted visual processing beyond simply preventing additional afterimages, we ran Experiment 12 (Control), where we presented participants with unmasked colour-inverted stimuli, using three exposure durations. Participants' discrimination improved over increasing exposure duration, but they did not exhibit a face-inversion effect for either discrimination or expression identification. In conclusion, our findings suggest that afterimage processing could have contributed to stimulus discrimination, but it is rather unlikely that it contributed to holistic or expression identification processes.

In conclusion, by using an LCD tachistoscope with sub-millisecond precision, we were able to describe how faces are processed by presenting them for extremely brief exposure durations. Using a range of very brief exposure durations, we identified a sequence of processing steps that require different amounts of minimal exposure to occur, and clarified that out of two processes that have been claimed to enhance perception – holistic processing and emotional processing – only the former does.

Our findings raise new questions. We found a consistent sequence of processing steps for face perception. However, do neural systems engage with this sequence of processing steps in the same way? Or does the neural activity that yields behavioural performance unfold in ways that are not revealed by behavioural measures? Specifically, might neural measures reveal face processing that occurs before (i.e. at shorter exposure durations than) conscious face perception? This question is relevant to elucidating whether some facial features can be processed in absence of awareness. If, for example, neural systems can distinguish between emotional expressions before faces are discriminated, we could suggest unconscious emotion processing. On the other hand, if neural measures of emotion processing are only present at exposure durations that are similar to or longer than those enabling face discrimination, this would be evidence against unconscious emotion processing. We address these questions in Chapter 5.

# CHAPTER 5

# 5 MINIMAL REQUIRED EXPOSURE REVEALS GRADED NEURAL PROCESSING OF FACIAL CONFIGURATION, EMOTIONAL EXPRESSION, AND ACCESS TO AWARENESS

## 5.1 Introduction

Faces convey a wealth of crucial information for communication and social interaction, as described in previous chapters. Are all facial features processed in one step or in a sequence of steps? We have provided evidence for a sequence of processing steps in face perception, by measuring how faces gain access to perception and awareness based on peripheral (Experiment 9) and foveal vision (Experiment 11). First, stimulus discrimination between intact and scrambled faces, secondly holistic face processing, and thirdly holistic expression identification, all of which required less than 6 ms of visual exposure to arise. Awareness, indexed by metacognitive sensitivity, also required less than 6 ms of visual exposure to arise, increasing from no sensitivity to high sensitivity as exposure durations increased from below 1 ms to around 6 ms. These results suggest that facial features are processed in sequence. However, it could be the case that they are processed simultaneously in the visual cortex but transmitted to other brain areas at different speeds, resulting in the sequence of behavioural indices described above. Measuring neural markers of face processing, emotion processing, and perceptual awareness should clarify this.

In Chapter 5, we ask three follow-up questions based on the findings summarised above: first, do neural systems engage differently with emotional expressions than with neutral expressions before they gain access to perceptual awareness? If they do, then emotion can be processed in absence of awareness and stimulus integration. Secondly, what is the minimal exposure duration required for neural systems to track conscious awareness of faces? Arguably, any neural or behavioural process that can occur before neural systems engage with perceptual awareness of faces would unveil unconscious processing. And thirdly, do neural systems engage differently with face and non-face stimuli before they are processed holistically or gain access to awareness? If they do, then face-specific processing may occur or begin in absence of holistic integration and perceptual awareness. To address these questions, we conducted two EEG experiments. We measured event-related potential (ERP) markers of visual processing, emotion processing, and perceptual awareness, and a large-scale neural integration marker. EEG presents two important advantages when studying perceptual awareness: due to its high temporal resolution, it is sensitive to awareness effects that may be too fast for other neuroimaging techniques to capture, and it is a vastly studied technique, with many well established markers of perceptual and emotional processes.

Probably the most robust and well-studied neural marker of face processing is the N170 component, a large negative potential peaking around the occipitotemporal areas, bilaterally, between 140 and 200 ms after face stimulus onset, normally accompanied by a third potential named the vertex positive potential (VPP), a large positive potential peaking around frontocentral areas (Rossion & Jacques, 2012). The N170 is sensitive to face and face-like stimuli, and belongs to the N1-component family, a group of potentials that respond to different aspects of visual processing. The amplitude of N170 is larger to faces than to any other stimulus category, including objects (Eimer, 2011). Source estimation studies have suggested that the N170 component (or its equivalent in magnetoencephalography or MEG, M170) has its source in the anterior (i.e. fusiform face area in BA37), posterior, and middle fusiform gyri (i.e. including the occipital face area in BA19; Deffke et al., 2007; Gao et al., 2019; Pizzagalli et al., 2002; Schweinberger et al., 2002; Shibata et al., 2002), a structure that has been associated with face processing by fMRI studies as well (Kanwisher et al., 1997, 1998; Yovel & Kanwisher, 2005). A number of studies has suggested that N170's sensitivity to faces reflects face-specific processing (Deffke et al., 2007; Rossion et al., 2000; Tanskanen et al., 2005). For example, studies using Mooney faces (black and white distorted images that resemble faces only when seen upright) have reported larger N170 responses when the Mooney faces were presented in an upright orientation than upside down (George et al., 2005; Jeffreys, 1993). This large N170 response to faces is also present when face images are distorted or their local features are rearranged, as long as the images can be interpreted as faces. However, isolated local features such as a nose or a mouth produce small-peak responses (Bentin et al., 1996; George et al., 1996; Itier & Taylor, 2002). In summary, N170 is a reliable neural marker of face processing and seems to be largely driven by an activation of neural representations of faces.

Importantly, there are other visually evoked potentials that respond to face stimuli, like the P1, a positive-voltage peak arising around 100 ms after stimulus onset but driven by different factors (Pratt, 2011). P1 does not specifically index face processing, but rather low-level visual processing associated with early sensory, perceptual, and attentional processing in the visual cortex (Hillyard & Anllo-Vento, 1998; Luck, 2014; Pratt, 2011). Measuring both P1 and N170 can be useful to describe how face images modulate low-level and high-level visual processing. For example, Tanskanen et al. (2005) employed MEG to measure M1 (equivalent to P1 in EEG) and M170 (equivalent to N170 in EEG) responses to face images that were either masked with narrow-band spatial frequency noise or not. When face images were masked, and therefore imperceptible, M1 was maximal in amplitude despite observers not seeing the faces. In contrast, when noise was at the lowest or highest spatial frequencies, therefore rendering faces perceptible, M1 was minimal and the M170 was maximal. These findings stress two ideas: that N170/M170 is more closely related to face processing and that P1/M1 is not sensitive to face processing per se but rather to low-level visual information. Interestingly, it has been claimed that N170 can also index processing of emotional facial expressions, though the evidence is inconsistent (against: e.g. Batty & Taylor, 2006; Todd et al., 2008; Curtis & Cicchetti, 2011; in favour: e.g. Hietanen & Astikainen, 2013; Qiu et al., 2017; Tian et al., 2018; Blau et al., 2007).

Emotion processing has been studied extensively with EEG. Two neural markers of emotion have been repeatedly documented: the early-posterior negativity (EPN) and the late positive potential (LPP). EPN is a negative-voltage wave found around occipitotemporal areas. It arises between 250 and 450 ms after stimulus onset and exhibits increased negativity in response to emotionally arousing stimuli compared to neutral stimuli (Bradley et al., 2007). This effect has been attributed to a stronger call for attentional resources (Herbert et al., 2006; Junghöfer et al., 2001; Kissler et al., 2009; Schupp et al., 2003). On the other hand, LPP is a positive-voltage slow wave found around parietocentral areas. It arises between 300 and 600 ms after stimulus onset and exhibits increased positivity in response to emotionally salient stimuli (Foti & Hajcak, 2008; MacNamara & Hajcak, 2010; Weinberg & Hajcak, 2010). LPP is associated with an extensive neural network encompassing the amygdala, anterior cingulate cortex, medial prefrontal cortices, parietal cortex, and visual cortex (Liu et al., 2012; Sabatinelli et al., 2007), and it indexes attentional resource allocation for emotion processing (Hajcak et al., 2009; Nordström & Wiens, 2012). It is important to note that no ERP component, including EPN and LPP, has been found to reflect a specific emotion (Hajcak et al., 2011). EPN and LPP mainly respond to emotion intensity and arousal (Eimer & Holmes, 2007; Hajcak et al., 2011). Therefore, emotion processing studies involving faces normally compare ERP responses to emotional expressions and neutral expressions as a control condition. Psychophysiological measures like heart rate, skin conductance, facial muscle activity, and pupil diameter are also sensitive to emotion intensity and arousal (Codispoti et al., 2007; Codispoti & Cesarei, 2007; Olofsson & Polich, 2007), but ERP components present the advantage of having very high temporal resolution and being less vulnerable to habituation, making EPN and LPP excellent neural markers for the study of emotion processing in briefly presented faces.

The study of consciousness and perceptual awareness with EEG has a long tradition (Koch et al., 2016). Several markers have been proposed. A valid marker of perceptual awareness must be sensitive to subjective awareness, which is commonly measured using a subjective scale that participants are asked to complete (e.g. the PAS). The two most widely studied ERP components that fulfil this criterion are the visual awareness negativity (VAN) and the late positivity (LP; Koivisto & Revonsuo, 2010). The VAN is a negative difference wave (aware – unaware) found around occipitotemporal areas, bilaterally, between 200 and 400 ms after stimulus onset (Eklund & Wiens, 2018, 2019; Koivisto & Grassini, 2016; Koivisto & Revonsuo, 2003). Many studies employing masking techniques to suppress visual stimuli from awareness have reported that VAN voltage turns more negative, in a linear fashion, as stimuli are subjectively seen more clearly (Eklund & Wiens, 2018; Koivisto et al., 2006; Koivisto & Grassini, 2016; Koivisto & Revonsuo, 2003; Wilenius & Revonsuo, 2007; Wilenius-Emet et al., 2004). On the other hand, LP is a positive difference wave (aware – unaware) found around frontocentral areas, between 300 and 600 ms after stimulus onset (i.e. in the P3 wave; Koivisto & Revonsuo, 2010). Many studies have reported, using similar masking paradigms, that LP correlates with awareness (Wilenius-Emet et al., 2004), in a non-linear fashion (Del Cul et al., 2007; Sergent et al., 2005). However, unlike VAN, which is believed to index changes in subjective experience (i.e. phenomenal consciousness), LP has been claimed to index conscious access, i.e. the ability to consciously access sensory information (Del Cul et al., 2007; Genetti et al., 2009; Lamy et al., 2008; Salti et al., 2012; Sergent et al., 2005). While it is still a matter of debate what specific aspects of consciousness VAN indexes, there seems to be consensus on LP as an index of conscious access. This interpretation of LP is based on the global neuronal workspace model of consciousness, proposed by Dehaene & Naccache (2001), which distinguishes between subliminal, preconscious, and conscious processing (Dehaene et al., 2006; Railo et al., 2011). In subliminal processing, a stimulus is processed but not consciously accessible. In preconscious processing, conscious access to a stimulus is possible but does not occur due to lack of top-down attention. In conscious processing, on the other hand, a stimulus is seen and reportable. Because perceptual awareness is believed by many to be an all-or-none phenomenon rather than a continuous one (Overgaard et al., 2006; Sergent & Dehaene, 2004), proponents of the global neuronal workspace model of consciousness have argued that LP – due to its non-linear modulation – reflects conscious access (Overgaard et al., 2006; Sergent & Dehaene, 2004; Windey & Cleeremans, 2015).

Not only ERP components may index holistic processing, nonetheless. Functional connectivity markers may index large-scale cortical changes associated with holistic processing. Unlike ERP components, functional connectivity markers measure statistical relations between multiple physiological signals. One EEG functional connectivity marker that has exhibited sensitivity to different conscious states, by presenting distinctive signatures for each of them, is the weighted symbolic mutual information (wSMI; King et al., 2013; Sitt et al., 2014). wSMI assesses the extent to which two EEG signals present joint non-random fluctuations, suggesting that they share common sources and share information. wSMI is sensitive to a broad range of functional relations and, crucially, it is especially sensitive to non-linear activations (Canales-Johnson et al., 2020; Imperatori et al., 2019). Therefore, wSMI is an excellent candidate to search for evidence of information sharing across the brain, given its holistic nature. If face and non-face stimuli involve differences in information sharing, with extremely brief visual presentations, wSMI should capture them.

In this Chapter, we study face processing, emotion processing, and perceptual awareness, by measuring EEG response to visual stimuli presented with the LCD tachistoscope described in Chapter 4. Therefore, we expand on the effects found in Experiment 9 by examining signal detection indices and neural markers associated with said processes.

## 5.2 Experiment 13

Our first question in Experiment 13 is whether neural systems engage differently with different emotional expressions before these gain access to awareness. Neural systems could discriminate facial expressions, based on their emotional content, with a minimal exposure duration that is shorter than the minimal exposure duration required for awareness as shown in signal detection indices. To test for this possibility, we measured EEG neural markers to track changes in face and emotion processing across relevant exposure durations. If neural markers of emotion processing require shorter exposure durations than behavioural emotion identification sensitivity, this will provide evidence for emotion processing in the absence of awareness.

Our second question is whether there is a minimal required exposure for neural systems to reflect perceptual awareness. A neural marker able to reflect awareness with a minimal required exposure that is consistently the same as the minimal exposure required for awareness (indexed by metacognitive sensitivity) should be a neural correlate of consciousness. If there is a neural marker that fulfils these criteria, what place does it take in the sequence of processing steps of face perception found in Chapter 4? If neural markers of perceptual awareness require a minimal exposure duration to discriminate between awareness reports, then such a duration may convey the minimal visual exposure required for perceptual awareness to arise.

### 5.2.1 Methods

#### 5.2.1.1 *Participants*

Thirty-six students of the Université Libre de Bruxelles provided informed consent and were paid €30 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Four participants were excluded from the analysis (see Analysis section). The remaining 32 participants (18 female; all right-handed) had a mean age of 24.3 ($SD_{age} = 4.9$; range: 18 – 31).

A retrospective power analysis, conducted using G*Power 3.1.9.7 (Faul et al., 2009), to test for a difference between conditions in a repeated-measures ANOVA, with a small to medium effect size ($\eta p^2 = 0.04$) and alpha of .05, aiming to achieve a statistical power of 95%, determined that a sample of 19 participants would be required. If a non-sphericity correction ε of .5 were to be added – as reported in the Results section, a number of tests violated this assumption – then a sample of 29 participants would be required. This analysis supports our initial decision of aiming at 32 participants per experiment.

#### 5.2.1.2 *Stimuli*

Stimuli were 60 human faces (20 fearful, 20, happy, and 20 neutral; the same 20 identities – 10 female – were used for the three categories; see Appendix F) taken from the RaFD. They were selected by applying the same criteria used in Experiment 9; their differences in expression identification[4] ($M_{fearful} = 92.2\%$; [$SD_{fearful} = 5.33$]; $M_{happy} = 99.15\%$

[4] A one-way ANOVA was conducted to compare expression identification between emotional expressions (fearful, happy, neutral) based on the norms published by Langner et al. (2010). The effect of emotional valence was significant ($F_{(2, 57)} = 11.13, p < .001, \eta^2p = .285$). Bonferroni-corrected pairwise comparisons revealed that happy expressions were easier to identify than fearful ($t(57) = -4.75, p < .001, d = 0.1$) and neutral expressions ($t(57) = 2.73, p = .025, d = 0.06$). However, we did not find a

[1.76]; $M_{\text{neutral}}$ = 95.15% [5.73]) and intensity[5] ($M_{\text{fearful}}$ = 4.22; [0.21]; $M_{\text{happy}}$ = 4.20 [0.26]; $M_{\text{neutral}}$ = 3.61 [0.25]) were minimised. The same image processing steps were followed, too.

#### *5.2.1.3 Procedure*

The procedure was very similar to that of Experiment 9, with a few differences: we used fearful, happy, and neutral expressions. They were all presented in upright orientation, and we only used three exposure durations: 1.7, 4.4, and 6.2 ms. We selected these exposure durations to measure neural response in the absence (1.7 ms) and presence of holistic face processing (4.4 ms), as well as when both holistic and emotion processing were very likely to have arisen (6.2 ms), according to the findings of Experiment 9.

Participants performed 40 practice trials followed by 1680 experimental trials. Emotional expression was blocked (105 trials/block), with block order counterbalanced across participants in an ABBABAAB BAABABBA order, where each block contained an equal number of emotional (fearful or happy, depending on the block) and neutral expressions. All face images were presented in upright orientation. Participants were given self-terminated breaks between blocks and a compulsory 15-minute break after completing 840 trials.

#### *5.2.1.4 EEG recording and pre-processing*

EEG data were recorded and digitised at a sampling rate of 512 Hz using a 64-channel Biosemi system with an elastic cap, in which electrodes were integrated at sites

significant difference in identification between fearful and neutral expressions $(\mathrm{t}(57) = -2.01, p = .146, d = 0.04)$. See Appendix G.

[5] A one-way ANOVA was conducted to compare emotional intensity between emotional expressions (fearful, happy, neutral) based on the norms published by Langner et al. (2010). The effect of emotional valence was significant $(F_{(2,\,57)} = 41.6, p < .001, \eta^2 p = .593)$. Bonferroni-corrected pairwise comparisons revealed that fearful expressions were rated as more intense than neutral expressions $(\mathrm{t}(57) = 7.99, p < .001, d = 0.04)$, but not than happy expressions $(\mathrm{t}(57) = 0.183, p = 1, d = 0)$. Likewise, happy expressions were rated as more intense than neutral expressions $(\mathrm{t}(57) = 7.8, p < .001, d = 0.04)$. See Appendix H.

conforming to the 10-20 system. All impedance values were kept below 50 kΩ. Scalp electrodes were referenced to Cz. The continuous EEG data was resampled to 256 Hz, then filtered leaving frequencies between 0.3 and 40 Hz, and finally epoched from 200 ms before to 600 ms after stimulus onset. An independent-component analysis (ICA) was run on the epoched EEG signal. Components attributed to eye blinks, ocular movements, heartbeat, and channel noise were taken out. Trials with voltage exceeding 150 mV were excluded from further analysis. On average, 1.8% of trials were removed. The EEG signal was then re-referenced to the average across all electrodes. Waveforms were then averaged for all electrodes. By eye inspection on canonical sites, we determined the following electrodes for each event-related potential (ERP) component of interest: P1 (Oz), left N170 (PO7), right N170 (PO8), left VAN (PO7), right VAN (PO8), VPP (FCz), left EPN (P7), right EPN (P8), LPP (Pz), and LP (Pz). Because emotion processing-related components (EPN and LPP) mainly respond to emotional intensity rather than valence (Hajcak et al., 2010, 2011), and because we wanted to preserve a balanced number of trials between conditions, we collapsed all components across trials, separately for emotional trials (fearful and happy expressions together) and neutral trials (neutral expressions). ERP components were collapsed across trials from the same stimulus condition regardless of behavioural performance in the task, except for our pre-registered analyses of the VAN and LP[6] (see Appendix I). These two components were collapsed across two groups of trials according to the trials' PAS ratings: awareness-present trials (PAS ratings of "vague impression", "almost clear experience", or "clear experience") and awareness-absent trials (PAS rating of "no experience"). To investigate changes in ERP components, mean amplitudes were computed within the following time windows: 105-135 ms (P1), 170-200 ms (N170 and VPP), 295-325 ms (EPN), 310-430 ms (LPP), 185-215 ms (VAN), and 320-520 ms (LP). These time windows were determined by eye inspection on the grand average plots, in specific time windows informed by previous studies, and did not involve exploratory statistical testing, as a way to reduce familywise error (Luck, 2014; Luck & Gaspelin, 2017).

[6] For Pre-registration, also see here: https://aspredicted.org/53sv5.pdf

#### *5.2.1.5 Signal detection analysis*

Signal detection analysis was performed in the same way as in Experiment 9.

#### *5.2.1.6 ERP analysis*

ERP analysis was performed on mean amplitude values of each ERP component at the specific electrodes and time windows stated above. We reconstructed each ERP's cortical source using Brainstorm (Tadel et al., 2011; latest version, released October 2019). To estimate the cortical source of an ERP component, we need to model the electromagnetic properties of the head and of the sensor array (forward model), and then estimate the brain sources that produced the EEG signal of interest (inverse model). The forward model was calculated using the OpenMEEG Boundary Element Method (Gramfort et al., 2010) on the cortical surface of a template MNI brain (ICBM152) with 1mm resolution. The inverse model was constrained using weighted minimum-norm estimation (wMNE; Baillet et al., 2001) to measure source activation in picoampere-meters. wMNE looks for a distribution of sources with the minimum current that can account for the EEG data. We corrected grand-averaged activation values by subtracting the mean of the baseline period (-200 to 0 ms before stimulus onset) and spatially smoothed with a 5-mm kernel. This procedure was applied separately for each ERP component, collapsing all conditions.

We excluded data from four participants who presented more than 15% of noisy electrodes (whose impedance values exceeded 50 k$\Omega$ throughout the experiment). They were excluded from both signal detection and EEG analyses. One additional participant did not provide any "almost clear experience" and "clear experience" PAS ratings and therefore their data were excluded from the VAN and LP analyses.

### 5.2.2 Behavioural results

#### 5.2.2.1 *Location sensitivity*

Neutral expressions appeared in both fearful and happy expression blocks. We entered location d' scores for neutral-expression trials into a preliminary repeated-measures 2 (Block-type: neutral-expression in fearful vs happy expression blocks) × 3 (exposure durations) repeated-measures ANOVA; there was no significant effect of block type $(F_{(1, 31)} = 0.358, p = .554, \eta p^2 = .011)$, nor an interaction of block type with exposure duration $(F_{(1.945, 60.291)} = 1.468, p = .239, \eta p^2 = .045)$. We also estimated Bayes factors: the results suggest the data are substantially better explained under the null hypothesis model $(BF_{01} = 6.440)$. We therefore collapsed neutral-expression trials into one condition, making three conditions in total.

In the main analysis, we measured location d' to confirm that these new participants displayed similar performance to participants in Experiment 9. To examine how conditions affected location discrimination, we entered location d' scores into a 3 (expression: fearful, happy, neutral) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration, whereby sensitivity increased with increasing durations $(F_{(1.68, 52.05)} = 342.395, p < .001, \eta p^2 = .917)$. Importantly, we found very similar location d' scores to the ones reported in Experiment 9 (Figure 5.1a). As expected, based on the findings of Experiment 9, we did not find a main effect of expression $(F_{(1.77, 54.86)} = 0.154, p = .832, \eta p^2 = .005)$, thus suggesting again that there is no greater perceptual sensitivity to emotional expressions than to neutral ones. The interaction between expression and exposure duration did not reach significance either $(F_{(3.54, 109.80)} = 1.833, p = .135, \eta p^2 = .056)$. Thus, location sensitivity performance replicated the findings obtained in equivalent conditions in Experiment 9.

To determine the shortest exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that above-chance identification was already present at the shortest exposure duration, 1.7 ms, for fearful expressions $(M = 0.120\ [0.285];\ t(31) = 2.39, p = .023, d = 0.420)$, neutral expressions $(M = 0.172\ [0.208];\ t(31) = 4.69, p < .001, d = 0.829)$, and happy expressions $(M = 0.070\ [0.204];\ t(31) = 1.95, p = .060, d = 0.345)$. Thus, above-chance performance was found at slightly shorter exposure durations than in Experiment 9.

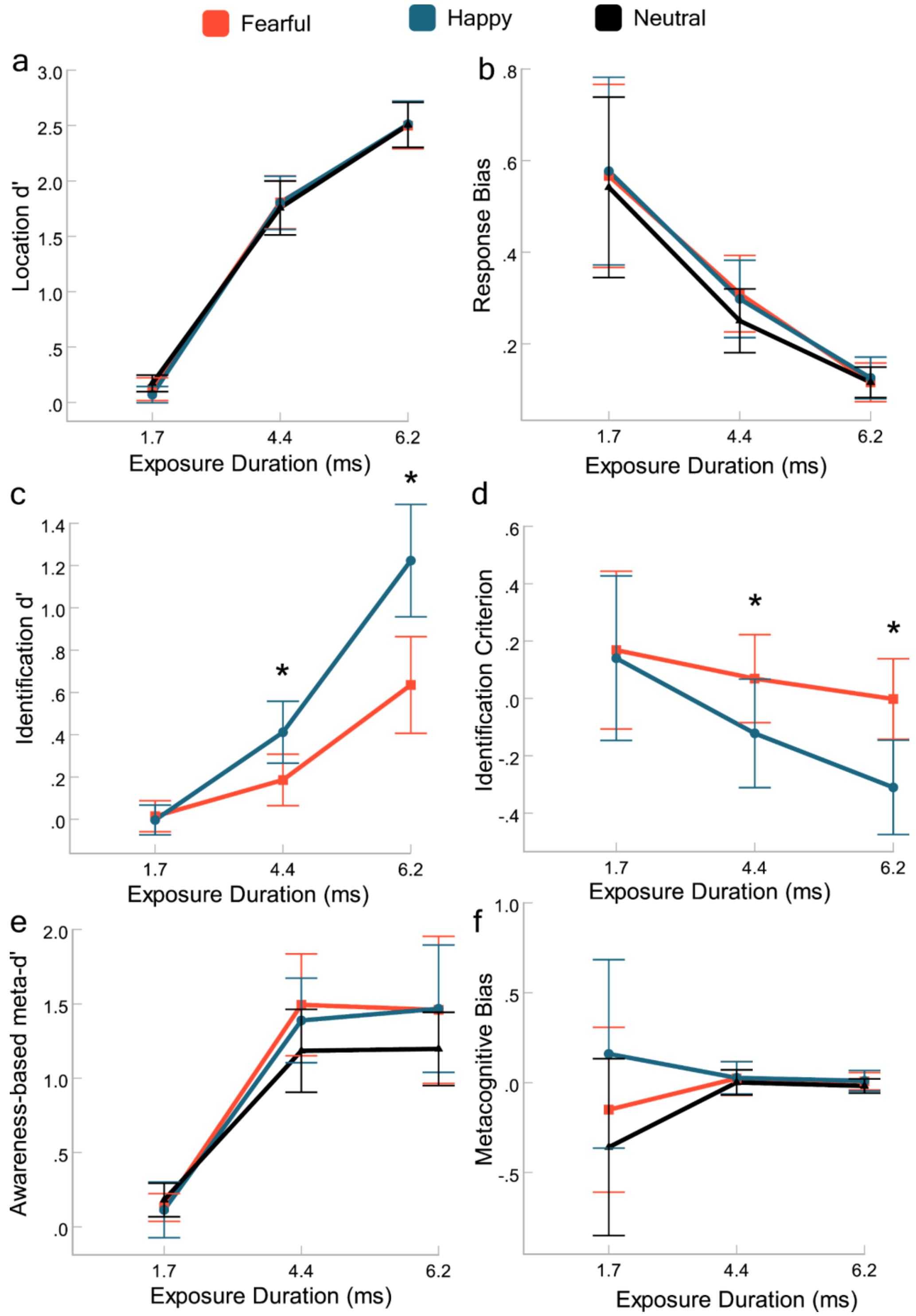



**Figure 5.1.** Behavioural results of Experiment 13. (a) Location sensitivity. Location d' increased with increasing exposure duration. There was no effect of emotional expression. (b) Absolute-value response bias scores for reporting location (bias toward either left or right). The amount of bias decreased as exposure duration increased, but there was no difference in amount of response bias between expressions. (c) Identification sensitivity

for expression. Identification d' increases with increasing exposure duration. A significant advantage in expression identification for happy expressions over fearful expressions arises by 4.4 ms of exposure. (d) Criterion scores for reporting expression. Criterion becomes more liberal with increasing exposure duration. A more liberal criterion for happy expressions over fearful expressions arises by 4.4 ms of exposure. (e) Awareness-based metacognitive sensitivity. Meta-d' increased with exposure duration but was unaffected by expression. (f) Metacognitive bias scores for reporting subjective awareness. Metacognitive bias was unaffected by exposure duration and expression. Asterisks index statistically significant differences between fearful and happy expressions. Error bars represent 95 CI.

#### 5.2.2.2 *Location response bias*

We examined whether participants' response bias for reporting face location varied across conditions by entering the absolute values of $C_{location}$ scores into a 3 (emotional expression: fearful, happy, neutral) × 3 (exposure durations) repeated-measures ANOVA. As in Experiment 9, response bias significantly decreased with increasing exposure duration $(F_{(1.09, 33.78)} = 20.109, p < .001, \eta p^2 = .393)$, indicating that as participants' ability to discriminate faces increased, they became less likely to exhibit a systematic bias in their preference to report one side or the other (Figure 5.1b). However, we did not find a main effect of expression $(F_{(1.93, 59.89)} = 2.093, p = .134, \eta p^2 = .063)$, suggesting that response bias was unaffected by the emotional content of the faces. The interaction between expression and exposure duration did not reach significance either $(F_{(3.38, 104.69)} = 0.571, p = .656, \eta p^2 = .018)$.

To assess whether the obtained data support the absence of an effect of emotional expression, we estimated the Bayes factor for this effect, which indicated strong evidence for the null hypothesis model $(BF_{01} = 16.978)$.

### 5.2.2.3 *Expression identification sensitivity*

We examined whether participants' sensitivity to identifying emotional expressions from their neutral counterparts varied across conditions by entering identification d' scores into a 2 (emotional expression: fearful, happy) × 3 (exposure durations) repeated-measures ANOVA (Figure 5.1c). As expected, we found a main effect of exposure duration $(F_{(1.38,\ 42.75)} = 62.9, p < .001, \eta p^2 = .670)$, indicating that sensitivity to emotional expression increased with increasing exposure duration. We also found a main effect of expression $(F_{(1,\ 31)} = 36.3, p < .001, \eta p^2 = .539)$, showing better sensitivity to happy expressions $(M = 0.544\ [0.625])$ than to fearful expressions $(M = 0.278\ [0.321])$. The interaction between expression and exposure duration also reached significance $(F_{(1.65,\ 51.01)} = 25.6, p < .001, \eta p^2 = .452)$. To determine at which of the three durations we used this advantage was present, we ran post hoc Bonferroni-corrected pairwise comparisons. They revealed significant advantages at 4.4 ms $(t(88.1) = -3.417, p = .014, d = -0.604)$ and 6.2 ms $(t(88.1) = -8.90, p < .001, d = -1.573)$ of exposure.

To determine the shortest exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance identification was 4.4 ms for fearful $(M = 0.186\ [0.338];\ t(31) = 3.11, p = .004, d = 0.550)$ and happy facial expressions $(M = 0.412\ [0.406];\ t(31) = 5.73, p < .001, d = 1.014)$. Neither expression exhibited above-chance performance with 1.7 ms of exposure, therefore we estimated Bayes factors to test whether the obtained data support this absence of an effect. Bayes factors indicated substantial evidence in favour of the null hypothesis model for fearful $(BF_{01} = 4.93)$ and happy expressions $(BF_{01} = 5.27)$ presented for 1.7 ms of exposure, thus supporting the finding that neither facial expression enjoyed above-chance performance when presented for 1.7 ms.

These findings, along with the metacognitive sensitivity findings reported below, will be relevant for our EEG analysis – if we were to find neural evidence of emotion processing with 1.7 ms of exposure by measuring neural markers, we could argue that such findings are evidence of unconscious emotion processing.

#### 5.2.2.4 *Expression identification criterion*

We examined whether participants' criterion for reporting fearful and happy expressions varied across conditions by entering $C_{identification}$ scores into a 2 (expression: fearful, happy) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.11, 34.32)} = 4.67, p = .034, \eta p^2 = .131$), indicating that identification criterion became more liberal across increasing exposure duration, i.e. participants became more willing to report an emotional expression (fearful or happy) than a neutral one as exposure duration increased (Figure 5.1d). We also found a main effect of emotional expression ($F_{(1, 31)} = 15.42, p < .001, \eta p^2 = .332$), indicating that happy expressions ($M = -0.097$ [0.226]) enjoyed a more liberal identification criterion than fearful expressions ($M = 0.079$ [0.086]). Finally, the interaction between both factors reached significance as well ($F_{(1.62, 50.26)} = 17.58, p < .001, \eta p^2 = .362$). To determine at which exposure durations happy expressions enjoyed a significantly more liberal criterion than fearful expressions, we ran post hoc Bonferroni-corrected pairwise comparisons. They revealed significant advantages at 4.4 ms ($t(54.7) = 3.637, p = .009, d = 0.643$) and 6.2 ms of exposure ($t(54.7) = 5.873, p < .001, d = 1.038$).

#### 5.2.2.5 *Awareness-based metacognitive sensitivity*

We examined whether awareness scores were sensitive to participants' location sensitivity scores by estimating meta-d', a measure of metacognitive sensitivity. To examine whether metacognitive sensitivity varied across conditions, we entered meta-d' scores into a 3 (expression: fearful, happy, neutral) × 3 (exposure durations) repeated-measures ANOVA (Figure 5.1e). A main effect of exposure duration indicated that meta-d' increased with increasing exposure duration ($F_{(1.69, 52.54)} = 55.970, p < .001, \eta p^2 = .644$). However, we did not find a main effect of expression ($F_{(1.87, 58.11)} = 1.297, p = .280, \eta p^2 = .040$), suggesting that emotional expression did not affect metacognitive sensitivity. To examine the duration interval in which awareness arose, we ran post hoc Bonferroni-corrected pairwise comparisons between each combination of exposure durations (collapsed across expressions). We found that meta-d' scores obtained with 4.4

ms were significantly higher than scores obtained with 1.7 ms $(t(62) = -9.09, p < .001, d = -1.607)$, whereas meta-d' scores obtained with 6.2 ms were not significantly higher than scores obtained with 4.4 ms $(t(62) = -0.143, p = .999, d = -0.025)$. This may suggest that 4.4 ms of exposure provided sufficient visual information for metacognitive sensitivity to exhibit a non-linear increase. Finally, we did not find an interaction between expression and exposure duration $(F_{(2.69,\,83.32)} = 0.934, p = .420, \eta p^2 = .029)$.

As described, we did not find a main effect of expression, therefore we calculated Bayes factors to test whether the obtained data support this absence of an effect. Bayes factors indicated strong evidence in favour of the null hypothesis model $(BF_{01} = 12.06)$, thus supporting the finding that no expression was prioritised by metacognitive sensitivity.

To determine the shortest exposure that exhibited above-chance metacognitive sensitivity (meta-d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found above-chance meta-d' was already present at the shortest exposure duration, 1.7 ms, for fearful $(M = 0.130\ [0.259];\ t(31) = 2.84, p = .008, d = 0.503)$ and neutral expressions $(M = 0.181\ [0.313];\ t(31) = 3.27, p = .003, d = 0.578)$; meta-d' was above chance at 4.4 ms for happy expressions $(M = 1.389\ [0.790];\ t(31) = 9.95, p < .001, d = 1.759)$. Although meta-d' scores were higher than zero at the shortest exposure duration used, they were not higher than location d' scores at the same exposure duration.

#### 5.2.2.6 *Metacognitive bias*

Metacognitive bias (meta-bias) is the tendency to give high confidence ratings regardless of actual performance. In this experiment, however, we used the PAS, a more exhaustive measure of awareness than confidence ratings (Sandberg et al., 2010). The PAS allows participants to rate their visual experience using four ratings covering from "no experience" to "clear experience" reports. Participants described their visual experience of a face and, therefore, metacognitive bias describes the tendency to describe one's visual experience as clear. To examine whether metacognitive bias varied across conditions, we entered meta-bias scores into a 3 (expression: fearful, happy, neutral) × 3 (exposure durations) repeated-measures ANOVA (Figure 5.1f). We did not find an effect of exposure duration ($F_{(1.05,\,32.51)} = 0.555, p = .470, \eta p^2 = .018$), or of facial expression ($F_{(1.94,\,60.19)} = 1.585, p = .214, \eta p^2 = .049$). The interaction between expression and exposure duration did not reach significance either ($F_{(1.97,\,61.17)} = 1.226, p = .30, \eta p^2 = .038$). These results suggest that the participants' tendency to report high confidence to characterise their experience was not affected by any of the conditions.

### 5.2.3 ERP results

#### *5.2.3.1 Early visual processing (P1)*

We examined how emotional expression and exposure duration conditions affected early visual processing by measuring voltage changes in P1. This component is measured over occipital regions. We entered mean voltage values into a 2 (expression: emotional, neutral) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.14,\,35.27)} = 154.18, p < .001, \eta p^2 = .833$), whereby P1 voltage increased with exposure duration, thus indicating that longer exposure durations involved greater early visual processing than shorter exposure durations (Figure 5.2). As expected, given the nature of this component, we did not find a main effect of expression ($F_{(1,\,31)} = 1.21, p = .280, \eta p^2 = .038$) and the interaction between expression and exposure duration did not reach significance either ($F_{(1.96,\,60.68)} = 2.03, p = .142, \eta p^2 = .061$). To assess whether the obtained data support the absence of an effect of emotional expression, we estimated a Bayes factor, which indicated substantial evidence for the null hypothesis model ($BF_{01} = 6.385$). These results indicate that P1 is sensitive to extremely brief visual exposure durations but is not sensitive to emotional expressions.

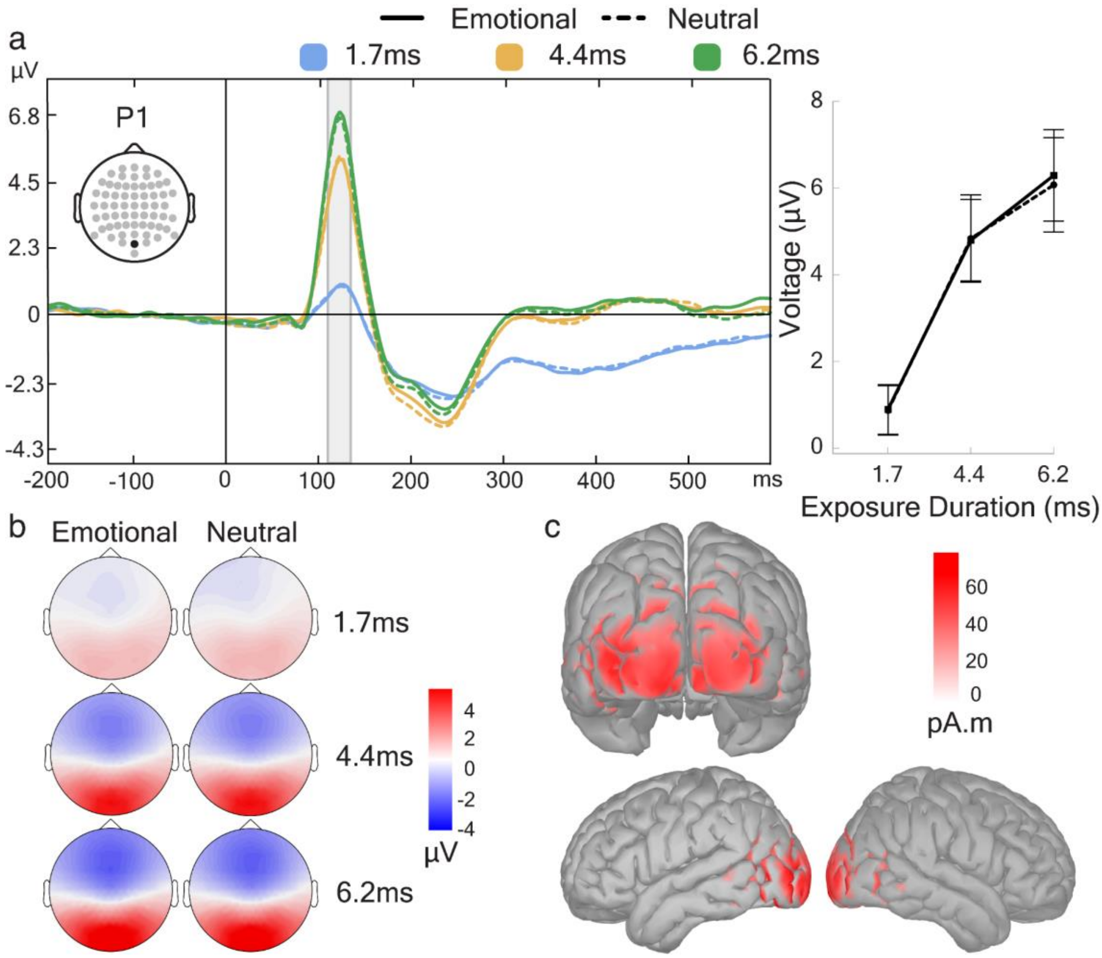



**Figure 5.2.** Early visual processing indexed by P1. (a) P1 response to emotional and neutral expressions across exposure durations. Left: P1 peak with averaged time window highlighted in grey (105 – 135 ms); right: voltage means. (b) Topographical distributions of P1 for each condition at the relevant time window show an increase in positive voltage over visual cortex across exposure durations. (c) Source estimation of P1 visually identified on cortical maps. Estimated current sources of P1 are localised around the primary visual cortex. Error bars represent 95% CI.

#### 5.2.3.2 *Face processing (N170/VPP)*

We examined how facial expression and exposure duration conditions affected face processing by measuring voltage changes in the N170/VPP complex. N170/VPP consists of two negative peaks (left and right N170) found in occipitotemporal areas, bilaterally, and one positive peak (vertex positive potential or VPP) found in frontocentral areas. We entered mean voltage values into a 2 (expression: emotional, neutral) × 3 (electrode site: left, right, central) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration $(F_{(1.10, 34.20)} = 43.195, p < .001, \eta p^2 = .582)$, whereby N170/VPP amplitude changed with increasing exposure duration, thus suggesting longer exposure durations involved greater face processing than shorter exposure durations (Figure 5.3). We also found a main effect of electrode site $(F_{(1.49, 46.24)} = 94.125, p < .001, \eta p^2 = .752)$, which was expected given that left $(M = -5.404\ [1.674])$ and right $(M = -5.554\ [1.523])$ N170 potentials are negative-voltage peaks whereas VPP is a positive-voltage peak $(M = 4.575\ [1.085])$. This effect confirmed that we measured the N170/VPP complex. We also found a main effect of expression $(F_{(1, 31)} = 1.696, p = .002, \eta p^2 = .274)$, indicating that neutral expressions $(M = -2.198\ [5.303])$ had significantly more negative voltage values than emotional expressions $(M = -2.058\ [5.134])$. However, the interaction between expression and exposure duration did not reach significance $(F_{(1.57, 48.65)} = 0.318, p = .676, \eta p^2 = .01)$. In addition, we found a significant interaction between expression and electrode site $(F_{(1.37, 42.32)} = 4.716, p = .025, \eta p^2 = .132)$, which was expected given electrode sites' opposite voltage polarities, and a significant interaction between electrode site and exposure duration $(F_{(1.74, 53.95)} = 56.074, p < .001, \eta p^2 = .644)$, also expected given that N170 sites and VPP site became more negative and positive, respectively, as exposure duration increased. The three-way interaction between expression, electrode site, and exposure duration did not reach significance. Together, these results indicate that both exposure duration and expression affect N170/VPP – greater duration increases its amplitude (more negative for N170 and more positive for VPP), whereas the presence of emotion decreases it (less negative for N170 and less positive for VPP).

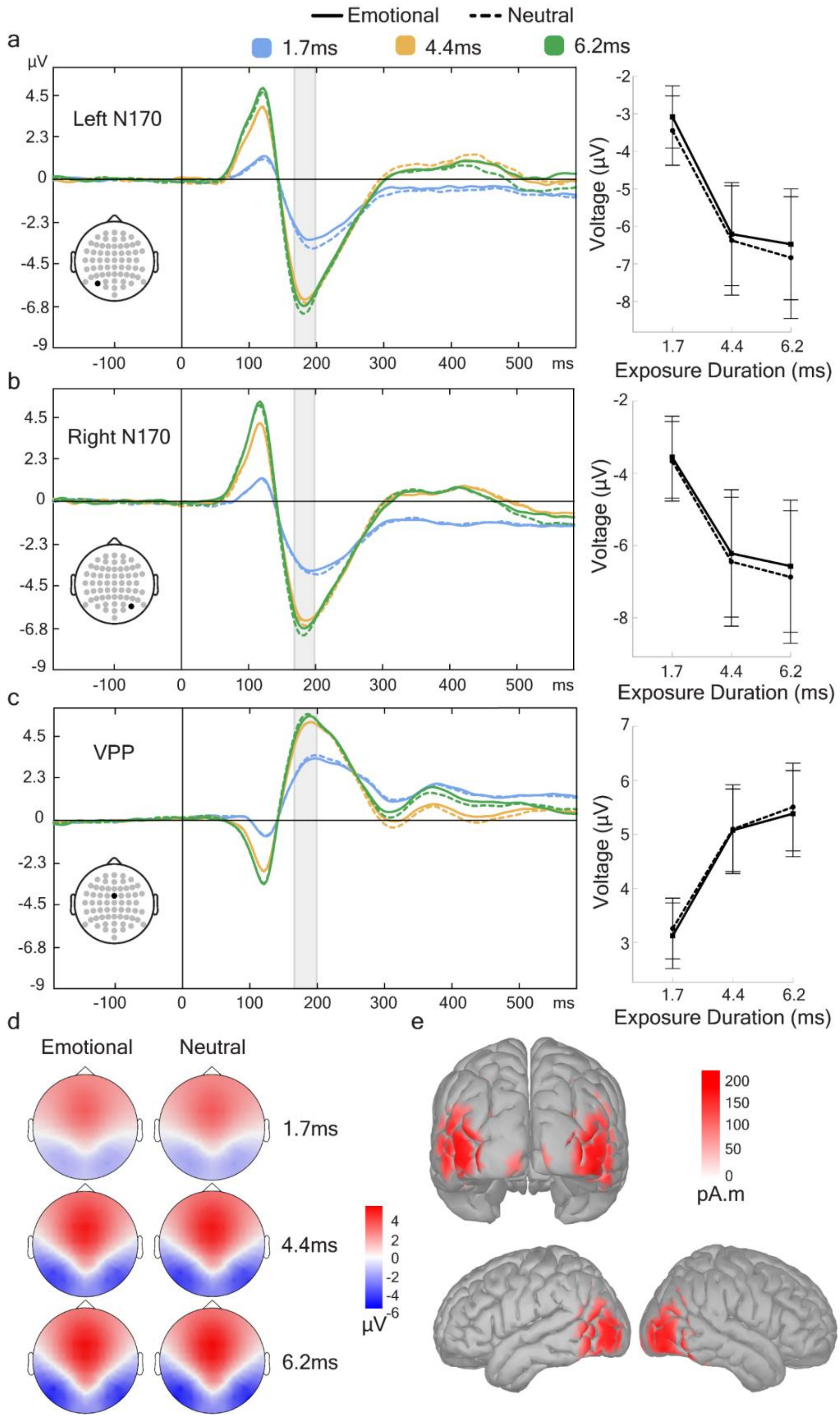



**Figure 5.3.** Face processing indexed by N170/VPP. (a) N170/VPP response to emotional and neutral facial expressions across increasing exposure durations. Left:

N170/VPP peaks with averaged time window highlighted in grey (170-200 ms); right: voltage means. (b) Topographical distributions of N170/VPP for each condition at the relevant time window show increase in negative voltage around posterior areas across exposure durations. (c) Source estimation of N170 visually identified on cortical maps. Estimated current sources of N170 are localised around the inferotemporal gyri and lateral occipital sulcus. Error bars represent 95% CI.

#### *5.2.3.3 Early emotion processing (EPN)*

We examined how emotional expression and exposure duration conditions affected early emotion processing by measuring voltage changes in the early posterior negativity (EPN) component. This component is measured over occipitotemporal regions, bilaterally. We entered mean voltage values into a 2 (expression: emotional[7], neutral) × 2 (electrode site: left, right) × 3 (exposure durations) repeated-measures ANOVA. We did not find an effect of exposure duration ($F_{(1.26, 39.18)} = 2.884, p = .089, \eta p^2 = .085$), indicating that EPN did not vary across exposure durations overall (Figure 5.4). Importantly, however, we found a main effect of expression ($F_{(1, 31)} = 6.042, p = .02, \eta p^2 = .163$), thus confirming EPN was sensitive to emotional content in faces, showing more negative values for emotional ($M = -2.166$ [0.374]) than neutral faces ($M = -1.981$ [0.444]). We did not find an effect of electrode site ($F_{(1, 31)} = 2.157, p = .152, \eta p^2 = .065$), indicating EPN responded to changes across conditions similarly in both hemispheres. Crucially, we found a significant interaction between emotional expression and exposure duration ($F_{(1.90, 58.86)} = 8.675, p < .001, \eta p^2 = .219$), suggesting EPN was sensitive to emotional content at specific

[7]As an exploratory analysis, we tested whether EPN could discriminate between fearful and happy expressions. To test this, we entered mean voltage values into a 2 (expression: fearful, happy) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of expression ($F_{(1, 31)} = 8.809, p = .006, \eta p^2 = .221$), indicating that happy expressions had significantly more negative voltage values than fearful expressions. We did not find a main effect of exposure duration ($F_{(1.38, 42.63)} = 2.839, p = .087, \eta p^2 = .084$). The interaction between expression and exposure duration did not reach significance either ($F_{(1.91, 59.07)} = 1.91, p = .160, \eta p^2 = .058$). These results might suggest that happy expressions are more emotionally intense given that they are less ambiguous and thus easier to recognise than negative expressions (Hugdahl et al., 1993; Leppänen & Hietanen, 2004; Smith & Rossit, 2018; Wells et al., 2016).

exposure durations. To test at what specific exposure durations EPN was sensitive to emotional content, we ran post hoc Bonferroni-corrected pairwise comparisons and found that EPN was significantly more negative for emotional expressions ($M_{6.2\text{ms}} = -2.793\ [2.580]$) than neutral expressions ($M_{6.2\text{ms}} = -2.070\ [3.423]$) only at the longest exposure duration ($t(87.6) = -4.009, p = .002, d = -0.709$), when both awareness and emotion identification are very likely; therefore, these results suggest that emotion processing requires a longer exposure duration than holistic face processing to occur. Less importantly, we found a significant interaction between electrode site and exposure duration ($F_{(1.49,\ 46.14)} = 5.44, p = .013, \eta p^2 = .149$), though no comparison of interest (i.e. between electrode sites for each exposure duration) reached significance. The three-way interaction between expression, electrode site, and exposure duration did not reach significance either ($F_{(2,\ 61.97)} = 0.038, p = .962, \eta p^2 = .001$). These results suggest that emotion processing arises with exposure durations that are greater than 4.4 ms but smaller than or equal to 6.2 ms.

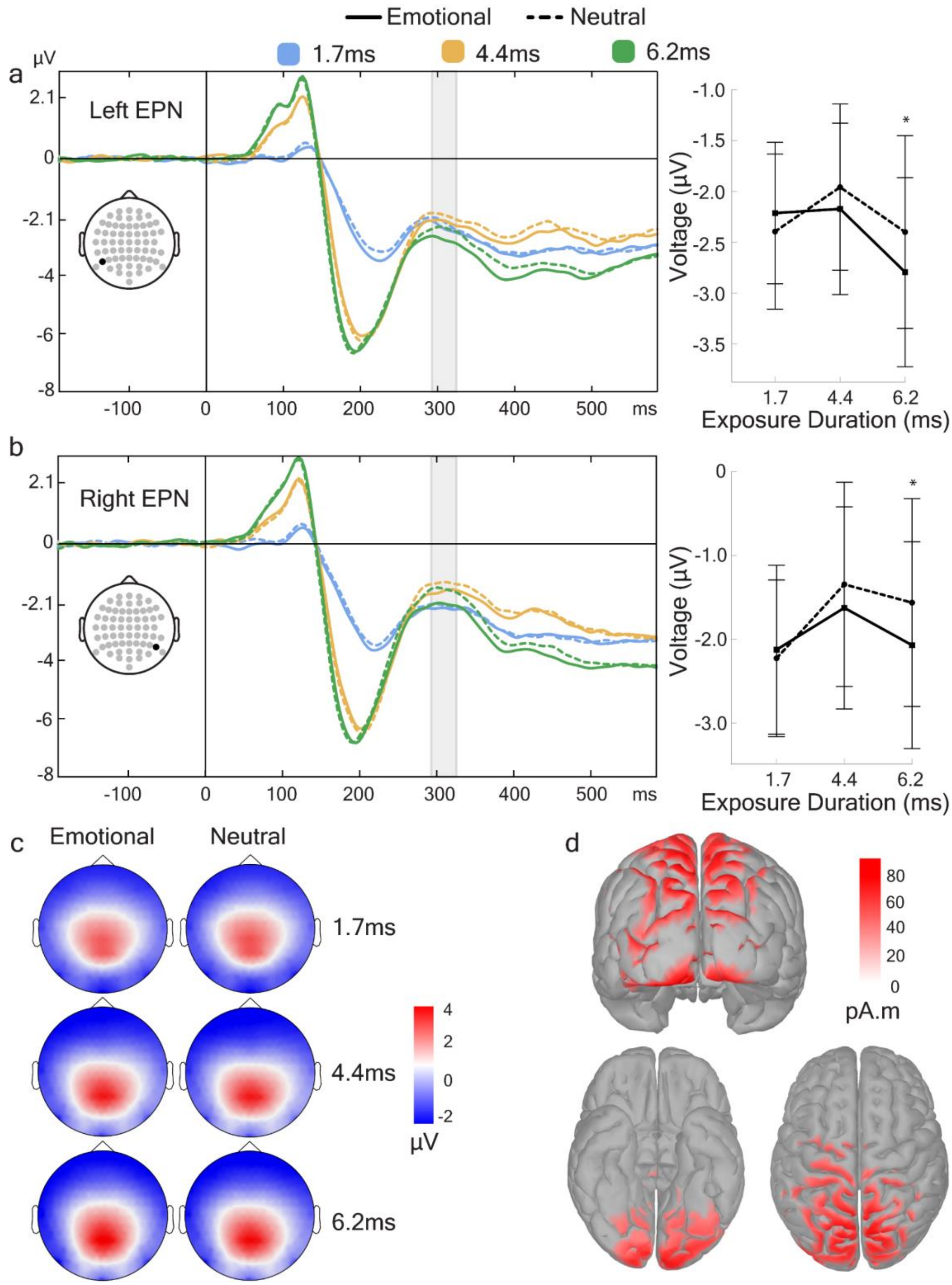



**Figure 5.4.** Early emotion processing indexed by EPN. (a) EPN response to emotional and neutral expressions across increasing exposure durations. Left: EPN peak with averaged time window highlighted in grey (295-325 ms); right: voltage means. Emotional expressions had significantly more negative voltage than neutral expressions at 6.2 ms of exposure. (b) Topographic distributions of EPN for each condition at the relevant time

window show increase in negative voltage for emotional expressions than neutral expressions across exposure durations. (c) Source estimation of EPN visually identified on cortical maps. Distributions of estimated current sources are localised around postcentral and parietocentral areas. Asterisks indicate statistically significant differences between conditions. Error bars represent 95% CI.

#### 5.2.3.4 *Late emotion processing (LPP)*

The late positive potential (LPP) is another ERP component that is sensitive to emotional content and intensity. This component is measured over postcentral areas. We entered mean voltage values into a 2 (expression: emotional[8], neutral) × 3 (exposure durations) repeated-measures ANOVA. We found an effect of exposure duration ($F_{(1.45,\,44.97)} = 60.735, p < .001, \eta p^2 = .662$), indicating that LPP became more positive in voltage as exposure duration increased (Figure 5.5). We did not find a main effect of expression ($F_{(1,\,31)} = 2.12, p = .156, \eta p^2 = .064$) – overall, LPP did not vary between emotional and neutral facial expressions. Crucially, however, we found a significant interaction between expression and exposure duration ($F_{(1.98,\,61.40)} = 9.80, p < .001, \eta p^2 = .240$). To test whether LPP was sensitive to emotional content at specific exposure durations, we ran post hoc Bonferroni-corrected pairwise comparisons and found that LPP was significantly more positive for emotional expressions ($M_{6.2\mathrm{ms}} = 6.328\ [2.878]$) than neutral expressions ($M_{6.2\mathrm{ms}} = 5.981\ [2.904]$) only at the longest exposure duration ($t(90.3) = 4.284,\ p < .001,\ d = 0.757$). Thus, these findings of LPP, a marker of late emotion processing, converge with those for the EPN, a marker of early emotion processing, arising by the longest exposure duration used.

---

[8] As an exploratory analysis, we tested whether LPP could discriminate between fearful and happy expressions. To test this, we entered mean voltage values into a 2 (expression: fearful, happy) × 3 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.49,\,46.10)} = 67.051, p < .001, \eta p^2 = .684$), suggesting that voltage turned more positive with increasing exposure duration. We did not find a main effect of expression ($F_{(1,\,31)} = 0.053, p = .819, \eta p^2 = .002$) and the interaction between expression and exposure duration did not reach significance either ($F_{(1.87,\,58.07)} = 0.074, p = .919, \eta p^2 = .002$). These results might suggest that LPP is sensitive to emotional information relative to neutral information, regardless of emotional valence.

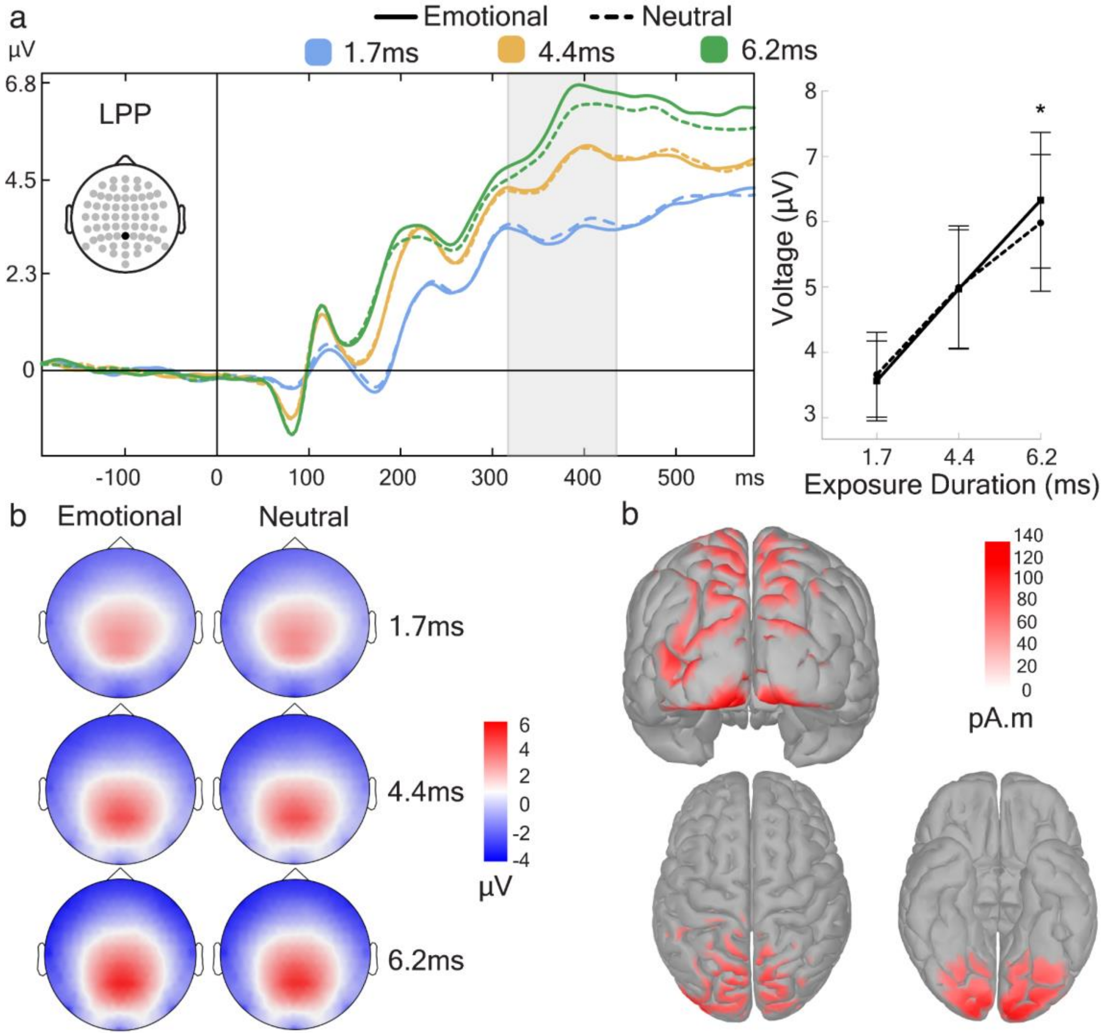



**Figure 5.5.** Late emotion processing indexed by LPP. (a) LPP response to emotional and neutral facial expressions across exposure durations. Left: LPP peak with averaged time window highlighted in grey (310-430 ms); right: voltage means. Emotional expressions had significantly more positive voltage than neutral expressions at 6.2 ms of exposure. (b) Topographic distributions of the LPP component for each condition at the relevant time window show increase in positive voltage around parietocentral areas across exposure durations. (c) Source estimation of the LPP component visually identified on cortical maps. Estimated current sources are localised around parietocentral and inferotemporal areas. Asterisks indicate statistically significant differences between conditions. Error bars represent 95% CI.

#### 5.2.3.5 *Visual awareness (VAN)*

The visual awareness negativity (VAN) is a voltage difference measured over occipitotemporal areas, bilaterally. The standard experimental manipulation to extract VAN consists on comparing trials where participants reported seeing a stimulus to those where they reported not seeing it. We examined how awareness, facial expression, and exposure duration conditions affected VAN. By using the distinction between trials with and without awareness based on PAS ratings (see EEG analysis section, above), we could find the exposure-duration range in which visual awareness arises. Since awareness (by including every rating except 'no experience') was present over 90% of trials at the longest exposure duration (Figure 5.6), we only included trials with the exposure durations of 1.7 ms and 4.4 ms in this analysis.

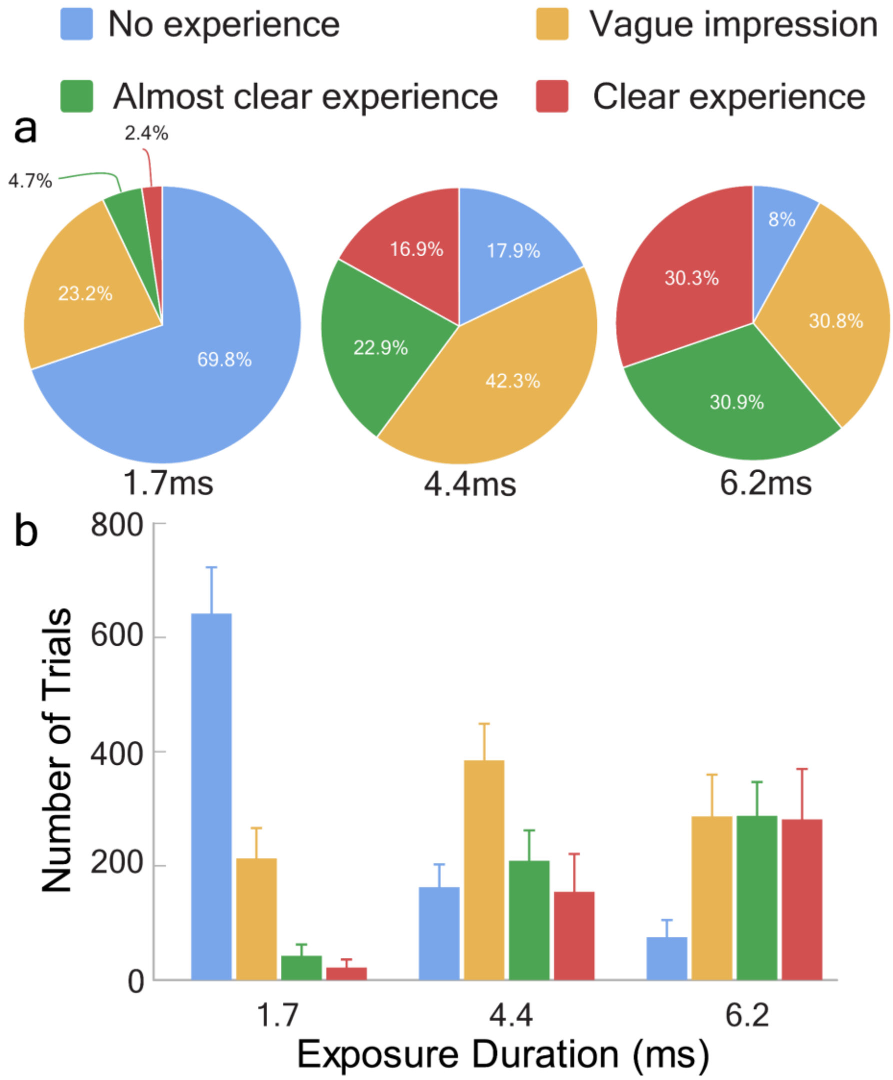



**Figure 5.6.** Proportion of Perceptual Awareness Scale (PAS) ratings. (a) Percentages represent overall numbers per exposure duration. At 1.7 ms, 69.8% of the trials were reported as awareness-absent trials ("no experience") and 30.2% were reported as awareness-present trials ("vague impression", "almost clear experience" or "clear experience"). At 4.4 ms, 17.9% of the trials were reported as awareness-absent and 82.1% were reported as awareness-present. At 6.2 ms, 8% of the trials were reported as awareness-absent and 92% were reported as awareness-present. (b) Mean number of trials per PAS rating. Error bars denote 95% CI, representing variability across subjects.

We entered mean voltage values into a 2 (awareness: aware, unaware) × 2 (expression: emotional, neutral) × 2 (electrode: left, right) × 2 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration

$(F_{(1, 30)} = 32.042, p < .001, \eta p^2 = .516)$, indicating that VAN turned more negative with increasing exposure duration (Figure 5.7). Importantly, we found a main effect of awareness $(F_{(1, 30)} = 17.755, p < .001, \eta p^2 = .372)$, thus suggesting VAN could distinguish between awareness-present $(M = -4.299\ [1.099])$ and awareness-absent trials $(M = -3.535\ [0.825])$. Unexpectedly, we also found a main effect of expression $(F_{(1, 30)} = 8.679, p = .006, \eta p^2 = .224)$, indicating that VAN was significantly less negative in voltage for emotional $(M = -3.747\ [1.073])$ than neutral expressions $(M = -4.087\ [1.002])$. We did not find an effect of electrode site $(F_{(1, 30)} = 1.719, p = .20, \eta p^2 = .054)$, indicating that VAN did not significantly vary between hemispheres. Crucially, the interaction between awareness and exposure duration reached significance $(F_{(1, 30)} = 10.062, p = .003, \eta p^2 = .251)$. To test whether awareness significantly modulated VAN at specific exposure durations, and thus answer our main question here – i.e. at which exposure duration we find a neural indication of awareness – we ran post hoc Bonferroni-corrected pairwise comparisons and found that VAN was significantly more negative in awareness-present than awareness-absent trials only at 4.4 ms of exposure $(t(45.4) = -5.21, p < .001, d = -0.935)$. This finding suggests that 4.4 ms of exposure may be sufficient for faces to reach awareness. We did not find significant interactions between expression and awareness $(F_{(1, 30)} = 2.424, p = .130, \eta p^2 = .075)$, expression and electrode site $(F_{(1, 30)} = 0.99, p = .755, \eta p^2 = .003)$, expression and exposure duration $(F_{(1, 30)} = 0.06, p = .938, \eta p^2 = 0)$, electrode site and exposure duration $(F_{(1, 30)} = 1.232, p = .276, \eta p^2 = .039)$, or any of the three-way interactions and four-way interaction (all p > .113).

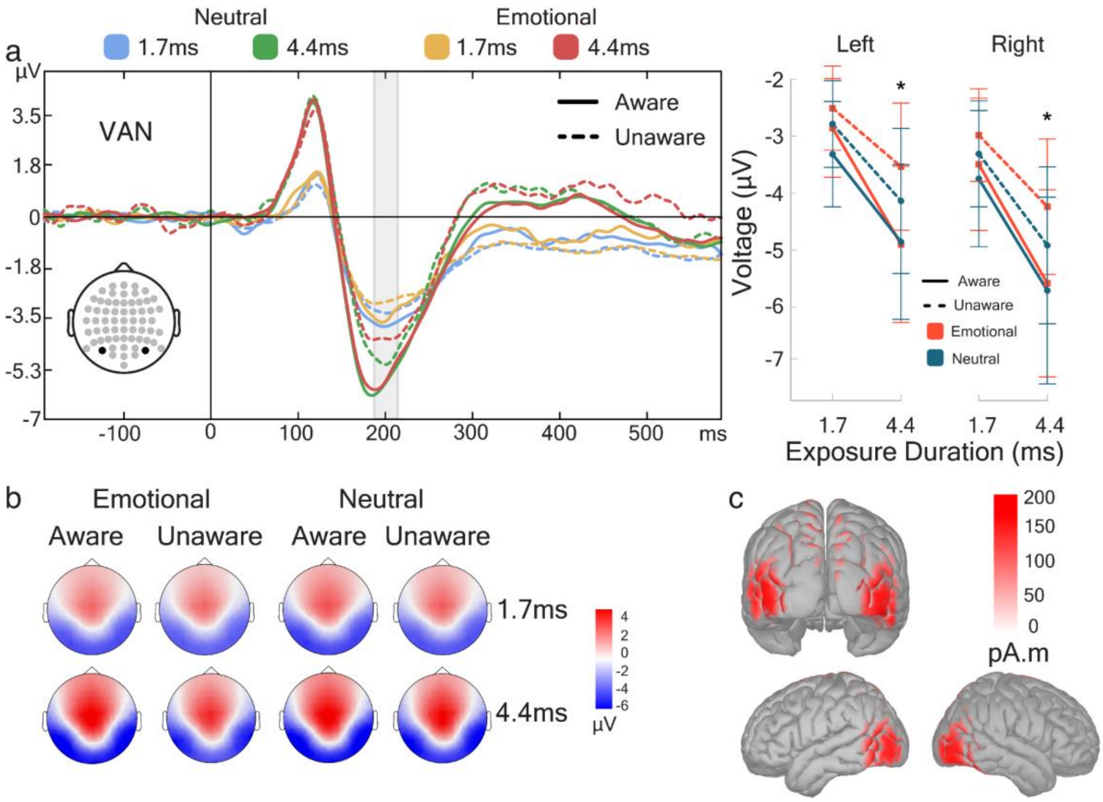



**Figure 5.7.** Visual awareness indexed by VAN. (a) VAN response to emotional and neutral facial expressions across exposure durations and awareness ratings. Left: averaged waveforms resulting from VAN in the left and right hemispheres. Averaged time window highlighted in grey (185-215 ms); right: voltage means. Awareness-present trials had significantly more positive voltage than awareness-absent trials at 4.4 ms of exposure. (b) Topographical distributions of VAN for each condition at the relevant time window show an increase in negative voltage in posterior areas across exposure durations across expressions, awareness ratings, and exposure durations. (c) Source estimation of VAN visually identified on cortical maps. Estimated current sources are localised around inferotemporal gyri and lateral occipital sulcus, across all conditions. Asterisks indicate statistically significant differences between awareness-present and awareness-absent trials. Error bars represent 95% CI.

### 5.2.3.6 *Conscious access (LP)*

The late positivity (LP) is a positive voltage enhancement in the P3 wave. This component is measured over parietooccipital areas, centrally. To test whether LP was sensitive to awareness, we sorted trials in the same manner we did with VAN. Then, we entered mean voltage values into a 2 (awareness: aware, unaware) × 2 (expression: emotional, neutral) × 2 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1,30)} = 27.518, p < .001, \eta p^2 = .478$), indicating that LP became more positive with increasing exposure duration (Figure 5.8). Importantly, we found a main effect of awareness ($F_{(1,30)} = 7.737, p = .009, \eta p^2 = .205$), with more positive LP voltage in awareness-present ($M = 3.597$ [0.872]) than awareness-absent trials ($M = 3.150$ [0.111]). Unexpectedly, we also found a main effect of expression ($F_{(1,30)} = 4.603, p = .04, \eta p^2 = .133$), indicating more positive voltage values for neutral ($M = 3.452$ [5.590]) than emotional expressions ($M = 3.295$ [0.736]). This effect is consistent with the (equally unexpected) effect of expression found for the VAN (i.e. greater enhancement for neutral expressions than for emotional ones). Crucially, the interaction between awareness and exposure duration reached significance ($F_{(1,30)} = 37.420, p < .001, \eta p^2 = .555$). *T*o test at which specific exposure durations LP can discriminate between awareness-present and awareness-absent trials, and thus answer our main question here – i.e. at which exposure duration we find a neural indication of awareness – we ran post hoc Bonferroni-corrected pairwise comparisons and found that LP was significantly more positive in awareness-present than awareness-absent trials only at 4.4 ms of exposure ($t(55) = 5.861, p < .001, d = 1.053$). This finding suggests that 4.4 ms of exposure may be sufficient for enabling conscious access to faces. Importantly, this finding converges with VAN findings. Neither the interaction between expression and awareness ($F_{\mathbf{(1,30)}} = 0.457, p = .504, \eta p^2 = .015$) nor the interaction between expression and exposure duration ($F_{(1,30)} = 0.377, p = .544, \eta p^2 = .012$) reached significance. The three-way interaction did reach significance ($F_{(1,30)} = 6.361, p = .017, \eta p^2 = .175$). This interaction seems to be driven by the fact that exposure duration modulated the effect of awareness more strongly for emotional than for neutral expressions. However, we did not examine this in greater detail as we did not have a specific hypothesis concerning such modulations.

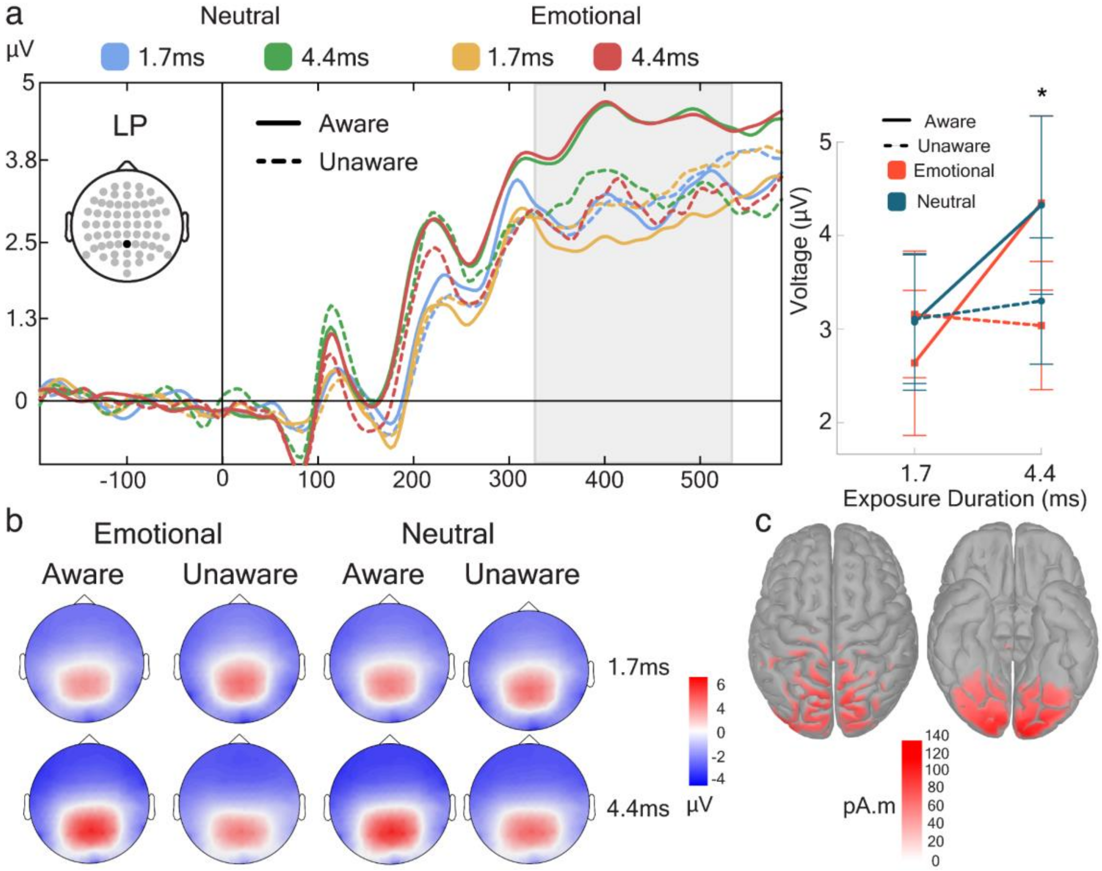



**Figure 5.8.** Conscious access indexed by LP. (a) LP response to emotional and neutral facial expressions across exposure durations and awareness ratings. Left: LP differences with averaged time window highlighted in grey (320-520 ms); right: voltage means. Awareness-present trials had significantly more positive voltage than awareness-absent trials at 4.4 ms of exposure. (b) Topographical distributions of the LP for each condition at the relevant time window show increase in positive voltage in parietocentral areas across exposure durations. (c) Source estimation of LP visually identified on cortical maps. Estimated current sources are localised around parietocentral and inferotemporal areas. Asterisks indicate statistically significant differences between awareness-present and awareness-absent trials. Error bars represent 95% CI.

### 5.2.4 Discussion

Are neural systems of emotion processing engaged before faces reach a level of processing that enables expression identification and reports of perceptual awareness? In this experiment, we measured neural markers of visual processing (P1), face processing (N170/VPP), emotion processing (EPN and LPP), and awareness (VAN and LP) with EEG, using three exposure durations that proved relevant according to signal detection analyses in Experiment 9: 1.7 ms (when holistic face processing and expression identification are absent), 4.4 ms (when holistic face processing arises), and 6.2 ms (when both expression identification and conscious awareness are very likely to have arisen). Signal detection analyses in Experiment 13 revealed a very similar sequence of processing steps: at 1.7 ms we found weak, but above-chance, stimulus discrimination; and at 4.4 ms above-chance emotion identification, and metacognitive sensitivity arose. By 6.2 ms, all these processes exhibited high sensitivity. If neural markers of emotion processing could distinguish between emotional and neutral expressions at shorter durations than those at which we find behavioural indications of emotion identification, this would provide evidence for unconscious emotion processing. Crucially, we found that both neural markers of emotion processing (EPN and LPP) discriminated between emotional and neutral expressions at 6.2 ms of exposure, meaning that emotion processing arose at some point between 4.4 ms and 6.2 ms of exposure. Together, neural markers and signal detection indices suggest that emotion processing does not arise before holistic face processing and awareness do.

How much visual exposure is required for faces to reach perceptual awareness? To address this question, we measured two neural markers of awareness (VAN and LP); both markers discriminated between awareness-present and awareness-absent trials when faces were presented for 4.4 ms, thus suggesting that perceptual awareness requires between 1.7 ms and 4.4 ms of visual exposure, alongside holistic face processing (as shown in Experiment 9) and expression identification (as shown in Experiment 13). However, what aspects of awareness VAN and LP specifically index is still a matter of debate (Koivisto & Revonsuo, 2010; Railo et al., 2011). Some studies have suggested that VAN may index phenomenal consciousness (i.e. subjective experience) whereas LP would index conscious access (i.e. the ability to consciously access and thus report sensory information; Eklund & Wiens, 2018, 2019; Jimenez et al., 2018; Koivisto et al., 2016; Wilenius-Emet et al., 2004). Regardless of this distinction's validity, the fact that both neural markers discriminated between awareness-present and awareness-absent trials by 4.4 ms of exposure may suggest that perceptual awareness emerges alongside face holistic processing. It is interesting to note that we also found an unexpected effect of emotion on both VAN and LP, which might suggest that these components are affected by other cognitive processes, too. For example, Hernández-Lorca et al. (2019), using Binocular Rivalry, found that VAN was modulated by awareness and emotional content, which they interpreted as evidence of an emotional bias in perceptual awareness. It has been also suggested that VAN, N170, and EPN may be affected by different but correlated underlying processes when studying the interaction between face perception, emotion processing, and awareness, as those three markers share very similar topographies and time windows (Wierzchoń et al., 2016). Future studies should expand on whether VAN specifically indexes awareness or other aspects involved in perception that may be necessary but not sufficient for awareness. Similarly, the effect of emotion on LP might have been driven by the LPP's sensitivity to emotion content; they have similar topographies and time windows, and both are part of the P3 wave (Hajcak et al., 2010; Wilenius & Revonsuo, 2007).

In conclusion, both neural markers and signal detection indices required roughly the same minimal exposure duration to discriminate emotional content in faces. So far, however, all experiments in this and the previous chapter have used only images of faces. To what extent are the processes and durations we found specific to faces? In Experiment 9, we found that holistic face processing, indexed by the face-inversion effect, required a minimal exposure duration of 4.4 ms to arise. Do neural systems of face processing require the same amount of visual exposure to discriminate a face from a non-face stimulus? In Experiment 13, we measured the N170/VPP complex, which is a neural marker of face processing. However, we did not include a control condition to determine by which exposure duration neural face processing arises. To determine the minimal required exposure of face processing, we need to measure the N170/VPP in response to face and non-face stimuli.

## 5.3 Experiment 14

In Experiment 13, we used signal detection indices and EEG neural markers to test whether emotion processing arises before perceptual awareness. We found that neural emotion processing arose with a longer exposure duration (6.2 ms) than perceptual awareness did (4.4 ms), thus indicating that emotion processing arises after participants become aware of faces. Therefore, perceptual awareness may be required for emotion processing to occur. But is this the case of face processing? Does face processing arise before perceptual awareness does? It could be the case that while emotion processing may require perceptual awareness to arise, face processing arises before perceptual awareness does.

In Experiment 14, our main question is whether neural systems can discriminate face from non-face stimuli before holistic processing and perceptual awareness arise. In Experiment 9, we found evidence of holistic face processing – indexed by the face-inversion effect in location discrimination – by 4.4 ms of exposure for stimuli presented peripherally, as in the experiments of this chapter. But may neural markers that are specific to face perception be evident even before holistic processing (a hallmark of specialised face processing) arises? In Experiment 13, we measured the N170 component, which is sensitive to face and face-like visual information. The N170 component belongs to the N1-family, a group of visually evoked potentials that are elicited over visual cortical areas in response to visual stimulation (Eimer, 2011). Therefore, we cannot tell whether the responses of N170 in Experiment 13 were specific to faces as any kind of visual processing could explain it. By comparing the ERP evoked by face and non-face stimuli we can find the exposure duration at which the response evoked by faces becomes an N170, distinguishable from the N1evoked by non-face stimuli. If N170 can discriminate between faces and non-faces before there is sufficient visual information available to elicit the face-inversion effect (i.e. around 4.4 ms), then neural systems can process facial information before holistic face processing arises. On the other hand, if the same minimal exposure duration is required to generate an N170 that discriminates between face and non-face stimuli as is required for the face-inversion effect to arise, then face processing may indeed require holistic processing to occur.

Measuring the response of N170 to face and non-face stimuli, as just described, should answer our main question. However, N170 is a local neural marker with its source in the fusiform gyri (Gao et al., 2019). Because neural face processing may recruit different neural systems to engage with face stimuli, we also measured a large-scale functional connectivity marker: weighted Symbolic Mutual Information (wSMI). wSMI is a marker of information sharing across the cortex. By measuring wSMI, we can address two questions: whether information sharing increases within the range of extremely brief exposure durations we used, and whether a difference in information sharing for face and non-face stimuli arises within this exposure-duration range. Importantly, wSMI offers a way to explore whether the information conveyed with these brief exposure durations can detect global changes in information sharing across the brain, a relevant characteristic in theories of consciousness.

### 5.3.1 Method

#### *5.3.1.1 Participants*

Thirty-eight students of the Université Libre de Bruxelles provided informed consent and were paid €30 for participation. All had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disorders. Six participants were excluded from the analysis due to EEG artefacts (see Analysis section). The remaining 32 participants (21 female; all right-handed) had a mean age of 23.2 ($SD_{age}$ = 3.2; range: 18 – 27).

#### *5.3.1.2 Stimuli*

Stimuli were 10 human faces (all with neutral expressions, in upright orientation; 5 female) taken from the RaFD and 10 highly recognisable objects taken from the Bank of Standardized Stimuli (BOSS; Brodeur et al., 2010, 2014). All images went through the same standardisation procedure described in Experiment 9, including the generation of scrambled images, which followed the same procedure as in Chapter 4. Scrambled and non-scrambled object images were additionally put inside a grey oval that resembled the faces' contour. The region between the oval's edge and the object inside was painted with a shade of grey equivalent to the mean RGB value of the 10 human faces' contours (see Appendix J).

#### 5.3.1.3 *Procedure*

The procedure was very similar to that of Experiment 9, but with four exposure durations, equally spaced on a logarithmic scale: 0.8, 1.4, 2.45, and 4.288 ms. We used a logarithmic scale to describe with higher precision how face processing arose at the shortest durations. Half the trials contained an intact face image on one side and its scrambled counterpart on the other side; the other half contained an intact object image on one side with its scrambled counterpart on the other side. Participants were asked to judge the location of the intact stimulus (left or right) and its category (face or object) with a single keypress. The 'left Control' and 'Left Shift' keys of a standard keyboard were used for 'left' reports, and the 'up arrow' and 'down arrow' keys were used for 'right' reports. Mapping of keys to stimulus category was counterbalanced across participants. Then they judged their visual experience (PAS). Participants performed 40 practice trials followed by 1120 experimental trials. Trial order was fully randomised. Participants were given self-terminated breaks every 70 trials and a compulsory 15-minute break after completing 560 trials.

#### 5.3.1.4 *EEG Recording and Pre-processing*

EEG recording and pre-processing was performed as in Experiment 13, with two differences: first, based on eye inspection of canonical sites, we selected slightly different time windows for the following ERP components: N170/VPP (175 – 205 ms), VAN (200 – 230 ms), and LP (300 – 585 ms). Second, mean amplitudes were not computed for EPN and LPP because we did not employ emotional expressions. For wSMI, we employed the time window used to measure N170/VPP.

#### 5.3.1.5 Signal Detection Analysis

Signal detection analysis of location sensitivity and response bias was performed as in Experiment 9. In addition, we measured stimulus category identification sensitivity – how well participants can discriminate faces from objects – by defining face trials as signal and object trials as noise. Then, we applied the equivalent signal detection analysis employed for expression identification sensitivity in Experiment 9 to obtain stimulus category identification d' and criterion scores.

#### 5.3.1.6 ERP Analysis

ERP analysis and source reconstruction were performed as in Experiment 13. We excluded data from six participants with more than 15% of noisy electrodes (whose impedance values exceeded 50 kΩ throughout the experiment). They were excluded from both signal detection and EEG analyses. One additional participant did not provide "almost clear experience" and "clear experience" PAS ratings and therefore their data were excluded from the VAN and LP analyses.

#### 5.3.1.7 EEG Neural Information Integration Analysis

We quantified the information flow between electrodes by calculating the weighted symbolic mutual information (wSMI). This index estimates to which extent two EEG signals exhibit non-random joint (i.e. correlated) fluctuations. Thus, wSMI has been proposed as a measure of neural information sharing (King et al., 2013; Sitt et al., 2014) and has three main advantages. First, it is a rapid and robust estimate of signals' entropy (i.e. statistical uncertainty in signal patterns), as it reduces the signal's length (i.e. dimensionality) by looking for qualitative or "symbolic" patterns of increase or decrease in the signal. Second, it efficiently detects high non-linear coupling (i.e. non-proportional relationships between neural signals) between EEG signals, as it has been shown with simulated (Imperatori et al., 2019) and experimental data (Canales-Johnson et al., 2020) that high non-linear coupling between EEG signals is neurocognitively meaningful, and it can be easily captured by wSMI (unlike with other traditional EEG connectivity measures such as phase synchronization). Third, it rejects spurious correlations between signals that share a common source, thus prioritising non-trivial pairs of symbols.

We calculated wSMI between each pair of electrodes (King et al., 2013), for each trial, after transforming the EEG signal into a sequence of discrete symbols defined by ordering of k time samples with a temporal separation between each pair (or τ). The symbolic transformation is determined by a fixed symbol size (k = 3, i.e. 3 samples represent a symbol) and the variable τ between samples (temporal distance between samples), thus determining the frequency range in which wSMI is estimated (Sitt et al., 2014). We chose τ = 32. The frequency specificity f of wSMI is related to k and τ as follows:

$$f = \frac{1000}{\tau * k}$$

This formula, with a kernel size k of 3 and τ values of 32, produced a sensitivity to frequencies in the range under 10 Hz. This range encompasses alpha frequencies, which are relevant for awareness (Lozano-Soldevilla & VanRullen, 2019; VanRullen & Macdonald, 2012), and also cover the frequency range of the N170/VPP complex (Eimer & Holmes, 2007; Itier & Taylor, 2004).

wSMI was estimated for each pair of transformed EEG signals by calculating the joint probability of each pair of symbols. The joint probability matrix was multiplied by binary weights to reduce spurious correlations between signals. The weights were set to zero for pairs of identical symbols, as these could have been elicited by a unique common source, and for opposite symbols (i.e. of in opposite direction), as these could reflect the two sides of a single electric dipole. The following formula calculates wSMI (in bits, but shown in arbitrary units or AU, with absolute values, in Figure 5.14):

$$wSMI(X,Y) = \frac{1}{\log(k!)} \sum_{x \in X} \sum_{y \in Y} w(x,y)\, p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

Here, x and y are symbols present in signals X and Y respectively; w(x,y) is the weight matrix and p(x,y) is the joint probability of co-occurrence of symbol x in signal X and symbol y in signal Y. Finally, p(x) and p(y) are the probabilities of those symbols in each signal and k! is the number of symbols used to normalise the mutual information by the signal's maximal entropy. Temporal evolution of wSMI was calculated using a 500ms sliding window with 2-ms time step, i.e. with a 96% overlap between two adjacent windows.

### 5.3.2 Behavioural results

#### *5.3.2.1 Location sensitivity*

We measured location sensitivity to confirm that our new participants displayed comparable performance to participants in the equivalent conditions of Experiment 9. To examine how conditions affected stimulus discrimination, we entered location d' scores into a 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration $(F_{(1.28,\ 39.75)} = 184.63, p < .001, \eta p^2 = .856)$, indicating that location d' scores increased with increasing exposure duration (Figure 5.9a). We also found a main effect of stimulus category $(F_{(1,\ 31)} = 5.63, p = .024, \eta p^2 = .154)$, with higher location d' scores for objects $(M = 0.709\ [0.777])$ than faces $(M = 0.567\ [0.751])$. Finally, the interaction between the two factors also reached significance $(F_{(2.06,\ 63.81)} = 8.92, p < .001, \eta p^2 = .224)$. To test whether location d' scores differed between stimulus categories at any exposure duration, we ran post hoc Bonferroni-corrected pairwise comparisons and found significantly higher location d' scores for objects over faces at 2.45 ms of exposure $(t(98.3) = -5.386, p < .001, d = -0.952)$. This advantage may be due to low-level visual differences such as stimulus size – no object could fill the oval entirely like faces, which may have led to contrast differences. Overall, these results indicate that location d' to both faces and objects increases with increasing exposure duration even in the brief range from

0.8 to 4.288 ms. Importantly, the d' scores we found resemble the ones found in Experiment 9.

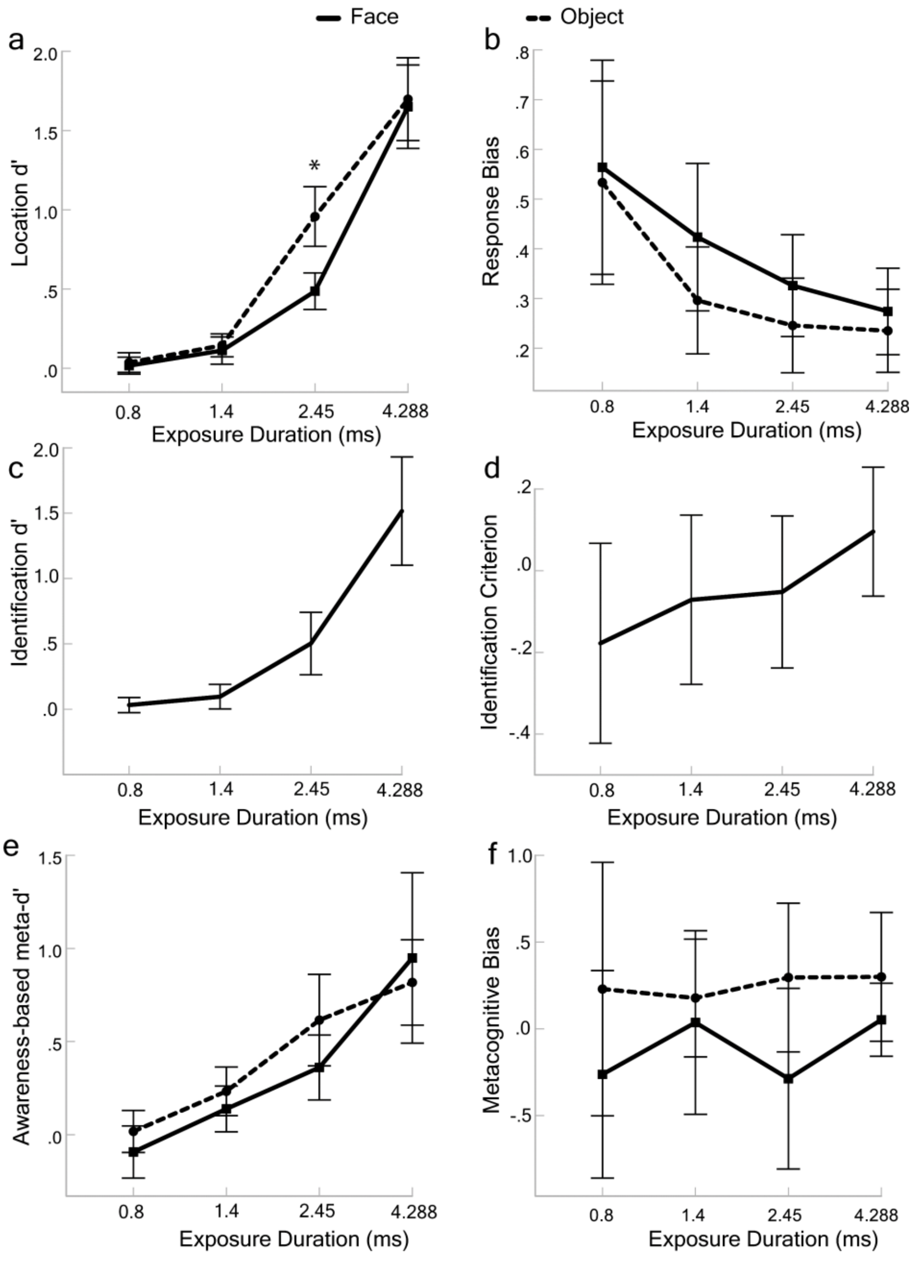



**Figure 5.9.** Behavioural results of Experiment 14. (a) Location sensitivity. Location d' increased with increasing exposure duration. A significant advantage for objects over faces was found at 2.45 ms of exposure. (b) Absolute-value location response bias scores for reporting location (bias toward either left or right). The amount of bias decreased as exposure duration increased, with greater bias for faces than objects. (c) Identification

sensitivity for stimulus category. Participants' ability to discriminate faces and objects increased with exposure duration. (d) Criterion scores for reporting stimulus category. Lower criterion indicates greater bias to reporting a face. Criterion does not significantly change with increasing exposure duration. (e) Awareness-based metacognitive sensitivity. Meta-d' increased with exposure duration but was unaffected by stimulus category. (f) Metacognitive bias scores for reporting subjective awareness. Metacognitive bias was unaffected by exposure duration and stimulus category. Asterisks index statistically significant differences between faces and objects. Error bars represent 95% CI.

To determine the minimal required exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance discrimination was 1.4 ms for both face $(M = 0.112\ [0.239];\ t(31) = 2.65, p = .013, d = 0.468)$ and object stimuli $(M = 0.145\ [0.201];\ t(31) = 4.08, p < .001, d = 0.721)$.

### 5.3.2.2 *Location response bias*

We examined whether participants' response bias for reporting stimulus location varied across conditions by entering the absolute values of $C_{identification}$ scores into a 2 (stimulus categories: face, object) × 4 (exposure durations) repeated-measures ANOVA. As in previous experiments, response bias significantly decreased with exposure duration (Figure 5.9b), as indicated by a main effect of exposure duration $(F_{(3,\ 39.57)} = 8.26, p = .004, \eta p^2 = .210)$. We also found a main effect of stimulus category $(F_{(1,\ 31)} = 8.60, p = .006, \eta p^2 = .217)$, indicating higher amount of response bias for faces $(M = 0.387\ [0.127])$ than objects $(M = 0.328\ [0.140])$. The interaction between these two factors did not reach significance $(F_{(2.68,\ 82.95)} = 1.79, p = .161, \eta p^2 = .055)$.

### 5.3.2.3 *Stimulus category identification sensitivity*

We examined whether participants' sensitivity to discriminating faces from objects varied across exposure durations by comparing identification d' scores in a one-way ANOVA (Figure 5.9c). As expected, identification d' increased alongside exposure duration ($F_{(3, 124)} = 32.4, p < .001, \eta p^2 = .440$).

To determine the minimal required exposure that exhibited above-chance performance (d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance stimulus category identification was 1.4 ms ($M = 0.096$ $[0.260]$; $t(31) = 2.09, p = .045, d = 0.369$).

### 5.3.2.4 *Stimulus category identification criterion*

We examined whether participants' criterion for reporting faces rather than objects varied across exposure durations by comparing criterion scores in a one-way ANOVA (Figure 5.9d). Lower scores indicated a greater tendency to report a face. Identification criterion did not vary across exposure durations ($F_{(3, 124)} = 1.30, p = .278, \eta p^2 = .030$).

### 5.3.2.5 *Awareness-based metacognitive sensitivity*

We examined whether awareness scores were sensitive to participants' location sensitivity scores by calculating meta-d', a measure of metacognitive sensitivity. To examine whether metacognitive sensitivity varied across conditions, we entered meta-d' scores into a 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA (Figure 5.9e). A main effect of exposure duration indicated that meta-d' increased with increasing exposure duration ($F_{(1.79, 55.36)} = 24.20, p < .001, \eta p^2 = .438$). However, we did not find a main effect of stimulus category ($F_{(1, 31)} = 0.902, p = .350, \eta p^2 = .003$), suggesting that stimulus category did not affect metacognitive sensitivity. We did not find an interaction between stimulus category and exposure duration either ($F_{(2.06,\ 63.96)} = 1.744, p = .182, \eta p^2 = .053$).

As described, we did not find a main effect of stimulus category, therefore we calculated Bayes factors to test whether the obtained data support this absence of an effect. Bayes factors indicated substantial evidence in favour of the null hypothesis model ($BF_{01} = 4.67$), supporting the finding that metacognitive sensitivity was not greater for either stimulus category.

To determine the minimal exposure that exhibited above-chance performance (meta-d' > 0), we ran a series of uncorrected one-sample t-tests against zero. We found that the earliest exposure duration that elicited above-chance metacognitive awareness was 1.4 ms for both face ($M = 0.139\ [0.341];\ t(31) = 2.3, p = .028, d = 0.407$) and object stimuli ($M = 0.233\ [0.362];\ t(31) = 3.64, p < .001, d = 0.644$).

### 5.3.2.6 *Metacognitive bias*

As described above, metacognitive bias is the tendency to give high confidence (or awareness) ratings regardless of actual performance. To examine whether metacognitive bias varied across conditions, we entered meta-bias scores into a 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA (Figure 5.9f). We did not find an effect of exposure duration ($F_{(2.05,\ 63.47)} = 0.419, p = .664, \eta p^2 = .013$), or of stimulus category ($F_{(1,\ 31)} = 1.661, p = .207, \eta p^2 = .051$). The interaction between stimulus category and exposure duration did not reach significance either ($F_{(2.12,\ 65.59)} = 0.392, p = .689, \eta p^2 = .012$). These results suggest that the participants' tendency to describe their subjective awareness was not affected by any of the conditions.

## 5.3.3 ERP results

### 5.3.3.1 *Early visual processing (P1)*

We examined how stimulus categories and exposure duration conditions affected early visual processing by measuring voltage changes in P1. This component is measured over occipital regions. We entered mean voltage values into a 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.52,\ 47.02)} = 81.86, p < .001, \eta p^2 = .725$), whereby P1 amplitude increased with increasing exposure duration, indicating that longer exposure durations involve more visual processing than shorter exposure durations (Figure 5.10). We did not find a main effect of stimulus category, suggesting that images of faces and objects did not affect early visual processing differently ($F_{(1,\ 31)} = 1.49, p = .232, \eta p^2 = .046$). The interaction between these factors reached significance ($F_{(2.84,\ 88.04)} = 2.88, p = .043, \eta p^2 = .085$), but post hoc Bonferroni-corrected pairwise comparisons did not reveal significant differences between stimulus categories at any exposure duration. These results indicate that P1 was sensitive to visual differences between extremely brief exposure durations.

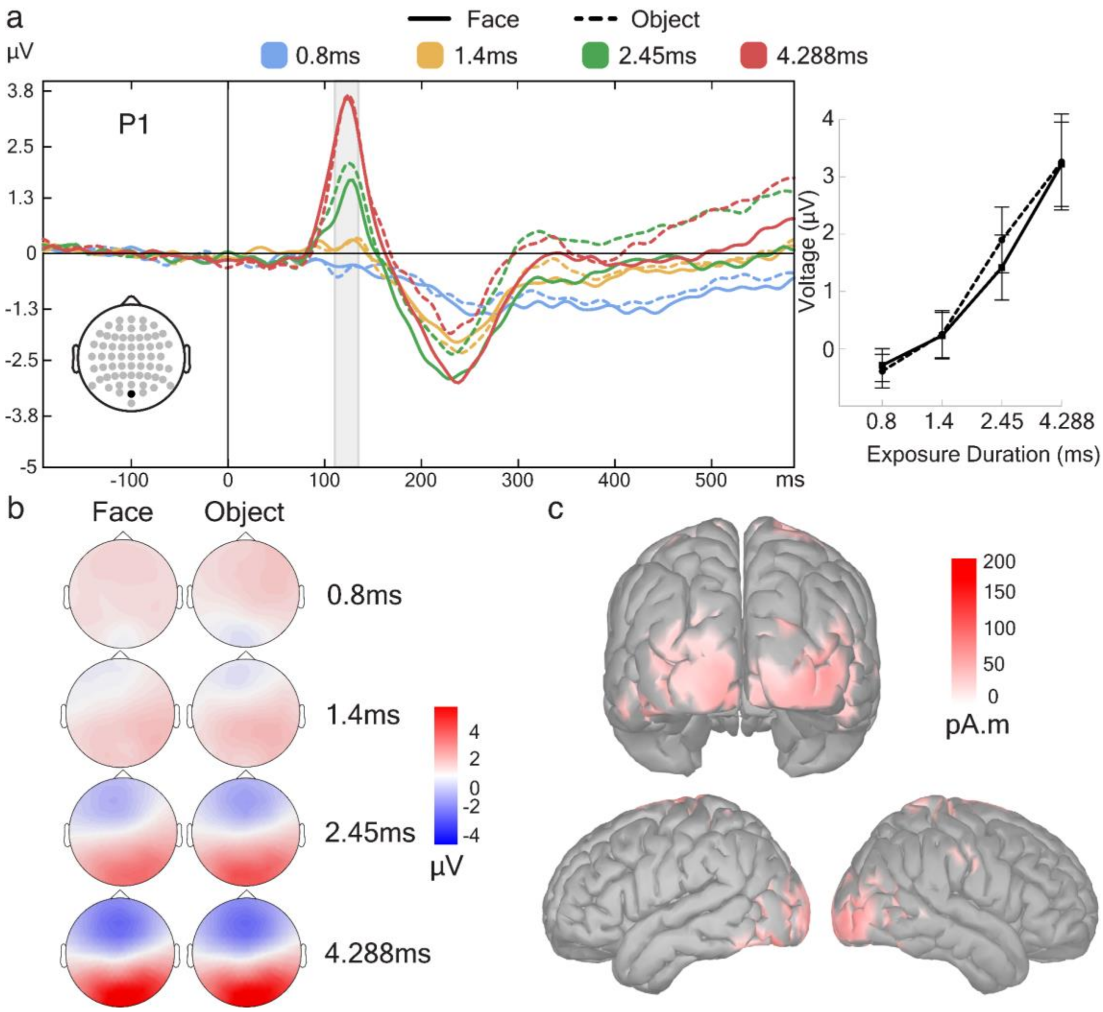



**Figure 5.10.** Early visual processing indexed by P1. (a) P1 response to face and object stimuli across exposure durations. Left: P1 peak with averaged time window highlighted in grey (105-135 ms); right: voltage means. P1 voltage increased with increasing exposure duration. (b) Topographical distributions of P1 for each condition at the relevant time window show increase in positive voltage in visual cortex across exposure durations. (c) Source estimation of P1 visually identified on cortical maps. Estimated current sources are localised around the visual cortex. Error bars represent 95% CI.

#### 5.3.3.2 *Face processing (N170/VPP)*

We examined how stimulus category and exposure duration conditions affected face processing by measuring voltage changes in N170/VPP. This component is measured over occipitotemporal regions, bilaterally, and over frontocentral regions. We entered mean voltage values into a 2 (stimulus categories: face, object) × 3 (electrode site: left, right, central) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.38,\,42.92)} = 45.681, p < .001, \eta p^2 = .596$), indicating that the magnitude of this component significantly increased with increasing exposure duration. We also found a main effect of electrode site ($F_{(1.31,\,40.59)} = 62.566, p < .001, \eta p^2 = .669$), which was expected given that left ($M = -3.225$ [1.510]) and right N170 peaks ($M = -3.394$ [1.603]) are negative in voltage whereas VPP is positive ($M = 2.789$ [1.050]). Like in Experiment 13, this effect helped confirm that we were measuring the N170/VPP complex. We did not find a main effect of stimulus category ($F_{(1,\,31)} = 0.005, p = .946, \eta p^2 = 0$), but crucially, we found a significant interaction between stimulus category and exposure duration ($F_{(2.45,\,75.89)} = 6.398, p = .001, \eta p^2 = .171$). To test whether N170/VPP was sensitive to faces at specific exposure durations, we ran post hoc Bonferroni-corrected pairwise comparisons and found that the N170/VPP was significantly greater in magnitude for face stimuli ($M_{\mathrm{N170:\,4.288ms}} = -5.093$ [0.307]) compared to object stimuli ($M_{\mathrm{N170:\,4.288ms}} = -4.290$ [0.159]) only at the longest exposure duration ($t(114) = -3.467, p = .021, d = -0.613$), thus indicating that 4.288 ms of visual exposure conveyed sufficient information for neural systems to distinguish between a face and a no-face stimuli (Figure 5.11). These results suggest that neural systems require around the same amount of visual exposure to trigger face processing as is required for the face-inversion effect to arise (see Experiment 9). Taken together, these findings indicate that neural face processing requires around 4 ms of exposure – and definitely no less than 2.45 ms where we see no difference.

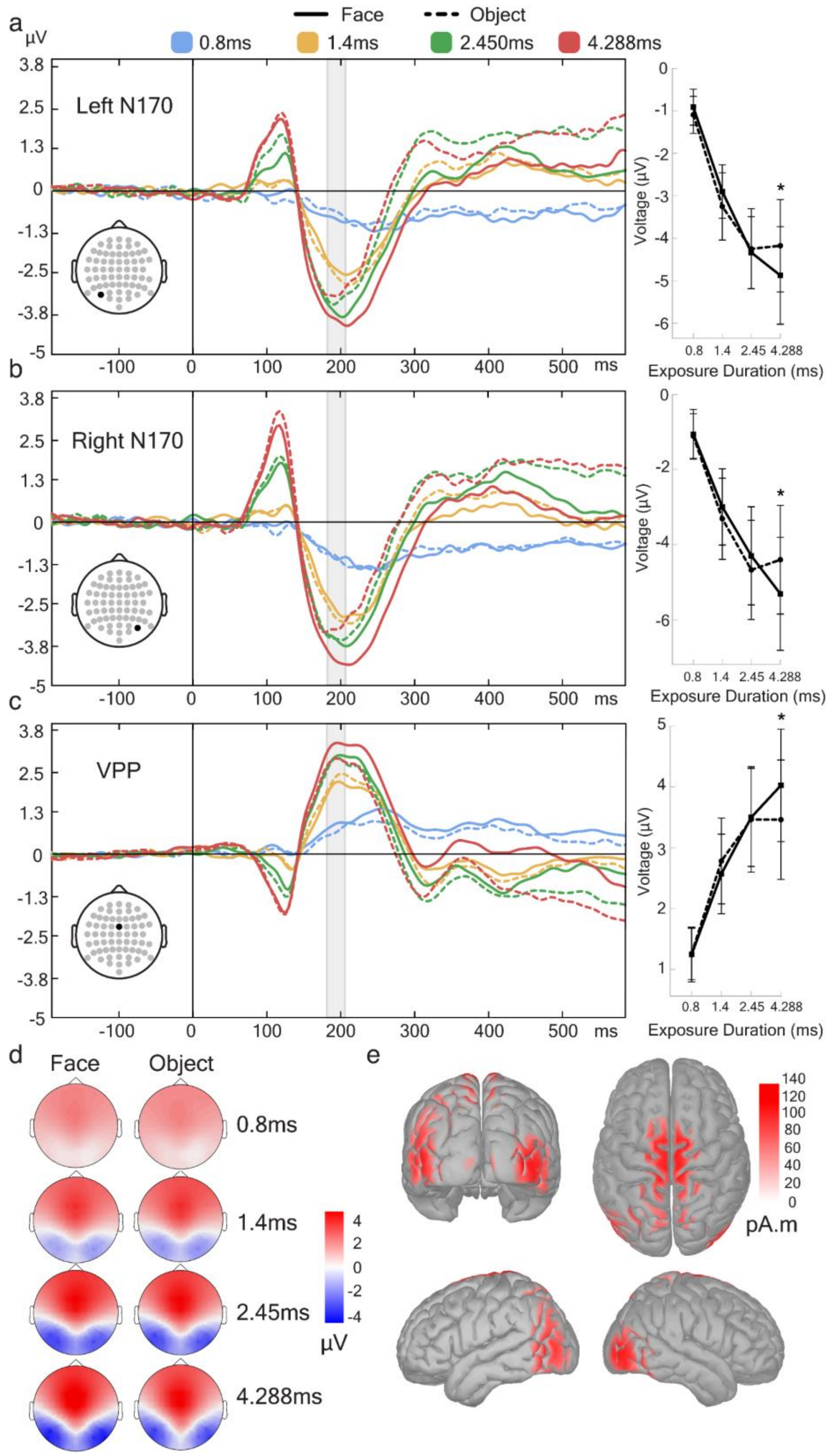



**Figure 5.11.** Face processing indexed by the N170/VPP. (a) N170/VPP response to face and object stimuli across increasing exposure durations. Left: N170/VPP peaks with averaged time window highlighted in grey (175-205 ms); right: voltage means. N170 and

VPP had significantly more negative and positive voltages, respectively, for faces than objects at 4.288 ms of exposure. (b) Topographical distributions of the N170/VPP complex for the relevant time window across stimulus category and exposure duration conditions. (c) Source estimation of N170/VPP visually identified on cortical maps. Estimated current sources are localised around frontocentral areas, the inferotemporal gyri, and the lateral occipital sulcus. Asterisks indicate statistically significant differences between conditions. Error bars represent 95% CI.

#### *5.3.3.3 Visual awareness (VAN)*

We examined whether VAN could distinguish between awareness-present and awareness-absent trials across exposure durations by sorting trials the same way we did in Experiment 13. This component is measured over occipitotemporal regions, bilaterally. We entered mean voltage values into a 2 (awareness: aware, unaware) × 2 (stimulus category: face, object) × 2 (electrode: left, right) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.94,\ 58.11)} = 18.86, p < .001, \eta p^2 = .386$), indicating that VAN became more negative with increasing exposure duration (Figure 5.12). We also found a main effect of stimulus category ($F_{(1,\ 30)} = 4.258, p = .048, \eta p^2 = .124$), which indicated that faces evoked more negative voltage means ($M = -2.607\ [1.076]$) than objects ($M = -2.270\ [0.931]$). This effect may be due to the fact that VAN was measured at the same electrode sites as N170, and at a later yet close time window. Importantly, we found a main effect of awareness ($F_{(1,\ 30)} = 19.612, p < .001, \eta p^2 = .395$) – awareness-present trials had more negative voltage means ($M = -2.766\ [1.044]$) than awareness-absent trials ($M = -2.112\ [0.880]$). Importantly, the interaction between awareness and exposure duration did not reach significance this time ($F_{(2.67,\ 80.19)} = 1.368, p = .260, \eta p^2 = .044$), suggesting that VAN distinguished between awareness-present and awareness-absent trials overall, in a linear fashion. This finding is in line with the view that VAN may be an index of phenomenal consciousness – arguably, such a marker should be sensitive to awareness reports regardless of conscious access to sensory information. Therefore, if VAN were a marker of phenomenal consciousness, it should correlate with awareness reports only in a linear way. Interestingly, the interaction between stimulus category and exposure duration reached significance ($F_{(2.49, 74.81)} = 3.708, p = .021, \eta p^2 = .110$). To test whether VAN was sensitive to stimulus categories at specific exposure durations, we ran post hoc Bonferroni-corrected pairwise comparisons and found that VAN was significantly more negative for faces than objects only at the longest exposure duration: 4.288 ms ($t(117) = -3.877, p = .005, d = -0.696$). This effect may have been driven by voltage changes in N170, as this component and VAN share electrode sites and have very similar temporal windows. We did not find a main effect of electrode site ($F_{(1, 30)} = 1.845, p = .185, \eta p^2 = .058$), suggesting that VAN changes did not differ between hemispheres. Finally, we did not find significant interactions between awareness and electrode site ($F_{(1, 30)} = 1.096, p = .303, \eta p^2 = .035$), stimulus category and awareness ($F_{(1, 30)} = 0.930, p = .343, \eta p^2 = .03$), electrode site and exposure duration ($F_{(1.95, 58.41)} = 1.269, p = .288, \eta p^2 = .041$), nor any three-way interaction and the four-way interaction. In conclusion, VAN was sensitive to awareness ratings provided by participants, irrespective of visual exposure.

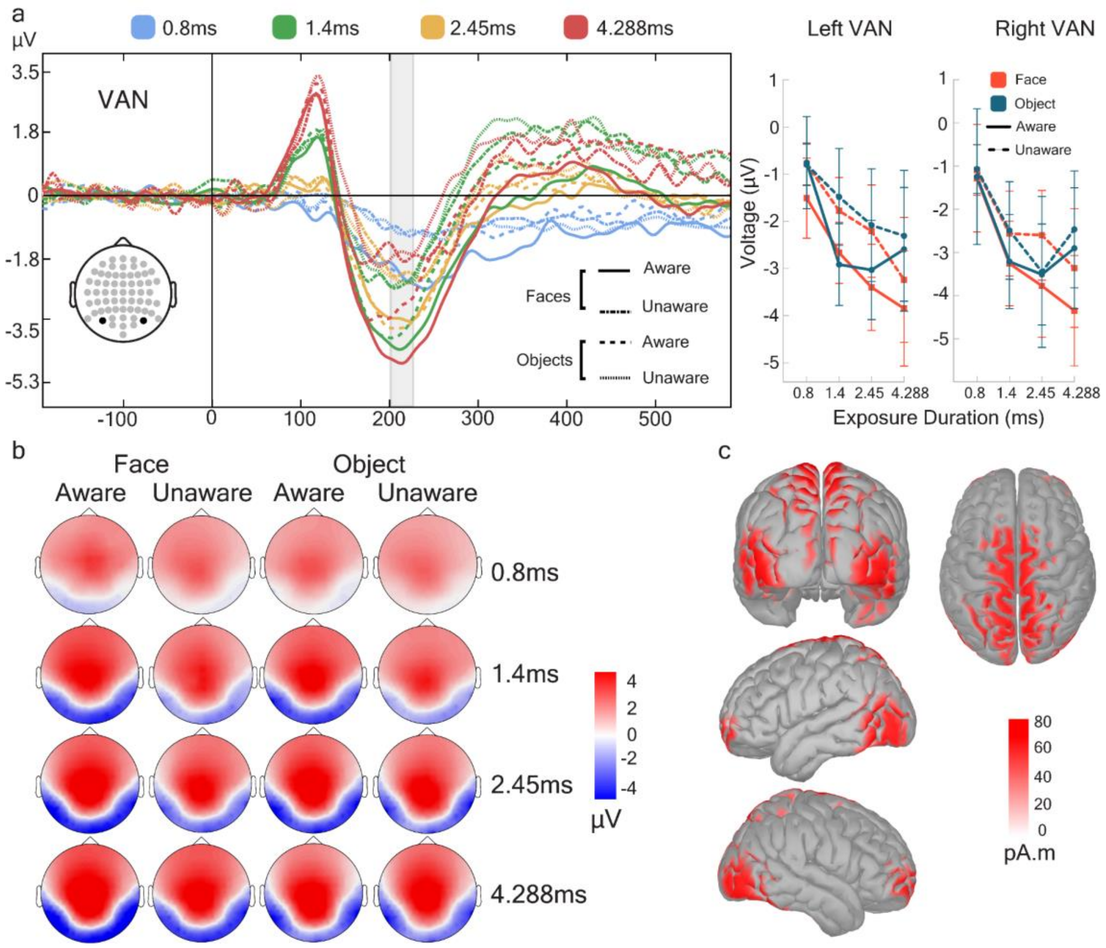



**Figure 5.12.** Visual awareness indexed by VAN. (a) VAN response to face and object stimuli across exposure durations and awareness reports. Left: VAN peak with averaged time window highlighted in grey (200-230 ms); right: voltage means. Awareness-present trials (solid lines), irrespective of stimulus category, had significantly more negative voltage than awareness-absent trials (dashed lines) at 4.288 ms of exposure. (b) Topographical distributions of the VAN for each condition at the relevant time window show increase in negative voltage on posterior areas across exposure durations. (c) Source estimation of VAN visually identified on cortical maps. Estimated current sources are localised around frontocentral areas, parietocentral areas, the inferotemporal gyri, and the lateral occipital sulcus. Error bars represent 95% CI.

### 5.3.3.4 *Conscious access (LP)*

We examined whether LP could distinguish between awareness-present and awareness-absent trials by sorting trials the same way we did in Experiment 13. This component is measured over parietooccipital regions. We entered mean voltage values into a 2 (awareness: aware, unaware) × 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(1.99, 59.78)} = 11.715, p < .001, \eta p^2 = .281$), indicating that LP voltage increased with increasing exposure duration (Figure 5.13). Importantly, we found a main effect of awareness ($F_{(1, 30)} = 4.826, p = .036, \eta p^2 = .139$), which indicates that LP voltage was significantly more positive in awareness-present trials ($M = 2.621$ [0.858]) than in awareness-absent trials ($M = 2.142$ [0.372]). Therefore, like in Experiment 13, LP was sensitive to awareness ratings. Crucially, the interaction between awareness and exposure duration reached significance, indicating that exposure duration may have significantly modulated LP's sensitivity to awareness ($F_{(2.33, 70.03)} = 5.393, p = .004, \eta p^2 = .152$). To test whether LP was sensitive to awareness ratings at specific exposure durations, we ran post hoc Bonferroni-corrected pairwise comparisons and found that LP was significantly more positive in awareness-present than awareness-absent trials only at the longest exposure duration: 4.288 ms ($t(87.3) = 3.683, p = .011, d = 0.661$). This finding suggests that neural systems require around said amount of visual exposure to elicit conscious access. We did not find a main effect of stimulus category ($F_{(1, 30)} = 0.168, p = .685, \eta p^2 = .006$) nor significant interactions between stimulus category and awareness ($F_{(1, 30)} = 20.69, p = .161, \eta p^2 = .065$), and between stimulus category and exposure duration ($F_{(2, 80.34)} = 0.319, p = .789, \eta p^2 = .011$). The three-way interaction between awareness, stimulus category, and exposure duration was not significant either ($F_{(2.32, 69.66)} = 1.764, p = .174, \eta p^2 = .056$).

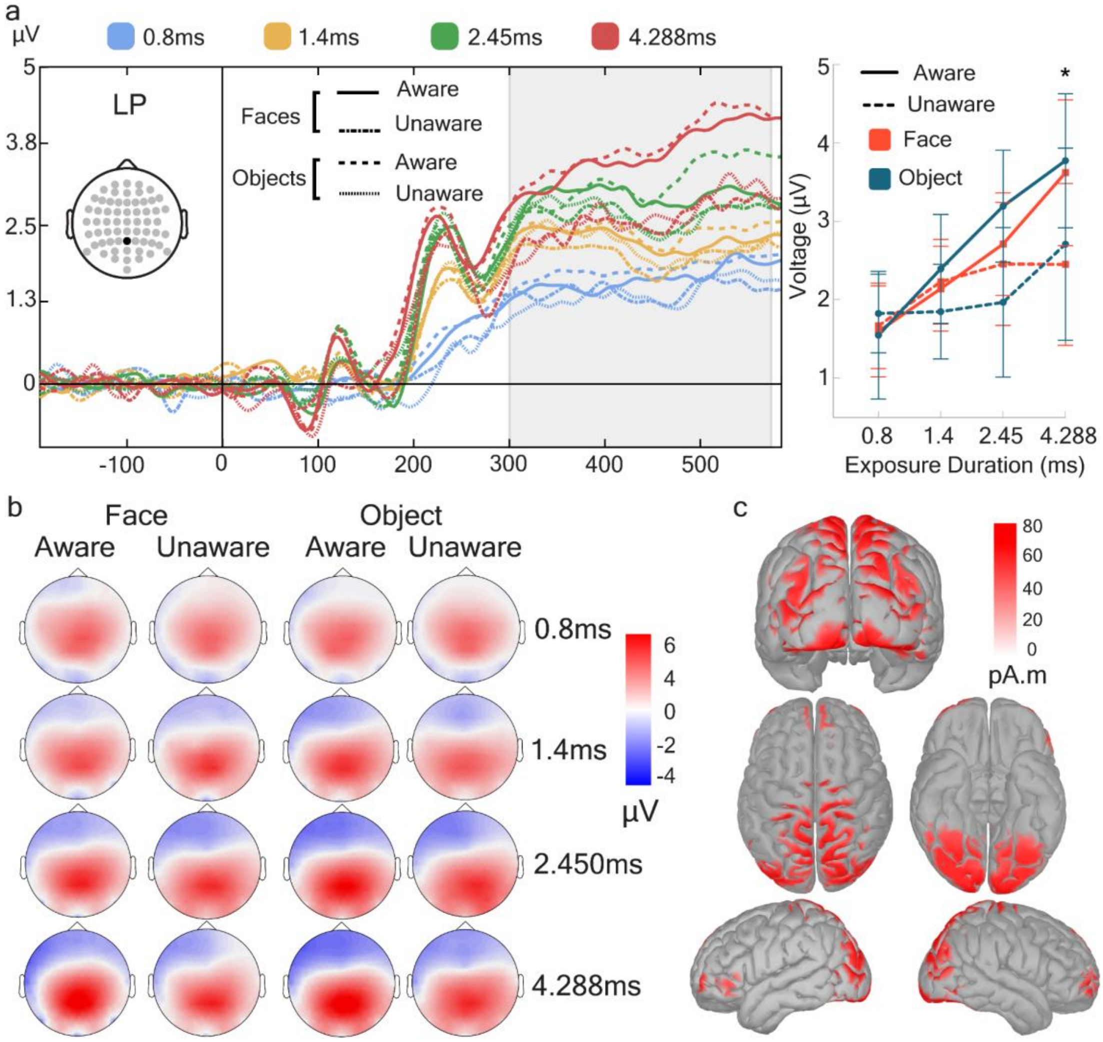



**Figure 5.13.** Conscious access indexed by LP. (a) LP response to face and object stimuli across exposure durations and awareness reports. Left: LP difference with averaged time window highlighted in grey (430-585 ms); right: voltage means. Awareness-present trials (solid lines) had significantly more positive voltage than awareness-absent trials (dashed lines) at 4.288 ms of exposure. (b) Topographical distributions of the LP for each condition at the relevant time window show increase in positive voltage on posterior areas across exposure durations. (c) Source estimation of LP visually identified on cortical maps. Estimated current sources are localised around postcentral, parietocentral, occipitotemporal, and inferotemporal areas. Asterisks indicate statistically significant differences between awareness-present and awareness-absent conditions. Error bars represent 95% CI.

### 5.3.4 EEG Neural information integration results

As a measure of neural information integration, we calculated wSMI, which estimates neural information sharing across the cortex. To determine whether wSMI could distinguish between faces and objects across exposure durations, we entered wSMI mean values into a 2 (stimulus category: face, object) × 4 (exposure durations) repeated-measures ANOVA. We found a main effect of exposure duration ($F_{(2.17, 67.13)} = 34.369, p < .001, \eta p^2 = .526$), indicating that large-scale information sharing increased with increasing exposure duration (Figure 5.14). However, we did not find a main effect of stimulus category ($F_{(1, 31)} = 0.042, p = .840, \eta p^2 = .001$), which may suggest that wSMI was not sensitive to stimulus categories with the exposure durations employed in Experiment 14. The interaction between the two factors was not significant either ($F_{(2.88, 89.28)} = 1.914, p < .135, \eta p^2 = .058$). To assess whether the obtained data support the absence of an effect of stimulus category, we estimated Bayes factors, which indicated substantial evidence for the null hypothesis model ($BF_{01} = 7.38$), thus supporting the absence of an effect. These results suggest that the exposure durations we employed were sufficiently long to elicit an increase in large-scale neural information integration, but not to cause different patterns of connectivity for each stimulus category.

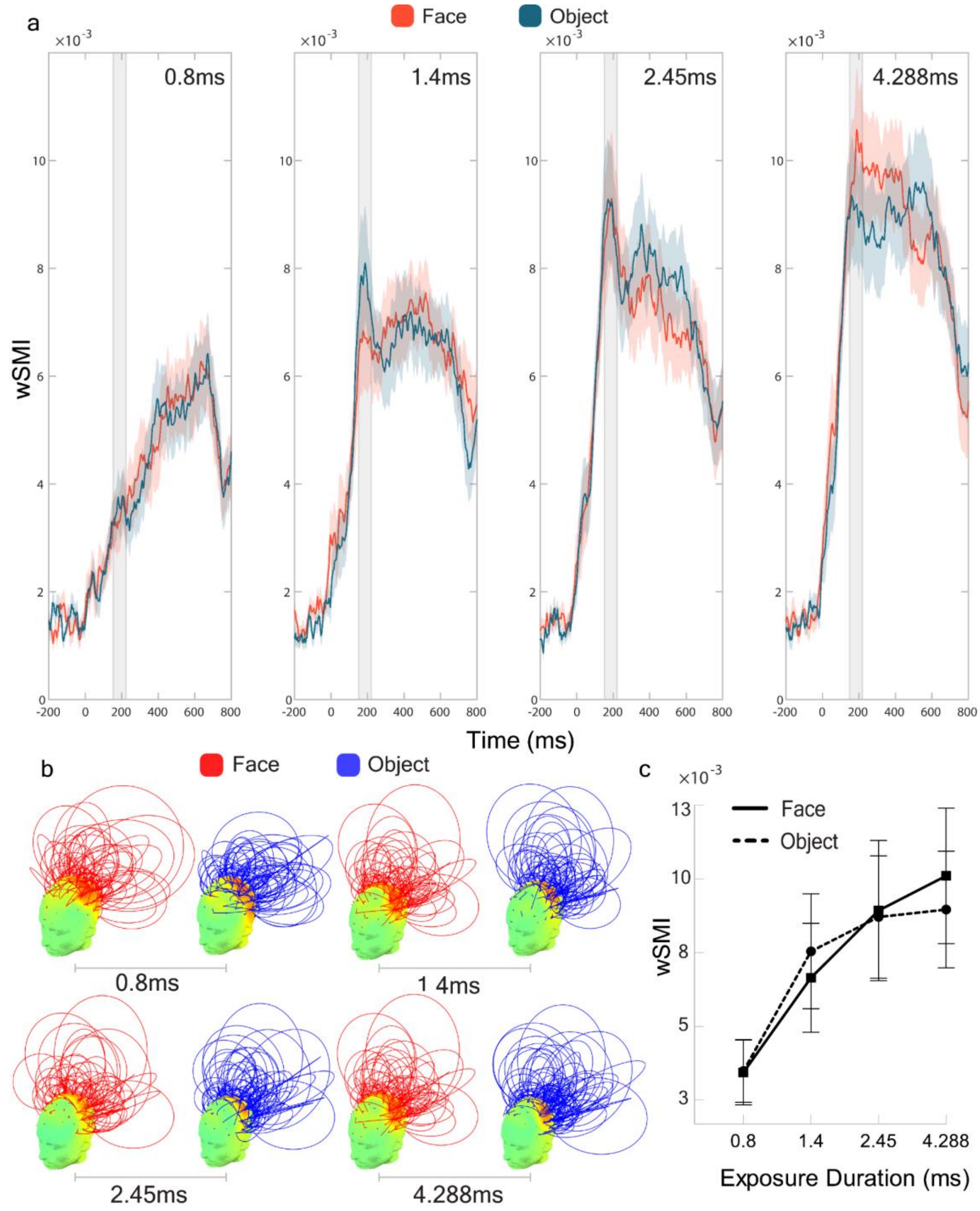



**Figure 5.14.** Neural information integration indexed by wSMI. Neural information sharing globally increased with increasing exposure durations. (a) wSMI increases across exposure durations, similarly for faces and objects, for the time window of interest. (b) Connectivity topographies for faces and objects across exposure durations. The height of an arc connecting two nodes indicates the strength of neural information integration between them. Faces and objects evoke similar patterns. Overall, the strength of functional connections increases with increasing exposure duration, but it is not

modulated by stimulus category. (c) wSMI means for each time window of interest. Higher scores indicate higher information sharing. Neural information sharing increased across exposure durations but did not distinguish between faces and objects. Error bars represent 95% CI.

### 5.3.5 Discussion

In this experiment, we measured neural markers of face processing to test whether they can discriminate face and non-face stimuli at shorter exposures than signal detection indices do. Crucially, we found that neural markers could discriminate between faces and objects at 4.288 ms of exposure, but not at shorter durations. This exposure duration roughly corresponds to the minimal exposure duration required for the face-inversion effect, as shown in Experiment 9. Therefore, our findings using neural markers here converge with findings from signal detection indices in Experiment 9, indicating that around four milliseconds of visual exposure are needed for the system to engage in face-specific processing.

What is the minimal exposure duration required for faces to access perceptual awareness? We found convergent evidence of faces (and objects) gaining access to perceptual awareness by 4.288 ms of exposure coming from signal detection indices (metacognitive sensitivity) and neural markers of conscious awareness (VAN and LP). Importantly, unlike in Experiment 13 where both VAN and LP discriminated between awareness-present and awareness-absent trials by 4.4 ms, in Experiment 14 only LP interacted with exposure duration and exhibited discrimination between awareness-present and awareness-absent trials by 4.288 ms of exposure. Although VAN exhibited a main effect of awareness, it did not interact with exposure duration. Importantly, these findings may suggest that in the sequence of processing steps of face perception, faces gain access to awareness as they become available for holistic processing.

Some studies have suggested that VAN indexes phenomenal consciousness (i.e. changes in subjective experience) whereas LP indexes conscious access (i.e. ability to report sensory information; Koivisto et al., 2016; Koivisto & Grassini, 2016; Koivisto & Revonsuo, 2010). Our findings in Experiment 14 may support this distinction and provide evidence of a minimal amount of visual exposure required for faces to gain access to awareness. Evidence from the N170/VPP complex and from signal detection analyses in this experiment converge, thereby suggesting that faces required around four milliseconds of exposure to gain access to perceptual awareness; this duration is similar to that required for holistic face processing in Experiment 9. This temporal convergence may suggest an underlying causal relationship between holistic face processing and awareness – e.g. awareness may be necessary for integration of visual features to allow or facilitate face recognition and subsequent emotional processing. We return to this point in the General Discussion.

In this experiment, we also measured wSMI, a neural marker of information integration, which indexes information sharing between different pairs of nodes across the brain cortex. As a measure of information sharing, wSMI may index global ignition in the brain and therefore broader mechanisms required for conscious access (King et al., 2013). We found that the narrow range of exposure durations used in Experiment 14 was sufficient to elicit global changes in information integration. However, we did not find differences in connectivity patterns between faces and objects. It may be the case that the exposure durations we used, from 0.8 to 4.288 ms, are simply too brief for the system to exhibit stimulus-category differences in information sharing. Alternatively, wSMI may not be sufficiently sensitive to capture those differences between stimulus categories. Future studies should explore whether wSMI can capture differences between contents of consciousness.

In summary, we found evidence with neural markers that face processing arises by four milliseconds of exposure. Similarly, we found evidence both with neural markers and signal detection indices that perceptual awareness of faces arises by around four milliseconds of exposure as well. Together, these findings suggest that face processing and awareness may arise together in face perception.

## 5.4 General Discussion

In Chapter 4, we presented participants with face images for exposure durations ranging from around half a millisecond to around six milliseconds. By measuring signal detection indices, we found evidence of a sequence of processing steps that unfolds in the early stages of face perception: stimulus discrimination, holistic face processing, and facial expression identification, in this order, all of which increased alongside metacognitive sensitivity, a measure of perceptual awareness. These findings fit with two viable descriptions of face perception: There are two possible accounts that would be consistent with these findings: on the one hand, it is possible that all facial features are indeed processed in sequence as visual information reaches the brain cortex. On the other hand, it is possible that facial features are processed simultaneously but transmitted to the rest of the brain at different speeds. These findings favour the former interpretation over the latter one.

In Chapter 5, we developed two EEG experiments to explore whether neural measures can distinguish emotional from unemotional facial expressions (Experiment 13) and faces from objects (Experiment 14) before relevant signal detection indices exhibit emotional and facial processing, respectively. Critically, in both experiments we additionally searched for the minimal exposure required for perceptual awareness.

### 5.4.1 Emotion processing and expression identification

Emotion processing was measured through expression identification and EEG evoked response (EPN and LPP; Experiment 13). While identification sensitivity suggests that facial expressions can be identified above chance with 4.4 ms of visual exposure, EEG markers suggest that emotion processing arises with a visual exposure somewhere between 4.4 and 6.2 ms. These findings have two key implications. First, that emotion processing cannot arise before holistic processing does. As shown in Experiment 9, holistic processing arises with a visual exposure of 4.4 ms, indexed by a face-inversion effect. And secondly, that emotion processing cannot arise in absence of awareness either, thus contradicting past claims about unconscious emotion processing (Hedger et al., 2015;

Pessoa et al., 2005; Yang et al., 2007; Yang & Yeh, 2018a). As shown by metacognitive sensitivity and ERP findings, perceptual awareness had reached high sensitivity with stimuli presented for 4.4 ms of exposure. In conclusion, neural systems do not engage differently with emotional and neutral expressions before gaining access to awareness.

Our findings may suggest that emotion processing of faces is carried out through a sequence of steps, with expression identification preceding emotion (intensity) processing. We acknowledge, however, that an expression identification task like ours could be completed by simply integrating facial features – physiological reactivity to emotion is expected but not required. Expression identification sensitivity showed a clear advantage of happy over fearful expressions despite differences in intensity between stimulus categories were minimised. This effect has been repeatedly reported in emotion recognition studies and has been attributed to happy expressions being easier to recognise than negative and neutral expressions, as it was also the case of our face stimuli (Calvo & Lundqvist, 2008; Goeleven et al., 2008; Langner et al., 2010; Svard et al., 2012). Future studies should address the question of whether expression identification is required for emotion processing and whether it is possible to recognise a face's expression without engaging neural emotion processing.

### 5.4.2 Face processing and visual integration

Face processing was indexed by the N170/VPP (Experiments 13 and 14). As shown by Experiment 9, a discrimination advantage in favour of upright faces over inverted faces (face-inversion effect) arises with a visual exposure of 4.4 ms, once both stimulus discrimination and expression identification sensitivity are above chance. Many studies have consistently reported that turning a face upside down disrupts face recognition (Farah et al., 1995; Goodrich & Yonelinas, 2019; Jiang et al., 2007; Stein et al., 2012; Yin, 1969), suggesting that face inversion disrupts the holistic configuration of a face. But are neural markers that are specific to face processing evident before holistic processing arises? N170/VPP indicates that face processing arises with a visual exposure of around 4.3 ms (Experiment 14), roughly the same minimal exposure required for the face-inversion effect, as reported in Experiment 9. These results elucidate the minimal required exposure for face-specific processing. While perceptual processing in general reached above-chance performance by 1.4 ms of visual exposure, face-specific systems required around 4.4 ms of visual exposure to discriminate between faces and objects.

What do the face-inversion effect and the N170/VPP complex specifically index? A number of studies have shown that the N170/VPP complex, whose main brain source is the fusiform gyrus (Deffke et al., 2007; Gao et al., 2019; Pizzagalli et al., 2002; Schweinberger et al., 2002; Shibata et al., 2002), is also sensitive to face inversion – with stronger amplitude and slight but consistently delayed latency for inverted faces (Bentin et al., 1996; Civile et al., 2018; Eimer, 2000; Heisz et al., 2006; Rossion et al., 2000). The fusiform gyrus exhibits increased activation, measured with fMRI, for upright over inverted faces (Kanwisher et al., 1998; Yovel & Kanwisher, 2005), which may suggest that both the face-inversion effect and the N170/VPP complex index holistic face processing. Thus, our findings can be taken as evidence of a consistent minimal required exposure for holistic face processing, which is greater than the exposure duration required for stimulus discrimination.

In Experiment 14, information sharing increased with increasing exposure durations ranging from 0.8 to 4.288 ms. This finding indicates that extremely small differences between visual exposures can significantly affect large-scale functional connectivity in the cortex. However, unlike N170/VPP, wSMI did not find differences in how faces and objects were processed. On the one hand, a longer exposure duration may be required for functional connectivity to capture high-level differences between faces and objects. A numerical trend that fits this interpretation was found. On the other hand, wSMI is a relatively new neural marker and it may be the case that it is not sensitive to such differences in visual content.

### 5.4.3 Consciousness: subjective awareness and conscious access

So far, our findings have presented a very consistent picture of how face perception occurs: A sequence of processing steps goes through stimulus discrimination, holistic processing, and emotional processing. However, what is the minimal exposure duration required for perceptual awareness to arise? In Experiment 9, we estimated metacognitive sensitivity by testing how awareness ratings predicted changes in location discrimination across exposure durations. While we did find evidence of better metacognitive sensitivity for upright over inverted faces overall, this advantage did not interact with exposure duration. Therefore, it is not possible to conclude whether this face-inversion effect demands a clear minimal required exposure.

In Experiments 13 and 14, we measured two neural markers of perceptual awareness, VAN and LP, and tested whether their sensitivity to subjective awareness emerges by a specific exposure duration. In Experiment 13, both VAN and LP could discriminate between awareness-present and awareness-absent trials when stimuli were presented for 4.4 ms, suggesting that the exposure duration required for holistic processing, found in Experiment 9, may be required for perceptual awareness. This conclusion is supported by metacognitive sensitivity, too – in Experiment 13, we found a non-linear increase in metacognitive sensitivity with an exposure duration of 4.4 ms. Interestingly, though, when using face and object stimuli, as in Experiment 14, VAN could only discriminate between awareness-present and awareness-absent trials overall. On the other hand, LP was sensitive to awareness, again indicating that around 4.288 ms of visual exposure may convey sufficient visual information to discriminate awareness-present from awareness-absent trials. It is important to note, however, that metacognitive sensitivity was above-chance by visual exposures of 1.4 ms in Experiment 14, therefore suggesting that some form of awareness may have arisen alongside perceptual sensitivity. Because the PAS asked participants to rate their experience in a broad way, it is possible that metacognitive sensitivity at short exposure durations reflected awareness of basic aspects of the intact stimulus – required for its discrimination from its scrambled counterpart – but not awareness of its identity as a face or object. Taken together, these findings may suggest that faces (and perhaps objects) require a minimal visual exposure duration of around 4 ms to reach perceptual awareness.

But how reliable are VAN and LP as neural markers of awareness? Many studies have reported that these two ERP components are sensitive to awareness, having manipulated stimulus visibility using masking techniques (Genetti et al., 2009; Koivisto et al., 2005, 2006; Koivisto & Revonsuo, 2007, 2008b; Railo & Koivisto, 2009), reduced contrast (Koivisto et al., 2008; Pins & Ffytche, 2003; Wilenius & Revonsuo, 2007), attentional blink (Koivisto & Revonsuo, 2008a; Sergent et al., 2005), change blindness tasks (Busch et al., 2009; Eimer & Mazza, 2005; Fernandez-Duque et al., 2003; Koivisto & Revonsuo, 2003; Niedeggen et al., 2001; Pourtois et al., 2006; Schankin & Wascher, 2007; Turatto et al., 2002), and bistable perception (Kaernbach et al., 1999; Roeber et al., 2008; Veser et al., 2008). Some researchers have argued that VAN may index phenomenal consciousness (i.e. subjective experience) whereas LP would index reflective consciousness (i.e. conscious access to sensory information). Based on this conceptual distinction (Block, 2007; Revonsuo, 2009), they have argued that because VAN has an earlier onset than LP, and because its topography is closer to visual cortices, it may index the subjective component of awareness, before information can be manipulated, verbalised, or even accessed (Koivisto et al., 2016; Koivisto & Revonsuo, 2010; Railo et al., 2011). On the other hand, other researchers have argued that VAN may only reflect subliminal or preconscious processing rather than consciousness, given that, unlike LP, VAN responds to unseen stimuli (Sergent et al., 2005) in a linear fashion (Del Cul et al., 2007). This argument is based on the assumption that awareness is an all-or-none phenomenon rather than a continuous one (Overgaard et al., 2006; Sergent & Dehaene, 2004), hence implying that a marker of consciousness should exhibit non-linear modulation. Meanwhile, LP has shown a non-linear modulation in previous studies, leading these researchers to propose it as marker of reflective consciousness or conscious access. This second interpretation is based on the tripartite distinction found in the global neuronal workspace model of consciousness, which distinguishes between subliminal (inaccessible information), preconscious (accessible but not attended), and conscious processing (attended and reportable; Dehaene et al., 2006; Dehaene & Naccache, 2001). According to this model, purely phenomenal consciousness should not exist (Railo et al., 2011). Regardless of what aspects of consciousness VAN and LP specifically index, we found that VAN was sensitive to awareness reports with a visual exposure duration of 4.4 ms in Experiment 13 but not in Experiment 14, where it was sensitive to awareness only overall, providing partial evidence that around 4 ms of visual exposure are required to gain access to awareness. LP, on the other hand, was sensitive to awareness reports with exposure durations of 4.4 ms and 4.288 ms in Experiment 13 and Experiment 14, respectively, thus providing consistent evidence that faces require around those exposure durations of visual information to gain access to awareness.

### 5.4.4 Conclusion

In conclusion, we found that face perception occurs in a sequence of processing steps, encompassing stimulus discrimination, holistic processing and perceptual awareness, and emotion processing, all of which arise, in order, with no more than 6 milliseconds of visual exposure.

# Chapter 6

# 6 GENERAL DISCUSSION

In this thesis, I have investigated how human faces gain access to perceptual awareness. More specifically, whether faces' access to awareness can be modulated by emotional expression, orientation, and gaze. Furthermore, I examined whether face perception, and the processes it involves, arises in a unitary manner, all at once, or whether different aspects are perceived sequentially. In this thesis, my findings suggest that whilst orientation and gaze do modulate faces' access to awareness, emotional expression does not. Importantly, my findings also suggest that face perception arises through a consistent sequence of processing steps.

## 6.1 The present work: a brief overview

Researchers have claimed that emotional faces, in particular negative ones such as fearful faces, gain access to awareness faster than other expressions. This is based on the finding that fearful expressions break through masking faster than positive and neutral expressions (Capitão et al., 2014; Gray et al., 2013; Yang et al., 2007; Yang & Yeh, 2018a). Similarly, researchers have claimed that faces' access to awareness is also prioritised by eye contact, based on the finding that faces making eye contact reach awareness faster than faces looking away (Akechi et al., 2014; Chen & Yeh, 2012; Seymour et al., 2016; Stein, Senju, et al., 2011). Additionally, researchers have found that upright faces reach awareness faster than inverted faces (Jiang et al., 2007; Stein, Hebart, et al., 2011). Because turning a face upside down disrupts holistic information, researchers have suggested that holistic processing could be behind some – if not all – of these effects.

Most of the studies that led researchers to make these claims employed a specific variant of the interocular suppression technique CFS, called Breaking CFS. As discussed in Chapter 1, b-CFS presents several limitations: (1) detection tasks in b-CFS studies, often used to measure when a stimulus breaks through CFS, could be confounded by identification processes; (2) stimulus reports in b-CFS studies could be confounded by post-perceptual factors; and (3) the fact that many reported effects have failed to replicate, and some of them have recently been attributed to low-level confounds.

These issues were addressed empirically in Chapters 2 and 3. To circumvent them, I developed a CFS procedure that presents participants with face stimuli for predefined exposure durations, thereby preventing them from controlling the amount of visual information they receive. This procedure requires participants to detect on which side a face was presented and to identify their gaze direction (Chapter 2) or emotional expression (Chapter 3) with a single keypress. By measuring bias-independent perceptual sensitivity and decision criterion for detection and identification, I could disentangle detection from identification, and control in a more reliable way than before for post-perceptual factors.

However, as explained in Chapter 1, masking techniques such as backward masking and CFS may introduce a different potential confound since we do not know what specific aspects of visual processing they interrupt. Furthermore, because masks replace stimuli, they may also interact differently with different stimulus categories. Therefore, extremely brief unmasked visual presentations may offer a more reliable approach that would allow researchers to search for the minimal exposure duration required for visual perception and awareness, including what factors can modulate faces' access to awareness. However, as explained in Chapter 1, due to hardware limitations, it is extremely difficult to present visual stimuli for sufficiently brief exposure durations to assess this.

These issues were addressed empirically in Chapters 4 and 5. By using an LCD tachistoscope that allows submillisecond presentations, I presented participants with unmasked face stimuli for extremely brief exposure durations, ranging from 0.6 to 6.2 ms. By using a modified task based on the one described above (see Chapter 4), I measured how perceptual and metacognitive sensitivity increased, and whether any of these measures was modulated by face orientation or emotional expression. Unlike perceptual sensitivity, which could theoretically increase as a function of visual exposure regardless of awareness, metacognitive sensitivity is an index of awareness, as it specifically measures subjective awareness' sensitivity to performance.

Finally, to test whether neural activity may reveal evidence of holistic face processing, emotion processing, and perceptual awareness with shorter exposures than the ones indicated by signal detection analyses, I used EEG to measure neural markers of these processes.

## 6.2 Gaze processing

### 6.2.1 Summary of findings

The original b-CFS finding reported by Stein, Senju, et al. (2011), where faces making eye contact were associated with shorter breakthrough times than faces looking away, did replicate, supporting the existence of an eye-contact effect, i.e. faces making eye contact are prioritised over faces looking away in their access to perceptual awareness. However, as argued in Chapter 1, b-CFS studies suffer from important methodological limitations: they do not distinguish between detection and criterion and are vulnerable to post-perceptual factors. In a follow-up experiment using a method that controlled for those issues in more stringent conditions, I found that said advantage of eye contact was present in detection and identification sensitivity, thus indicating that the eye-contact effect just described is due to perceptual sensitivity rather than criterion differences.

However, the eye-contact effect was not disrupted by turning the face upside down, suggesting that the eye-contact effect may rely on low-level rather than high-level visual processing. This finding is in line with past b-CFS reports (Akechi et al., 2014; Seymour et al., 2016; Stein, Senju, et al., 2011).

### 6.2.2 Theoretical implications and future directions

Gaze is considered a crucial social cue; it is essential for guessing other people's intentions and actions. Gaze processing is commonly impaired in psychiatric disorders that involve social cognition limitations, such as autism (Akechi et al., 2014). Because of the relevance of gaze processing in social cognition, it could be hypothesised that the eye-contact effect relies on holistic information. However, eye-contact has been shown to be processed by a subcortical pathway that involves the superior colliculus, pulvinar, and amygdala (Senju & Johnson, 2009a), and not in the cerebral cortex, thus supporting the notion that holistic processing may not be required for its detection. Future studies should explore whether the impaired eye-contact effect in autism is due to low-level visual processing, or rather, due to other forms of integration that the face-inversion effect does not index.

## 6.3 Emotion processing

### 6.3.1 Summary of findings

I ran eleven experiments that involved emotion, described in Chapters 3, 4, and 5. Only one CFS experiment – Experiment 6 in Chapter 3 – found evidence of emotional expressions gaining prioritised access to awareness. Contrary to our expectations, I found an advantage of happy over fearful expressions. Because no other experiment found an effect of emotional expression on detection (CFS experiments) or discrimination (tachistoscope experiments), I argue that this finding may be due to low-level visual confounds or rather, given the number of experiments, a statistical fluke. To my knowledge, only Stein & Sterzer (2012) have found an advantage of happy expressions over others, in a b-CFS study. Employing schematic faces, they found that this advantage was driven by a low-level confound in the mouths' contours.

Expression identification revealed a consistent advantage of happy expressions over fearful and neutral expressions, both in CFS and tachistoscope experiments. This finding is in line with aspects that are intrinsic to facial expressions, as reported in the KDEF (Calvo & Lundqvist, 2008; Goeleven et al., 2008) and RaFD (Langner et al., 2010) stimuli norms: happy faces are more easily recognised than fearful and neutral expressions.

Importantly, the identification advantage of happy expressions over other expressions was accompanied by more liberal identification criteria associated with happy faces than the other expressions. This finding indicates, as argued in Chapters 1 and 3, that reporting different emotional expressions may involve different – and consistent – decision criteria, which stresses the importance of measuring detection and identification separately. If they are not measured separately, and identification processes confound detection processes, post-perceptual factors that affect identification (e.g. decision criterion) may also confound detection reports. While this may not have been the case in our experiments, given that we never found more liberal criterion for reporting fearful expressions – the kinds of expressions reported faster in b-CFS studies (e.g. Yang et al., 2007) – our experiments show that criterion effects are not just a theoretical possibility, and provide a stringent effort to address whether emotion modulates faces' access to awareness.

Moreover, the tachistoscope experiments indicated that emotion processing (indexed by EEG markers EPN and LPP) requires a longer minimal exposure duration to arise than holistic face processing (indexed by a behavioural face-inversion effect and EEG component N170/VPP) and awareness (indexed by SDT measure meta-d', as well as EEG markers VAN and LP), which suggests that emotion processing of faces requires face processing and awareness to arise.

### 6.3.2 Theoretical implications and future directions

This thesis produced two crucial findings regarding emotion processing of faces: that the emotional content of a face does not modulate how faces gain access to awareness, and that emotion processing probably arises after holistic face processing and awareness do.

Our experiments did not directly test whether emotion effects (should they have been found) may be due to low-level confounds. Importantly, though, our face stimuli were equated in luminance and their differences in expression identification were minimised. If the advantage effect for fearful expressions claimed in past reports (Capitão et al., 2014; Oliver et al., 2015; Yang et al., 2007; Zhan et al., 2015) is indeed due to low-level features, one should not expect to find said effect when using our stimuli. Our findings at least indirectly support the claim that emotion effects in b-CFS studies may have been due to low-level confounds (Gray et al., 2013; Hedger et al., 2015) and therefore call for a re-examination of related findings as those that psychiatric symptoms such as anxiety (Capitão et al., 2014) and depression (Sterzer et al., 2011) may prioritise certain emotional expressions' access to awareness. Further studies could explore whether those findings hold up under the better-controlled conditions of the methods described in this thesis.

If neural processing of emotion arises at longer exposure durations than holistic face processing and perceptual awareness, then emotion processing may require these two processes to occur, or rather, may be an independent process that requires a longer visual exposure to collect sufficient information to arise. Our findings support the notion that emotion processing arises after holistic face processing and perceptual awareness, but they are not sufficient to suggest that these two processes are a necessary condition for emotion processing to arise. Future studies should explore this possibility.

In addition, our findings suggest that different indices of emotion processing may actually index different processing steps within emotion processing. Above-chance expression identification arose at shorter exposure durations than neural markers of emotion processing. It is important to note, however, that an expression identification task could be done without fully engaging with emotion processing, by simply processing the visual features that are relevant to response mapping. Psychopaths, for example, can identify expressions above chance, despite exhibiting substantial limitations in that area (Hastings et al., 2008). Importantly, the neural markers of emotion processing EPN and LPP are specifically sensitive to emotion intensity and arousal (Hajcak et al., 2011), regardless of the emotional expression observed. Therefore, our findings may suggest that expression identification may not necessarily index emotion processing but might be a necessary condition for neural emotion processing of faces. Arguably, emotional expressions convey their valence and intensity through their configuration. If this is indeed the case, one would expect expression identification to arise before emotional intensity processing.

Multiple psychiatric disorders involve impairments in emotion processing and regulation. For instance, it has been shown that patients with borderline personality disorder (Fenske et al., 2015; Kaiser et al., 2019; Kleindienst et al., 2019) and social anxiety (Chen et al., 2019; Maoz et al., 2016; Park et al., 2016; Yoon & Zinbarg, 2008) perceive facial expressions in a biased way, especially when the face's expression is ambiguous. In such cases, they tend to describe the expression as more threatening and negative than healthy participants. Exploring the relationship between expression identification and emotion processing may shed light on the nature of such psychiatric ailments, specifically on whether their problems in engaging with emotional faces are driven by biased face identification or impaired emotional intensity processing.

## 6.4 Holistic face processing

### 6.4.1 Summary of findings

I found evidence of holistic face processing, indexed by the face-inversion effect, in most experiments; this consistent finding suggests that upright faces gain access to awareness faster due to their holistic configuration. This finding is consistent with previous claims that upright faces are processed holistically, whereas inversion disrupts their holistic information and therefore their recognition (Farah et al., 1995; Kanwisher et al., 1998; Rakover & Teucher, 1997; Yin, 1969; Yovel & Kanwisher, 2005). In addition, this finding is in line with b-CFS studies, which have shown that upright faces enter awareness faster than inverted faces (Akechi et al., 2015; Gayet & Stein, 2017; Jiang et al., 2007; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Hebart, et al., 2011). However, two experiments – Experiments 7 and 8 in Chapter 3 – did not find a face-inversion effect of detection. I see two possible explanations. First, it could be the case that the face-inversion effect is not as consistent among people as it is believed to be – some people may exhibit it while others do not. If this were the case, those two experiments may have had samples with an over-representation of the latter group. Secondly, it could be the case that some people manage to suppress their holistic processing when performing a face detection task. Arguably, a participant could perform a detection task by simply discriminating differences in contrast between two screen locations, whereas an identification task – because identification (recognition) requires integration – could not be performed by doing so alone. The fact that I always found a face-inversion effect for identification sensitivity supports this second explanation. Furthermore, I also found a face-inversion effect in every tachistoscope experiment that involved a face inversion manipulation. In these experiments, however, the detection tasks were based on discrimination, hence they were discrimination tasks in essence. Participants had to compare an intact face with a scrambled face. Unlike comparing a screen location containing a face with a screen location containing nothing (as in all CFS experiments except Experiment 7), the detection task used in the tachistoscope experiments was more difficult as it relied on discriminating intact faces from scrambles.

The tachistoscope experiments revealed a face-inversion effect both when stimulus discrimination relied on peripheral (Experiment 9) and foveal vision (Experiment 11), arising by around the same exposure duration. This suggests that despite differences in image resolution, face-specific holistic processes prioritised upright faces in their access to awareness similarly in both experiments. Crucially, the shortest exposure duration by which the face-inversion effect arose was also sufficiently long for the neural marker of face processing (N170/VPP) to successfully discriminate between face and non-face stimuli (Experiment 14). Taken together, these results suggest that holistic face processing does modulate how faces gain access to perceptual awareness.

I also found a face-inversion effect for a measure of metacognitive sensitivity. It is theoretically possible that holistic face processing may prioritise upright over inverted faces without impacting how faces gain access to awareness. However, the face-inversion effect found in metacognitive sensitivity suggests that holistic processing provided an advantage in perceptual awareness, too.

### 6.4.2 Theoretical implications and future directions

The overwhelming evidence for holistic face processing found in this thesis indicates that upright faces reach perceptual awareness faster than inverted faces, probably due to holistic processing. These findings replicate a long list of reports that have claimed that faces' access to awareness is prioritised thanks to holistic processing (Akechi et al., 2015; Gayet & Stein, 2017; Jiang et al., 2007; Kobylka et al., 2017; Moors, Wagemans, & de-Wit, 2016; Stein, Hebart, et al., 2011). Thus, these findings may suggest that faces enjoy unconscious processing – a frequent interpretation of the results of b-CFS studies. An alternative interpretation, however, is that holistic face processing triggers access to perceptual awareness. This latter interpretation may be supported by the fact that I found that holistic face processing and perceptual awareness arise by the same predefined exposure durations in the tachistoscope experiments. Even though the temporal co-occurrence of two processes may support such an interpretation, it is not sufficient to make a causal claim.

## 6.5 Perceptual awareness

### 6.5.1 Summary of findings

In the tachistoscope experiments (Chapters 4 and 5), perceptual awareness was measured psychophysically through metacognitive sensitivity, and neurophysiologically through the EEG markers VAN and LP. These experiments revealed that around four milliseconds of exposure are required for a non-linear increase in metacognitive sensitivity, and for VAN and LP's sensitivity to awareness reports to arise. Therefore, one could argue that 4 milliseconds of exposure convey sufficient facial information for the emergence of perceptual awareness of faces. However, it is important to note some inconsistencies in the data. First, metacognitive sensitivity exhibited a linear increase across exposure durations in three experiments (Experiments 9, 11, and 14), whereas it exhibited a non-linear increase only in one experiment (Experiment 13). And second, even though VAN exhibited overall sensitivity to subjective awareness in both EEG experiments (Experiments 13 and 14), its sensitivity did not arise by a specific exposure duration in one (Experiment 13).

Metacognitive sensitivity measures how sensitive participants' subjective experience is to objective perceptual performance. The two experiments that found a linear increase in metacognitive sensitivity with increasing exposure duration used upright and inverted faces, presented for seven equally spaced exposure durations, either peripherally (Experiment 9) or foveally (Experiment 11). On the other hand, the experiment where metacognitive sensitivity increased non-linearly used only upright faces and three exposure durations that were not equally spaced (they were selected based on results from those two previous experiments). One possible explanation for this inconsistency is that using three exposure durations, more widely spaced, created a clearer contrast between conditions' visibility. Participants may have experienced a discrete and perhaps dichotomous perception, where they felt closer to having seen the stimulus (4.4 and 6.2 ms) or not (1.7 ms). Additionally, this perhaps dichotomously experienced exposure conditions may have induced higher consistency in subjective reports – participants' awareness reports may have been more consistent per exposure duration in this experiment than in the other two experiments that used seven exposure durations. Another possible explanation for said inconsistency is that participants were more effective at describing their visual experience when the holistic information in faces was not manipulated (no orientation manipulation), in which case the non-linear increase in metacognitive sensitivity may have indexed conscious access.

VAN, as mentioned above, was sensitive to awareness, overall, in both EEG experiments (Experiment 13 and 14). However, its sensitivity to awareness arose by a specific exposure duration in only one experiment (Experiment 13). As described in Chapter 5, it may be the case that VAN was confounded by stimulus category in the second EEG experiment (Experiment 14). Because VAN has very similar topography and time window than N170, its voltage response could have been modulated by faces and objects differently, thereby diluting a possibly smaller effect of awareness.

LP, on the other hand, was consistent in its effects. It revealed in both EEG experiments a sensitivity to subjective awareness arising by around 4.4 milliseconds of visual exposure.

### 6.5.2 Theoretical implications and future directions

Our findings suggest that around four milliseconds of exposure can convey sufficient information for perceptual awareness of faces. While the required exposure duration may differ if the physical properties of the stimuli (e.g. contrast, spatial frequency) were altered, the same sequence of processing steps described above should still be found. This poses two crucial questions: does differential processing of different stimulus features imply that one can become aware of them separately, and at different exposure durations? And are the same measures of awareness capable of capturing multiple steps in which different stimulus features are processed and reach awareness sequentially?

#### *6.5.2.1 The path towards perceptual awareness: a global workspace view*

Our findings can be interpreted in the light of the Global Workspace Theory (GWT) of consciousness, originally proposed by Baars (1993, 2002, 2017) and subsequently expanded into the global neuronal workspace model (Dehaene et al., 2006, 2014; Dehaene & Naccache, 2001; Mashour et al., 2020). GWT proposes that conscious access emerges as a function of global availability of information in the brain. A stimulus gains access to awareness when it activates a set of 'central workspace' neurons in parietal, prefrontal, and cingulate cortices, which enable broadcasting to many other areas. However, a stimulus may not gain access to awareness if it is weak and hence triggers insufficient bottom-up activation, thus remaining subliminal. Conversely, a stimulus may carry sufficient bottom-up activation but lack top-down attentional amplification, thus remaining preconscious – i.e. potentially gaining access to consciousness if attended.

Multiple studies, using different experimental paradigms, have suggested that LP (i.e. the positive modulation of P3 wave) indexes conscious access (Genetti et al., 2009; Koivisto et al., 2005; Koivisto & Revonsuo, 2008a; Pourtois et al., 2006; Railo & Koivisto, 2009; Sergent et al., 2005; Veser et al., 2008), as discussed in Chapter 5. Our findings may thus indicate that around 4 milliseconds of visual exposure convey sufficient bottom-up activation for a face to stimulate 'central workspace' neurons, thereby providing faces with access to awareness. Because shorter exposure durations should then have conveyed insufficient bottom-up activation for conscious access, our findings may indicate that any perceptual process that exhibited above-chance performance with less than 4 milliseconds of visual exposure must have been performed subliminally. Both stimulus discrimination and expression identification fulfil this criterion as they exhibited very low but above-chance sensitivity at shorter exposure durations; therefore, they may have been processed subliminally.

Similarly, any perceptual process that arises with more than 4 milliseconds of visual exposure should be consciously accessible. This might be the case of emotion processing of faces, as described in Chapter 5. However, none of our experiments was adequate to test for this specific possibility. The PAS categories used asked participants how clear their visual experience of the face was. Arguably, the subjective impression of a face, as a general stimulus, may require less visual exposure to become clear than the subjective impression of an emotional expression in a face, a question that our PAS did not ask. Future studies could address this by determining the minimal exposure duration of perceptual awareness when the PAS asks about the clarity of the faces' expression instead. If conscious access requires top-down attentional modulation as predicted by GWT, then the way the PAS question is defined becomes relevant as it may direct participants' attention to different stimulus features.

While LP has been repeatedly associated with conscious access, what VAN specifically indexes is less clear. Some GWT proponents suggest that VAN indexes preconscious processing that may subsequently lead to conscious access (Del Cul et al., 2007; Sergent et al., 2005). The fact that VAN is normally found around 100 ms before

LP, and that its topography is closer to visual cortices, may support this interpretation. On the other hand, other authors suggest that VAN – and not LP specifically – indexes conscious access alongside phenomenal consciousness (Eklund & Wiens, 2018; Koivisto et al., 2016; Koivisto & Revonsuo, 2010). The tachistoscope experiments did not involve manipulations to test whether VAN and LP index the same or different aspects of awareness. However, the fact that these two neural markers showed sensitivity to subjective awareness by the same minimal exposure duration (4.4 ms; Experiment 13) suggests that this exposure duration may convey a critical amount of information for perceptual awareness of faces.

Surprisingly, VAN and LP were sensitive to subjective awareness by the same minimal exposure duration, 4.4 ms, in Experiment 13. This convergence may suggest that these two markers index the same aspects of visual awareness, or rather, that they may index different aspects of visual awareness that arise with the same exposure duration. Arguably, these two markers may be driven by predominantly different neural sources: VAN may be mainly associated with temporooccipital cortices (Koivisto et al., 2016; Shafto & Pitts, 2015; Chapter 14) whereas LP with parietocentral cortices (Babiloni et al., 2006; Railo et al., 2015; Chapter 14). Future studies should explore which neural networks underly each of these ERP components and what role they play in faces' access to awareness.

#### *6.5.2.2 The role of perceptual awareness: integration?*

Our findings indicate that indices of holistic face processing, neural face processing and perceptual awareness, are engaged with facial information by the same visual exposure duration: 4.4 ms. While this co-occurrence may simply indicate that these three processes require very similar bottom-up activation to arise, it could also suggest a causal relationship between them. For example, perceptual awareness could be a requirement for visual integration of faces, as occurs in holistic face processing. However, none of our experiments were designed to address this question.

## 6.6 Face perception: are faces processed in sequence or at once?

The tachistoscope experiments have provided a consistent picture of face perception as a process that unfolds in a sequence of processing steps: (1) stimulus discrimination, (2) holistic face processing along with (3) perceptual awareness, and (4) emotion processing. Altogether, these findings may indicate that stimulus discrimination and perhaps holistic face processing arise before perceptual awareness does. On the other hand, emotion processing may arise after faces have already reached awareness, thus suggesting that the visual information contained in faces needs to be discriminated and integrated before its emotional content is extracted. This claim is also supported by the CFS experiments (Chapter 3), which indicate that holistic face processing modulates faces' breakthrough times, whereas emotion does not.

# 7 REFERENCES


Abir, Y., Sklar, A. Y., Dotsch, R., Todorov, A., & Hassin, R. R. (2018). The determinants of consciousness of human faces. *Nature Human Behaviour*, *2*(3), 194–199. https://doi.org/10.1038/s41562-017-0266-3

Abraham, W. T., & Russell, D. W. (2008). Statistical Power Analysis in Psychological Research. *Social and Personality Psychology Compass*, *2*(1), 283–301. https://doi.org/10.1111/j.1751-9004.2007.00052.x

Agaoglu, S., Agaoglu, M. N., Breitmeyer, B., & Ogmen, H. (2015). A statistical perspective to visual masking. *Vision Research*, *115*, 23–39. https://doi.org/10.1016/j.visres.2015.07.003

Akechi, H., Stein, T., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2015). Preferential awareness of protofacial stimuli in autism. *Cognition*, *143*, 129–134. https://doi.org/10.1016/j.cognition.2015.06.016

Akechi, H., Stein, T., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2014). Absence of Preferential Unconscious Processing of Eye Contact in Adolescents With Autism Spectrum Disorder. *Autism Research*, *7*(5), 590–597. https://doi.org/10.1002/aur.1397

Alpers, G. W., & Gerdes, A. B. M. (2007). Here is looking at you: Emotional faces predominate in binocular rivalry. *Emotion*, *7*(3), 495–506. https://doi.org/10.1037/1528-3542.7.3.495

Alsius, A., & Munhall, K. G. (2013). Detection of Audiovisual Speech Correspondences Without Visual Awareness. *Psychological Science*, *24*(4), 423–431. https://doi.org/10.1177/0956797612457378

Amihai, I., Deouell, L., & Bentin, S. (2011). Conscious awareness is necessary for processing race and gender information from faces. *Consciousness and Cognition*, *20*(2), 269–279. https://doi.org/10.1016/j.concog.2010.08.004

Ashwin, C., Hietanen, J. K., & Baron-Cohen, S. (2015). Atypical integration of social cues for orienting to gaze direction in adults with autism. *Molecular Autism*, *6*(1), 5. https://doi.org/10.1186/2040-2392-6-5

Augusto, L. M. (2016). Lost in dissociation: The main paradigms in unconscious cognition. *Consciousness and Cognition*, *42*, 293–310. https://doi.org/10.1016/j.concog.2016.04.004

Axelrod, V., Bar, M., & Rees, G. (2015). Exploring the unconscious using faces. *Trends in Cognitive Sciences*, *19*(1), 35–45. https://doi.org/10.1016/j.tics.2014.11.003

Axelrod, V., & Rees, G. (2014). Conscious awareness is required for holistic face processing. *Consciousness and Cognition*, *27*, 233–245. https://doi.org/10.1016/j.concog.2014.05.004

Baars, B. J. (1993). *A Cognitive Theory of Consciousness* (Reprint edition). Cambridge University Press.

Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. https://doi.org/10.1016/S1364-6613(00)01819-2

Baars, B. J. (2017). The Global Workspace Theory of Consciousness. In *The Blackwell Companion to Consciousness* (pp. 227–242). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119132363.ch16

Babiloni, C., Vecchio, F., Miriello, M., Romani, G. L., & Rossini, P. M. (2006). Visuo-spatial Consciousness and Parieto-occipital Areas: A High-resolution EEG Study. *Cerebral Cortex*, *16*(1), 37–46. https://doi.org/10.1093/cercor/bhi082

Baden, T., Euler, T., & Berens, P. (2020). Understanding the retinal basis of vision across species. *Nature Reviews Neuroscience*, *21*(1), 5–20. https://doi.org/10.1038/s41583-019-0242-1

Baillet, S., Mosher, J. C., & Leahy, R. M. (2001). Electromagnetic brain mapping. *IEEE Signal Processing Magazine*, *18*(6), 14–30. https://doi.org/10.1109/79.962275

Balconi, M., & Mazza, G. (2009). Consciousness and emotion: ERP modulation and attentive vs. pre-attentive elaboration of emotional facial expressions by backward masking. *Motivation and Emotion*, *33*(2), 113–124. https://doi.org/10.1007/s11031-009-9122-8

Baroni, F., van Kempen, J., Kawasaki, H., Kovach, C. K., Oya, H., Howard, M. A., Adolphs, R., & Tsuchiya, N. (2017). Intracranial markers of conscious face perception in humans. *NeuroImage*, *162*, 322–343. https://doi.org/10.1016/j.neuroimage.2017.08.074

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science*, *9*(2), 207–220. https://doi.org/10.1111/j.1467-7687.2006.00480.x

Bayliss, A. P., Bartlett, J., Naughtin, C. K., & Kritikos, A. (2011). A direct link between gaze perception and social attention. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 634–644. https://doi.org/10.1037/a0020559

Bekinschtein, T. A., Shalom, D. E., Forcato, C., Herrera, M., Coleman, M. R., Manes, F. F., & Sigman, M. (2009). Classical conditioning in the vegetative and minimally conscious state. *Nature Neuroscience*, *12*(10), 1343–1349. https://doi.org/10.1038/nn.2391

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

Biderman, N., & Mudrik, L. (2018). Evidence for Implicit—But Not Unconscious—Processing of Object-Scene Relations. *Psychological Science*, *29*(2), 266–277. https://doi.org/10.1177/0956797617735745

Bishop, S. J., Jenkins, R., & Lawrence, A. D. (2007). Neural Processing of Fearful Faces: Effects of Anxiety are Gated by Perceptual Capacity Limitations. *Cerebral Cortex*, *17*(7), 1595–1603. https://doi.org/10.1093/cercor/bhl070

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13–21. https://doi.org/10.1038/nrn701

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions*, *3*(1), 1–13. https://doi.org/10.1186/1744-9081-3-7

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*(5–6), 481–499. https://doi.org/10.1017/S0140525X07002786

Block, N., Carmel, D., Fleming, S. M., Kentridge, R. W., Koch, C., Lamme, V. A. F., Lau, H., & Rosenthal, D. (2014). Consciousness science: Real progress and lingering misconceptions. *Trends in Cognitive Sciences*, *18*(11), 556–557. https://doi.org/10.1016/j.tics.2014.09.004

Boly, M., Faymonville, M.-E., Schnakers, C., Peigneux, P., Lambermont, B., Phillips, C., Lancellotti, P., Luxen, A., Lamy, M., Moonen, G., Maquet, P., & Laureys, S.

(2008). Perception of pain in the minimally conscious state with PET activation: An observational study. *The Lancet Neurology*, *7*(11), 1013–1020. https://doi.org/10.1016/S1474-4422(08)70219-9
Bradley, M. M., Hamby, S., Löw, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology*, *44*(3), 364–373. https://doi.org/10.1111/j.1469-8986.2007.00520.x
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357
Breitmeyer, B. G., & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & Psychophysics*, *62*(8), 1572–1595. https://doi.org/10.3758/BF03212157
Breitmeyer, B. G., & Öğmen, H. (2006). Methods, applications, and findings in visual pattern masking. In *Visual Masking*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780198530671.001.0001/acprof-9780198530671-chapter-2
Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a New Set of 480 Normative Photos of Objects to Be Used as Visual Stimuli in Cognitive Research. *PLOS ONE*, *5*(5), e10773. https://doi.org/10.1371/journal.pone.0010773
Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos. *PLOS ONE*, *9*(9), e106953. https://doi.org/10.1371/journal.pone.0106953
Bunge, M. (1991). A skeptic's beliefs and disbeliefs. *New Ideas in Psychology*, *9*(2), 131–149. https://doi.org/10.1016/0732-118X(91)90017-G
Burkhouse, K. L., Owens, M., Feurer, C., Sosoo, E., Kudinova, A., & Gibb, B. E. (2017). Increased neural and pupillary reactivity to emotional faces in adolescents with current and remitted major depressive disorder. *Social Cognitive and Affective Neuroscience*, *12*(5), 783–792. https://doi.org/10.1093/scan/nsw184
Busch, N. A., Fründ, I., & Herrmann, C. S. (2009). Electrophysiological Evidence for Different Types of Change Detection and Change Blindness. *Journal of Cognitive Neuroscience*, *22*(8), 1852–1869. https://doi.org/10.1162/jocn.2009.21294
Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods*, *40*(1), 109–115. https://doi.org/10.3758/BRM.40.1.109
Canales-Johnson, A., Billig, A. J., Olivares, F., Gonzalez, A., Garcia, M. del C., Silva, W., Vaucheret, E., Ciraolo, C., Mikulan, E., Ibanez, A., Huepe, D., Chennu, S., & Bekinschtein, T. A. (2020). Dissociable Neural Information Dynamics of Perceptual Integration and Differentiation during Bistable Perception. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhaa058
Capitão, L. P., Underdown, S. J. V., Vile, S., Yang, E., Harmer, C. J., & Murphy, S. E. (2014). Anxiety increases breakthrough of threat stimuli in continuous flash suppression. *Emotion*, *14*(6), 1027–1036. https://doi.org/10.1037/a0037801
Carlson, J. M., & Reinke, K. S. (2008). Masked fearful faces modulate the orienting of covert spatial attention. *Emotion*, *8*(4), 522–529. https://doi.org/10.1037/a0012653
Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*(1), 1–29. https://doi.org/10.1016/S0010-0277(01)00162-7

Caruana, N., Inkley, C., Zein, M. E., & Seymour, K. (2019). No influence of eye gaze on emotional face processing in the absence of conscious awareness. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-52728-y
Caruana, N., Stein, T., Watson, T., Williams, N., & Seymour, K. (2019). Intact prioritisation of unconscious face processing in schizophrenia. *Cognitive Neuropsychiatry*, *24*(2), 135–151. https://doi.org/10.1080/13546805.2019.1590189
Chalmers, D. J. (1997). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage*, *73*, 80–94. https://doi.org/10.1016/j.neuroimage.2013.01.054
Chen, J., Milne, K., Dayman, J., & Kemps, E. (2019). Interpretation bias and social anxiety: Does interpretation bias mediate the relationship between trait social anxiety and state anxiety responses? *Cognition and Emotion*, *33*(4), 630–645. https://doi.org/10.1080/02699931.2018.1476323
Chen, Y.-C., & Yeh, S.-L. (2012). Look into my eyes and I will see you: Unconscious processing of human gaze. *Consciousness and Cognition*, *21*(4), 1703–1710. https://doi.org/10.1016/j.concog.2012.10.001
Cioffi, F. (1985). Psychoanalysis, Pseudo-Science and Testability. In G. Currie & A. Musgrave (Eds.), *Popper and the Human Sciences* (pp. 13–44). Springer Netherlands. https://doi.org/10.1007/978-94-009-5093-1_2
Cioffi, F. (1998). *Freud and the Question of Pseudoscience*. Open Court.
Civile, C., Elchlepp, H., McLaren, R., Galang, C. M., Lavric, A., & McLaren, I. P. L. (2018). The effect of scrambling upright and inverted faces on the N170: *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/1747021817744455
Civile, C., McLaren, R. P., & McLaren, I. P. L. (2014). The Face Inversion Effect—Parts and Wholes: Individual Features and Their Configuration. *Quarterly Journal of Experimental Psychology*, *67*(4), 728–746. https://doi.org/10.1080/17470218.2013.828315
Codispoti, M., & Cesarei, A. D. (2007). Arousal and attention: Picture size and emotional reactions. *Psychophysiology*, *44*(5), 680–686. https://doi.org/10.1111/j.1469-8986.2007.00545.x
Codispoti, M., Ferrari, V., & Bradley, M. M. (2007). Repetition and Event-related Potentials: Distinguishing Early and Late Processes in Affective Picture Perception. *Journal of Cognitive Neuroscience*, *19*(4), 577–586. https://doi.org/10.1162/jocn.2007.19.4.577
Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x
Cohen, M. A., Alvarez, G. A., & Nakayama, K. (2011). Natural-Scene Perception Requires Attention. *Psychological Science*, *22*(9), 1165–1172. https://doi.org/10.1177/0956797611419168
Corallo, G., Sackur, J., Dehaene, S., & Sigman, M. (2008). Limits on Introspection: Distorted Subjective Time During the Dual-Task Bottleneck. *Psychological Science*, *19*(11), 1110–1117. https://doi.org/10.1111/j.1467-9280.2008.02211.x

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and Cognition*, *18*(2), 375–382. https://doi.org/10.1016/j.concog.2009.02.003

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, *2*, 263–275.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, *8*(2), 97–107. https://doi.org/10.1093/cercor/8.2.97

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., Laureys, S., & Owen, A. M. (2011). Bedside detection of awareness in the vegetative state: A cohort study. *The Lancet*, *378*(9809), 2088–2094. https://doi.org/10.1016/S0140-6736(11)61224-5

Curtis, W. J., & Cicchetti, D. (2011). Affective facial expression processing in young children who have experienced maltreatment during the first year of life: An event-related potential study. *Development and Psychopathology*, *23*(2), 373–395. https://doi.org/10.1017/S0954579411000125

Daury, N. (2011). Influence of Gaze Direction on Face Recognition: A Sensitive Effect. *Psychologica Belgica*, *51*(2), 139–153. https://doi.org/10.5334/pb-51-2-139

Davis, C., & Kim, J. (2011). What's in a mask? Information masking with forward and backward visual masks. *The Quarterly Journal of Experimental Psychology*, *64*(10), 1990–2002. https://doi.org/10.1080/17470218.2011.573150

Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage*, *35*(4), 1495–1501. https://doi.org/10.1016/j.neuroimage.2007.01.034

Dehaene, S. (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts* (Reprint edition). Penguin Random House USA.

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. https://doi.org/10.1016/j.tics.2006.03.007

Dehaene, S., Charles, L., King, J.-R., & Marti, S. (2014). Toward a computational theory of conscious processing. *Current Opinion in Neurobiology*, *25*, 76–84. https://doi.org/10.1016/j.conb.2013.12.005

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*(1), 1–37. https://doi.org/10.1016/S0010-0277(00)00123-2

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLOS Biology*, *5*(10), e260. https://doi.org/10.1371/journal.pbio.0050260

Descartes, R. (1641). *Meditationes de Prima Philosophia, in qua Dei existentia et animæ immortalitas demonstratur*. Unknown.

Diano, M., Celeghin, A., Bagnis, A., & Tamietto, M. (2017). Amygdala Response to Emotional Stimuli without Awareness: Facts and Interpretations. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02029

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467–477. https://doi.org/10.1038/nrn3027

Doi, H., & Shinohara, K. (2016). Emotional faces influence numerosity estimation without awareness. *Cognitive Processing*, *17*(4), 389–397. https://doi.org/10.1007/s10339-016-0774-5

Dupierrix, E., de Boisferon, A. H., Méary, D., Lee, K., Quinn, P. C., Di Giorgio, E., Simion, F., Tomonaga, M., & Pascalis, O. (2014). Preference for human eyes in human infants. *Journal of Experimental Child Psychology*, *123*, 138–146. https://doi.org/10.1016/j.jecp.2013.12.010

Edwards, S. G., & Bayliss, A. P. (2019). Seeing eye-to-eye: Social gaze interactions influence gaze direction identification. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-019-01671-1

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport*, *11*(10), 2319–2324.

Eimer, M. (2011). *The Face-Sensitive N170 Component of the Event-Related Brain Potential.* Oxford Handbook of Face Perception. https://doi.org/10.1093/oxfordhb/9780199559053.013.0017

Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, *45*(1), 15–31. https://doi.org/10.1016/j.neuropsychologia.2006.04.022

Eimer, M., & Mazza, V. (2005). Electrophysiological correlates of change detection. *Psychophysiology*, *42*(3), 328–342. https://doi.org/10.1111/j.1469-8986.2005.00285.x

Eklund, R., & Wiens, S. (2018). Visual awareness negativity is an early neural correlate of awareness: A preregistered study with two Gabor sizes. *Cognitive, Affective, & Behavioral Neuroscience*, *18*(1), 176–188. https://doi.org/10.3758/s13415-018-0562-z

Eklund, R., & Wiens, S. (2019). Auditory awareness negativity is an electrophysiological correlate of awareness in an auditory threshold task. *Consciousness and Cognition*, *71*, 70–78. https://doi.org/10.1016/j.concog.2019.03.008

Erdelyi, M. H. (1986). Experimental indeterminacies in the dissociation paradigm of subliminal perception. *Behavioral and Brain Sciences*, *9*(1), 30–31. https://doi.org/10.1017/S0140525X00021348

Esteves, F., Dimberg, U., & öhman, A. (1994). Automatically elicited fear: Conditioned skin conductance responses to masked facial expressions. *Cognition and Emotion*, *8*(5), 393–413. https://doi.org/10.1080/02699939408408949

Esteves, F., & Öhman, A. (1993). Masking the face: Recognition of emotional facial expressions as a function of the parameters of backward masking. *Scandinavian Journal of Psychology*, *34*(1), 1–18. https://doi.org/10.1111/j.1467-9450.1993.tb01096.x

Evans, K. K., & Treisman, A. (2005). Perception of Objects in Natural Scenes: Is It Really Attention Free? *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476–1492. https://doi.org/10.1037/0096-1523.31.6.1476

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A Limit to the Speed of Processing in Ultra-Rapid Visual Categorization of Novel Natural Scenes. *Journal of Cognitive Neuroscience*, *13*(2), 171–180. https://doi.org/10.1162/089892901564234

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 628–634. https://doi.org/10.1037/0096-1523.21.3.628

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fenske, S., Lis, S., Liebke, L., Niedtfeld, I., Kirsch, P., & Mier, D. (2015). Emotion recognition in borderline personality disorder: Effects of emotional information on negative bias. *Borderline Personality Disorder and Emotion Dysregulation*, *2*(1), 10. https://doi.org/10.1186/s40479-015-0031-z

Fernandez-Duque, D., Grossi, G., Thornton, I. M., & Neville, H. J. (2003). Representation of Change: Separate Electrophysiological Markers of Attention, Awareness, and Implicit Processing. *Journal of Cognitive Neuroscience*, *15*(4), 491–507. https://doi.org/10.1162/089892903321662895

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating Introspective Accuracy to Individual Differences in Brain Structure. *Science*, *329*(5998), 1541–1543. https://doi.org/10.1126/science.1191883

Foti, D., & Hajcak, G. (2008). Deconstructing Reappraisal: Descriptions Preceding Arousing Pictures Modulate the Subsequent Neural Response. *Journal of Cognitive Neuroscience*, *20*(6), 977–988. https://doi.org/10.1162/jocn.2008.20066

Fox, E. (2002). Processing emotional facial expressions: The role of anxiety and awareness. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(1), 52–63. https://doi.org/10.3758/CABN.2.1.52

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial Expressions of Emotion: Are Angry Faces Detected More Efficiently? *Cognition and Emotion*, *14*(1), 61–92. https://doi.org/10.1080/026999300378996

Freud, S. (1915). *The Unconscious*. S.E.

Freud, S. (1927). *The Ego and the Id*. Hogarth Press.

Gao, C., Conte, S., Richards, J. E., Xie, W., & Hanayik, T. (2019). The neural sources of N170: Understanding timing of activation in face-selective areas. *Psychophysiology*, *56*(6), e13336. https://doi.org/10.1111/psyp.13336

Gayet, S., Douw, I., van der Burg, V., Van der Stigchel, S., & Paffen, C. L. E. (2020). Hide and seek: Directing top-down attention is not sufficient for accelerating conscious access. *Cortex*, *122*, 235–252. https://doi.org/10.1016/j.cortex.2018.08.027

Gayet, S., Paffen, C. L. E., Belopolsky, A. V., Theeuwes, J., & Van der Stigchel, S. (2016). Visual input signaling threat gains preferential access to awareness in a breaking continuous flash suppression paradigm. *Cognition*, *149*, 77–83. https://doi.org/10.1016/j.cognition.2016.01.009

Gayet, S., & Stein, T. (2017). Between-Subject Variability in the Breaking Continuous Flash Suppression Paradigm: Potential Causes, Consequences, and Solutions. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00437

Gayet, S., Van der Stigchel, S., & Paffen, C. L. E. (2014). Breaking continuous flash suppression: Competing for consciousness on the pre-semantic battlefield. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00460

Gelbard-Sagiv, H., Faivre, N., Mudrik, L., & Koch, C. (2016). Low-level awareness accompanies “unconscious” high-level processing during continuous flash suppression. *Journal of Vision*, *16*(1), 3–3. https://doi.org/10.1167/16.1.3

Genetti, M., Khateb, A., Heinzer, S., Michel, C. M., & Pegna, A. J. (2009). Temporal dynamics of awareness for facial identity revealed with ERP. *Brain and Cognition*, *69*(2), 296–305. https://doi.org/10.1016/j.bandc.2008.08.008

Geng, H., Zhang, S., Li, Q., Tao, R., & Xu, S. (2012). Dissociations of subliminal and supraliminal self-face from other-face processing: Behavioral and ERP evidence. *Neuropsychologia*, *50*(12), 2933–2942. https://doi.org/10.1016/j.neuropsychologia.2012.07.040

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research*, *4*(2), 65–76. https://doi.org/10.1016/0926-6410(95)00045-3

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. *Cognitive Brain Research*, *24*(3), 663–673. https://doi.org/10.1016/j.cogbrainres.2005.03.017

Getov, S., Kanai, R., Bahrami, B., & Rees, G. (2015). Human brain structure predicts individual differences in preconscious evaluation of facial dominance and trustworthiness. *Social Cognitive and Affective Neuroscience*, *10*(5), 690–699. https://doi.org/10.1093/scan/nsu103

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Hughes, H. C., & Cipolli, C. (2013). Processing of invisible social cues. *Consciousness and Cognition*, *22*(3), 765–770. https://doi.org/10.1016/j.concog.2013.05.002

Gobbini, M. I., Gors, J. D., Halchenko, Y. O., Rogers, C., Guntupalli, J. S., Hughes, H., & Cipolli, C. (2013). Prioritized Detection of Personally Familiar Faces. *PLOS ONE*, *8*(6), e66620. https://doi.org/10.1371/journal.pone.0066620

Goeleven, E., Raedt, R. D., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion*, *22*(6), 1094–1118. https://doi.org/10.1080/02699930701626582

Goldstein, A. G., & Papageorge, J. (1980). Judgments of facial attractiveness in the absence of eye movements. *Bulletin of the Psychonomic Society*, *15*(4), 269–270. https://doi.org/10.3758/BF03334529

Goodrich, R. I., & Yonelinas, A. P. (2019). The effects of face inversion on perceiving- and sensing-based change detection. *Journal of Experimental Psychology: General*, No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/xge0000618

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual Following and Pattern Discrimination of Face-like Stimuli by Newborn Infants. *Pediatrics*, *56*(4), 544–549.

Goupil, L., & Bekinschtein, T. (2012). Cognitive processing during the transition to sleep. *Archives Italiennes de Biologie*, *150*(2/3), 140–154. https://doi.org/10.4449/aib.v150i2.1247

Gramfort, A., Papadopoulo, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *BioMedical Engineering OnLine*, *9*(1), 45. https://doi.org/10.1186/1475-925X-9-45

Gray, K. L. H., Adams, W. J., Hedger, N., Newton, K. E., & Garner, M. (2013). Faces and awareness: Low-level, not emotional factors determine perceptual dominance. *Emotion*, *13*(3), 537–544. https://doi.org/10.1037/a0031403

Greene, M. R., & Oliva, A. (2009). The Briefest of Glances: The Time Course of Natural Scene Understanding. *Psychological Science*. http://journals.sagepub.com/doi/10.1111/j.1467-9280.2009.02316.x

Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The Functional Neuroanatomy of Human Face Perception. *Annual Review of Vision Science*, *3*(1), 167–196. https://doi.org/10.1146/annurev-vision-102016-061214

Gruber, J., Siegel, E. H., Purcell, A. L., Earls, H. A., Cooper, G., & Barrett, L. F. (2016). Unseen positive and negative affective information influences social perception in bipolar I disorder and healthy adults. *Journal of Affective Disorders*, *192*, 191–198. https://doi.org/10.1016/j.jad.2015.12.037

Hajcak, G., Dunning, J. P., & Foti, D. (2009). Motivated and controlled attention to emotion: Time-course of the late positive potential. *Clinical Neurophysiology*, *120*(3), 505–510. https://doi.org/10.1016/j.clinph.2008.11.028

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-Related Potentials, Emotion, and Emotion Regulation: An Integrative Review. *Developmental Neuropsychology*, *35*(2), 129–155. https://doi.org/10.1080/87565640903526504

Hajcak, G., Weinberg, A., MacNamara, A., & Foti, D. (2011). ERPs and the Study of Emotion. In *The Oxford Handbook of Event-Related Potential Components*. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780195374148.001.0001/oxfordhb-9780195374148-e-016

Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V., & Hari, R. (2000). Cognitive Response Profile of the Human Fusiform Face Area as Determined by MEG. *Cerebral Cortex*, *10*(1), 69–81. https://doi.org/10.1093/cercor/10.1.69

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, *54*(6), 917–924. https://doi.org/10.1037/0022-3514.54.6.917

Hastings, M. E., Tangney, J. P., & Stuewig, J. (2008). Psychopathy and identification of facial expressions of emotion. *Personality and Individual Differences*, *44*(7), 1474–1483. https://doi.org/10.1016/j.paid.2008.01.004

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, *51*(1), 59–67. https://doi.org/10.1016/S0006-3223(01)01330-0

Hedger, N., Adams, W. J., & Garner, M. (2015). Fearful faces have a sensory advantage in the competition for awareness. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1748–1757. https://doi.org/10.1037/xhp0000127

Hedger, N., Garner, M., & Adams, W. J. (2019). Do emotional faces capture attention, and does this depend on awareness? Evidence from the visual probe paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(6), 790–802. https://doi.org/10.1037/xhp0000640

Hedger, N., Gray, K. L. H., Garner, M., & Adams, W. J. (2016). Are visual threats prioritized without awareness? A critical review and meta-analysis involving 3 behavioral paradigms and 2696 observers. *Psychological Bulletin*, *142*(9), 934–968. https://doi.org/10.1037/bul0000054

Heisz, J. J., Watter, S., & Shedden, J. M. (2006a). Progressive N170 habituation to unattended repeated faces. *Vision Research*, *46*(1), 47–56. https://doi.org/10.1016/j.visres.2005.09.028

Heisz, J. J., Watter, S., & Shedden, J. M. (2006b). Automatic face identity encoding at the N170. *Vision Research*, *46*(28), 4604–4614. https://doi.org/10.1016/j.visres.2006.09.026

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, *43*(2), 197–206. https://doi.org/10.1111/j.1469-8986.2006.00385.x

Hernández-Lorca, M., Sandberg, K., Kessel, D., Fernández-Folgueiras, U., Overgaard, M., & Carretié, L. (2019). Binocular rivalry and emotion: Implications for neural

correlates of consciousness and emotional biases in conscious perception. *Cortex*, *120*, 539–555. https://doi.org/10.1016/j.cortex.2019.08.003
Heyman, T., Maerten, A.-S., Vankrunkelsven, H., Voorspoels, W., & Moors, P. (2019). Sound-Symbolism Effects in the Absence of Awareness: A Replication Study. *Psychological Science*, *30*(11), 1638–1647. https://doi.org/10.1177/0956797619875482
Heyman, T., & Moors, P. (2014). Frequent Words Do Not Break Continuous Flash Suppression Differently from Infrequent or Nonexistent Words: Implications for Semantic Processing of Words in the Absence of Awareness. *PLOS ONE*, *9*(8), e104719. https://doi.org/10.1371/journal.pone.0104719
Hietanen, J. K., & Astikainen, P. (2013). N170 response to facial expressions is modulated by the affective congruency between the emotional expression and preceding affective picture. *Biological Psychology*, *92*(2), 114–124. https://doi.org/10.1016/j.biopsycho.2012.10.005
Hilgard, E. R. (1980). Consciousness in Contemporary Psychology. *Annual Review of Psychology*, *31*(1), 1–28. https://doi.org/10.1146/annurev.ps.31.020180.000245
Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences*, *95*(3), 781–787. https://doi.org/10.1073/pnas.95.3.781
Hixon, J. G., & Swann, W. B. (1993). When does introspection bear fruit? Self-reflection, self-insight, and interpersonal choices. *Journal of Personality and Social Psychology*, *64*(1), 35–43. https://doi.org/10.1037/0022-3514.64.1.35
Hobson, J. A. (2007, May). *States of consciousness: Normal and abnormal variation*. The Cambridge Handbook of Consciousness; Cambridge University Press. https://doi.org/10.1017/CBO9780511816789.017
Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134. https://doi.org/10.1111/1467-9280.00024
Hugdahl, K., Iversen, P. M., & Johnsen, B. H. (1993). Laterality for Facial Expressions: Does the Sex of the Subject Interact with the Sex of the Stimulus Face? *Cortex*, *29*(2), 325–331. https://doi.org/10.1016/S0010-9452(13)80185-2
Hung, S.-M., Nieh, C.-H., & Hsieh, P.-J. (2016). Unconscious processing of facial attractiveness: Invisible attractive faces orient visual attention. *Scientific Reports*, *6*(1), 1–8. https://doi.org/10.1038/srep37117
Husserl, E. (1913). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie* (1 edition). Verlag von Max Niemeyer.
Imperatori, L. S., Betta, M., Cecchetti, L., Canales-Johnson, A., Ricciardi, E., Siclari, F., Pietrini, P., Chennu, S., & Bernardi, G. (2019). EEG functional connectivity metrics wPLI and wSMI account for distinct types of brain functional interactions. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-45289-7
Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 604–610. https://doi.org/10.1037/0096-1523.7.3.604
Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews*, *33*(6), 843–863. https://doi.org/10.1016/j.neubiorev.2009.02.004
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*(2), 353–372. https://doi.org/10.1006/nimg.2001.0982
Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal Differences between Object and Face Processing Using ERPs. *Cerebral Cortex*, *14*(2), 132–142. https://doi.org/10.1093/cercor/bhg111
Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052
James, W. (1890). *The Principles of Psychology*. Henry Holt.
JASP Team. (2020). *JASP* (Version 0.12.2) [Computer software]. https://jasp-stats.org/
Jeffreys, D. A. (1993). The influence of stimulus orientation on the vertex positive scalp potential evoked by faces. *Experimental Brain Research*, *96*(1), 163–172. https://doi.org/10.1007/BF00230449
Jennings, B. J., Yu, Y., & Kingdom, F. A. A. (2017). The role of spatial frequency in emotional face classification. *Attention, Perception, & Psychophysics*, *79*(6), 1573–1577. https://doi.org/10.3758/s13414-017-1377-7
Jiang, Yi, Costello, P., & He, S. (2007). Processing of Invisible Stimuli: Advantage of Upright Faces and Recognizable Words in Overcoming Interocular Suppression. *Psychological Science*, *18*(4), 349–355. https://doi.org/10.1111/j.1467-9280.2007.01902.x
Jiang, Yi, & He, S. (2006). Cortical Responses to Invisible Faces: Dissociating Subsystems for Facial-Information Processing. *Current Biology*, *16*(20), 2023–2029. https://doi.org/10.1016/j.cub.2006.08.084
Jiang, Yi, Shannon, R. W., Vizueta, N., Bernat, E. M., Patrick, C. J., & He, S. (2009). Dynamics of processing invisible faces in the brain: Automatic neural encoding of facial expression information. *NeuroImage*, *44*(3), 1171–1177. https://doi.org/10.1016/j.neuroimage.2008.09.038
Jiang, Yunpeng, Wu, X., Saab, R., Xiao, Y., & Gao, X. (2018). Time course of influence on the allocation of attentional resources caused by unconscious fearful faces. *Neuropsychologia*, *113*, 104–110. https://doi.org/10.1016/j.neuropsychologia.2018.04.001
Jimenez, M., Grassini, S., Montoro, P. R., Luna, D., & Koivisto, M. (2018). Neural correlates of visual awareness at stimulus low vs. High-levels of processing. *Neuropsychologia*, *121*, 144–152. https://doi.org/10.1016/j.neuropsychologia.2018.11.001
Joubert, O. R., Rousselet, G. A., Fize, D., & Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Research*, *47*(26), 3286–3297. https://doi.org/10.1016/j.visres.2007.09.013
Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*(2), 175–178. https://doi.org/10.1111/1469-8986.3820175
Jusyte, A., Mayer, S. V., Künzel, E., Hautzinger, M., & Schönenberg, M. (2015). Unemotional traits predict early processing deficit for fearful expressions in young violent offenders: An investigation using continuous flash suppression. *Psychological Medicine*, *45*(2), 285–297. https://doi.org/10.1017/S0033291714001287
Kaernbach, C., Schröger, E., Jacobsen, T., & Roeber, U. (1999). Effects of consciousness on human brain waves following binocular rivalry. *NeuroReport*, *10*(4), 713–716.

Kaiser, D., Jacob, G. A., van Zutphen, L., Siep, N., Sprenger, A., Tuschen-Caffier, B., Senft, A., Arntz, A., & Domes, G. (2019). Biased Attention to Facial Expressions of Ambiguous Emotions in Borderline Personality Disorder: An Eye-Tracking Study. *Journal of Personality Disorders*, *33*(5), 671-S8. https://doi.org/10.1521/pedi_2019_33_363

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), B1–B11. https://doi.org/10.1016/S0010-0277(98)00035-3

Kihara, K., & Takeda, Y. (2019). The Role of Low-Spatial Frequency Components in the Processing of Deceptive Faces: A Study Using Artificial Face Models. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01468

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible 'invisible'. *Trends in Cognitive Sciences*, *9*(8), 381–388. https://doi.org/10.1016/j.tics.2005.06.012

King, J.-R., Sitt, J. D., Faugeras, F., Rohaut, B., El Karoui, I., Cohen, L., Naccache, L., & Dehaene, S. (2013). Information Sharing in the Brain Indexes Consciousness in Noncommunicative Patients. *Current Biology*, *23*(19), 1914–1919. https://doi.org/10.1016/j.cub.2013.07.075

Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing—An ERP study. *Biological Psychology*, *80*(1), 75–83. https://doi.org/10.1016/j.biopsycho.2008.03.004

Kleindienst, N., Hauschild, S., Liebke, L., Thome, J., Bertsch, K., Hensel, S., & Lis, S. (2019). A negative bias in decoding positive social cues characterizes emotion processing in patients with symptom-remitted Borderline Personality Disorder. *Borderline Personality Disorder and Emotion Dysregulation*, *6*(1), 17. https://doi.org/10.1186/s40479-019-0114-3

Kobylka, F., Persike, M., & Meinhardt, G. (2017a). Object Localization Does Not Imply Awareness of Object Category at the Break of Continuous Flash Suppression. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00312

Kobylka, F., Persike, M., & Meinhardt, G. (2017b). Object localization does not imply awareness of object category at the break of continuous flash suppression. *Frontiers in Human Neuroscience*, *11*.

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

Koivisto, M., & Grassini, S. (2016). Neural processing around 200ms after stimulus-onset correlates with subjective visual awareness. *Neuropsychologia*, *84*, 235–243. https://doi.org/10.1016/j.neuropsychologia.2016.02.024

Koivisto, M., Lähteenmäki, M., Sørensen, T. A., Vangkilde, S., Overgaard, M., & Revonsuo, A. (2008). The earliest electrophysiological correlate of visual awareness? *Brain and Cognition*, *66*(1), 91–103. https://doi.org/10.1016/j.bandc.2007.05.010

Koivisto, M., & Revonsuo, A. (2003). An ERP study of change detection, change blindness, and visual awareness. *Psychophysiology*, *40*(3), 423–429. https://doi.org/10.1111/1469-8986.00044

Koivisto, M., & Revonsuo, A. (2007). Electrophysiological correlates of visual consciousness and selective attention. *Neuroreport*, *18*(8), 753–756. https://doi.org/10.1097/WNR.0b013e3280c143c8

Koivisto, M., & Revonsuo, A. (2008a). Comparison of event-related potentials in attentional blink and repetition blindness. *Brain Research*, *1189*, 115–126. https://doi.org/10.1016/j.brainres.2007.10.082

Koivisto, M., & Revonsuo, A. (2008b). The role of selective attention in visual awareness of stimulus features: Electrophysiological studies. *Cognitive, Affective, & Behavioral Neuroscience*, *8*(2), 195–210. https://doi.org/10.3758/CABN.8.2.195

Koivisto, M., & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness. *Neuroscience & Biobehavioral Reviews*, *34*(6), 922–934. https://doi.org/10.1016/j.neubiorev.2009.12.002

Koivisto, M., Revonsuo, A., & Lehtonen, M. (2006). Independence of Visual Awareness from the Scope of Attention: An Electrophysiological Study. *Cerebral Cortex*, *16*(3), 415–424. https://doi.org/10.1093/cercor/bhi121

Koivisto, M., Revonsuo, A., & Salminen, N. (2005). Independence of visual awareness from attention at early processing stages. *NeuroReport*, *16*(8), 817–821.

Koivisto, M., Salminen-Vaparanta, N., Grassini, S., & Revonsuo, A. (2016). Subjective visual awareness emerges prior to P3. *European Journal of Neuroscience*, *43*(12), 1601–1611. https://doi.org/10.1111/ejn.13264

Kring, A. M., Siegel, E. H., & Barrett, L. F. (2014). Unseen Affective Faces Influence Person Perception Judgments in Schizophrenia. *Clinical Psychological Science*, *2*(4), 443–454. https://doi.org/10.1177/2167702614536161

Krysko, K. M., & Rutherford, M. D. (2009). The face in the crowd effect: Threat-detection advantage with perceptually intermediate distractors. *Visual Cognition*, *17*(8), 1205–1217. https://doi.org/10.1080/13506280902767789

Kuhn, T. S. (1977). *Essential Tension: Selected Studies in Scientific Tradition and Change*. University of Chicago Press.

Kwon, M.-K., Setoodehnia, M., Baek, J., Luck, S. J., & Oakes, L. M. (2016). The development of visual search in infancy: Attention to faces versus salience. *Developmental Psychology*, *52*(4), 537–555. https://doi.org/10.1037/dev0000080

Lacan, J. (1970). *Le Séminaire de Jacques Lacan. Livre V. Les formations de l'inconscient*. Seuil.

Ladouceur, C. D., Silk, J. S., Dahl, R. E., Ostapenko, L., Kronhaus, D. M., & Phillips, M. L. (2009). Fearful faces influence attentional control processes in anxious youth and adults. *Emotion (Washington, D.C.)*, *9*(6), 855–864. https://doi.org/10.1037/a0017747

Lamy, D., Salti, M., & Bar-Haim, Y. (2008). Neural Correlates of Subjective Awareness and Unconscious Processing: An ERP Study. *Journal of Cognitive Neuroscience*, *21*(7), 1435–1446. https://doi.org/10.1162/jocn.2009.21064

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & Knippenberg, A. van. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Langton, S. R. H., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*(1), 330–342. https://doi.org/10.1016/j.cognition.2007.07.012

Lashley, K. S. (1923). The behavioristic interpretation of consciousness. I. *Psychological Review*, *30*(4), 237–272. https://doi.org/10.1037/h0073839

Lazarov, A., Ben-Zion, Z., Shamai, D., Pine, D. S., & Bar-Haim, Y. (2018). Free viewing of sad and happy faces in depression: A potential target for attention bias modification. *Journal of Affective Disorders*, *238*, 94–100. https://doi.org/10.1016/j.jad.2018.05.047
Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *51*(3), 449–473. https://doi.org/10.1080/713755777
Leder, Helmut, & Bruce, V. (2000). When Inverted Faces are Recognized: The Role of Configural Information in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *53*(2), 513–536. https://doi.org/10.1080/713755889
Leopold, D. A., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, *124*(3), 233–251. https://doi.org/10.1037/a0019460
Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research*, *69*(1), 22–29. https://doi.org/10.1007/s00426-003-0157-2
Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596–9601. https://doi.org/10.1073/pnas.092277599
Li, Y., & Li, S. (2015). Contour integration, attentional cuing, and conscious awareness: An investigation on the processing of collinear and orthogonal contours. *Journal of Vision*, *15*(16), 10–10. https://doi.org/10.1167/15.16.10
Lin, Z., & He, S. (2009). Seeing the invisible: The scope and limits of unconscious processing in binocular rivalry. *Progress in Neurobiology*, *87*(4), 195–211. https://doi.org/10.1016/j.pneurobio.2008.09.002
Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). The many faces of research on face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1634–1637. https://doi.org/10.1098/rstb.2010.0386
Liu, J., Higuchi, M., Marantz, A., & Kanwisher, N. (2000). The selectivity of the occipitotemporal M170 for faces. *NeuroReport*, *11*(2), 337.
Liu, X., & Tanaka, J. W. (2019). Holistic perception of faces in 17 milliseconds: Evidence from three measures. *Journal of Vision*, *19*(10), 92–92. https://doi.org/10.1167/19.10.92
Liu, Y., Huang, H., McGinnis-Deweese, M., Keil, A., & Ding, M. (2012). Neural Substrate of the Late Positive Potential in Emotional Processing. *Journal of Neuroscience*, *32*(42), 14563–14572. https://doi.org/10.1523/JNEUROSCI.3109-12.2012
Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*(4853), 740–749. https://doi.org/10.1126/science.3283936
Lozano-Soldevilla, D., & VanRullen, R. (2019). The Hidden Spatial Dimension of Alpha: 10-Hz Perceptual Echoes Propagate as Periodic Traveling Waves in the Human Brain. *Cell Reports*, *26*(2), 374-380.e4. https://doi.org/10.1016/j.celrep.2018.12.058
Lu, S. T., Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Lounasmaa, O. V., Sams, M., & Vilkman, V. (1991). Seeing faces activates three separate areas outside the occipital visual cortex in man. *Neuroscience*, *43*(2), 287–290. https://doi.org/10.1016/0306-4522(91)90293-W
Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique* (second edition edition). A Bradford Book.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Macmillan, N. A., & Creelman, C. D. (2004). *Detection Theory: A User's Guide* (2 edition). Lawrence Erlbaum.

MacNamara, A., & Hajcak, G. (2010). Distinct electrocortical and behavioral evidence for increased attention to threat in generalized anxiety disorder. *Depression and Anxiety*, *27*(3), 234–243. https://doi.org/10.1002/da.20679

Madipakkam, A. R., Rothkirch, M., Dziobek, I., & Sterzer, P. (2019). Access to awareness of direct gaze is related to autistic traits. *Psychological Medicine*, *49*(6), 980–986. https://doi.org/10.1017/S0033291718001630

Madipakkam, A. R., Rothkirch, M., Guggenmos, M., Heinz, A., & Sterzer, P. (2015). Gaze Direction Modulates the Relation between Neural Responses to Faces and Visual Awareness. *Journal of Neuroscience*, *35*(39), 13287–13299. https://doi.org/10.1523/JNEUROSCI.0815-15.2015

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Maniscalco, B., & Lau, H. (2014). Signal detection theory analysis of type 1 and type 2 data: Meta-d', response-specific meta-d', and the unequal variance SDT model. In *The cognitive neuroscience of metacognition* (pp. 25–66). Springer-Verlag Publishing. https://doi.org/10.1007/978-3-642-45190-4_3

Maniscalco, B., & Lau, H. (2016). The signal processing architecture underlying subjective reports of sensory awareness. *Neuroscience of Consciousness*, *2016*(1). https://doi.org/10.1093/nc/niw002

Maoz, K., Eldar, S., Stoddard, J., Pine, D. S., Leibenluft, E., & Bar-Haim, Y. (2016). Angry-happy interpretations of ambiguous faces in social anxiety disorder. *Psychiatry Research*, *241*, 122–127. https://doi.org/10.1016/j.psychres.2016.04.100

Marti, S., Sackur, J., Sigman, M., & Dehaene, S. (2010). Mapping introspection's blind spot: Reconstruction of dual-task phenomenology using quantified introspection. *Cognition*, *115*(2), 303–313. https://doi.org/10.1016/j.cognition.2010.01.003

Mashour, G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron*, *105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

McGovern, K., & Baars, B. J. (2007, May). *Cognitive theories of consciousness*. The Cambridge Handbook of Consciousness. https://doi.org/10.1017/CBO9780511816789.009

McKone, E., & Yovel, G. (2009). Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing. *Psychonomic Bulletin & Review*, *16*(5), 778–797. https://doi.org/10.3758/PBR.16.5.778

Merigan, W. H., & Maunsell, J. H. R. (1993). How Parallel are the Primate Visual Pathways? *Annual Review of Neuroscience*, *16*(1), 369–402. https://doi.org/10.1146/annurev.ne.16.030193.002101

Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Éditions Gallimard.

Mermillod, M., Vuilleumier, P., Peyrin, C., Alleysson, D., & Marendaz, C. (2009). The importance of low spatial frequency information for recognising fearful facial

expressions. *Connection Science*, *21*(1), 75–83. https://doi.org/10.1080/09540090802213974
Milders, M., Sahraie, A., & Logan, S. (2008). Minimum presentation time for masked facial expression discrimination. *Cognition and Emotion*, *22*(1), 63–82. https://doi.org/10.1080/02699930701273849
Miller, G. A. (1962). *Psychology: The science of mental life* (pp. xi, 388). Harper & Row.
Mitsudo, T., Kamio, Y., Goto, Y., Nakashima, T., & Tobimatsu, S. (2011). Neural responses in the occipital cortex to unrecognizable faces. *Clinical Neurophysiology*, *122*(4), 708–718. https://doi.org/10.1016/j.clinph.2010.10.004
Mogg, K., Garner, M., & Bradley, B. P. (2007). Anxiety and orienting of gaze to angry and fearful faces. *Biological Psychology*, *76*(3), 163–169. https://doi.org/10.1016/j.biopsycho.2007.07.005
Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., Owen, A. M., & Laureys, S. (2010). Willful Modulation of Brain Activity in Disorders of Consciousness. *New England Journal of Medicine*, *362*(7), 579–589. https://doi.org/10.1056/NEJMoa0905370
Moors, P., Boelens, D., van Overwalle, J., & Wagemans, J. (2016). Scene Integration Without Awareness: No Conclusive Evidence for Processing Scene Congruency During Continuous Flash Suppression. *Psychological Science*, *27*(7), 945–956. https://doi.org/10.1177/0956797616642525
Moors, P., & Hesselmann, G. (2018). A critical reexamination of doing arithmetic nonconsciously. *Psychonomic Bulletin & Review*, *25*(1), 472–481. https://doi.org/10.3758/s13423-017-1292-x
Moors, P., Wagemans, J., & de-Wit, L. (2016). Faces in commonly experienced configurations enter awareness faster due to their curvature relative to fixation. *PeerJ*, *4*, e1565. https://doi.org/10.7717/peerj.1565
Moors, P., Wagemans, J., van Ee, R., & de-Wit, L. (2016). No evidence for surface organization in Kanizsa configurations during continuous flash suppression. *Attention, Perception, & Psychophysics*, *78*(3), 902–914. https://doi.org/10.3758/s13414-015-1043-x
Morris, J. S., Öhman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, *393*(6684), 467–470. https://doi.org/10.1038/30976
Mudrik, L., Breska, A., Lamy, D., & Deouell, L. Y. (2011). Integration Without Awareness: Expanding the Limits of Unconscious Processing. *Psychological Science*, *22*(6), 764–770. https://doi.org/10.1177/0956797611408736
Mundy, P., & Newell, L. (2007). Attention, Joint Attention, and Social Cognition. *Current Directions in Psychological Science*, *16*(5), 269–274. https://doi.org/10.1111/j.1467-8721.2007.00518.x
Nakamura, K., & Kawabata, H. (2018). Preferential access to awareness of attractive faces in a breaking continuous flash suppression paradigm. *Consciousness and Cognition*, *65*, 71–82. https://doi.org/10.1016/j.concog.2018.07.010
Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: Single trial analysis of concurrent EEG–fMRI. *NeuroImage*, *86*, 492–502. https://doi.org/10.1016/j.neuroimage.2013.10.047
Niedeggen, M., Wichmann, P., & Stoerig, P. (2001). Change blindness and time to consciousness. *European Journal of Neuroscience*, *14*(10), 1719–1726. https://doi.org/10.1046/j.0953-816x.2001.01785.x

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Noel, J.-P., Blanke, O., Serino, A., & Salomon, R. (2017). Interplay between Narrative and Bodily Self in Access to Consciousness: No Difference between Self- and Non-self Attributes. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00072

Nordström, H., & Wiens, S. (2012). Emotional event-related potentials are larger to figures than scenes but are similarly reduced by inattention. *BMC Neuroscience*, *13*(1), 49. https://doi.org/10.1186/1471-2202-13-49

Noreika, V., Canales-Johnson, A., Johnson, A., Arnatkevičiūtė, A., Koh, J., Chennu, S., & Bekinschtein, T. A. (2019). Wakefulness state modulates conscious access: Suppression of auditory detection in the transition to sleep. *BioRxiv*, 155705. https://doi.org/10.1101/155705

Nummenmaa, L., & Calder, A. J. (2009). Neural mechanisms of social attention. *Trends in Cognitive Sciences*, *13*(3), 135–143. https://doi.org/10.1016/j.tics.2008.12.006

Oliver, L. D., Mao, A., & Mitchell, D. G. V. (2015). "Blindsight" and subjective awareness of fearful faces: Inversion reverses the deficits in fear perception associated with core psychopathic traits. *Cognition and Emotion*, *29*(7), 1256–1277. https://doi.org/10.1080/02699931.2014.976182

Olofsson, J. K., & Polich, J. (2007). Affective visual event-related potentials: Arousal, repetition, and time-on-task. *Biological Psychology*, *75*(1), 101–108. https://doi.org/10.1016/j.biopsycho.2006.12.006

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Overgaard, M., Rote, J., Mouridsen, K., & Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during a visual task. *Consciousness and Cognition*, *15*(4), 700–708. https://doi.org/10.1016/j.concog.2006.04.002

Paffen, C. L. E., Gayet, S., Heilbron, M., & Stigchel, S. V. der. (2018). Attention-based perceptual learning does not affect access to awareness. *Journal of Vision*, *18*(3), 7–7. https://doi.org/10.1167/18.3.7

Palmer, C. J., Caruana, N., Clifford, C. W. G., & Seymour, K. J. (2018). Perceptual integration of head and eye cues to gaze direction in schizophrenia. *Royal Society Open Science*, *5*(12), 180885. https://doi.org/10.1098/rsos.180885

Park, G., Vasey, M. W., Kim, G., Hu, D. D., & Thayer, J. F. (2016). Trait Anxiety Is Associated with Negative Interpretations When Resolving Valence Ambiguity of Surprised Faces. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01164

Pereira, G. (1554). *Antoniana Margarita* (Vol. 1–2). Unknown.

Pessoa, L., & Adolphs, R. (2010). Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nature Reviews Neuroscience*, *11*(11), 773–782. https://doi.org/10.1038/nrn2920

Pessoa, L., Japee, S., & Ungerleider, L. G. (2005). Visual Awareness and the Detection of Fearful Faces. *Emotion*, *5*(2), 243–247. https://doi.org/10.1037/1528-3542.5.2.243

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion Facilitates Perception and Potentiates the Perceptual Benefits of Attention: *Psychological Science*. http://journals.sagepub.com/doi/10.1111/j.1467-9280.2006.01701.x

Piepers, D., & Robbins, R. (2012). A Review and Clarification of the Terms "holistic," "configural," and "relational" in the Face Perception Literature. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00559

Pins, D., & Ffytche, D. (2003). The Neural Correlates of Conscious Vision. *Cerebral Cortex*, *13*(5), 461–474. https://doi.org/10.1093/cercor/13.5.461

Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective Judgments of Faces Modulate Early Activity (~160 ms) within the Fusiform Gyri. *NeuroImage*, *16*(3, Part A), 663–677. https://doi.org/10.1006/nimg.2002.1126

Popper, K. (1963). *Conjectures and Refutations*. Routledge.

Pourtois, G., De Pretto, M., Hauert, C.-A., & Vuilleumier, P. (2006). Time Course of Brain Activity during Change Blindness and Change Awareness: Performance is Predicted by Neural Events before Change Onset. *Journal of Cognitive Neuroscience*, *18*(12), 2108–2129. https://doi.org/10.1162/jocn.2006.18.12.2108

Pratt, H. (2011). Sensory ERP Components. *The Oxford Handbook of Event-Related Potential Components*. https://doi.org/10.1093/oxfordhb/9780195374148.013.0050

Pratt, J. B. (1922). Behaviorism and Consciousness. *The Journal of Philosophy*, *19*(22), 596–604. JSTOR. https://doi.org/10.2307/2939512

Qiu, R., Wang, H., & Fu, S. (2017). N170 Reveals the Categorical Perception Effect of Emotional Valence. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02056

Rabagliati, H., Robertson, A., & Carmel, D. (2018). The importance of awareness for understanding language. *Journal of Experimental Psychology: General*, *147*(2), 190–208. https://doi.org/10.1037/xge0000348

Railo, H., & Koivisto, M. (2009). The electrophysiological correlates of stimulus visibility and metacontrast masking. *Consciousness and Cognition*, *18*(3), 794–803. https://doi.org/10.1016/j.concog.2009.01.006

Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious perception: A review of event-related potential correlates of visual consciousness. *Consciousness and Cognition*, *20*(3), 972–983. https://doi.org/10.1016/j.concog.2011.03.019

Railo, H., Revonsuo, A., & Koivisto, M. (2015). Behavioral and electrophysiological evidence for fast emergence of visual consciousness. *Neuroscience of Consciousness*, *2015*(1). https://doi.org/10.1093/nc/niv004

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics*, *59*(5), 752–761. https://doi.org/10.3758/BF03206021

Ramsøy, T. Z., & Overgaard, M. (2004). Introspection and subliminal perception. *Phenomenology and the Cognitive Sciences*, *3*(1), 1–23. https://doi.org/10.1023/B:PHEN.0000041900.30172.e8

Reddy, L., Wilken, P., & Koch, C. (2004). Face-gender discrimination is possible in the near-absence of attention. *Journal of Vision*, *4*(2), 4–4. https://doi.org/10.1167/4.2.4

Revonsuo, A. (2009). *Inner Presence: Consciousness as a Biological Phenomenon*. MIT Press.

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, *49*(23), 2856–2861. https://doi.org/10.1016/j.visres.2009.08.025

Roeber, U., Widmann, A., Trujillo-Barreto, N. J., Herrmann, C. S., O'Shea, R. P., & Schröger, E. (2008). Early correlates of visual awareness in the human brain: Time and place from event-related brain potentials. *Journal of Vision*, *8*(3), 21–21. https://doi.org/10.1167/8.3.21

Rossion, B., Gauthier, I., Tarr, M., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuroreport*, *11*(1), 69–72.

Rossion, Bruno. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139–253. https://doi.org/10.1080/13506285.2013.772929

Rossion, Bruno, & Jacques, C. (2012). The N170: Understanding the time course of face perception in the human brain. In *The Oxford handbook of event-related potential components* (pp. 115–141). Oxford University Press.

Rossion, Bruno, 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*(3), 1609–1624. https://doi.org/10.1016/j.neuroimage.2003.07.010

Rousselet, G., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition*, *12*(6), 852–877. https://doi.org/10.1080/13506280444000553

Sabatinelli, D., Lang, P. J., Keil, A., & Bradley, M. M. (2007). Emotional Perception: Correlation of Functional MRI and Event-Related Potentials. *Cerebral Cortex*, *17*(5), 1085–1091. https://doi.org/10.1093/cercor/bhl017

Salomon, R., Lim, M., Herbelin, B., Hesselmann, G., & Blanke, O. (2013). Posing for awareness: Proprioception modulates access to visual consciousness in a continuous flash suppression task. *Journal of Vision*, *13*(7), 2–2. https://doi.org/10.1167/13.7.2

Salti, M., Bar-Haim, Y., & Lamy, D. (2012). The P3 component of the ERP reflects conscious perception, not confidence. *Consciousness and Cognition*, *21*(2), 961–968. https://doi.org/10.1016/j.concog.2012.01.012

Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J., & Lounasmaa, O. V. (1997). Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience*, *77*(1), 49–55. https://doi.org/10.1016/S0306-4522(96)00419-8

Sandberg, K., & Overgaard, M. (2015). Using the perceptual awareness scale (PAS). In *Behavioral methods in consciousness research*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199688890.001.0001/acprof-9780199688890-chapter-11

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? *Consciousness and Cognition*, *19*(4), 1069–1078. https://doi.org/10.1016/j.concog.2009.12.013

Sato, W., Kochiyama, T., Minemoto, K., Sawada, R., & Fushiki, T. (2019). Amygdala activation during unconscious visual processing of food. *Scientific Reports*, *9*(1), 1–14. https://doi.org/10.1038/s41598-019-43733-2

Sato, W., Okada, T., & Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. *Experimental Brain Research*, *183*(1), 87–94. https://doi.org/10.1007/s00221-007-1025-x

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. https://doi.org/10.1038/253265a0

Schankin, A., & Wascher, E. (2007). Electrophysiological correlates of stimulus processing in change blindness. *Experimental Brain Research*, *183*(1), 95–105. https://doi.org/10.1007/s00221-007-1023-z

Schlossmacher, I., Junghöfer, M., Straube, T., & Bruchmann, M. (2017). No differential effects to facial expressions under continuous flash suppression: An event-related potentials study. *NeuroImage*, *163*, 276–285. https://doi.org/10.1016/j.neuroimage.2017.09.034

Schmidt, T., & Vorberg, D. (2006). Criteria for unconscious cognition: Three types of dissociation. *Perception & Psychophysics*, *68*(3), 489–504. https://doi.org/10.3758/BF03193692

Schupp, H., Junghöfer, M., Weike, A., & Hamm, A. (2003). Attention and emotion: An ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, *14*(8), 1107–1110.

Schweinberger, S. R., Pickering, E. C., Jentzsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Cognitive Brain Research*, *14*(3), 398–409. https://doi.org/10.1016/S0926-6410(02)00142-8

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial–relational information in faces. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(4), 904–915. https://doi.org/10.1037/0096-1523.22.4.904

Searle, J. R. (1992). *The Rediscovery of the Mind*. MIT Press.

Senju, A., & Hasegawa, T. (2005). Direct gaze captures visuospatial attention. *Visual Cognition*, *12*(1), 127–144. https://doi.org/10.1080/13506280444000157

Senju, A., & Johnson, M. H. (2009a). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. https://doi.org/10.1016/j.tics.2008.11.009

Senju, A., & Johnson, M. H. (2009b). Atypical eye contact in autism: Models, mechanisms and development. *Neuroscience & Biobehavioral Reviews*, *33*(8), 1204–1214. https://doi.org/10.1016/j.neubiorev.2009.06.001

Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, *8*(10), 1391–1400. https://doi.org/10.1038/nn1549

Sergent, C., & Dehaene, S. (2004). Is Consciousness a Gradual Phenomenon?: Evidence for an All-or-None Bifurcation During the Attentional Blink. *Psychological Science*, *15*(11), 720–728. https://doi.org/10.1111/j.0956-7976.2004.00748.x

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioural and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*(8), 314–321. https://doi.org/10.1016/j.tics.2008.04.008

Seymour, K., Rhodes, G., Stein, T., & Langdon, R. (2016). Intact unconscious processing of eye contact in schizophrenia. *Schizophrenia Research: Cognition*, *3*, 15–19. https://doi.org/10.1016/j.scog.2015.11.001

Shafto, J. P., & Pitts, M. A. (2015). Neural Signatures of Conscious Face Perception in an Inattentional Blindness Paradigm. *Journal of Neuroscience*, *35*(31), 10940–10948. https://doi.org/10.1523/JNEUROSCI.0145-15.2015

Shanks, D. R. (2017). Regressive research: The pitfalls of post hoc data selection in the study of unconscious mental processes. *Psychonomic Bulletin & Review*, *24*(3), 752–775. https://doi.org/10.3758/s13423-016-1170-y

Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S., & Ono, T. (2002). Generators of Visual Evoked Potentials for Faces and Eyes in the Human Brain as Determined by Dipole Localization. *Brain Topography*, *15*(1), 51–63. https://doi.org/10.1023/A:1019944607316

Shimojo, S., Kamitani, Y., & Nishida, S. (2001). Afterimage of Perceptually Filled-in Surface. *Science*, *293*(5535), 1677–1680. https://doi.org/10.1126/science.1060161

Sitt, J. D., King, J.-R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., & Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain*, *137*(8), 2258–2270. https://doi.org/10.1093/brain/awu141

Skinner, B. F. (1974). *About behaviorism*. Knopf.

Sklar, A. Y., Levy, N., Goldstein, A., Mandel, R., Maril, A., & Hassin, R. R. (2012). Reading and doing arithmetic nonconsciously. *Proceedings of the National Academy of Sciences*, *109*(48), 19614–19619. https://doi.org/10.1073/pnas.1211645109

Smith, F. W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLOS ONE*, *13*(5), e0197160. https://doi.org/10.1371/journal.pone.0197160

Snodgrass, M. (2004). The dissociation paradigm and its discontents: How can unconscious perception or memory be inferred? *Consciousness and Cognition*, *13*(1), 107–116. https://doi.org/10.1016/j.concog.2003.11.001

Song, C., & Yao, H. (2016). Unconscious processing of invisible visual stimuli. *Scientific Reports*, *6*(1), 1–6. https://doi.org/10.1038/srep38917

Sperdin, H. F., Repnow, M., Herzog, M. H., & Landis, T. (2013). An LCD tachistoscope with submillisecond precision. *Behavior Research Methods*, *45*(4), 1347–1357. https://doi.org/10.3758/s13428-012-0311-0

Stein, T. (2019). The Breaking Continuous Flash Suppression paradigm. In G. Hesselmann, *Transitions between consciousness and unconsciousness* (pp. 1–38). Routledge.

Stein, T., Awad, D., Gayet, S., & Peelen, M. V. (2018). Unconscious processing of facial dominance: The role of low-level factors in access to awareness. *Journal of Experimental Psychology: General*, *147*(11), e1–e13. https://doi.org/10.1037/xge0000521

Stein, T., End, A., & Sterzer, P. (2014). Own-race and own-age biases facilitate visual awareness of faces under interocular suppression. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00582

Stein, T., Grubb, C., Bertrand, M., Suh, S. M., & Verosky, S. C. (2017). No impact of affective person knowledge on visual awareness: Evidence from binocular rivalry and continuous flash suppression. *Emotion*, *17*(8), 1199–1207. https://doi.org/10.1037/emo0000305

Stein, T., Hebart, M. N., & Sterzer, P. (2011). Breaking Continuous Flash Suppression: A New Measure of Unconscious Processing during Interocular Suppression? *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00167

Stein, T., Reeder, R. R., & Peelen, M. V. (2016). Privileged access to awareness for faces and objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 788–798. https://doi.org/10.1037/xhp0000188

Stein, T., Senju, A., Peelen, M. V., & Sterzer, P. (2011). Eye contact facilitates awareness of faces during interocular suppression. *Cognition*, *119*(2), 307–311. https://doi.org/10.1016/j.cognition.2011.01.008

Stein, T., Seymour, K., Hebart, M. N., & Sterzer, P. (2014). Rapid Fear Detection Relies on High Spatial Frequencies. *Psychological Science*, *25*(2), 566–574. https://doi.org/10.1177/0956797613512509

Stein, T., & Sterzer, P. (2012a). Not just another face in the crowd: Detecting emotional schematic faces during continuous flash suppression. *Emotion*, *12*(5), 988–996. https://doi.org/10.1037/a0026944

Stein, T., & Sterzer, P. (2012b). Not just another face in the crowd: Detecting emotional schematic faces during continuous flash suppression. *Emotion*, *12*(5), 988–996. https://doi.org/10.1037/a0026944

Stein, T., Sterzer, P., & Peelen, M. V. (2012). Privileged detection of conspecifics: Evidence from inversion effects during continuous flash suppression. *Cognition*, *125*(1), 64–79. https://doi.org/10.1016/j.cognition.2012.06.005

Sterzer, P., Hilgenfeldt, T., Freudenberg, P., Bermpohl, F., & Adli, M. (2011). Access of emotional information to visual awareness in patients with major depressive disorder. *Psychological Medicine*, *41*(8), 1615–1624. https://doi.org/10.1017/S0033291710002540

Sterzer, Philipp, Jalkanen, L., & Rees, G. (2009). Electromagnetic responses to invisible face stimuli during binocular suppression. *NeuroImage*, *46*(3), 803–808. https://doi.org/10.1016/j.neuroimage.2009.02.046

Stevner, A. B. A., Vidaurre, D., Cabral, J., Rapuano, K., Nielsen, S. F. V., Tagliazucchi, E., Laufs, H., Vuust, P., Deco, G., Woolrich, M. W., Van Someren, E., & Kringelbach, M. L. (2019). Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep. *Nature Communications*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-08934-3

Stewart, L. H., Ajina, S., Getov, S., Bahrami, B., Todorov, A., & Rees, G. (2012). Unconscious evaluation of faces on social dimensions. *Journal of Experimental Psychology: General*, *141*(4), 715–727. https://doi.org/10.1037/a0027950

Stirrat, M., & Perrett, D. I. (2010). Valid Facial Cues to Cooperation and Trust: Male Facial Width and Trustworthiness. *Psychological Science*, *21*(3), 349–354. https://doi.org/10.1177/0956797610362647

Svard, J., Wiens, S., & Fischer, H. (2012). Superior Recognition Performance for Happy Masked and Unmasked Faces in Both Younger and Older Adults. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00520

Sylvers, P. D., Brennan, P. A., & Lilienfeld, S. O. (2011). Psychopathic Traits and Preattentive Threat Processing in Children: A Novel Test of the Fearlessness Hypothesis. *Psychological Science*, *22*(10), 1280–1287. https://doi.org/10.1177/0956797611420730

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). *Brainstorm: A User-Friendly Application for MEG/EEG Analysis* [Research Article]. Computational Intelligence and Neuroscience. https://doi.org/10.1155/2011/879716

Tamietto, M., & de Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience*, *11*(10), 697–709. https://doi.org/10.1038/nrn2889

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), 225–245. https://doi.org/10.1080/14640749308401045

Tanaka, J. W., & Gordon, I. (2011). *Features, Configuration, and Holistic Face Processing*. Oxford Handbook of Face Perception. https://doi.org/10.1093/oxfordhb/9780199559053.013.0010

Tanaka, J. W., & Simonyi, D. (2016). The "Parts and Wholes" of Face Recognition: A Review of the Literature: *Quarterly Journal of Experimental Psychology*. https://journals.sagepub.com/doi/10.1080/17470218.2016.1146780

Tanskanen, T., Näsänen, R., Montez, T., Päällysaho, J., & Hari, R. (2005). Face Recognition and Cortical Responses Show Similar Sensitivity to Noise Spatial Frequency. *Cerebral Cortex*, *15*(5), 526–534. https://doi.org/10.1093/cercor/bhh152

The jamovi project. (2020). *Jamovi* (Version 1.2.17) [Computer software]. Retrieved from https://www.jamovi.org

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Tian, J., Wang, J., Xia, T., Zhao, W., Xu, Q., & He, W. (2018). The influence of spatial frequency content on facial expression processing: An ERP study using rapid serial visual presentation. *Scientific Reports*, *8*(1), 1–8. https://doi.org/10.1038/s41598-018-20467-1

Tian, Y., Zhang, H., Pang, Y., & Lin, J. (2018). Classification for Single-Trial N170 During Responding to Facial Picture With Emotion. *Frontiers in Computational Neuroscience*, *12*. https://doi.org/10.3389/fncom.2018.00068

Todd, R. M., Lewis, M. D., Meusel, L.-A., & Zelazo, P. D. (2008). The time course of social-emotional processing in early childhood: ERP responses to facial affect and personal familiarity in a Go-Nogo task. *Neuropsychologia*, *46*(2), 595–613. https://doi.org/10.1016/j.neuropsychologia.2007.10.011

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience*, *8*(8), 1096. https://doi.org/10.1038/nn1500

Tsuchiya, N., Moradi, F., Felsen, C., Yamazaki, M., & Adolphs, R. (2009). Intact rapid detection of fearful faces in the absence of the amygdala. *Nature Neuroscience*, *12*(10), 1224–1225. https://doi.org/10.1038/nn.2380

Turatto, M., Angrilli, A., Mazza, V., Umiltà, C., & Driver, J. (2002). Looking without seeing the background change: Electrophysiological correlates of change detection versus change blindness. *Cognition*, *84*(1), B1–B10. https://doi.org/10.1016/S0010-0277(02)00016-1

Vadillo, M. A., Konstantinidis, E., & Shanks, D. R. (2016). Underpowered samples, false negatives, and unconscious learning. *Psychonomic Bulletin & Review*, *23*(1), 87–102. https://doi.org/10.3758/s13423-015-0892-6

Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica*, *61*(3), 259–273. https://doi.org/10.1016/0001-6918(86)90085-5

van Doorn, J., van den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Raj, A., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers,

E.-J. (2019). *The JASP Guidelines for Conducting and Reporting a Bayesian Analysis* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/yqxfr
VanRullen, R., & Macdonald, J. S. P. (2012). Perceptual Echoes at 10 Hz in the Human Brain. *Current Biology*, *22*(11), 995–999. https://doi.org/10.1016/j.cub.2012.03.050
Vermeulen, T., Lauwers, T., Van Diermen, L., Sabbe, B. G., van der Mast, R. C., & Giltay, E. J. (2019). Cognitive Deficits in Older Adults With Psychotic Depression: A Meta-Analysis. *The American Journal of Geriatric Psychiatry*, *27*(12), 1334–1344. https://doi.org/10.1016/j.jagp.2019.07.011
Veser, S., O'Shea, R. P., Schröger, E., Trujillo-Barreto, N. J., & Roeber, U. (2008). Early correlates of visual awareness following orientation and colour rivalry. *Vision Research*, *48*(22), 2359–2369. https://doi.org/10.1016/j.visres.2008.07.024
Vetter, P., Badde, S., Phelps, E. A., & Carrasco, M. (2019). Emotional faces guide the eyes in the absence of awareness. *ELife*, *8*, e43467. https://doi.org/10.7554/eLife.43467
Vieira, J. B., Wen, S., Oliver, L. D., & Mitchell, D. G. V. (2017). Enhanced conscious processing and blindsight-like detection of fear-conditioned stimuli under continuous flash suppression. *Experimental Brain Research*, *235*(11), 3333–3344. https://doi.org/10.1007/s00221-017-5064-7
Vizueta, N., Patrick, C. J., Jiang, Y., Thomas, K. M., & He, S. (2012). Dispositional fear, negative affectivity, and neuroimaging response to visually suppressed emotional faces. *NeuroImage*, *59*(1), 761–771. https://doi.org/10.1016/j.neuroimage.2011.07.015
Vrancken, L., Germeys, F., & Verfaillie, K. (2017). Holistic integration of gaze cues in visual face and body perception: Evidence from the composite design. *Journal of Vision*, *17*(1), 24–24. https://doi.org/10.1167/17.1.24
Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*(6), 624–631. https://doi.org/10.1038/nn1057
Wang, L., Weng, X., & He, S. (2012). Perceptual Grouping without Awareness: Superiority of Kanizsa Triangle in Breaking Interocular Suppression. *PLOS ONE*, *7*(6), e40106. https://doi.org/10.1371/journal.pone.0040106
Watanabe, S., Kakigi, R., Koyama, S., & Kirino, E. (1999). Human face perception traced by magneto- and electro-encephalography. *Cognitive Brain Research*, *8*(2), 125–142. https://doi.org/10.1016/S0926-6410(99)00013-0
Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, *114*(4), 522–536. https://doi.org/10.1037/0021-843X.114.4.522
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158–177. https://doi.org/10.1037/h0074428
Weinberg, A., & Hajcak, G. (2010). Beyond good and evil: The time-course of neural activity elicited by specific picture content. *Emotion*, *10*(6), 767–782. https://doi.org/10.1037/a0020242
Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). VISUAL CAPACITY IN THE HEMIANOPIC FIELD FOLLOWING A RESTRICTED OCCIPITAL ABLATION. *Brain*, *97*(1), 709–728. https://doi.org/10.1093/brain/97.1.709

Wells, L. J., Gillespie, S. M., & Rotshtein, P. (2016). Identification of Emotional Facial Expressions: Effects of Expression, Intensity, and Sex on Eye Gaze. *PLOS ONE*, *11*(12), e0168307. https://doi.org/10.1371/journal.pone.0168307

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked Presentations of Emotional Facial Expressions Modulate Amygdala Activity without Explicit Knowledge. *Journal of Neuroscience*, *18*(1), 411–418. https://doi.org/10.1523/JNEUROSCI.18-01-00411.1998

Wierzchoń, M., Wronka, E., Paulewicz, B., & Szczepanowski, R. (2016). Post-Decision Wagering Affects Metacognitive Awareness of Emotional Stimuli: An Event Related Potential Study. *PLOS ONE*, *11*(8), e0159516. https://doi.org/10.1371/journal.pone.0159516

Wilenius, M. E., & Revonsuo, A. T. (2007). Timing of the earliest ERP correlate of visual awareness. *Psychophysiology*, *44*(5), 703–710. https://doi.org/10.1111/j.1469-8986.2007.00546.x

Wilenius-Emet, M., Revonsuo, A., & Ojanen, V. (2004). An electrophysiological correlate of human visual awareness. *Neuroscience Letters*, *354*(1), 38–41. https://doi.org/10.1016/j.neulet.2003.09.060

Willenbockel, V., Lepore, F., Nguyen, D. K., Bouthillier, A., & Gosselin, F. (2012). Spatial Frequency Tuning during the Conscious and Non-Conscious Perception of Emotional Facial Expressions – An Intracranial ERP Study. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00237

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Wilson, E., & MacLeod, C. (2003). Contrasting two accounts of anxiety-linked attentional bias: Selective attention to varying levels of stimulus threat intensity. *Journal of Abnormal Psychology*, *112*(2), 212–218. https://doi.org/10.1037/0021-843X.112.2.212

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*(1), 117–142. https://doi.org/10.1037/0033-2909.116.1.117

Wilson, T. D., & Dunn, E. W. (2004). Self-Knowledge: Its Limits, Value, and Potential for Improvement. *Annual Review of Psychology*, *55*(1), 493–518. https://doi.org/10.1146/annurev.psych.55.090902.141954

Windey, B., & Cleeremans, A. (2015). Consciousness as a graded and an all-or-none phenomenon: A conceptual analysis. *Consciousness and Cognition*, *35*, 185–191. https://doi.org/10.1016/j.concog.2015.03.002

Wormwood, J. B., Siegel, E. H., Kopec, J., Quigley, K. S., & Barrett, L. F. (2019). You are what I feel: A test of the affective realism hypothesis. *Emotion*, *19*(5), 788–798. https://doi.org/10.1037/emo0000484

Xu, S., Zhang, S., & Geng, H. (2011). Gaze-induced joint attention persists under high perceptual load and does not depend on awareness. *Vision Research*, *51*(18), 2048–2056. https://doi.org/10.1016/j.visres.2011.07.023

Yang, E., & Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of Vision*, *12*(3), 8–8. https://doi.org/10.1167/12.3.8

Yang, E., Brascamp, J., Kang, M.-S., & Blake, R. (2014). On the use of continuous flash suppression for the study of visual processing outside of awareness. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00724

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*(4), 882–886. https://doi.org/10.1037/1528-3542.7.4.882

Yang, Y.-H., & Yeh, S.-L. (2011). Accessing the meaning of invisible words. *Consciousness and Cognition*, *20*(2), 223–233. https://doi.org/10.1016/j.concog.2010.07.005

Yang, Y.-H., & Yeh, S.-L. (2018a). Can emotional content be extracted under interocular suppression? *PLOS ONE*, *13*(11), e0206799. https://doi.org/10.1371/journal.pone.0206799

Yang, Y.-H., & Yeh, S.-L. (2018b). Unconscious processing of facial expression as revealed by affective priming under continuous flash suppression. *Psychonomic Bulletin & Review*, *25*(6), 2215–2223. https://doi.org/10.3758/s13423-018-1437-6

Ye, X., He, S., Hu, Y., Yu, Y. Q., & Wang, K. (2014). Interference between Conscious and Unconscious Facial Expression Information. *PLOS ONE*, *9*(8), e105156. https://doi.org/10.1371/journal.pone.0105156

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. https://doi.org/10.1037/h0027474

Yokoyama, T., Noguchi, Y., & Kita, S. (2013). Unconscious processing of direct gaze: Evidence from an ERP study. *Neuropsychologia*, *51*(7), 1161–1168. https://doi.org/10.1016/j.neuropsychologia.2013.04.002

Yoon, K. L., & Zinbarg, R. E. (2008). Interpreting neutral faces as threatening is a default mode for socially anxious individuals. *Journal of Abnormal Psychology*, *117*(3), 680–685. https://doi.org/10.1037/0021-843X.117.3.680

Yovel, G., & Kanwisher, N. (2004). Face Perception: Domain Specific, Not Process Specific. *Neuron*, *44*(5), 889–898. https://doi.org/10.1016/j.neuron.2004.11.018

Yovel, G., & Kanwisher, N. (2005). The Neural Basis of the Behavioral Face-Inversion Effect. *Current Biology*, *15*(24), 2256–2262. https://doi.org/10.1016/j.cub.2005.10.072

Yovel, G., Sadeh, B., Podlipsky, I., Hendler, T., & Zhdanov, A. (2008). The face-selective ERP component (N170) is correlated with the face-selective areas in the fusiform gyrus (FFA) and the superior temporal sulcus (fSTS) but not the occipital face area (OFA): A simultaneous fMRI-EEG study. *Journal of Vision*, *8*(6), 401–401. https://doi.org/10.1167/8.6.401

Yuan, J., Hu, X., Chen, J., Bodenhausen, G. V., & Fu, S. (2019). One of us? How facial and symbolic cues to own- versus other-race membership influence access to perceptual awareness. *Cognition*, *184*, 19–27. https://doi.org/10.1016/j.cognition.2018.12.003

Yuan, J., Hu, X., Lu, Y., Bodenhausen, G. V., & Fu, S. (2017). Invisible own- and other-race faces presented under continuous flash suppression produce affective response biases. *Consciousness and Cognition*, *48*, 273–282. https://doi.org/10.1016/j.concog.2016.12.012

Zaki, J. (2013). Cue Integration: A Common Framework for Social Cognition and Physical Perception. *Perspectives on Psychological Science*, *8*(3), 296–312. https://doi.org/10.1177/1745691613475454

Zhan, M., Hortensius, R., & Gelder, B. de. (2015). The Body as a Tool for Anger Awareness—Differential Effects of Angry Facial and Bodily Expressions on Suppression from Awareness. *PLOS ONE*, *10*(10), e0139768. https://doi.org/10.1371/journal.pone.0139768

Zhou, G., Zhang, L., Liu, J., Yang, J., & Qu, Z. (2010). Specificity of face processing without awareness. *Consciousness and Cognition*, *19*(1), 408–412. https://doi.org/10.1016/j.concog.2009.12.009

# 8 APPENDICES

**Appendix A.** Experiment 2 Pre-registration

ASPREDICTED

## Unconscious detection of eye gaze: a Signal Detection study (#23394)

**Created:** 05/14/2019 01:52 AM (PT)
**Public:** 04/12/2020 10:32 AM (PT)

**Author(s)**
Renzo Lanfranco (University of Edinburgh) - Renzo.Lanfranco@ed.ac.uk
Timo Stein (University of Amsterdam) - T.Stein@uva.nl
Hugh Rabagliati (University of Edinburgh) - Hugh.Rabagliati@ed.ac.uk
David Carmel (Victoria University of Wellington) - david.carmel@vuw.ac.nz

**1) Have any data been collected for this study already?**
No, no data have been collected for this study yet.

**2) What's the main question being asked or hypothesis being tested in this study?**
It has been shown in the past that people have a visual perception preference for direct-gaze faces than for averted-gaze faces. This finding has been replicated when suppressing faces with continuous flash suppression, a masking technique, which has been tested by measuring detection reaction times. However, detection is a process that combines perceptual sensitivity and decision criterion.
Main questions: Are people more sensitive to direct-gaze faces than averted-gaze faces? Can previous detection-related reaction times be explained by differences in decision criterion?
Hypothesis: If previous findings are due to decision criterion, then we should find no differences in perceptual sensitivity for direct-gaze and averted-gaze faces along with a significantly more liberal decision criterion for direct-gaze faces.

**3) Describe the key dependent variable(s) specifying how they will be measured.**
The main dependent variables are location sensitivity (d') and gaze orientation decision criterion (C). The former is a measure of sensitivity for side detection as defined by Signal Detection Theory. The latter, a measure of bias in responses for different eye-gaze faces.

**4) How many and which conditions will participants be assigned to?**
There will be 28 within-subject conditions divided into 3 factors: face orientation (upright or inverted), gaze orientation (direct or averted), and predefined exposure times (7).

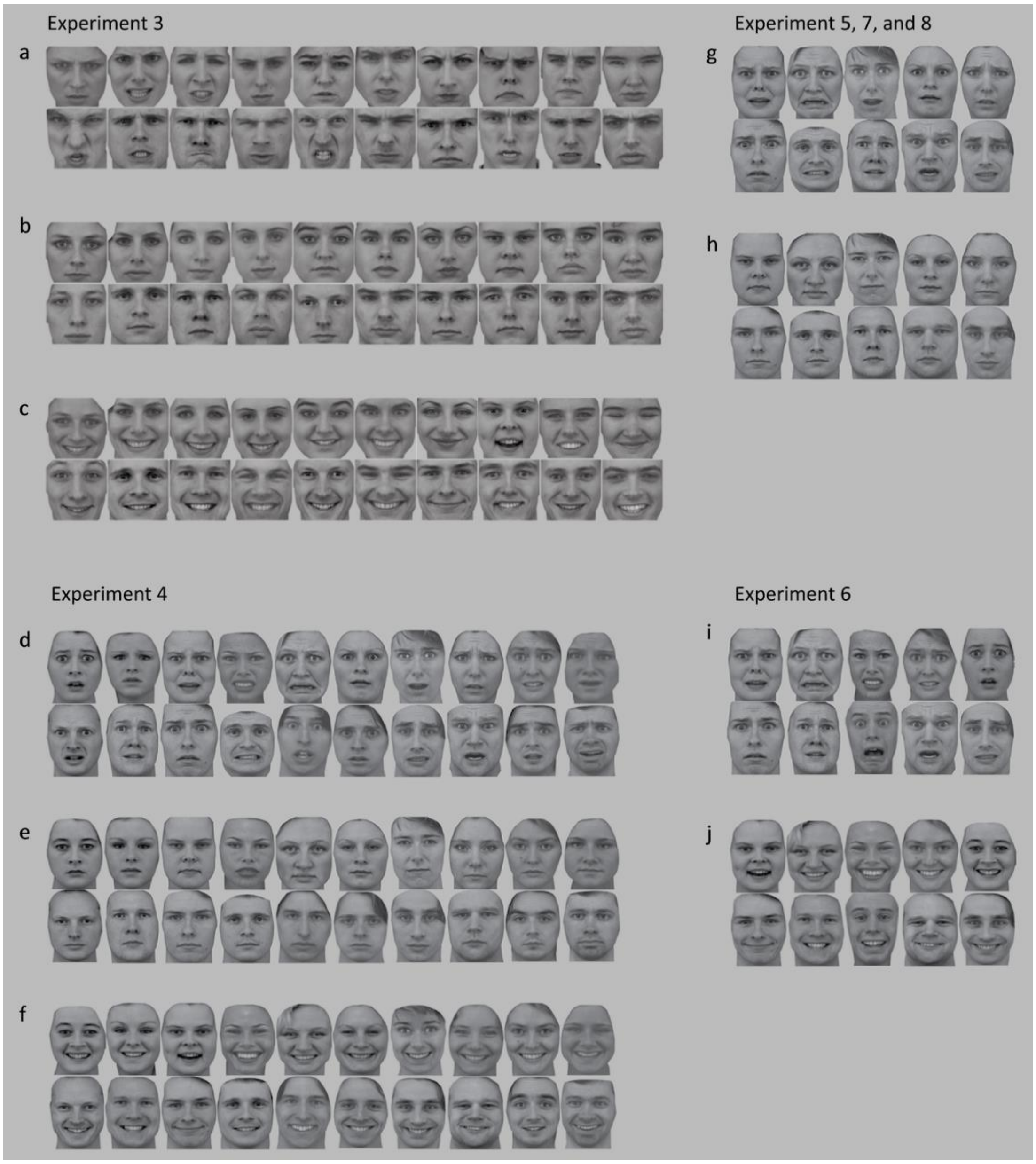
**5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.**
1. Independent repeated-measures ANOVAs will be run to analyse (1) location sensitivity (d'), gaze orientation sensitivity (d'), response bias, and eye gaze orientation criterion (C).
2. Bayes factors will be estimated whenever null effects are found in order to test whether the data is better explained by the null hypothesis model or the alternative hypothesis model.

**6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.**
1. Participants whose responses are random will be excluded, i.e. whenever there is no significant effect of exposure time for location sensitivity.
2. Participants who present more than 10% than no-response trials.

**7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.**
We aim for a sample size of 32 participants per experiment.

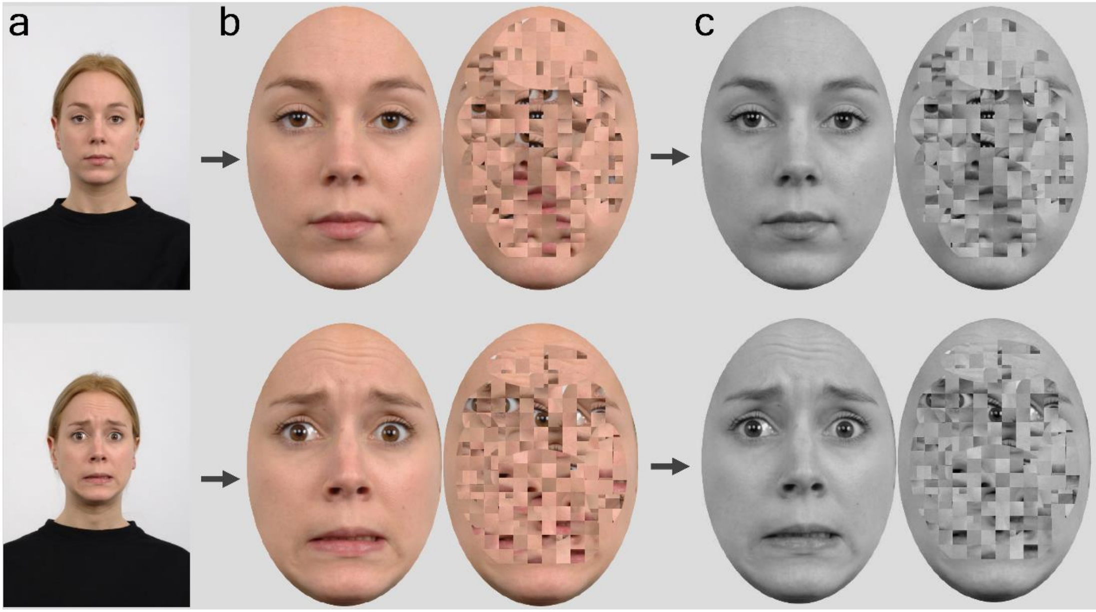
**8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)**
Nothing else to pre-register.

**Appendix B.** KDEF Face stimuli selected and standardised for experiments. Experiment 3: (a) angry, (b) neutral faces, and (c) happy faces. Experiment 4: (d) fearful, (e) neutral, and (f) happy faces. Experiment 5: (g) fearful and (h) neutral faces. Experiment 6: (i) fearful and (j) happy faces. Experiments 7 and 8 employed the same face stimuli of Experiment 5.

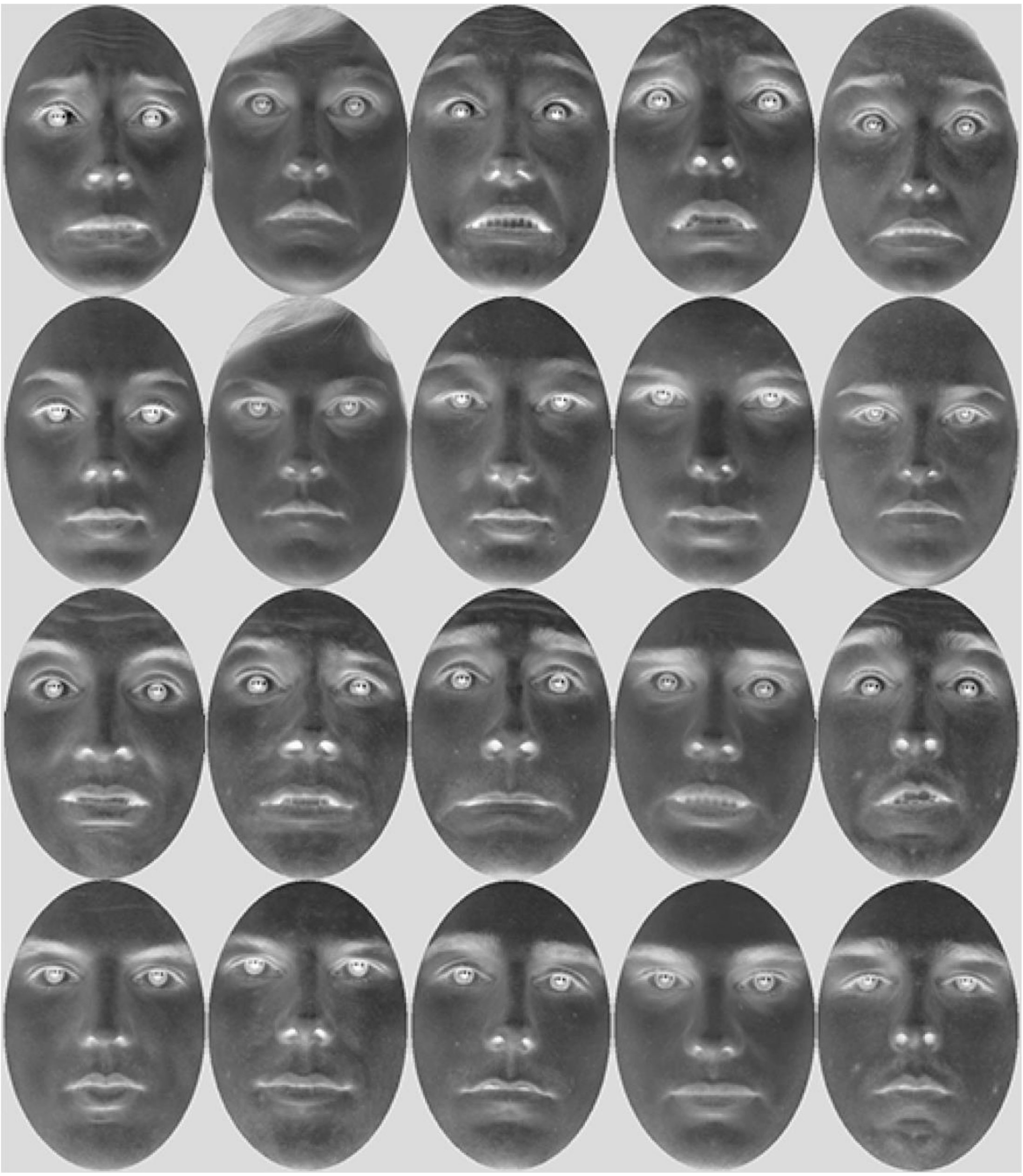


**Appendix C.** RaFD intact faces used in Experiments 9 and 11: fearful and neutral expressions.

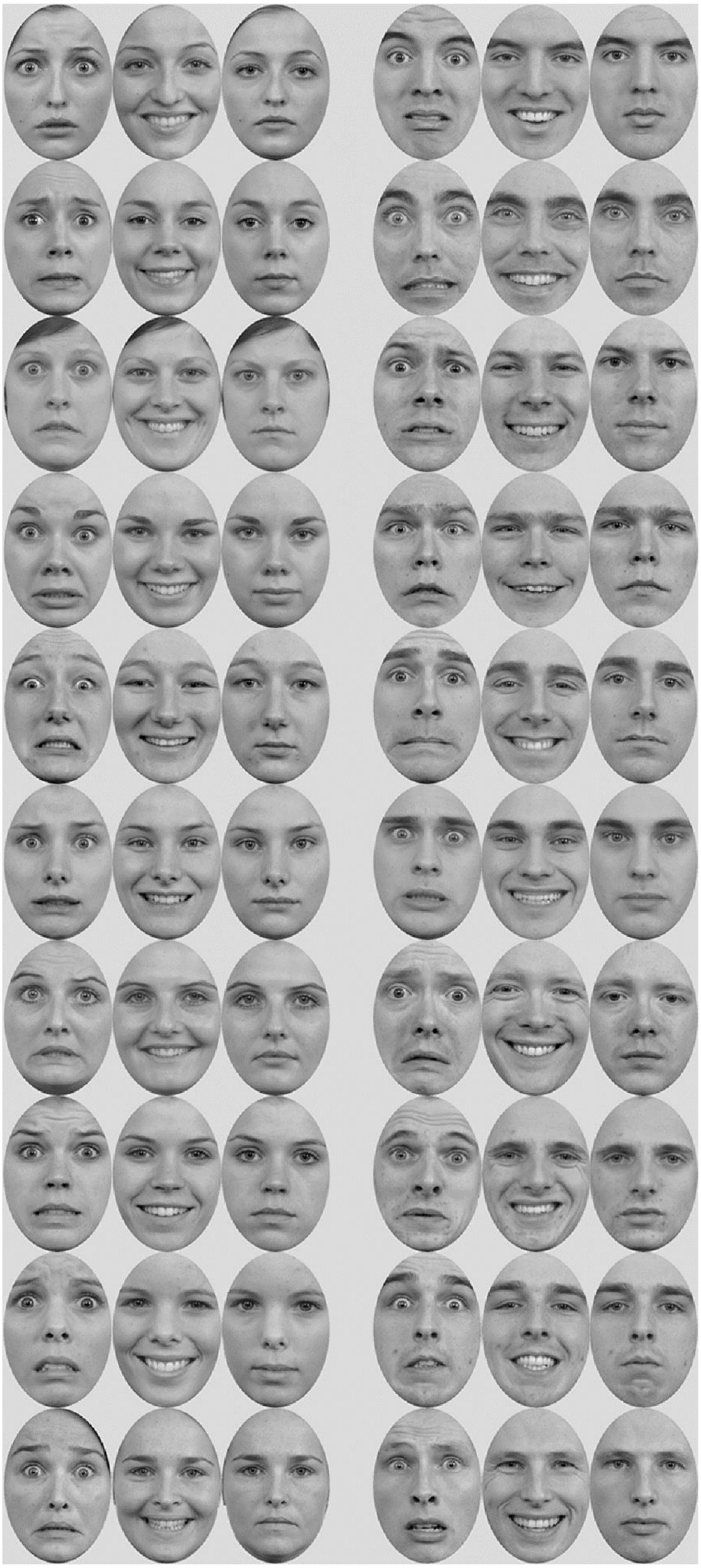
**Appendix D.** Standardisation of face stimuli. (a) Regular stimuli taken from RaFD. (b) Main facial features are left. The rest is cropped out. (c) Result of standardisation procedure after equating for luminance and contrast with Matlab SHINE toolbox.



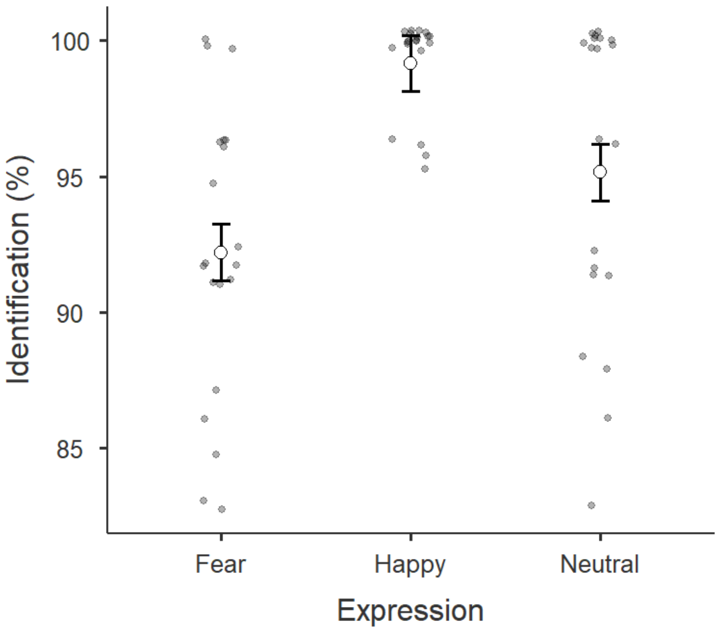
**Appendix E.** Intact colour-inverted faces used in Experiments 10 and 12 to emulate afterimages.



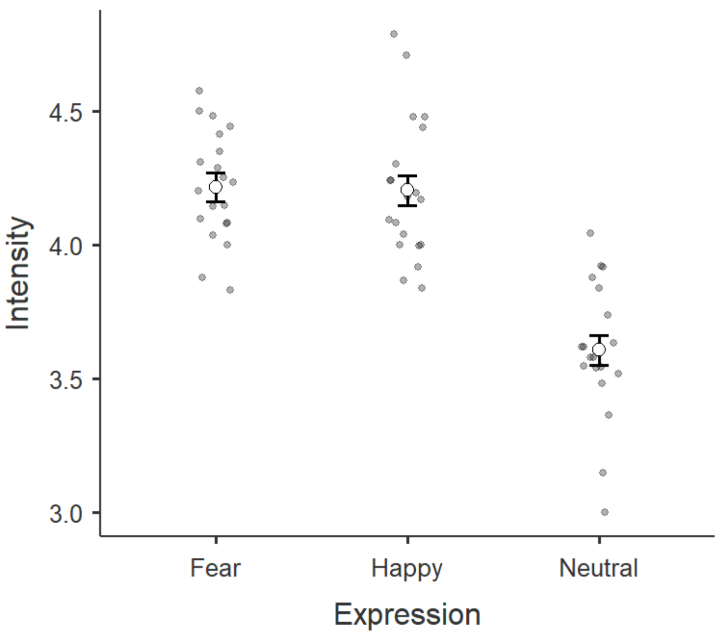
**Appendix F.** RaFD stimuli used in Experiment 13: fearful, happy, and neutral faces.



**Appendix G.** Expression identification/agreement (%) scores of RaFD stimuli selected for Experiment 13, based on RaFD validation norms, between emotional expressions (fearful, happy, neutral).




**Appendix H.** Expression intensity scores of RaFD stimuli selected for Experiment 13, based on RaFD validation norms, between emotional expressions (fearful, happy, neutral).



**Appendix I.** Experiments 13 and 14 Pre-registration for VAN and LP analysis

AsPredicted

## The minimal exposure duration required for reflective consciousness (#27805)

**Created:** 09/12/2019 05:10 PM (PT)
**Public:** 04/12/2020 05:58 PM (PT)

**Author(s)**
Renzo Lanfranco (University of Edinburgh) - Renzo.Lanfranco@ed.ac.uk
Andres Canales-Johnson (University of Cambridge) - afc37@cam.ac.uk
Axel Cleeremans (Université Libre de Bruxelles) - axcleer@ulb.ac.be
Hugh Rabagliati (University of Edinburgh) - hugh.rabagliati@ed.ac.uk
David Carmel (Victoria University of Wellington) - david.carmel@vuw.ac.nz

**1) Have any data been collected for this study already?**
It's complicated. We have already collected some data but explain in Question 8 why readers may consider this a valid pre-registration nevertheless.

**2) What's the main question being asked or hypothesis being tested in this study?**
Two ERPs have been related to consciousness. First, the visual awareness negativity (VAN) has been related to phenomenal consciousness. Second, the late positivity (LP) has been related to reflective consciousness. If this is the case, in a visual discrimination task there should be a minimal exposure duration required for LP to discriminate between aware and unaware images while there should not be one for VAN.

We ran an experiment using an LCD tachistoscope that can present visual stimuli for microseconds. In one experiment, we presented participants with images of faces (and their scrambled counterparts), one of each on each side of the screen, and of objects in the same fashion. We found that VAN responds to the subjective awareness rating of the participant even when there was no perceptual discrimination. We also found that LP can only discriminate between aware and unaware trials when the exposure duration was 4.3 milliseconds or higher.

Here, we are pre-registering the analyses we will run for a second experiment. In this new experiment, we used emotional facial expressions instead:
1. If VAN is related to phenomenal consciousness, it should not depend on the exposure duration of the stimulus, even if they are emotional faces.
2. If LP is related to reflective consciousness, it should require a minimal exposure duration to be able to discriminate between "aware" and "unaware" faces.
3. If emotional valence can facilitate reflective awareness, then LP amplitude should be higher for emotional than non-emotional faces once it can distinguish between "aware" and "unaware" trials.

**3) Describe the key dependent variable(s) specifying how they will be measured.**
It is voltage measured with EEG electrodes. We have defined two ERPs of interest, which will be measured at two different topographic locations on the scalp. Both respond to consciousness-related aspects of visual processing.
There are also other dependent variables that are relevant but their results are not new: location sensitivity, emotion identification sensitivity, emotion decision criterion, metacognitive sensitivity, metacognitive bias, and perceptual awareness.

**4) How many and which conditions will participants be assigned to?**
This particular experiment involved presenting emotional expressions (neutral or emotional - the latter could be fearful or happy) at three different exposure durations (1.7, 4.4, and 6.2 ms). Therefore, there were 6 conditions in a repeated-measures within subjects design.

**5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.**
This time, the main hypothesis is about VAN and LP and about what aspects of consciousness they mark. We will separate each trial by their perceptual awareness scale (PAS) subjective score. Trials with a PAS = 1 ("no experience") will be classified as the "unaware" trials while the PAS > 1 trials ("vague impression", "almost clear experience", "clear experience") will be classified as the "aware" trials. We will compare:
1) VAN mean voltage between aware and unaware trials for each face expression condition
2) The same as 1 but for LP.

**6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.**
Exclusion criteria:
1. Participants whose EEG signal required >10 channels to be interpolated
2. Participants who required >5% of trials to be rejected due to EEG artifacts.

**7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.**
40 participants were tested. 32 of them were included in the analysis. This is because 8 of them either presented >5% of rejected trials or >10 channels that needed interpolation.

**8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)**
Originally, these two experiments were ran with a completely different purpose. The objective was to measure with psychophysics the minimal exposure

durations required for (1) perceptual discrimination and metacognitive sensitivity of faces and emotional expressions, and (2) neural markers of face and emotion processing (N170, EPN, and LPP).

Now, we thought about using the same experiments for a very different purpose. To use the PAS scores originally thought to measure metacognitive sensitivity in order to categorise trials as aware or unaware and look whether there is a minimal exposure duration of reflective consciousness.

**Appendix J.** Neutral faces and highly recognisable objects used in Experiment 14.