Université de Montréal

Visual information processing during conscious and non-conscious face perception

par

Verena Willenbockel

Département de psychologie

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures

en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.) en psychologie

option sciences cognitives et neuropsychologie

Septembre, 2012

Résumé

Les stimuli naturels projetés sur nos rétines nous fournissent de l'information visuelle riche. Cette information varie le long de propriétés de « bas niveau » telles que la luminance, le contraste, et les fréquences spatiales. Alors qu'une partie de cette information atteint notre conscience, une autre partie est traitée dans le cerveau sans que nous en soyons conscients. Les propriétés de l'information influençant l'activité cérébrale et le comportement de manière consciente *versus* non-consciente demeurent toutefois peu connues. Cette question a été examinée dans les deux derniers articles de la présente thèse, en exploitant les techniques psychophysiques développées dans les deux premiers articles.

Le premier article présente la boîte à outils SHINE (*spectrum, histogram, and intensity normalization and equalization*), développée afin de permettre le contrôle des propriétés de bas niveau de l'image dans MATLAB. Le deuxième article décrit et valide la technique dite des bulles fréquentielles, qui a été utilisée tout au long des études de cette thèse pour révéler les fréquences spatiales utilisées dans diverses tâches de perception des visages. Cette technique offre les avantages d'une haute résolution au niveau des fréquences spatiales ainsi que d'un faible biais expérimental. Le troisième et le quatrième article portent sur le traitement des fréquences spatiales en fonction de la conscience. Dans le premier cas, la méthode des bulles fréquentielles a été utilisée avec l'amorçage par répétition masquée dans le but d'identifier les fréquences spatiales corrélées avec les réponses comportementales des observateurs lors de la perception du genre de visages présentés de façon consciente *versus* non-consciente. Les résultats montrent que les mêmes fréquences spatiales influencent de façon significative les temps de réponse dans les deux conditions de conscience, mais dans des sens opposés. Dans le dernier article, la méthode des bulles fréquentielles a été combinée à des enregistrements intracrâniens et au Continuous Flash Suppression (Tsuchiya & Koch, 2005), dans le but de cartographier les fréquences spatiales qui modulent l'activation de structures spécifiques du cerveau (l'insula et l'amygdale) lors de la perception consciente *versus* non-consciente des expressions faciales émotionnelles. Dans les deux régions, les résultats montrent que la perception non-consciente s'effectue plus rapidement et s'appuie davantage sur les basses fréquences spatiales que la perception consciente.

La contribution de cette thèse est donc double. D'une part, des contributions méthodologiques à la recherche en perception visuelle sont apportées par l'introduction de la boîte à outils SHINE ainsi que de la technique des bulles fréquentielles. D'autre part, des indications sur les « corrélats de la conscience » sont fournies à l'aide de deux approches différentes.

Mots-clés : conscience visuelle, fréquences spatiales, perception des visages

Abstract

Natural stimuli impinging on our retinas provide us with a wealth of visual information. This information varies along "low-level" features, such as luminance, contrast, and spatial frequency (SF). Whereas some of this information reaches our awareness, some of it is processed in the brain without us ever becoming aware of it (i.e., non-consciously). A remaining question is precisely which SFs influence brain activation and behavior consciously vs. non-consciously. The aim of this thesis was to address this question using state-of the-art psychophysical techniques.

The first article introduces the SHINE (spectrum, histogram, and intensity normalization and equalization) toolbox for controlling low-level image properties in MATLAB. The second article describes and validates the SF Bubbles technique, which was used throughout the studies in this thesis to map SF tuning for various face perception tasks with a high SF resolution and low experimental bias. The third and fourth articles focus on SF processing as a function of awareness. In the former, SF Bubbles was employed together with repetition priming and masking to investigate which SFs are correlated with observers' behavioral responses during conscious vs. non-conscious face-gender perception. The results show that the same SFs significantly influenced response times in both prime awareness conditions but, surprisingly, in opposite ways. In the latter, SF Bubbles was combined with intracranial recordings from awake human patients and Continuous Flash Suppression (Tsuchiya & Koch, 2005). This allowed us to map the SFs that modulate activation in specific brain structures (the insula and the amygdala) during the conscious vs. non-conscious perception of emotional facial expressions. The results for both regions demonstrate that non-conscious perception relied on low SFs more and was faster than conscious perception.

The contribution made in this thesis is thus two-fold: methodological contributions to visual perception research are made by introducing the SHINE toolbox and the SF Bubbles technique, and insights into the "informational correlates" of consciousness are provided from two different angles.

Keywords: face perception, spatial frequency, visual consciousness

Table of contents

List of tables

Chapter 2: First article

Chapter 4: Third article

Chapter 5: Fourth article

List of figures

Chapter 2: First article

Chapter 3: Second article

Chapter 4: Third article

Chapter 5: Fourth article

Chapter 6: General discussion

List of symbols and abbreviations

CFS: continuous flash suppression

CI: classification image

cpf: cycles per face (width)

cpi: cycles per image

CSF: contrast sensitivity function

CV: classification vector

EEG: electroencephalography

ERP: event-related potential

FFA: fusiform face area

FFT: fast Fourier transform

IFFT: inverse fast Fourier transform

IT: inferotemporal cortex

FIE: face inversion effect

fMRI: functional magnetic resonance imaging

LGN: lateral geniculate nucleus

MEG: magnetoencephalography

MT: middle temporal area

NCC: neuronal correlates of consciousness

NCE: negative compatibility effect

OFC: orbitofrontal cortex

RMS: root mean square

RMSE: root-mean-square error

RT: response time

SF: spatial frequency

σ: standard deviation

SSIM: structural similarity

V1: primary visual cortex

Dedication

Für meine Eltern und meinen Bruder

Acknowledgments

My first and primary thank you goes to my supervisors, Frédéric Gosselin and Franco Lepore, without whose support and guidance this thesis would not have been possible. Thank you profoundly for your time and good advice, for giving me the freedom to choose my projects, encouraging me to present my research at international conferences, and financial support. I am very grateful for everything you have taught me during the last four years. Frédéric, thank you for your detailed and quick constructive feedback, not only on various drafts of my thesis, manuscripts, and project proposals but also on my MATLAB code—I have learned a lot from you.

My deep gratitude also goes to my former supervisor, Jim Tanaka, for continuing to collaborate with me during my time in Montreal. Thank you, Jim, for your helpful advice, patiently waiting for me to finish some projects I started in Victoria, financial support (especially to present SHINE at VSS 2010), and introducing me to whole networks of great people and exciting research.

I would also like to express my gratitude to my co-authors. I especially thank Caroline Blais and Daniel Fiset who have helped me in various ways over the last years and without whom I could not have started my PhD here.

Many thanks to my colleagues at the CERNEC. Merci Alex, Catherine, Colin, Cynthia, Frédéric, Giulia, Karine, Marie-Hélène, Mélanie, Mercédès, Nic, Sacha, Sylvain, and Xavier for helping me out with French countless times, participating in my experiments, introducing me to PhD comics... I also thank everyone else at the CERNEC who helped me solve technical problems (Ping, thanks a million!) or brightened up my days, especially Latifa, Lucia, and Sophie. A special thank you to Luke for the nice welcome and to Maria van der Knaap.

I am very grateful to everyone who proofread versions of my manuscripts and parts of this thesis: Caroline, Bärbel, Catherine, Chris, Giulia, Helen, Jen, Julianna, Lara, Luke, Peggy, Sarah, and Sam—I sincerely appreciate it. I especially thank Quoc Vuong for reading all thesis chapters and for the many helpful comments and suggestions.

I would also like to thank everyone who helped me with testing participants. Josiane, thanks for your time and for keeping participants motivated through thousands of Bubbles trials. I thank Yamila for translating my experimental instructions into Spanish at very short notice.

A very big thank you goes out to all those who participated in my experiments.

On a more personal note, I would like to say *danke* to Bärbel Knäuper and her lovely family for the nice dinners, kids' birthday parties, a weekend in the Laurentians, the table at which I am writing these acknowledgments, good advice, housesitting opportunities, and many other things. Your place was like an oasis for me in the hot Montreal summers, and I wrote quite a few pages of this thesis there. Thank you so much for making this city a "home away from home" for me. I am greatly indebted to Anthony Marley for introducing us.

Thank you also to my other friends in Montreal for sharing delicious food, listening during the tougher times, movie nights, music, games of pool, practicing karate, and far too many other things to mention here: Andrea & Mark, Atoosa, Caroline, Lara & Peter, Latifa, Lucia, Marie, Nav, and my roommates Han, Melissa, and Sophie. A special thank you to Janet Cattanach and Stan Highway. Thank you also to Aryel and Sophia for making volunteering at the RVH and MNH so fun. Thanks to all of you I got the necessary distraction but also the energy to work a lot.

I am very grateful to Steffi for helping me find an apartment when I first moved to Montreal and to my friends who came for a visit and provided me with nice excuses to explore more of Quebec. Thanks Jason, Julianna, Megan, Quoc, Sarah, and Tanja. I also warmly thank all my friends who offered me a place to stay on my way home or to conferences (or just to escape my thesis writing): Caro & Daniel, Dani, Jam, Kristin, Lena & Raul, Magdalena, Marie, Peggy, Quoc, Sarah & Jason, Susan & Mark, and Tanja—thank you for the fun times in Canada, France, Germany, Ireland, the Netherlands, UK, and US.

Thank you also to my other friends from out west (especially Colin, Eri, Greg, Hayley, Heather, Jenn, Karen, Marina, Nathan, Ryan, and Sam) and from home (especially Britta, Jessica, Laura, and Melli) for keeping in touch, birthday surprises, wonderful memories from Salsbury Way, movie nights, ice cream… and giving me energy for the "last 5k of my PhD marathon". I would like to say a special thank you to Matt, Mirjam—who has been the best friend I could ask for since kindergarten—and Waltraut.

Finally, the most important thank you straight from my heart goes to my family. I am more grateful than words can say for the loving support from my parents, grandparents, and my brother. Vielen lieben Dank für alles.

Chapter 1: General introduction

"The question is not what you look at but what you see."

Henry David Thoreau

We humans are conscious beings. According to the philosopher John Searle "'Consciousness' refers to those states of *sentience* or *awareness* that typically begin when we wake from a dreamless sleep and continue through the day until we fall asleep again, die, go into a coma or otherwise become 'unconscious'" (Searle, 1998, p. 381; see also Koch, 2004). From a first-person perspective, conscious experience is probably highly familiar to all of us. However, trying to understand what exactly makes humans conscious beings from the third-person perspective is a different story. Consciousness has sparked philosophers' and researchers' interest for centuries (see, e.g., Ludwig, 2007, for an overview on the ancient Mind-Body Problem), but despite immense progress in understanding the workings of the human brain, there is still no scientific explanation or widely accepted formal definition of it (e.g., Dehaene & Changeux, 2011; Koch, 2004). Conscious experience has been referred to as "at once the most familiar thing in the world and the most mysterious" (Chalmers, 1995, p. 80).

The main challenge is to explain the qualitative nature of conscious experiences—the "what it is like" aspect—also called *qualia* (Lewis, 1929). Frequently used examples of a quale are the redness of red or the painfulness of pain. Is it possible to fully explain qualia in terms of physical processes, such as certain neuronal firing patterns? If so, how and why do physical processes cause qualia? It is not clear how to answer or directly tackle these questions with scientific approaches, and hence they have been referred to as the *hard problem* of consciousness (Chalmers, 1995, 1996; but see Dennett, 1996). Some philosophers have argued that the hard problem will never be solvable (e.g., Chalmers, 1995, 1996; Levine, 1983; Nagel, 1974)—i.e., that there will remain an "explanatory gap" (Levine, 1983). This led many scientists to believe that consciousness is a philosophical issue, beyond scientific study.

However, consciousness research has seen a lot of interest and optimism since the late 1990s (e.g., Crick & Koch, 1998, 2003; Searle, 1998). A number of labs have begun to examine the issue from a different angle using new neuroimaging techniques (e.g., Dehaene et al., 1998; Sahraie et al., 1997). The focus shifted to the so-called easy problems: What specifically characterizes conscious perception? What does it correlate with at the neural level? Although it is still unclear how to find out what *causes* consciousness, it is widely agreed that there are material *correlates* of consciousness and that those can be discovered with scientific approaches. This research could have important clinical implications—e.g., for

anesthesia, coma, vegetative-state, or minimally conscious patients (Dehaene & Changeux, 2011)—and is regarded by many as helpful for coming up with a final theory of consciousness (Koch, 2004).

Broadly, there are two complementary lines of research that are linked with different meanings of the term "consciousness" (e.g., Rees, Kreiman, & Koch, 2002; Tononi & Koch, 2008): One approach consists of examining the neural events associated with changes in the overall level of consciousness (e.g., sleep, vegetative states, and seizures). Here, the term is used in the intransitive sense of *being conscious* vs. *not being conscious*. The other approach focuses on the *contents* of consciousness; wakefulness is considered a prerequisite. In the latter context, researchers have searched for the neuronal correlate(s) of consciousness (NCC) defined as "the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept" (Koch, 2004, p. 16; see Block, 2005, for the suggestion of at least two distinct NCCs). Here, the term consciousness is used in the transitive sense of being conscious *of* X vs. not being conscious *of* X. In this thesis, consciousness will be used in the latter, transitive sense.

Many studies in the latter context have involved studying the *non-conscious* perception of stimuli (see, e.g., Kouider & Dehaene, 2007, for a review). It is thought that at any given moment, we perceive a subset of the sensory input consciously, while a considerable amount of information is also processed non-consciously, that is, without ever reaching our awareness (e.g., see Dehaene & Changeux, 2011). Non-consciously processed stimuli have been found to influence observers' behavioral responses (e.g., see Kouider & Dehaene, 2007, for a review) and brain activation at several levels (for a recent review, see Van Gaal & Lamme, 2011). For example, activation to non-consciously processed, "unseen" visual stimuli was observed in subcortical structures such as the amygdala (e.g., Jiang & He, 2006; Naccache et al., 2005; see Tamietto & De Gelder, 2010, for a review), in primary visual cortex (V1; e.g., Haynes & Rees, 2005), and also in high-level cortical areas (e.g., Fang & He, 2005; Kouider, Eger, Dolan, & Henson, 2009). These and similar findings have raised the question of what distinguishes non-conscious from conscious perception.

In studies designed to investigate the NCC in neurotypical participants, often one condition is set up in which stimuli are presented above the participant's awareness threshold and one, minimally different, condition in which those stimuli fall below the threshold (e.g.,

Dehaene et al., 2001; see Dehaene & Changeux, 2011, for a review). The threshold is typically operationally defined by subjective reports (e.g., visibility ratings) or objective measures (e.g., chance performance in detection or forced-choice discrimination tasks). Which measures are optimal continues to be debated but recently it has been argued that subjective and objective measures are largely consistent (Dehaene & Changeux, 2011). Contrasting conscious and non-conscious perception using a variety of techniques, such as neuroimaging, electrophysiology and behavioral methods, is thought to provide important insights into whether conscious and non-conscious processing differ merely quantitatively—i.e., whether one is basically a weaker form of the other—or qualitatively—i.e., whether they are mediated by distinct processes.

Most of the studies comparing conscious and non-conscious processing so far have focused on the visual domain. Two likely reasons for this are that normal-sighted humans tend to rely a lot on their sense of vision, and that the visual system is the most thoroughly studied sensory system so far (e.g., Koch, 2004). Moreover, visual stimuli can be systematically manipulated relatively easily. Several psychophysical techniques have been developed to reliably suppress visual stimuli from awareness (i.e., render them "invisible"; see Kim & Blake, 2005, for a review; Tsuchiya & Koch, 2005). Some of these techniques do so by diverting the observer's top-down attention[1], whereas others reduce the bottom-up stimulus strength (Dehaene & Changeux, 2011; Dehaene et al., 2006). For example, two commonly used techniques that do the latter are masking (e.g., Breitmeyer & Öğmen, 2006; Kouider & Dehaene, 2007) and Continuous Flash Suppression (CFS; Tsuchiya & Koch, 2005). In masking paradigms, a stimulus is rendered invisible by presenting it briefly and immediately followed by a second stimulus (the mask) that interrupts processing of the first. In CFS, a static stimulus is presented to one eye and rendered invisible by flashing high-contrast dynamic noise to the other eye. Each technique has its own strengths and weaknesses (see, e.g., Kim & Blake, 2005); thus, ideally, one would use a variety of methods to try to collect converging evidence. In Chapters 4 and 5 of this thesis, we used masking and CFS, respectively, to study visual consciousness.

[1] The discussion of the relation between attention and consciousness is beyond the scope of this thesis; see, e.g., Finkbeiner and Palermo, 2009; Koch and Tsuchiya, 2007; Lamme, 2003; Naccache, Blandin, and Dehaene, 2002, for different views.

Contrasting conscious and non-conscious visual perception has shed some light on the neural markers of consciousness (for a recent review see Dehaene & Changeux, 2011). Dehaene and Changeux reviewed the results obtained with a variety of awareness-manipulating techniques and recording methods. The latter included functional magnetic resonance imaging (fMRI), which has a high spatial resolution, surface event-related potentials (ERPs), which have a high temporal resolution, and intracranial ERPs (recorded from patients undergoing epilepsy monitoring), which have both a high spatial and high temporal resolution. The authors concluded that the findings converge in suggesting that conscious (vs. non-conscious) perception is characterized by increases in activity approximately 300–500 ms after stimulus onset, activation of several distributed cortical regions that include bilateral parietal cortex and prefrontal cortex, and synchronized activity at beta (13–30 Hz) and gamma (50–100 Hz) frequencies across distant regions in the cortex in the aforementioned time window.

Despite a lot of progress with revealing the neural correlates of visual consciousness, however, little is known yet about other correlates and characteristics of visual awareness. Specifically, which aspects of a complex visual stimulus form our contents of consciousness in a given situation, and which aspects remain non-consciously processed? For example, is coarse information processed quickly and non-consciously to enable us to respond rapidly in dangerous situations, whereas detailed information is not available during non-conscious perception? Studies on conscious vs. non-conscious perception have typically traced the "footprints" of complex stimuli in the brain that are ecologically important to us, such as faces or words (e.g., Dehaene et al., 2001; Jiang & He, 2006). However, each of these stimuli is comprised of a wealth of low-level information, such as variations in luminance, contrast, color, and spatial frequency (SF; see below). Which subsets of the available information modulate brain responses under different awareness conditions—i.e., represent the "informational correlates" of consciousness—remains largely unknown.

Investigating the informational correlates of consciousness could provide important insights into whether qualitative or quantitative differences exist between conscious and non-conscious perception from an information-processing perspective. For instance, a qualitative difference would be that detailed information of an image is only available consciously, whereas coarse information influences observers' responses independently of awareness. A

quantitative difference would be that exactly the same information is processed consciously and non-consciously but with greater sensitivity in one of the conditions. Despite growing support for qualitative differences in observers' response patterns under different awareness conditions (e.g., Barbot & Kouider, 2012; Eimer & Schlaghecken, 2002; Frings & Wentura, 2005; Snodgrass & Shevrin, 2006), the more parsimonious, quantitative view has remained difficult to fully refute (e.g., Holender & Duscherer, 2004; Perruchet & Vinter, 2002; Verleger et al., 2004; see also Hannula, Simons, & Cohen, 2005, for a review). Thus, it might be insightful to investigate this issue from a complementary, information-processing perspective.

One relatively straightforward approach to reveal the informational correlates of consciousness could consist of mapping the correlations between specific stimulus information and observers' responses (as can be done with classification image techniques; e.g., Eckstein & Ahumada, 2002; Gosselin & Schyns, 2001) as a function of awareness in a behavioral study. One could, however, also take this idea a step further and map the correlations between stimulus information and observers' brain responses with a high spatial and millisecond temporal resolution as provided, for example, by intracranial ERPs. "Zooming in" by tracing exactly which aspects of a complex stimulus elicit brain activation during different stages of conscious vs. non-conscious perception could possibly further clarify our picture about the NCC. Arguably, this approach would bring us as close as possible to mapping the brain activation associated with specific qualia: Some philosophers have defended the view that "the phenomenal character of a perceptual experience is entirely determined by the experience's propositional content—that is, by what it represents" (Byrne, 2001, p. 199).

Which type of information would be a promising candidate for the informational correlates of consciousness? As a first step, one might want to consider information that is universally present in all visual stimuli and that has previously been shown to be used by the visual system to construct our percepts. One promising candidate for such information is the SF content of images: Analogous to how auditory signals can be broken down into tone frequencies, visual signals can be decomposed into SFs (e.g., Campbell & Robson, 1968; De Valois & De Valois, 1990). The SF gives the number of light-dark transitions per unit of distance, e.g., per image width or per degree of visual angle (see, e.g., Sowden & Schyns, 2006). SFs have been referred to as the "building blocks" of visual representations (e.g.,

Schyns, Petro, & Smith, 2009; see Morrison & Schyns, 2001, and Ruiz-Soler & Beltran, 2006, for reviews; for more information about the SF content of images, see Chapter 2).

Psychophysical experiments have indicated that the visual system comprises approximately four to six SF "channels" (for reviews, see De Valois & De Valois, 1990; Morrison & Schyns, 2001). A channel is broadly defined as consisting of all neurons that take part in the processing of a stimulus that contains energy at specific SFs (De Valois & De Valois, 1990). Neurons can in this context be considered as filters that respond maximally to a certain SF and less to SFs that are more and more different from this optimal frequency (i.e., they are selectively *tuned* to certain SFs). The SF range that activates a channel gives the channel bandwidth, which is typically specified in octaves (one octave corresponds to a doubling of frequency—e.g., from 2 to 4 cycles per degree, from 4 to 8 cycles per degree, and so forth). It has been observed that channel bandwidths lay mostly between 1.0 and 1.5 octaves (e.g., De Valois, Albrecht, & Thorell, 1982; Simpson & McFadden, 2005; Wilson & Wilkinson, 1997).

Together the different SF channels are thought to define the contrast sensitivity function (CSF) of the visual system. The CSF describes the relationship between SF and sensitivity to luminance contrast—i.e., how much contrast as a function of SF is required to perceive a stimulus. For humans, the peak of the CSF is at 6 cycles per degree of visual angle, and the CSF is shaped like an inverted U-function (i.e., sensitivity falls off at the extrema of the SF spectrum; De Valois & De Valois, 1990). Given that SFs are a universal property of all visual stimuli and are used by the visual system, they appear to be an important property to consider if one wants to shed light on the informational correlates of consciousness.

As one might anticipate, this brings us to the main goal of this thesis: to shed light on the puzzle of consciousness by investigating which SF information of socially relevant stimuli correlates with conscious vs. non-conscious perception.

The focus will be on the perception of face stimuli, which are particularly well suited for several reasons. For instance, faces are natural stimuli with high behavioral significance in our everyday interactions. Using face images allows for a number of different real-world categorization tasks (e.g., face identification, face-gender classification, emotion categorization) to be performed on the same set of physical stimuli (e.g., see Gosselin & Schyns, 2001). Also, previous studies have suggested that faces can be processed non-

consciously (e.g., Finkbeiner & Palermo, 2009; Jiang & He, 2006; Jiang et al., 2009; Smith, 2012). Finally, faces are complex stimuli that contain a broad spectrum of SFs. Thus, they comprise a promising set of stimuli to uncover potential shifts in information use as a function of awareness and to shed light on the nature of the differences between conscious and non-conscious perception in general.

The following sections of this General Introduction provide background information about the physiological properties of the visual system with an emphasis on SF processing, theories on visual perception, some of which attribute different functional roles to different SFs, and a brief summary of influential neurobiological theories on consciousness. Afterwards, I will describe previous work on SF processing and awareness in the context of face perception, and further outline the structure of this thesis.

Overview of SF processing in the brain

Extracting information from visual scenes is a seemingly effortless task for most individuals with intact eyesight, something most of us do "all the time" and often probably without giving it much thought. However, the neural mechanisms underlying visual perception are incredibly complex, as one might begin to realize when looking at the complex connectivity patterns between over 30 visual brain areas (e.g., Felleman & Van Essen, 1991). It is impossible to give a detailed description of all the different levels of visual processing in just a few pages of this thesis. However, the following section contains a coarse, condensed overview on the main processing pathways in the visual system, with an emphasis on SF processing. The section serves the purpose to facilitate understanding of the following sections on theories of visual categorization and recognition, and theories on consciousness. For more detailed information on the visual system in general see, e.g., Kandel, Schwartz, and Jessell (2000), and for more information related to the neural substrate of SF processing see, e.g., De Valois and De Valois (1990).

Visual information processing begins after light has reached the retina at the back of the eye. There, light patterns are transduced into electrochemical signals through photoreceptor cells and encoded as a function of position, time, and wavelength (e.g., Nassi & Callaway, 2009; Rodieck, 1998). The output of the photoreceptors is processed and condensed in a number of complex steps by retinal cells of various types (in total, at least 80 different

types forming approximately 20 separate circuits; Nassi & Callaway, 2009). Basic sensory cues are extracted, such as temporal frequency, SF, luminance, and color contrasts. The retinal output is conveyed through different types of ganglion cells, the axons of which form the optic nerve. The receptive field of ganglion cells (i.e., the region in visual space where a change in stimulation can change the cell's activity) has a concentric, center-surround structure that allows them to serve as broad SF filters (Van Essen, Anderson, & Felleman, 1992; see De Valois & De Valois, 1990, for a review and illustrations).

Most of the retinal output is conveyed through three parallel pathways to the lateral geniculate nucleus (LGN) of the thalamus and from there on to V1 (see Nassi & Callaway, 2009, for a review): The parvocellular pathway originates from midget ganglion cells, which have relatively small receptive fields and make up approximately 70% of cells projecting to the LGN. Their axons reach the parvocellular LGN layers, which pass on signals to V1 (layers 4Cß and 6). Broadly speaking, the parvocellular pathway is characterized by low contrast sensitivity, slow conduction velocities, and sensitivity to low temporal and high SFs. The magnocellular pathway originates from parasol ganglion cells, which have relatively large receptive fields and constitute approximately 10% of cells conveying signals to the LGN. Their axons terminate in the magnocellular LGN layers, from where information is projected to V1 (layers 4Cα and 6). Broadly speaking, the magnocellular pathway is characterized by high contrast sensitivity, fast conduction velocity, and sensitivity to high temporal frequencies and low SFs. The koniocellular pathway consists, at least partly, of bistratified ganglion cells that make up about 8% of the cells that project to the LGN. They terminate in the koniocellular LGN layers, from where cells project to V1 (layer 1 and the cytochrome oxidase blobs of layer 2/3). Less has been discovered so far about this pathway, but it has been suggested that it is sensitive to short-wavelength information and that some of its response properties lay in-between those of the magno- and parvocellular pathways (e.g., see Briggs & Usrey, 2011, for a review).

Besides the parvo-, magno-, and koniocellular pathways, there are a number of smaller pathways that leave the eye (see, e.g., Martin & Solomon, 2011). It has also been suggested that there is a direct subcortical route from the retina via the superior colliculus and the pulvinar nucleus of the thalamus to the amygdalae, which are located deep within the medial temporal lobes (see Johnson, 2005; Tamietto & De Gelder, 2010, for reviews; but see Pessoa

& Adolphs, 2010). This pathway has been suggested to convey low-SF information quickly and largely automatically (e.g., Vuilleumier et al., 2003).

The magno-, parvo-, and koniocellular pathways converge in V1 (striate cortex). There, more complex information is extracted than at the retinal and LGN levels. Whereas retinal ganglion and LGN cells have rather broad SF tuning characterized by steep decreases in sensitivity to high SFs but only gentle decreases in sensitivity to low SFs, V1 cells show steep decreases in sensitivity to both high and low SFs (i.e., band-pass tuning properties; e.g., De Valois & De Valois, 1990; Henriksson, Nurminen, Hyvärinen, & Vanni, 2008). Moreover, in contrast to ganglion and LGN cells, V1 cells are orientation-selective. In V1, information is represented in overlapping functional maps, with orientation, ocular dominance and SF columns (Nassi & Callaway, 2009). Because V1 cells are the first to show sufficiently narrow SF and orientation tuning, it has been suggested that the substrate for SF channels is located in the cortex (De Valois & De Valois, 1990). In fact, the connected area V2 and several areas beyond have also been found to show band-pass SF tuning (e.g., Henriksson et al., 2008). After integrating information from the parvo-, magno-, and koniocellular pathways, new output streams originate from V1 and V2 that project to a "mosaic" of extrastriate areas (Van Essen et al., 1992).

More specifically, V1 and V2 outputs represent the beginning of the dorsal and ventral streams (Ungerleider & Mishkin, 1982). The dorsal pathway projects to the dorsomedial area, middle temporal area (V5/MT), and to the posterior parietal cortex. The ventral pathway goes through V4 to interconnected extrastriate areas in the temporal cortex, including the inferotemporal cortex (IT) and the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997). It is thought that the streams stay largely segregated all the way into motor-related frontal areas but that they are interconnected at several sites. The two streams have been suggested to play different functional roles: The dorsal pathway has been referred to as the "where" or vision-for-action pathway, whereas the ventral pathway has been associated with object recognition ("what") and vision-for-perception (e.g., Milner & Goodale, 1995; see also Nassi & Callaway, 2009, for a review).

It is thought that the relative contributions of the magno- and parvocellular pathways to those streams differ. The dorsal pathway was found to mainly receive magnocellular input (thus, mainly low SFs); the ventral stream was found to receive both parvo- and magnocellular

inputs (thus, high and low SFs) (Ferrera, Nealey, & Maunsell, 1994; Merigan & Maunsell, 1993). Consistent with this idea, area V5/MT in the dorsal stream was found to show low-pass SF tuning (Henriksson et al., 2008), and the results from several fMRI studies that used face images as stimuli appear consistent with the suggestion that high-SF processing involves ventral stream areas (Eger, Schyns, & Kleinschmidt, 2004; Iidaka, Yamashita, Kashikura, & Yonekura, 2004; Rotshtein, Vuilleumier, Winston, Driver, & Dolan, 2007; Vuilleumier, Armony, Driver, & Dolan, 2003; see De Gardelle & Kouider, 2010).

One might get the wrong impression that visual information processing is essentially hierarchical; however, information does not just flow bottom-up but also top-down. There are many feedback connections at all levels in the visual system (except the retina). Overall, it is thought that key principles by which the visual system is organized include hierarchical, modular, parallel, and recurrent processing; and SF processing was found to be an efficient strategy for the visual system to condense complex input relatively early on (e.g., Olshausen & Field, 1996).

Theories on visual processing

Although the architecture of the visual system has been rather well known since the late 1980s (Bullier, 2001b), many open questions remain regarding the processes underlying our seemingly effortless perception of our complex surroundings. One challenge has been to disentangle the influences of bottom-up factors (e.g., physical properties of the stimuli) and top-down factors (e.g., the observer's expectations and experience). It is not yet known, for example, how exactly SF information is integrated to allow us to quickly categorize and recognize complex objects and scenes. Two main hypotheses have been put forward in this context—the *coarse-to-fine hypothesis* and the *flexible usage hypothesis* (e.g., see Morrison & Schyns, 2001, for a review)—which will be the focus of this section.

A number of models on visual perception have suggested that the visual system integrates SF information in a fixed temporal order (e.g., Bar, 2003; Bullier, 2001a, b; Marr, 1982; Watt, 1987; see also Hochstein & Ahissar, 2002; for a review, see Hegdé, 2008). Specifically, it has been proposed that the processing of coarse information, carried by low SFs, precedes the processing of detailed information, carried by high SFs (i.e., this is the coarse-to-fine hypothesis). In this framework, different functional roles have been attributed to

different SFs (see, e.g., Bar, 2003; Bullier, 2001a, b): It is thought that the low-SF "gist" of the input, processed rapidly up to higher visual areas, serves as a basis for deriving a top-down hypothesis that helps to narrow down the number of possible interpretations of the input. High SFs then come into play to flesh out the details and allow for fine-grained categorizations. This is considered to be an efficient strategy because the low-SF information presumably provides a stable "frame" of the image before the noisier high-SF content is extracted and fills in the details.

The models differ, however, in their proposals of the specific areas involved. In the model by Bullier (2001a, b) low-SF information is first projected rapidly via the magnocellular pathway to areas of the dorsal stream, such as V5/MT. The computations done in these areas are then projected back to V1/V2, where they guide the processing of high-SF information that arrives through the slower parvocellular pathway. V1/V2 are thus thought to serve as "active blackboards" that integrate low- and high-SF information. The theory by Bar (2003, 2004; Bar et al., 2006) also postulates that quick low-SF processing is achieved via the magnocellular pathway and dorsal stream, while high SFs are processed more slowly via the parvocellular pathway and ventral stream. However, central roles are attributed to the orbitofrontal cortex (OFC) and IT. In the OFC, the "initial guess" about the most likely interpretation of the input is derived, which is then back-projected to IT. The OFC prediction is thought to influence the ongoing processing of high SFs along the ventral stream.

Support for coarse-to-fine processing in general has come from psychophysical (e.g., Schyns & Oliva, 1994; see Morrison & Schyns, 2001, for a review), electrophysiological (e.g., Bredfeldt & Ringach, 2002; Frazor et al., 2004; Mazer et al., 2002), and neuroimaging studies (e.g., Goffaux et al., 2011) using a variety of stimuli, including lines, dots, and gratings (Hughes, Nozawa, & Kitterle, 1996; Mihaylova, Stomonyakov, & Vassilev, 1999; Musselwhite & Jeffreys, 1985; Parker & Dutch, 1987; Watt, 1987), faces (Halit, De Haan, Schyns, & Johnson, 2006; McCarthy, Puce, Belger, & Allison, 1999; Vlamings, Goffaux, & Kemner, 2009), and scenes (Parker, Lishman, & Hughes, 1992, 1997; Peyrin, Mermillod, Chokron, & Marendaz, 2006; Schyns & Oliva, 1994).

However, as already indicated, an alternative to the coarse-to-fine hypothesis has also been proposed. Oliva and Schyns (1997; Schyns & Oliva, 1999) put forward the flexible SF usage hypothesis, according to which SF integration is not fixed but can proceed in a flexible

order, determined by top-down factors and stimulus properties (for reviews see Morrison & Schyns, 2001; Ruiz-Soler & Beltran, 2006; Sowden & Schyns, 2006). Such factors include attention (e.g., Özgen, Sowden, Schyns, & Daoutis, 2005), sensitization (e.g., Oliva & Schyns, 1997), and task demands (e.g., Schyns & Oliva, 1999).

For instance, Oliva and Schyns (1997) presented observers with scenes of either low or high SFs that were combined with noise at the other SFs. Observers were instructed to perform a 4-choice scene categorization task. Afterwards, *hybrid* scenes were displayed, which consisted of a low-SF scene combined with another high-SF scene (e.g., highway + city or vice versa). It was found that observers who had previously seen the low-SF (+ noise) scenes, typically perceived the low-SF component of the hybrids, whereas those who saw the high-SF scenes (+ noise) perceived the high-SF component. In another study with hybrid stimuli, Schyns and Oliva (1999) showed that SF usage can be influenced by task demands: When participants were asked to judge whether a hybrid face was expressive or not, observers used the high SFs, when asked to categorize expressions, they used the low-SFs, and when asked to judge facial gender, they were not biased toward either low or high SFs. These results suggest that there is some flexibility in SF usage.

The flexible usage hypothesis is not necessarily inconsistent with the fact that coarse-to-fine processing has been observed in several studies, because for certain situations—in which SF integration from low to high is the most efficient strategy—the flexible usage hypothesis would also predict that SF processing proceeds from low to high. The hypothesis is also not necessarily inconsistent with the finding that at very early levels of processing (e.g., before V1) low SFs are processed faster than high SFs: It is possible that this bias does not constrain SF usage at higher processing levels involved in categorization and recognition (Morrison & Schyns, 2001).

To test the coarse-to-fine vs. flexible usage hypothesis, one might want to investigate SF processing with a higher SF resolution (not just low vs. high SFs) and with a high spatial and temporal brain resolution. Although it was not the main objective of the studies in this thesis, some of the experiments included here did touch on this issue. For instance, in Chapter 3, we examined if different SFs underlie the identification of upright vs. inverted faces, which is feasible according to the flexible usage hypothesis but would not be predicted by the coarse-to-fine hypothesis. In the same chapter, we also re-examined with a high SF resolution

whether different SFs are used as a function of categorization task on the same physical stimuli. Moreover, in Chapter 5, we investigated SF tuning in an intracranial ERP study, which allowed us to see for individual brain regions how SF processing unfolds over time.

The models presented in this section will be important with regard to our hypothesis about putative interactions between SF processing and awareness.

Theories on consciousness

It is thought that much of the visual processing described in the previous sections occurs without our awareness. For instance, Bar (2003) suggested for his model that until the most likely interpretation of the input is chosen, processing proceeds non-consciously. How then does a stimulus become the content of our consciousness? A variety of cognitive and neurobiological theories on consciousness have been proposed (see De Gardelle & Kouider, 2009, and Kouider, 2009, for reviews, respectively). In this section, the focus will be on neurobiological theories, some of which, however, are based on cognitive ones (e.g., the Global Neuronal Workspace theory). The aim is to provide a brief sketch of five influential theories in the field of research on the NCC. As Kouider (2009) points out, those theories have in common that they propose that consciousness is related to neural activity at the cerebral level (as opposed to more localist or more globalist theories).

Milner and Goodale (1995) proposed the *Duplex Vision Theory,* according to which the ventral and dorsal streams have evolved for the perception of objects and the control of actions directed at those objects, respectively (see the *Overview of SF processing in the brain* section). Processing in the dorsal stream has been found to be very quick and automatic, in contrast to processing in the ventral stream. Hence, according to Milner and Goodale, the dorsal stream subserves non-conscious perception whereas the ventral stream constitutes the NCC (see also Crick & Koch, 2003; Fang & He, 2005; but see Kouider et al., 2009; Tapia & Breitmeyer, 2011).

Edelman and Tononi (2000) proposed the *Reentrant Dynamic Core Theory*. This theory rests on the assumption that the brain is a dynamic complex system with individual groups of neurons carrying out their functions, while communicating with other groups of neurons in the "dynamic core". The dynamic core consists of widely distributed groups of neurons in the cortex and the thalamus. Consciousness is thought to result from the interactivity in the thalamocortical network, rather than from specific brain areas or types of neurons themselves. It is thus a

relatively globalist theory of consciousness (Kouider, 2009).

Dehaene and Naccache (2001) put forward the *Global Neuronal Workspace* theory, which is an extension of Baar's Global Workspace theory. It suggests that information is first non-consciously conveyed in a bottom-up fashion to specialized regions (e.g., V4 for color processing, V5/MT for movement processing). Information can reach awareness if, due to top-down amplification through attention, the respective specialized processes are made available to a broad network of brain areas through long-distance connections (i.e., to the *global workspace*). It is thought that the information that is globally and flexibly available in the workspace forms the content of consciousness.

Lamme (2003) suggested that consciousness is associated with three stages of processing. The first stage involves bottom-up processing to extrastriate areas and parietal and temporal cortices (the "feed-forward sweep"); as also suggested by Dehaene and Naccache (2001), this first stage is thought to occur without the observer's awareness. In the second stage, the information is fed back to the early visual cortex. According to Lamme, the recurrent interactions between early and later visual areas are specifically linked with awareness. Third, global interactions are thought to take place that involve the frontoparietal network and language areas (e.g., enabling us to report the stimuli).

Zeki (2003) proposed that instead of a single consciousness, there are several consciousnesses. For instance, conscious perception of color reflects neural activity in area V4; conscious perception of motion reflects activity in V5/MT. According to his account, conscious perception (microconsciousness) occurs when a certain activation threshold is exceeded in the respective specialized brain region. Binding of the different attributes happens afterwards (macroconsciousness).

As can easily be seen, the different theories on consciousness diverge in many aspects. However, many of them converge in suggesting that non-conscious processing precedes conscious processing.

Interactions between SF processing and awareness during face perception

The previous three sections have outlined that (a) SF processing plays a central role during visual processing in the human brain; (b) one strategy for the quick recognition of complex stimuli might consist of extracting low SFs before high SFs; and (c) several neurobiological theories of consciousness suggest that non-conscious processing occurs before

conscious processing. Can we derive hypotheses from this about the informational correlates of consciousness?

One hypothesis that researchers have recently begun to test is that low SFs can be extracted independently of awareness, whereas high SFs might require a slower, conscious analysis (De Gardelle & Kouider, 2010; Khalid, Finkbeiner, König, & Ansorge, in press). This hypothesis would fit well with the temporal dynamics proposed for low- and high-SF processing—with low SFs being processed quickly along the magnocellular pathway and high SFs more slowly along the parvocellular pathway—and the temporal dynamics for non-conscious and conscious processing—with the former preceding the latter (e.g., Dehaene & Naccache, 2001; Lamme, 2003; Zeki, 2003). It would also be in accordance with the different roles attributed to low and high SFs in models like the one by Bar (2003). Moreover, it would be consistent with the suggestion that the ventral stream, which was found to receive both parvo- and magnocellular inputs, constitutes the NCC, whereas the dorsal stream, which mainly receives magnocellular inputs, subserves non-conscious processing (Milner & Goodale, 1995). Finally, it would fit with the idea that low SFs are quickly and non-consciously conveyed along a subcortical route to the amygdala (e.g., Vuilleumier et al., 2003; see Johnson, 2005, and Tamietto & De Gelder, 2010, for reviews).

To investigate potential interactions between SF processing and awareness, De Gardelle and Kouider (2010) employed a masking paradigm with hybrid faces, which were displayed for varying durations (43–300 ms) to create different visibility conditions. Observers were asked to perform a fame judgment task on visible full-SF-spectrum target faces displayed following the hybrid and mask stimuli on each trial. The results revealed that both low- and high-SF information from unseen hybrids influenced observers' response times to the targets (to a comparable extent). Thus, De Gardelle and Kouider's results did not support the aforementioned hypothesis; instead, the authors interpreted their findings in terms of the flexible SF usage hypothesis, which is consistent with high SFs being used quickly and non-consciously.

In contrast, Khalid and colleagues did find support for the hypothesis that low- but not high SFs of unseen stimuli can influence observers' responses. In their study, low-pass or high-pass filtered male and female faces were presented, followed by a mask and a full-SF-spectrum target. The observers were asked to judge the facial gender of the target. The results

showed that response times to the target were influenced by low-pass but not high-pass filtered unseen stimuli. Khalid et al. interpreted their findings as support for the existence of a subcortical pathway to the amygdala.

Although these studies (discussed in more detail in Chapter 4) have provided a first step toward investigating the informational correlates of consciousness in the SF domain, several open questions remain. More research is obviously needed to explain the difference between the findings. Also, both of these studies only investigated SF usage at a very coarse resolution (low vs. high SFs) with largely arbitrary cutoff frequencies. This represents a strong limitation because the SF spectrum is continuous, and the visual system has been shown to comprise several SF channels. Dividing the SF range into "halves" is typically done when sketching out the main visual processing pathways and illustrating theories about object recognition, but it can be problematic in the context of the aforementioned studies because the precise SF tuning of the different pathways is far from clear. The results obtained so far thus leave open the question of which individual SFs modulate observers' responses under conscious vs. non-conscious conditions.

The first aim of this thesis is to offer a new method that can be used to reveal SF tuning for a given task with a high SF resolution and low experimental bias. The second aim is to employ this method to shed light on SF tuning during face perception in general. The third aim is to investigate the informational correlates of consciousness in the context of face perception. This represents a rather exploratory (vs. strictly hypothesis-driven) endeavor but was motivated by the idea that qualitative differences in SF tuning may be revealed as a function of awareness. These qualitative differences might take the form of a range of relatively high SFs being only significant for conscious but not non-conscious processing, whereas lower SFs are significant independently of awareness.

Structure of this thesis

The following two chapters serve as an introduction to low-level image processing with a focus on the processing of SF information during face perception. The methods that form the basis for the following studies are introduced in those chapters. Specifically, Chapter 2 describes the image preprocessing tools we developed for creating well-controlled stimulus sets (the SHINE toolbox). Chapter 3 introduces and validates a novel SF sampling approach

(the SF Bubbles technique), which was employed to reveal the precise SF tuning curves for various face perception tasks throughout this thesis.

Chapters 4 and 5, which are the main chapters, present two studies on SF processing and face perception as a function of awareness. The first one is a behavioral study, in which we investigated SF tuning during conscious and non-conscious face-gender perception. The second one is an intracranial ERP study in which we mapped—with a high temporal and a high spatial resolution—the SFs correlated with brain activation during the conscious and non-conscious perception of emotional faces.

Each of the four chapters consists of a published or submitted manuscript, including a summary of our findings (Abstract), introduction, methods, results/discussion, and references.

Chapter 6 discusses the implications of this work and avenues for future research.

Chapter 2: First article

Article published:

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,* 671–684.

Controlling low-level image properties: The SHINE toolbox

Verena Willenbockel[1], Javid Sadr[2], Daniel Fiset[1], Greg O. Horne[3], Frédéric Gosselin[1], & James W. Tanaka[3]

[1]CERNEC, Département de Psychologie, Université de Montréal, Montréal, QC, Canada
[2]Department of Psychology, University of Massachusetts Boston, Boston, MA, USA
[3]Department of Psychology, University of Victoria, Victoria, BC, Canada

Abstract

Visual perception can be influenced by top-down processes related to the observer's goals and expectations, as well as by bottom-up processes related to low-level stimulus attributes, such as luminance, contrast, and spatial frequency. When using different physical stimuli across psychological conditions, one faces the problem of disentangling the contributions of low- and high-level factors. Here, we make available the SHINE (Spectrum, Histogram, and Intensity Normalization and Equalization) toolbox for MATLAB, which we have found useful for controlling a number of image properties separately or simultaneously. The toolbox features functions for specifying the (rotational average of the) Fourier amplitude spectra, for normalizing and scaling mean luminance and contrast, and for exact histogram specification optimized for perceptual visual quality. SHINE can thus be employed for parametrically modifying a number of image properties or for equating them across stimuli to minimize potential low-level confounds in studies on higher-level processes.

Keywords: visual perception, low-level image properties, luminance, histogram matching, spatial frequency

Introduction

Using identical stimuli across experimental conditions has the tremendous advantage that one does not have to worry about potential low-level confounds, such as differences in luminance, contrast, or spatial frequency, when studying higher-level visual processes. For example, Tanaka and Curran (2001) presented the same set of dog and bird images to a group of dog experts and to a group of bird experts. The differences in brain activity observed when the experts categorized objects in their domain of expertise relative to when they categorized objects outside their domain could thus not be due to stimulus characteristics. In some studies on visual perception, however, it is impossible to use the same physical stimuli across psychological conditions. For instance, studies in which the domain-specificity versus domain-generality accounts of face processing are examined typically involve comparisons between faces and non-face objects, such as comparisons of the inversion effect for faces and houses (e.g., Yin, 1969). When comparing effects across categories, one encounters the problem that there might be overall low-level differences between stimulus types or differences in the amount of within-category variation, which could potentially result in biases unrelated to the higher level processes meant to be studied. In general, when using different stimuli across conditions, one has to be careful to disentangle low-level and high-level factors (Fründ, Busch, Körner, Schadow, & Herrmann, 2007; Itier & Taylor, 2004; Luck, 2005; Rousselet, Macé, Thorpe, & Fabre-Thorpe, 2007; Sadr & Sinha, 2001, 2004; VanRullen & Thorpe, 2001).

Variations in low-level properties are thought to have contributed to controversies in the literature (e.g., Bentin et al., 2007; Dakin, Hess, Ledgeway, & Achtman, 2002; Hershler & Hochstein, 2006; Rousselet, Pernet, Bennett, & Sekuler, 2008; Thierry, Martin, Downing, & Pegna, 2007a, 2007b; VanRullen, 2006), and the importance of avoiding low-level confounds is an important and recurrent issue for a variety of research fields and measuring methods, such as eye tracking, functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG). Inconsistent findings were obtained, for instance, about the earliest real event-related potential (ERP) differences between face and object perception. Some studies have been focused on the N170 component—a negative deflection occurring approximately 170 ms after stimulus onset—as the first marker of face processing (e.g., Carmel & Bentin, 2002; Rossion et al., 2000;

Rousselet, Husk, Bennett, & Sekuler, 2008), whereas other studies revealed earlier differences between face and object processing, possibly as early as 50–80 ms after stimulus onset (e.g., George, Jemel, Fiori, & Renault, 1997; Seeck et al., 1997). Because the studies differ in the degree to which image properties were controlled across face and non-face categories, the different findings might, at least in part, be explained by differences in low-level influences (Rousselet, Husk, et al., 2008). Luck (2005) gave the general advice that one should ”never assume that a small physical stimulus difference cannot explain an ERP effect” (p. 74).

In studies in which it is impossible to use identical stimuli across conditions, it can thus be important to match certain image properties across the stimulus set. Luminance is one property of interest, because early visual processes have been found to be sensitive to luminance variations, indicated by the modulation of early ERP components, such as the P1 (e.g., Johannes, Münte, Heinze, & Mangun, 1995). In some studies, this has been accounted for by equating stimuli in terms of mean luminance and contrast (for recent examples, see Finkbeiner & Palermo, 2009; Hardee, Thompson, & Puce, 2008; Liang, Zebrowitz, & Aharon, 2009; Zion-Golumbic, Golan, Anaki, & Bentin, 2008). It is possible to go a step further and precisely match the luminance histograms—which give the number of pixels at each luminance level—across images, thereby equating not only the means and standard deviations of the luminance distributions but also their shape. Exact histogram matching addresses the finding that certain neural mechanisms are sensitive to luminance histogram skewness (Olman, Boyaci, Fang, & Doerschner, 2008)—in particular, mechanisms involved in estimating surface properties (Motoyoshi, Nishida, Sharan, & Adelson, 2007) or texture discrimination (Chubb, Landy, & Econopouly, 2004).

Besides luminance, one might want to equate the images’ spatial frequency content. Broadly speaking, low spatial frequencies represent the coarse information in an image, such as luminance blobs and blurred shapes, whereas high spatial frequencies carry the fine-grained information, such as the precise shape of an object. There is evidence that during early visual processing, the input is analyzed at multiple spatial frequencies by a number of channels, each tuned to a specific range (see De Valois & De Valois, 1990, for a review). Findings indicate that there are differences in sensitivity to specific spatial frequencies both between different visual areas and between the right and left hemispheres (e.g., Ivry & Robertson, 1998). Variations in the spatial frequency domain could, for example, lead to differences in stimulus

detectability (Campbell & Robson, 1968; Gold, Bennett, & Sekuler, 1999; Honey, Kirchner, VanRullen, 2008). Carefully controlling luminance information and the energy at different spatial frequencies (i.e., Fourier amplitude spectra) should thus be of interest to a variety of studies designed to investigate high-level visual processes. However, numerous examples can be found in the literature where low-level properties remained uncontrolled, and to our knowledge, there is no standard program for systematically matching image properties across experimental stimuli.

Many researchers use MATLAB (The MathWorks, Natick, USA) for image preprocessing, running experiments, and for data analysis. In recent years, the number of applications for MATLAB has increased enormously, and labs have shared their tools for experimentation to save time and reach a higher degree of standardization in the field. Besides several commercially available MATLAB toolboxes, such as the Image Processing Toolbox and the Signal Processing Toolbox, there are a number of freely available ones, such as the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), the Eyelink Toolbox (Cornelissen, Peters, & Palmer, 2002), EEGLAB (Delorme & Makeig, 2004), and the Biopsychology Toolbox (Rose, Otto, & Dittrich, 2008).

Here, we make accessible another MATLAB toolbox that we have found useful for controlling a number of low-level image properties in studies on visual perception. Specifically, our SHINE (Spectrum, Histogram, and Intensity Normalization and Equalization) toolbox, which was written using functions from the Image Processing Toolbox of MATLAB, includes different equalization approaches that can be applied together or separately, depending on the requirements of the experiment. SHINE features functions for specifying the Fourier amplitude spectra of images or scaling the rotational average of the spectra only (i.e., the energy at each spatial frequency averaged across orientations). It also includes functions for exactly specifying the luminance histograms or for normalizing and scaling means and standard deviations of the luminance distributions without affecting their shape. The program offers ways to equate the luminance properties across stimuli separately for the foregrounds and the backgrounds and to preserve perceptual image quality as much as possible. Although histogram specification and Fourier amplitude specification affect each other, we discovered that by using an iterative approach, a high degree of joint matching of the low-level properties of interest could be reached in many cases. SHINE also features tools for plotting the Fourier

amplitude spectra or the average energy at each spatial frequency to verify the output or to monitor the ecological low-level variations.

In the following sections, we provide an overview of the individual SHINE functions and give examples regarding how to work with the toolbox, which can be downloaded at www.mapageweb.umontreal.ca/gosselif/shine. The toolbox has successfully been used by us, as well as by others (e.g., Fiset, Blais, Gosselin, Bub, & Tanaka, 2008; Williams, Willenbockel, & Gauthier, 2009) in studies on visual perception to equate (or parametrically vary) low-level image properties across experimental stimuli. Moreover, prior to establishing the standard version of SHINE presented here, in a number of other behavioral and neuroscientific studies earlier or partial forms of these preprocessing steps were successfully used to achieve strict stimulus control (e.g., Adolphs et al., 2005; Liu, Harris, & Kanwisher, 2002; Loschky et al., 2007; Mack, Gauthier, Sadr, & Palmeri, 2008; Sadr & Sinha, 2001, 2004; Xu, Liu, & Kanwisher, 2005).

Method

The main m-file is *SHINE*, which calls the functions for the individual adjustment steps, such as histogram matching or Fourier amplitude matching. In *SHINE,* one can specify the parameters—for instance, the type of matching desired, the number of iterations, and whether to perform the luminance adjustment on the whole image or separately for selective regions, such as foreground and background. The individual functions called by *SHINE* are described in the following sections, and Figure 1 provides an overview of the structure of the toolbox.

Normalizing and scaling mean luminance and contrast

In some cases, it might be desirable to equate the images in terms of mean luminance and contrast (i.e., the standard deviation of the luminance distribution) only, instead of specifying the exact histogram shape and thereby equating the means and standard deviations automatically. The former typically results in higher image quality because no noise is introduced, as it is in most cases of exact histogram matching. The SHINE function *lumMatch* can be used for a simple normalization:

$$Z = (X - m) / s,$$

where X is the input image matrix containing the original grayscale values—that is, $X = \lfloor 255 (L - L_{min}) / (L_{max} - L_{min}) \rfloor$, where L is the luminance matrix as measured on the computer monitor, and L_{min} is the minimum and L_{max} the maximum luminance that can be displayed on the monitor; see the SHINE and monitor calibration section. m is the mean of the original values, and s is the standard deviation. For simplicity, we will refer to the grayscale values contained in X as *luminance values* throughout the article. The desired mean (M) and desired standard deviation (S) can then be applied to obtain the output image matrix E containing the scaled luminance values:

$$E = ZS + M.$$

The default values for M and S are obtained by averaging the means of all input stimuli and the standard deviations of all input stimuli, respectively.

As with exact histogram matching (see the *Luminance histogram matching* section), there are cases in which normalization on the whole image will give the best results and cases in which it might be advantageous to consider the foreground and the background separately. Using the function *separate*, SHINE transforms the input image into a binary template from which it extracts which pixels belong to the foreground and which belong to the background. A simple example illustrates why separate equalization might be important: Assume one wanted to equate a stimulus consisting of a relatively bright object on a midgray uniform background with a stimulus depicting a relatively dark object on the same midgray background. If the luminance adjustment was applied to the whole image, the first stimulus would end up with a darker background than the second image, so that the background itself would contain diagnostic information. In some cases, applying the matching step separately to selective regions might therefore be preferable.

Luminance histogram matching

The SHINE toolbox includes the function *histMatch* that exactly matches the luminance histograms of a number of source images with a specified target histogram. Specifically, it calls the function *match,* which transforms one luminance distribution into another by remapping the pixel values to control how frequently they occur relative to others. The average luminance distribution of the input set, which is computed using *avgHist* and then transformed into a sorted list of luminance values using *hist2list,* serves as the default target

(Figures 2a–d). Alternatively, it is possible to provide the *match* function with another target—for example, by using the function *tarhist*, which sorts the luminance values of the input images in ascending order and then averages the values across images (i.e., the darkest values across all images, the second darkest, etc.) to obtain the target (Figures 3a and b).

In general, two strategies can be applied for histogram matching: One can exactly match the histograms, or one can approximately match them by using the best sub-optimal pointwise transformation of the luminance values. The former is what SHINE does (the built-in MATLAB function *histeq* does the latter). Even though exact histogram matching comes at the expense of increased noise in the image, we observed that in many cases image quality is still very high after applying SHINE (see Figures 3a–c for examples and the *Joint matching of certain low-level properties* section for image quality measures). Several ways of exact histogram matching have been proposed and evaluated in terms of visual quality of the result and computational complexity (e.g., Avanaki, 2009; Bevilacqua & Azzari, 2007; Coltuc, Bolon, & Chassery, 2006; Morovic, Shaw, & Sun, 2002; Rolland, Vo, Bloss, & Abbey, 2000).

SHINE follows the exact global histogram-matching approach described in Table I (see also Avanaki, 2009; Rolland et al., 2000). Basically, the pixels of the source image and the pixels of the target distribution are sorted separately by their luminance value from darkest to lightest. The source pixel with the smallest value is then assigned the smallest value of the target, the source pixel with the second smallest value is assigned the second smallest value of the target, and so forth. Ambiguity arises when a number of source pixels with the same luminance value have to be broken down into two or more groups to match the target: It has to be decided which pixels get new values. For example, this would be the case if the two darkest pixels of the source image had the same luminance value (e.g., 0), but the target histogram contained only one pixel with a value of 0. So which of the two source pixels should be assigned a new value? The *match* function of SHINE uses random assignment (as was applied in Williams et al., 2009), but different approaches have been suggested (Coltuc et al., 2006; Wan & Shi, 2007). In all cases, the histogram-matching step produces a set of images that are made up of the same pixels, only with a different arrangement (Figure 3). In this way, mean luminance, contrast, and all other characteristics of the histogram (e.g., skew) are equated across stimuli.

In order to generate output images that not only are exactly equated in histograms, but also retain the structure of the original image as much as possible, the *histMatch* function includes the method developed by Avanaki (2009): It allows for iteratively optimizing the structural similarity (SSIM) index (Wang, Bovik, Sheikh, & Simoncelli, 2004; see the *Joint matching of certain low-level properties* section for more details on the SSIM index) between the original and histogram-matched images. In each iteration, the *match* function is applied to obtain an image with the specified target histogram. The visual quality of the histogram-matched image is measured using the SSIM index, and an SSIM gradient with respect to that image is computed using the *ssim_sens* function[1]. The gradient is then employed to adaptively increase the SSIM of the output (see Avanaki, 2009, for details). The algorithm is summarized in Table II. Using SSIM gradient ascent, the perceptual image quality can be improved considerably relative to basic histogram matching alone, whereas the histogram remains exactly as specified. This benefit, however, comes at the cost of increased runtime, which might make the SSIM optimization option unsuitable for the matching of very large image sets.

SHINE also includes the option for applying the histogram matching selectively to subsets of pixels—for example, separately to the foreground and the background. This can improve output quality in some cases, because it prevents object pixels from ending up in the background and vice versa (e.g., one might want to choose this option for stimuli displaying a single object on a uniform or noisy background; see the *Specifying Fourier amplitude spectra* section).

Specifying Fourier amplitude spectra

Using Fourier analysis, any complex two-dimensional image can be broken down into the sum of a set of sinusoidal gratings defined by four parameters: spatial frequency, orientation, amplitude, and phase. The spatial frequency refers to the width of the gratings' bars and can be specified as the number of light:dark cycles per image. Orientation refers to the angle of the light and dark bars, specified in degrees counter-clockwise from vertical; amplitude is given by the difference in luminance between the lightest and the darkest parts of the grating; and phase refers to the position of the sinusoid relative to some reference point.

After performing a Fourier transform on an image (e.g., with the built-in MATLAB function *fft2*), one can obtain two components: the amplitude spectrum and the phase spectrum[2]. The amplitude spectrum specifies the amplitude of each constituent grating at a particular spatial frequency and orientation. The phase spectrum specifies the phase of each grating at a particular spatial frequency and orientation. If all of the gratings at the corresponding phases and amplitudes were summed, they would exactly result in the original image.

SHINE includes the function *specMatch*, which matches the amplitude spectrum of the source image with a specified target spectrum. Each source image is submitted to a fast Fourier transform (FFT), and the spectrum is shifted so that the low spatial frequencies occupy the central region. If not specified otherwise, the average spectrum is obtained across all input images and serves as the default target. The phase of the original image is then combined with the target amplitude spectrum and back-transformed from the frequency into the spatial domain using an inverse fast Fourier transform (IFFT).

This approach matches the source spectrum exactly with the target spectrum at each spatial frequency and orientation (Figure 4c). The output images typically look a bit "cloudy" and their visual quality depends on the similarity of the input images; for example, equating the spectra of a number of faces will likely yield better results than equating the spectra across different categories, such as faces and cars. If image quality is a concern, we propose to use the *sfMatch* function described below.

Specifying the rotational average amplitude at each spatial frequency

The function *sfMatch* applies a more lenient approach for equating the Fourier amplitudes across stimuli. Unlike *specMatch*, it preserves the amplitude distribution across orientations, while ensuring that the rotational average amplitudes for a given spatial frequency are equated between images. The initial steps are the same as those described above for *specMatch*: The source images are subjected to an FFT, and the output is shifted so that the low spatial frequencies are in the center of the spectrum. If not specified otherwise, the target spectrum is obtained by averaging across input spectra. A coefficient is then computed for each spatial frequency: The amplitudes across orientations at the given spatial frequency are summed, separately for the source and the target spectrum. For each spatial frequency, the

resulting target sum is divided by the corresponding source sum to obtain a coefficient. Afterward, the amplitude at each spatial frequency and orientation is multiplied with the respective coefficient. The phase of the original image is combined with the modified amplitude spectrum and back-transformed into the spatial domain using an IFFT. As a result, the rotational average of the amplitude spectrum is equated between images (besides slight rounding errors that may arise), whereas the distribution of amplitudes among orientations is preserved. Although this controls overall amplitude, it often results in higher similarity of the output image with the source image than would strictly equating the amplitudes at each spatial frequency and each orientation (Figures 4a–c).

Rescaling of luminance values after the IFFT

It is possible that after the Fourier amplitudes are equated and the IFFT is applied, the luminance values of the resulting images are shifted out of the desired range or out of the range that can be displayed—for example, below 0 or above 255 for 8-bit images. In this case, the negative values and the ones larger than 255 will be clipped to 0 and 255, respectively. This results in a change of the actual final luminance and contrast but usually only a small change in the amplitude spectrum. One might want to rescale the luminance values after Fourier amplitude matching in such a way that all or at least the majority of the values fall back into the range of 0 to 255 (Sadr & Sinha, 2001, 2004). This is not done separately for each image, because individual rescaling would result in images that no longer match the properties one set out to normalize in the first place. The *rescale* function of SHINE therefore first obtains the full range of luminance values coming out of the IFFT, for all of the images, and then computes one set of rescaling values to apply to each image. Specifically, after going through all of the images and calculating the lowest and highest luminance values, there are two options: It is possible to rescale all images with the same linear scaling parameters so that absolutely all pixel values for all images are in the range of 0 to 255 or on average, the smallest and largest pixel values are rescaled to 0 and 255, but some clipping is still allowed to occur in the final product. The second alternative gives a slightly imperfect result, which is why the default SHINE is set to the first alternative. However, the second approach squashes the images' luminance less than the first one and so might give nicer-looking images with a little more contrast.

Joint matching of certain low-level properties

Theoretically, one could perfectly match all low-level properties across an image set; however, then the images would be identical. One challenge is to equate certain properties while preserving others. SHINE attempts to preserve the phase information in the image while matching the luminance histograms and Fourier amplitudes. This is associated with at least two problems: First, phase can also be affected by histogram matching (through the introduction of noise), and the visual quality of the image is thus decreased; second, histogram matching and Fourier amplitude matching affect each other. That is, Fourier amplitude matching performed after histogram matching will distort the histograms again to some extent and vice versa.

To measure the decrease in visual quality, we added two functions for image quality assessment to SHINE. The *getRMSE* function computes the root-mean-square error (RMSE), which is a simple and widely used metric for quantifying the visibility of differences between a distorted and a reference image. Because it is not always well matched to perceived visual quality, we also included the SSIM index (Wang et al., 2004) computed by Wang, Bovic, and Simoncelli's (2003) *ssim_index* function. The SSIM index has been developed for quality assessment based on the degradation of structural information. If the input images are identical, the SSIM index is 1, and if they are uncorrelated, it is very small. Although it will be impossible to completely avoid alterations to the visual quality, these two quality assessment functions provide means for comparing the different matching options objectively and could be helpful for finding the appropriate low-level equalization procedure (or parameters).

In order to address the problem of histogram and Fourier amplitude specification affecting each other, we developed and implemented an iterative approach: The histogram and Fourier amplitude-matching steps can be performed a number of times, whereby the respective target is recalculated at each iteration. Using this iterative strategy, we were able to reach a high degree of joint matching in previous studies using SHINE (e.g., Williams et al., 2009). Figures 5–6c illustrate the increasing degree of joint equalization over iterations for a few typical image sets. Specifically, Figure 5 shows one car, one chair, and one face stimulus from the set used by Williams and colleagues (2009) with the corresponding histograms after 0, 1, and 3 iterations. One iteration consisted of *histMatch* followed by *sfMatch* (with default

rescaling of the luminance values). It can be seen that the histograms gradually became more similar to each other.

In Figures 6a–c, we used the same face and car images as in Figure 5 (Set 1) and the scenes depicted in Figure 3b (Set 2). We ran 20 iterations separately on the two sets, whereby one iteration consisted of *histMatch* followed by *specMatch* (with default rescaling) or vice versa. The histogram matching was applied separately to the foreground and the background for the images in Set 1 and to the whole image for those in Set 2. To trace the increasing degree of joint equalization of amplitude spectra and luminance histograms, we computed the RMSE between the amplitude spectra of the two images in each set for each of the iterations, and similarly, we computed the RMSE between the normalized histograms of each image pair. As a comparison, for the original images, the RMSE between the amplitude spectra was 30.7394 for Set 1 and 21.3955 for Set 2; the RMSE between the normalized histograms was 0.0069 for Set 1 and 0.0046 for Set 2. As an image quality measure, we computed the mean SSIM index between the original image and the SHINE output for each set at the respective iteration. The results show that the individual matching steps worked precisely because the property that was matched last (e.g., amplitude spectra for *histMatch* followed by *specMatch*) was nearly perfectly equated for all iterations. The property that was matched first was altered less as the number of iterations increased, thus converging toward 0. Clearly, the greatest improvements in joint equalization were reached within the first four iterations, independently of equalization order. The SSIM index, which initially equaled 1, remained fairly high and constant through Iterations 1–20. For Set 1, it equaled .76 at Iteration 1 and .74 at Iteration 20, when *histMatch* was performed first. When *specMatch* was applied first, it equaled .80 and .78 for Iterations 1 and 20, respectively. Similarly, for Set 2, the SSIM index was .8 (Iterations 1–20) when the histograms were equated first and remained at .81 for all iterations when the spectra were matched first. Depending on the input set and the desired degree of matching, one might want to choose different numbers of iterations and might want to apply the steps in a different order.

Plotting functions and image statistics

The Image Processing Toolbox of MATLAB features a function for plotting the luminance histogram of a given image (the *imhist* function). However, it does not provide

functions for directly plotting the rotational average of the Fourier amplitude spectrum or the amplitude spectrum itself. SHINE includes these two plotting functions (*sfPlot* and *spectrumPlot*, respectively), which can be useful for checking the SHINE output, or for assessing low-level variations in the source images. Moreover, SHINE features a function for computing a number of image statistics across an image set (the *imstats* function).

Applying SHINE to color images

SHINE was originally designed for the preprocessing of grayscale luminance images, but there are ways to apply it to color images as well. In particular, the luminance-matching functions can be directly applied to color images if the images are in a color space that allows for separating luminance as one dimension, as does the HSL color space. One would then equate the histograms—or just the mean and standard deviation of the luminance distribution—of the respective layer without altering the hues and saturation of the image.

For equating the Fourier amplitudes, one could use the RGB color space and do a separate matching of amplitude spectra for each of the color layers. The basic matching step would thus be analogous to how it is described for one-layer grayscale images but would have to be applied three times for each RGB color image.

SHINE and monitor calibration

The matching functions of SHINE (except *histMatch*) assume that the relationship between the stored luminance values that make up the image matrix X and the luminance intensity produced at the face of the screen is linear. However, on a typical computer-monitor system, this is not the case. When using LCDs, one can correct for the nonlinearity by setting the *gamma* parameter to 1. When using a CRT monitor, however, this will not suffice. To meet the linearity assumption of SHINE, it is important in such a case to perform gamma correction after equating the low-level properties of a set of images expressed as Xs (see the *Normalizing and scaling mean luminance and contrast* section). For example, the inverse power function converting the luminance intensity (L) into RGB values (V)—

$$V \approx (L - k)^{\frac{1}{gamma}},$$

where gamma is a positive constant usually close to 2.5 and *k* is a constant that corresponds to the minimum luminance of the monitor—can be applied to correct for the nonlinearity. Gamma correction is a common practice and can be done in various ways (e.g., Metha, Vingrys, & Badcock, 1993; Pelli & Zhang, 1991; Poynton, 1993; Stanislaw & Olzak, 1990).

Summary: The default SHINE

The default version of SHINE first exactly equates the Fourier amplitude spectra, whereby the average spectrum obtained across all input images serves as the target. Rescaling ensures that the luminance values stay in the desired range and are not clipped. Afterward, SHINE equates the luminance histograms across all input images, so that each output image has the same luminance distribution as the average of the input sets. Fourier amplitude and histogram matching are performed iteratively a number of times specified by the researcher and depending on the requirements of the experiment to obtain a high degree of joint matching of Fourier amplitudes and luminance properties. Demos of the main program and individual functions are available online at www.mapageweb.umontreal.ca/gosselif/shine.

Discussion

The SHINE toolbox contains a number of MATLAB functions for controlling low-level image properties, such as luminance, contrast, and spatial frequency. Specifically, it can be used for specifying the (rotational average of the) Fourier amplitude spectra, for exact luminance histogram specification, and for normalizing and scaling mean luminance and contrast. SHINE offers ways to apply the luminance adjustments selectively to a subset of pixels (e.g., separately to the foreground and the background), and it includes the option to perform the histogram and Fourier amplitude specification iteratively a number of times to reach a high degree of joint matching of luminance and spatial frequency properties between source and target.

The main motivation behind SHINE was to provide tools that can easily be applied (even by novice MATLAB users) for equating a number of low-level properties across a stimulus set in order to minimize low-level confounds in studies on higher level processing. It has been shown that early vision is sensitive to variations in luminance and spatial frequency content (see De Valois & De Valois, 1990, for a review), and several researchers have recently emphasized the importance of disentangling low- and high-level factors when using different

physical stimuli across psychological conditions (e.g., Luck, 2005; Rousselet, Husk, et al., 2008; Sadr & Sinha, 2001, 2004). Controlled image normalization is an important issue in various research fields, including psychophysics and visual search, and for different brain imaging techniques, such as fMRI, MEG, and EEG. SHINE, which allows for controlling both luminance and spatial frequency characteristics of images with great precision, should thus be useful in a variety of experiments in order to minimize low-level confounds.

To our knowledge, there is no other commonly available program or toolbox for directly equating low-level image properties across a set of images. Photo-editing software, such as Photoshop, allows for individual luminance adjustments, but it will hardly be possible to use it for exact histogram matching of a large number of complex images. Furthermore, photo-editing programs typically do not provide tools for matching images in the Fourier domain. The Image Processing Toolbox of MATLAB, which SHINE is based on, contains the function *histeq* for histogram equalization; however, it serves the purpose of contrast enhancement and does not perform exact histogram matching as does the *match* function made available here. Furthermore, SHINE extends the toolbox by including the implementation of a state-of-the-art method to optimize the structural similarity between original and histogram-matched images (Avanaki, 2009), as well as tools for the direct manipulation and visualization of the Fourier amplitude spectra and the average energy at each spatial frequency.

Knebel, Toepel, Hudry, Le Coutre, and Murray (2008) proposed a different approach for generating controlled stimulus sets, which one might want to apply in combination with SHINE. Specifically, they introduced a method for spectral distance optimization, which does not alter physical image properties but, rather, identifies the subgroups of images that are spectrally most similar. For example, given 100 images of Category A and 100 images of Category B, their method could be applied to select 40 images per category so that the spectral differences between the subsets are minimized. Functions from the SHINE toolbox could then be used to match the selected images exactly; one could apply SHINE to all 80 images at once or to a smaller group (e.g., 10 images per category) at a time. As a result, the low-level properties would be equated across categories, whereas the alterations introduced by the exact-matching steps of SHINE would be minimized by choosing the most similar images in the first place.

Functions from the SHINE toolbox have recently been applied in a study comparing the spatial frequency sensitivity of face and object processing (Williams et al., 2009). The rotational average of the Fourier amplitude spectra and the luminance histograms were matched across a total of 45 images from three categories (faces, cars, and chairs; Experiment 1). SHINE was also applied in a study on race categorization, where 100 Caucasian, 100 Asian, and 100 African American faces were matched in terms of histograms and energy at each spatial frequency (Fiset et al., 2008). Furthermore, SHINE has been employed for parametrically varying the luminance properties in a study on the relative influence of luminance and featural information in race categorization (Willenbockel, Fiset, & Tanaka, 2008). Specifically, using the histogram specification function of SHINE, five luminance levels were created, ranging from the original distribution of an African American face over the average distribution to the original distribution of a Caucasian face. All of the levels were then applied to both Caucasian and African American faces. Likewise, a number of other researchers have used earlier or partial forms of these processing steps to create well-controlled image sets (e.g., Adolphs et al., 2005; Liu et al., 2002; Loschky et al., 2007; Mack et al., 2008; Sadr & Sinha, 2001, 2004; Willenbockel et al., 2010; Xu et al., 2005).

The SHINE approach has thus been successfully used for both equating and parametrically varying low-level properties, while the structure contained in the images was largely preserved. We hope that other laboratories will find SHINE as useful as we have for minimizing potential low-level confounds in studies on higher-level processes.

Acknowledgments

We would like to thank Alireza Avanaki for providing the *ssim_sens* function and two anonymous reviewers for their helpful suggestions.

Footnotes

[1] The *ssim_sens* function was obtained from A. N. Avanaki.

[2] *fft2* outputs the sin and cosin coefficients, which can be transformed into amplitudes using `abs(fft2(an_image))` and in phases using `angle(fft2(an_image))`.

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P. G., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*, 68–72.

Avanaki, A. N. (2009). Exact global histogram specification optimized for structural similarity. *Optical Review, 16*, 613–621.

Bentin, S., Taylor, M. J., Rousselet, G. A., Itier, R. J., Caldara, R., Schyns, P. G., Jacques, C., & Rossion, B. (2007). Controlling interstimulus perceptual variance does not abolish N170 face sensitivity. *Nature Neuroscience, 10*, 801–802.

Bevilacqua, A., & Azzari, P. (2007). A high performance exact histogram specification algorithm. In R. Cucchiara (Ed.), *14th International Conference on Image Analysis and Processing* (ICIAP'07) (pp. 623–628). Los Alamitos, CA: IEEE Computer Society Press.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433–436.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197*, 551–566.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition, 83*, 1–29.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research, 44*, 3223–3232.

Coltuc, D., Bolon, P., & Chassery, J.-M. (2006). Exact histogram specification. *IEEE Transactions on Image Processing*, *15*, 1143–1152.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*, 613–617.

Dakin, S. C., Hess, R. F., Ledgeway, T., & Achtman, R. L. (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Current Biology, 12*, R476–R477.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9–21.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. New York, NY: Oxford University Press.

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious processing: A comparison of face and nonface stimuli. *Psychological Science, 20*, 42–51.

Fiset, D., Blais, C., Gosselin, F., Bub, D., & Tanaka, J. W. (2008). Potent features for the categorization of Caucasian, African American, and Asian faces in Caucasian observers [Abstract]. *Journal of Vision*, *8*(6), 258a.

Fründ, I., Busch, N. A., Körner, U., Schadow, J., & Herrmann, C. S. (2007). EEG oscillations in the gamma and alpha range respond differently to spatial frequency. *Vision Research, 47*, 2086–2098.

George, N., Jemel, B., Fiori, N., & Renault, B. (1997). Face and shape repetition effects in humans: A spatio-temporal ERP study. *NeuroReport, 8*, 1417–1423.

Gold, J. G., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39,* 3537–3560.

Hardee, J. E., Thompson, J. C., & Puce, A. (2008). The left amygdala knows fear: Laterality in the amygdala response to fearful eyes. *Social Cognitive & Affective Neuroscience, 3*, 47–54.

Hershler, O., & Hochstein, S. (2006). With a careful look: Still no low-level confound to face pop-out. *Vision Research, 46*, 3028–3035.

Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of Vision, 8*, 1–13.

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex, 14*, 132–142.

Ivry, R., & Robertson, L. C. (1998). *The two sides of perception*. Cambridge, MA: MIT Press.

Johannes, S., Münte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research, 2,* 189–205.

Knebel, J.-F., Toepel, U., Hudry, J., Le Coutre, J., & Murray, M. M. (2008). Generating controlled image sets in cognitive neuroscience research. *Brain Topography, 20*, 284–289.

Liang, X., Zebrowitz, L. A., & Aharon, I. (2009). Effective connectivity between amygdala and orbitofrontal cortex differentiates the perception of facial expressions. *Social Neuroscience, 4*, 185–196.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*, 910–916.

Loschky, L. C., Sethi, A., Simons, D. J., Pydimari, T. W., Ochs, D., & Corbeille, J. L. (2007). The importance of information localization in scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1431–1450.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT press.

Mack, M. L., Gauthier, I., Sadr, J., & Palmeri, T. J. (2008). Object detection and basic-level categorization: Sometimes you know it is there before you know what it is. *Psychological Bulletin & Review, 15*, 28–35.

Metha, A. B., Vingrys, A. J., & Badcock, D. R. (1993). Calibration of a color monitor for visual psychophysics. *Behavior Research Methods, Instruments, & Computers, 25*, 371–383.

Morovic, J., Shaw, J., & Sun, P.-L. (2002). A fast, non-iterative and exact histogram matching algorithm. *Pattern Recognition Letters, 23*, 127–135.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447*, 206–209.

Olman, C., Boyaci, H., Fang, F., & Doerschner, K. (2008). V1 responses to different types of luminance histogram contrast [Abstract]. *Journal of Vision*, *8*(6), 345a.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31*, 1337–1350.

Poynton, C. A. (1993). "Gamma" and its disguises: The nonlinear mappings of intensity in perception, CRTs, film and video. *SMPTE Journal, 102*, 1099–1108.

Rolland, J. P., Vo, V., Bloss, B., & Abbey, C. K. (2000). Fast algorithms for histogram matching: Application to texture synthesis. *Journal of Electronic Imaging, 9*, 39–45.

Rose, J., Otto, T., & Dittrich, L. (2008). The Biopsychology-Toolbox: A free, open-source Matlab-toolbox for the control of behavioral experiments. *Journal of Neuroscience Methods, 175*, 104–107.

Rossion, B., Gauthier, I., Tarr, M. J., 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*, 69–74.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision, 8*, 1–18.

Rousselet, G. A., Macé, M. J. M., Thorpe, S. J., & Fabre-Thorpe, M. (2007). Limits of event-related potential differences in tracking object processing speed. *Journal of Cognitive Neuroscience, 19*, 1241–1258.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., & Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience, 9*, 98.

Sadr, J., & Sinha, P. (2001). *Exploring object perception with Random Image Structure Evolution*. MIT Artificial Intelligence Laboratory Memo No. 2001–6.

Sadr, J., & Sinha, P. (2004). Object recognition and Random Image Structure Evolution. *Cognitive Science, 28,* 259–287.

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., Landis, T., & Schomer, D. L. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *NeuroReport, 8*, 2749–2754.

Stanislaw, H., & Olzak, L. A. (1990). Parametric methods for gamma and inverse gamma correction with extensions to halftoning. *Behavior Research Methods, Instruments, & Computers, 22*, 402–408.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12*, 43–47.

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007a). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience, 10*, 505–511.

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007b). Is the N170 sensitive to the human face or to several intertwined perceptual and conceptual factors? *Nature Neuroscience, 10,* 802–803.

VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research, 46*, 3017–3027.

VanRullen, R., & Thorpe, S. J. (2001). The time course of visual processing: From early perception to decision-making. *Journal of Cognitive Neuroscience, 13*, 454–461.

Wan, Y., & Shi, D. (2007). Joint exact histogram specification and image enhancement through the wavelet transform. *IEEE Transactions on Image Processing, 16,* 2245–2250.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2003). *The SSIM index for image quality assessment*. Retrieved January 27, 2010, from www.ece.uwaterloo.ca/~z70wang/research/ssim/

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13,* 600–612.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C. Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Willenbockel, V., Fiset, D., & Tanaka, J. W. (2008). *The role of luminance and facial features in race categorization*. Poster presented at the North West Cognition and Memory 10th Annual Meeting, Seattle, USA.

Williams, N. R., Willenbockel, V., & Gauthier, I. (2009). Sensitivity to spatial frequency and orientation content is not specific to face perception. *Vision Research, 49*, 2353–2362.

Xu, Y., Liu, J., & Kanwisher, N. (2005). The M170 is selective for faces, not for expertise. *Neuropsychologia, 43*, 588–587.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Zion-Golumbic, E., Golan, T., Anaki, D., & Bentin, S. (2008). Human face preference in gamma-frequency EEG activity. *NeuroImage, 39*, 1980–1987.

Table captions

Table I. Algorithm for exact global histogram specification

Table II. SSIM optimization algorithm of Avanaki (2009). μ is a positive constant denoting the step size, ∇_y represents gradient with respect to image Y, and N is the number of pixels in X. SSIM(X,Y) is a measure of perceptual similarity of X and Y. The convergence criterion in Step 3 may be given by desired output quality [i.e., SSIM(X,Y) > threshold)] or by a certain number of iterations (or a combination thereof). The updated X_{new} in Step 4 is more similar to X, because each pixel of Y was altered so that the SSIM was increased. The histogram of the output image remains exactly as specified.

Figure captions

Figure 1. Overview of SHINE toolbox functions and their calling structure. For example, the *SHINE* function calls the *specMatch* function, which calls the *rescale* function.

Figure 2. Illustration of the basic luminance histogram-matching approach using simple patterns. The histograms show luminance (in arbitrary linear units [ALU]) on the x-axis and the number of pixels on the y-axis. The target histogram was computed by averaging the histograms of the two input images. a) Histogram matching was performed on two uniform surfaces (100 ALU and 200 ALU) of slightly different sizes centered on a uniform gray background (127 ALU). The output surfaces contain 50% of 100-ALU pixels and 50% of 200-ALU pixels in randomized order. The background remained unaltered. b) Two surfaces of the same size as in panel a) with 50% 100-ALU and 50% 200-ALU pixels served as input. The output is identical to the input. c) Illustration with surfaces similar to panel b) but with different amounts of dark and light pixels. d) Illustration with input surfaces of three different luminance levels (100, 200, and 230 ALU) with different numbers of pixels of each.

Figure 3. Illustration of the *match* function of SHINE. a) Two base face images with their luminance histograms (left) and the corresponding SHINE output images with their matched histograms (right). The target histogram was obtained using the function *tarhist*. b) Histogram matching was performed on two natural scenes. As in panel a), the function *tarhist* was used to obtain the target histogram. c) Histogram specification is illustrated for a face and a greeble. Here, the target was computed independently of the two input images by averaging the histograms of several faces (not shown) using *avgHist*. As a result, the greeble was altered more than the face when applying the target.

Figure 4. Illustration of the *sfMatch* and *specMatch* functions of SHINE. a) Three source images (a car, a chair, and a face) are shown with their Fourier spectra displayed as polar plots, where energy is plotted as a function of spatial frequency (distance from the origin) and orientation (angle). The log-log plot on the right depicts the rotational average of the spectra (i.e., the energy at each spatial frequency, in cycles per image, cpi, averaged across orientations). b) Using *sfMatch*, the rotational average of the Fourier spectra was equated while the energy distribution across orientations was preserved (see the text for details). c) Using *specMatch*, the Fourier spectra were equated on spatial frequencies and orientations.

The output in panels b) and c) is shown after the rescaling of the luminance values so that absolutely all pixels of the three images are in the range of 0–255.

Figure 5. Iterative histogram and Fourier amplitude matching. The histograms of three sample images (a car, a chair, and a face) are shown before applying SHINE (top), after one iteration (middle), and after three iterations (bottom). Each iteration consisted of equating the histograms using *histMatch* (with the average histogram serving as the target) and then equating the rotational average of the Fourier spectra using *sfMatch*. The latter step altered the histograms; however, after a number of iterations, the histograms typically converge toward the target.

Figure 6. The iterative equalization approach illustrated for two typical image pairs (Sets 1 and 2). Left: a) For each set, the root-mean-square error (RMSE) between the two images' Fourier spectra is plotted over Iterations 1–20. The dotted line shows the results for Fourier amplitude spectrum matching (*specMatch*) followed by histogram matching (*histMatch*) and the full line for *histMatch* followed by *specMatch*. b) Analogous to a), the RMSE between histograms is shown for each image pair over iterations. c) The average structural similarity (SSIM) index between the original images and the corresponding SHINE output is plotted for each set over Iterations 1–20. Right: The output images of Set 1 (left column) and Set 2 (right column) are shown after 1 and after 20 iterations of *histMatch* followed by *specMatch* or vice versa.

Tables

Table I: Algorithm for exact global histogram specification

Step 0	Let the target histogram be $H = \{h_0, h_1, \ldots, h_{D-1}\}$, where D is the number of possible luminance values (e.g., 256 in an 8-bit image). It is assumed that $\sum_{i=0}^{D-1} h_i = N$, where N equals the number of pixels in the image. If this assumption is not met, scale H (and perhaps round some h_i) to satisfy this assumption.
Step 1	Sort the pixels of the source image by their luminance value in ascending order (and in the cases in which the values are the same, randomize the pixels).
Step 2	Starting from the first pixel on the sorted list, assign the first h_0 pixels a new value of 0. Continue by assigning the next h_1 pixels a new value of 1, and so on until all pixels are allotted their new values.

Table II: SSIM optimization algorithm of Avanaki (2009)

Step 0	Let X show the original input image. Set $X_{new} = X$.
Step 1	Apply the *match* function to X_{new} to obtain image Y with given histogram H.
Step 2	Compute $\nabla_y \text{SSIM}(X,Y)$ and $\text{SSIM}(X,Y)$ using the *ssim_sens* function.
Step 3	If convergence is reached, then break.
Step 4	Set $X_{new} = Y + \mu N \nabla_y \text{SSIM}(X,Y)$ and go to Step 1.
Step 5	Output Y.

Figures

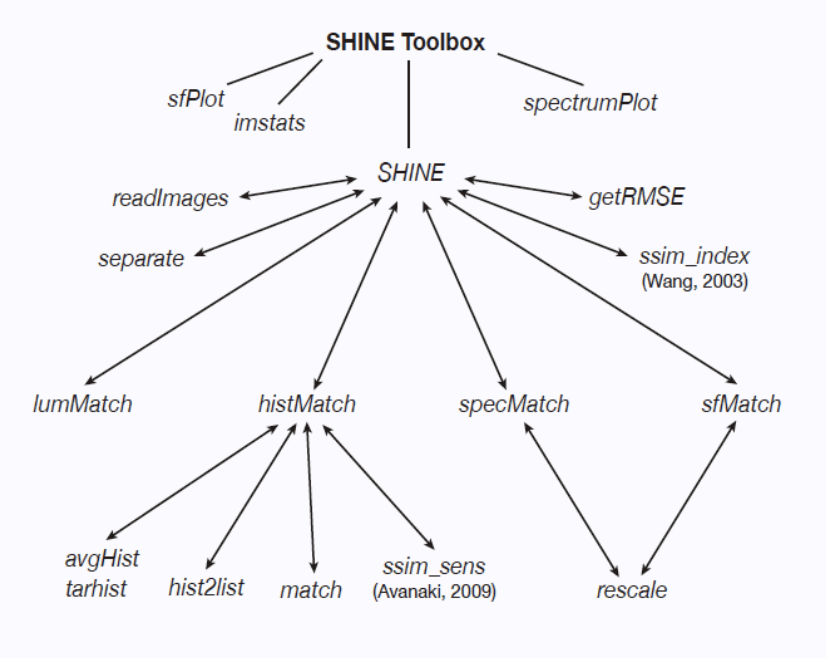


Figure 1: Overview of SHINE toolbox functions

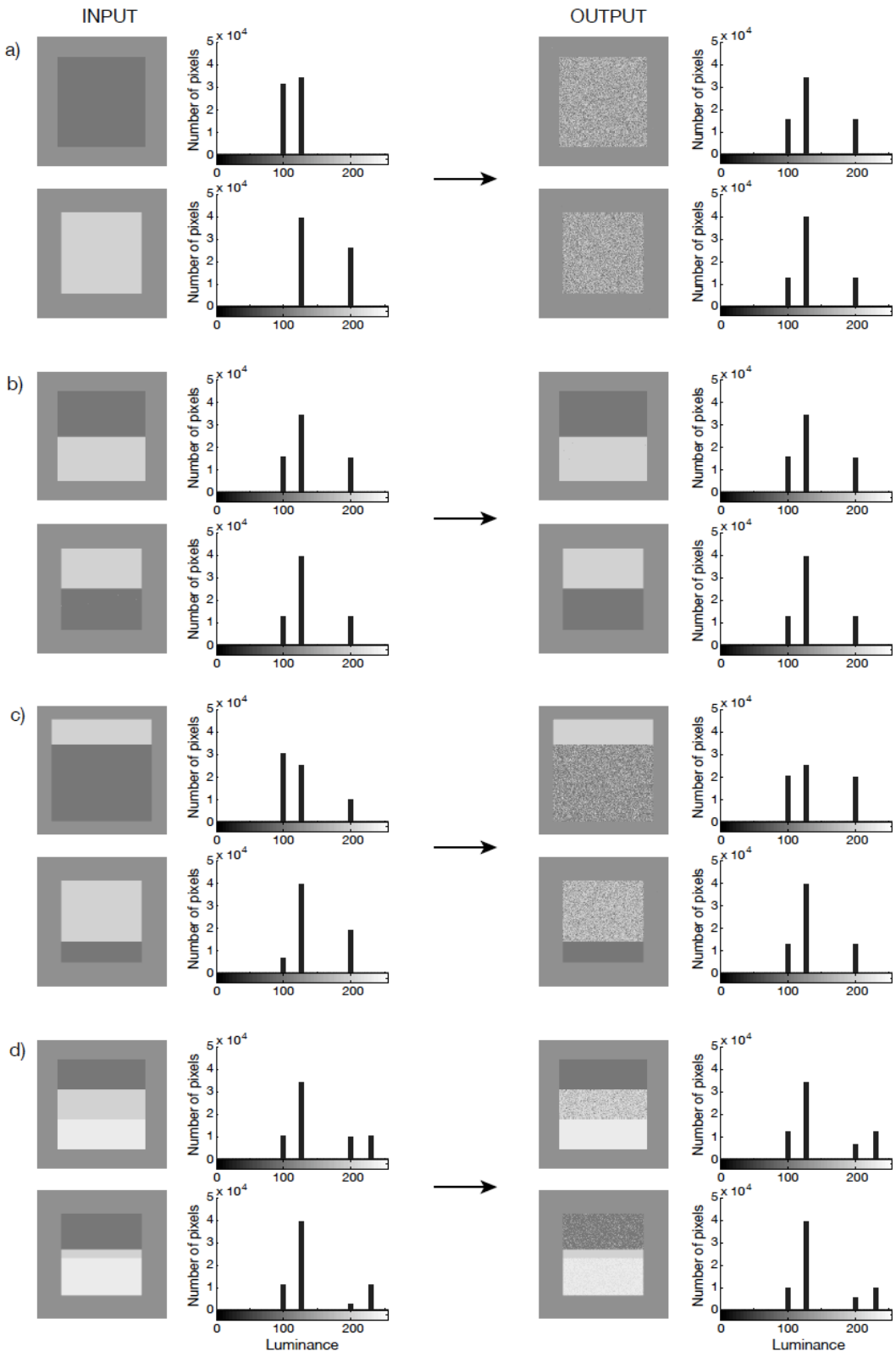


Figure 2: Illustration of luminance histogram matching

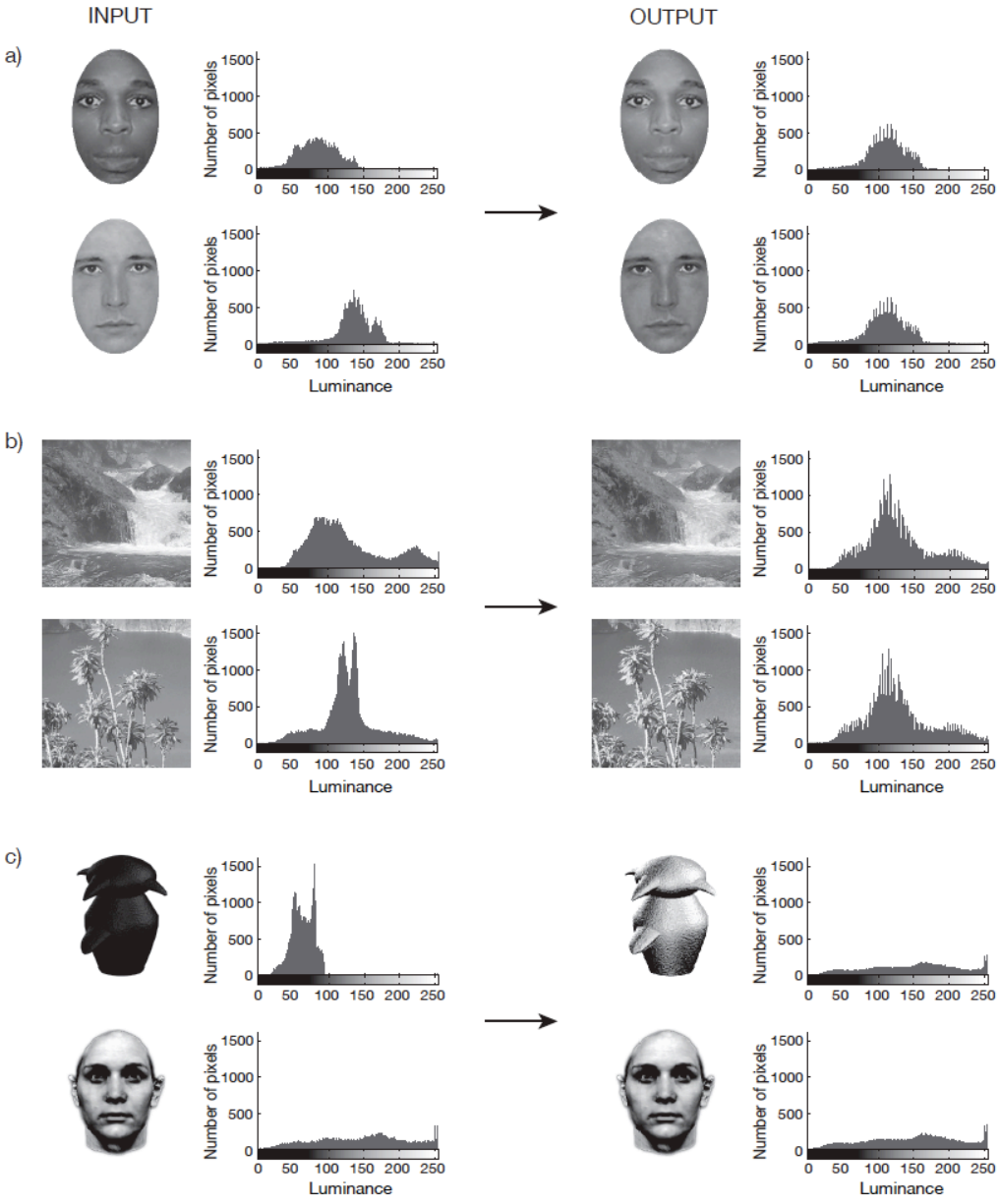


Figure 3: Illustration of the *match* function of SHINE

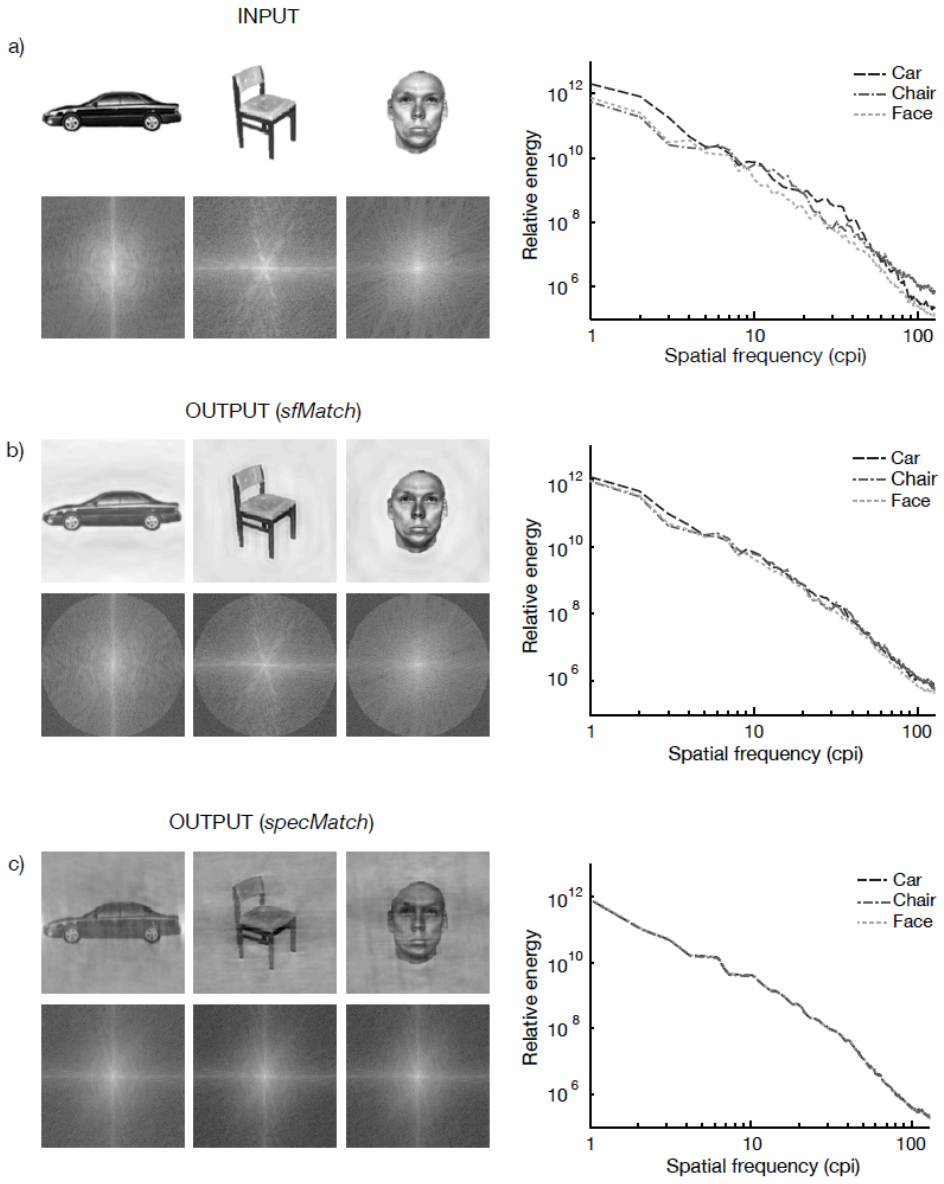


Figure 4: Illustration of the *sfMatch* and *specMatch* functions of SHINE

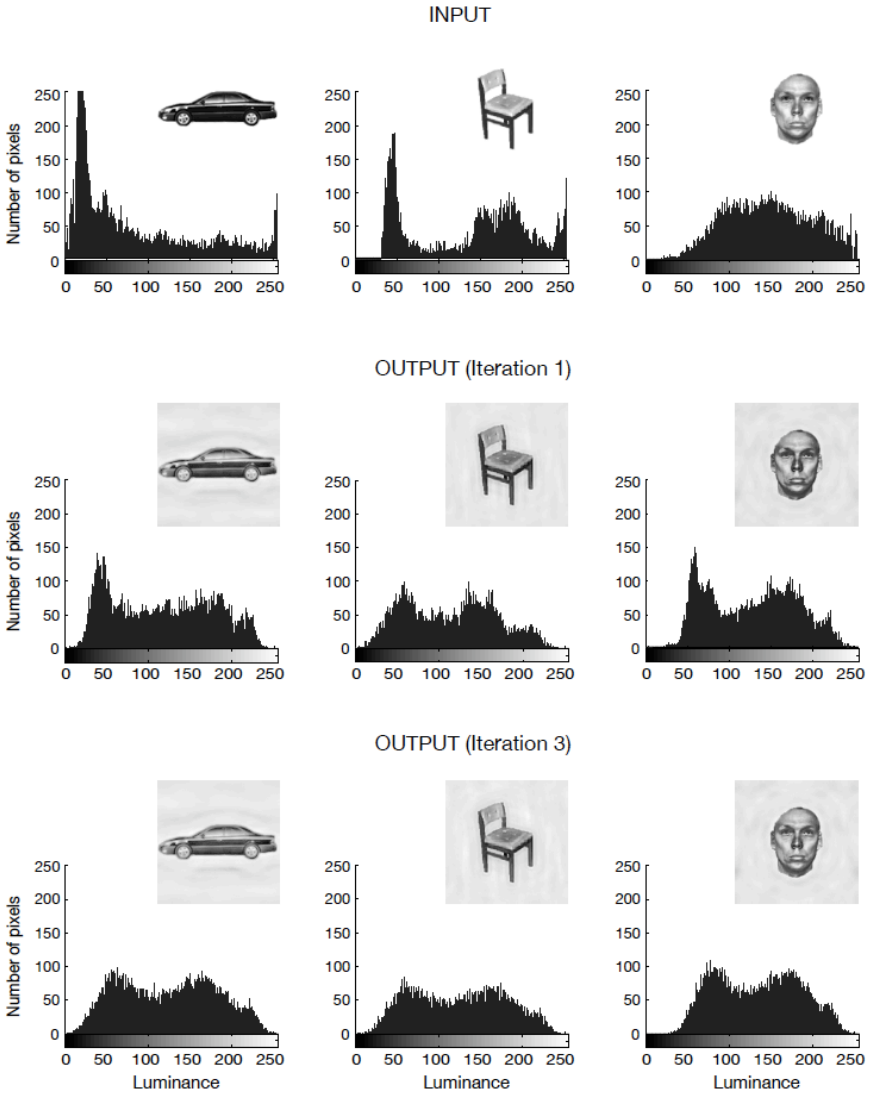


Figure 5: Iterative histogram and Fourier amplitude matching

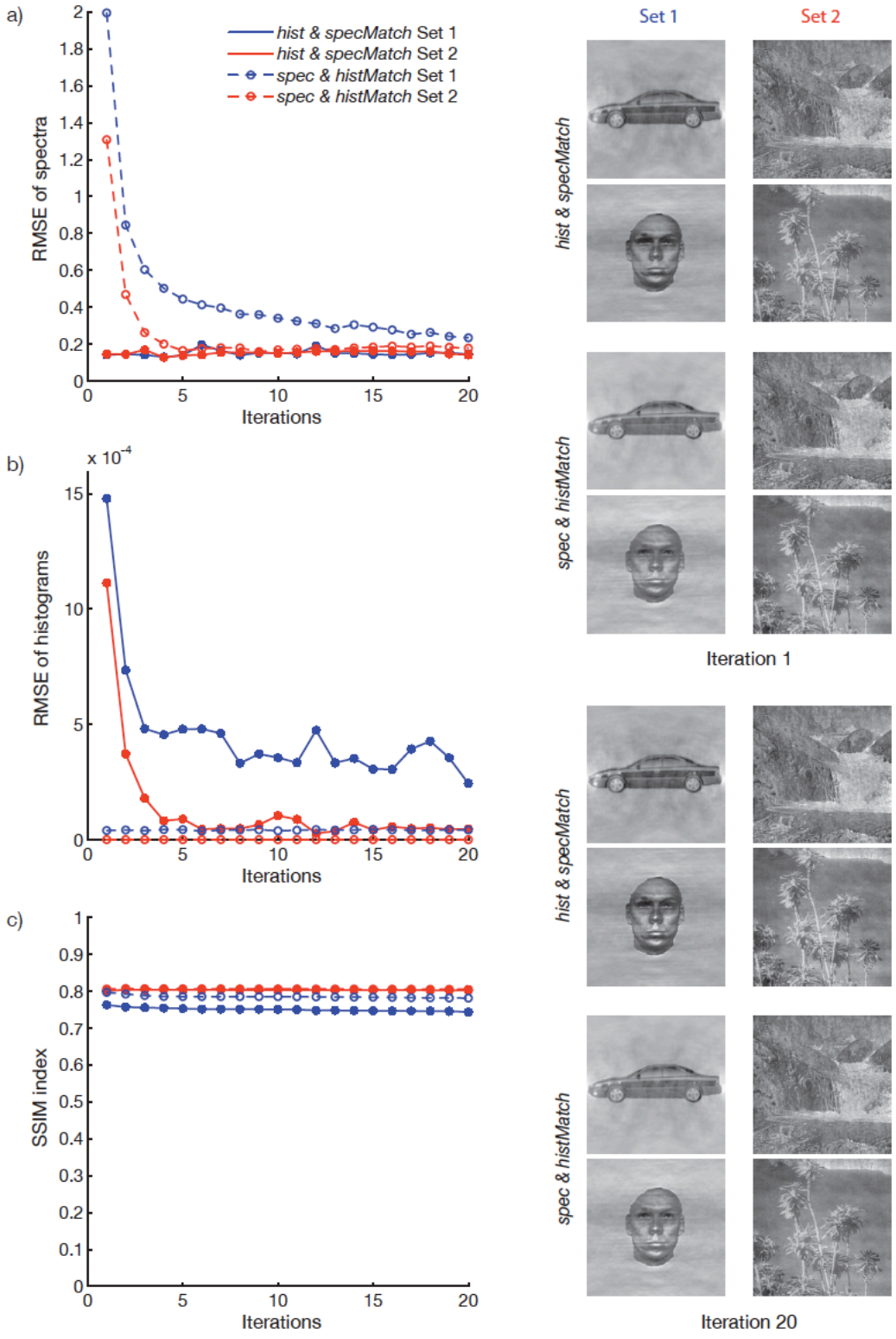


Figure 6: Iterative equalization and image quality

Transitional text # 1

Chapter 2 has provided background information on low-level image processing and introduced the SHINE toolbox for creating controlled stimulus sets. In the remainder of this thesis, SHINE functions are employed to equate luminance and/or the spectral content of our stimuli in studies on face perception. In Chapter 3, the SHINE article is not cited because it was published after the article that constitutes the chapter was accepted for publication; however, SHINE functions were employed to equate the low-level properties within the two sets of faces used in that study.

Chapter 3 serves two purposes. First, like Chapter 2, it makes a methodological contribution to visual perception research: It describes and validates the SF Bubbles technique, which allows for high-resolution, unbiased SF sampling. SF Bubbles is used in all the following studies of this thesis. Second, Chapter 3 serves as an introduction to SF tuning during human face perception. The main objective of the study was to reveal the precise SF tuning curves for the conscious identification of faces presented at different orientations using behavioral measures.

Portions of the research in Chapter 3 constitute my Masters thesis at the University of Victoria, BC (see *Author note*). This is why the findings of this project are given a relatively small weight in this thesis. The manuscript is included here, however, because the project continued on during my PhD (e.g., we added an experiment, re-analyzed the data, and revised the manuscript for publication), and because SF Bubbles is an integral part of the following studies on face perception and awareness in Chapters 4 and 5.

Chapter 3: Second article

Article published:

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Does face inversion change spatial frequency tuning?

Verena Willenbockel[1,3], Daniel Fiset[1,3], Alan Chauvin[2], Caroline Blais[3], Martin Arguin[3], James W. Tanaka[1], Daniel N. Bub[1], & Frédéric Gosselin[3]

[1]Department of Psychology, University of Victoria, Victoria, BC, Canada
[2]URECA, UFR de Psychologie, Université Ch. De Gaulle Lille III, Villeneuve d'Ascq, France
[3]CERNEC, Département de Psychologie, Université de Montréal, Montréal, QC, Canada

Abstract

We examined spatial frequency (SF) tuning of upright and inverted face identification using an SF variant of the Bubbles technique (Gosselin & Schyns, 2001). In Experiment 1, we validated the SF Bubbles technique in a plaid detection task. In Experiments 2a–c, the SFs used for identifying upright and inverted inner facial features were investigated. Although a clear inversion effect was present (mean accuracy was 24% higher and response times 455 ms shorter for upright faces), SF tunings were remarkably similar in both orientation conditions (mean r = .98; an SF band of 1.9 octaves centered at 9.8 cycles per face width for faces of about 6˚). In Experiments 3a and b, we demonstrated that our technique is sensitive to both subtle bottom-up and top-down induced changes in SF tuning, suggesting that the null results of Experiments 2a–c are real. The most parsimonious explanation of the findings is provided by the quantitative account of the face inversion effect: The same information is used for identifying upright and inverted inner facial features, but processing has greater sensitivity with the former.

Keywords: face perception, identification, spatial frequency, inversion effect

Introduction

Human adults are able to recognize thousands of faces rapidly and effortlessly, as long as the faces are seen upright (e.g., Maurer, Le Grand, & Mondloch, 2002). When stimuli are rotated by 180° in the picture plane, face recognition accuracy decreases considerably (e.g., Hochberg & Galper, 1967; Yin, 1969) and response latencies increase (e.g., Diamond & Carey, 1986). This drop in recognition performance was found to be disproportionately larger for faces than for other mono-oriented objects (e.g., houses and airplanes) in untrained individuals and is commonly referred to as the *face inversion effect* (FIE; e.g., Boutet, Collin, & Faubert, 2003; Leder & Carbon, 2006; Robbins & McKone, 2007; Yin, 1969; reviews in Rossion & Gauthier, 2002, and Valentine, 1988). The FIE can be observed for unfamiliar and familiar faces (Collishaw & Hole, 2000; Rock, 1974; Yarmey, 1971), for full faces (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006, Experiment 1) or inner facial features (faces shown through an elliptical aperture; e.g., Goffaux & Rossion, 2006, Experiments 2 and 3; Robbins & McKone, 2003; Sekuler, Gaspar, Gold, & Bennett, 2004), and in a variety of experimental conditions; for instance, it has been demonstrated using blocked or randomized presentation of upright and inverted faces, in old/new recognition tasks, and in matching tasks with or without delay (for reviews, see Rossion, 2008; Rossion & Gauthier, 2002). The FIE has been described as one of the most robust phenomena in the face processing literature.

Given that upright and inverted faces have the same complexity and are almost identical in their low-level properties, such as luminance, contrast, and spatial frequencies (only phase information differs), the difficulty of recognizing upside-down faces cannot easily be attributed to stimulus properties per se. Typically, the FIE has been attributed to qualitative processing differences between upright and inverted faces, that is, inversion is thought to disrupt certain face processes more than others (e.g., Rossion, 2008). In particular, it has been suggested that inversion mainly impairs the perception of the relative distances between facial features (Diamond & Carey, 1986; Leder & Bruce, 1998, 2000; Rhodes, 1988) or holistic processing—"the simultaneous integration of the multiple features of a face into a single perceptual representation" (Rossion, 2008, p. 275; see also Farah, Tanaka, & Drain, 1995; Sergent, 1984; Tanaka & Farah, 1993). Several other studies have demonstrated differential effects of inversion on the processing of the local facial features themselves versus their

configuration or integration into a holistic representation (e.g., Bartlett & Searcy, 1993; Barton, Keenan, & Bass, 2001; Collishaw & Hole, 2000; Freire, Lee, & Symons, 2000; Leder, Candrian, Huber, & Bruce, 2001; Le Grand, Mondloch, Maurer, & Brent, 2001; Rhodes, Brake, & Atkinson, 1993; Rossion & Boremanse, 2008; Tanaka & Sengco, 1997; Thompson, 1980; Young, Hellawell, & Hay, 1987), thus supporting the qualitative view of the FIE.

Despite a wide acceptance of the qualitative account, it has also been proposed that the FIE arises from quantitative processing differences. Using a response classification method (Ahumada & Lovell, 1971; Eckstein & Ahumada, 2002), Sekuler et al. (2004) found that the eye and eyebrow regions contain the features correlated linearly with the discrimination of both upright and inverted faces, and that the correlations associated with upright faces were greater than those associated with inverted faces. These findings are in line with previous behavioral studies that identified the eye and eyebrow regions as most important for face identification (e.g., Davies, Ellis, & Shepherd, 1977; Gold, Sekuler, & Bennett, 2004; Gosselin & Schyns, 2001, 2005; Haig, 1985, 1986; Schyns, Bonnar, & Gosselin, 2002; Shepherd, Davies, & Ellis, 1981; Walker-Smith, Gale, & Findlay, 1977). They are also consistent with eye-tracking data demonstrating that gaze is especially attracted by the eyes of both upright and inverted faces (e.g., Williams & Henderson, 2007; but see Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006, for a demonstration of more random global scan patterns and more fixations on lower face parts in the inverted condition). Furthermore, Sekuler and colleagues reported that contributions of nonlinear processes to performance were small and similar for both face orientations. The authors concluded from their classification image results that a quantitative, not a qualitative, difference underlies the FIE (see also Gaspar, Bennett, & Sekuler, 2008; Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Valentine, 1988; Yovel & Kanwisher, 2004; but see Rossion, 2008)—specifically, that the same facial information is processed with less sensitivity when faces are upside-down.

The classification images of Sekuler et al. (2004), however, are insufficient to determine whether the same cues subtend upright and inverted face recognition beyond the spatial domain. For example, they cannot rule out that the eye and eyebrow regions revealed in the upright and inverted classification images hide different patterns of spatial frequency (SF) use (e.g., 8 cycles per face width for upright eye and eyebrow regions and 16 cycles per face

width for the same regions but inverted). It thus remains possible that qualitative processing differences can be found in the SF domain.

Spatial frequency tuning for face identification

The human visual system analyzes the complex luminance variations that make up the visual stimulus with discrete channels, each tuned to a specific SF range (see De Valois & De Valois, 1990, for a review). High SFs represent the fine-grained information in a stimulus, such as the eyelashes or edges of the mouth, and low SFs convey coarse information, such as luminance blobs and blurred shapes (see Morrison & Schyns, 2001, and Ruiz-Soler & Beltran, 2006, for reviews). Several studies found that a narrow band of intermediate SFs centered between 7 and 16 cycles per face (cpf) is particularly important for upright face recognition (e.g., Costen, Parker, & Craw, 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999).

For the SF tuning of inverted face identification, mixed findings have been obtained. Collishaw and Hole (2000) demonstrated that blurred faces (i.e., faces containing only low-SF information) could still be recognized above chance level, unless they were presented upside-down. Because inversion of a scrambled face (a condition of comparable difficulty) did not decrease performance below chance, the findings were attributed to a disruption of the processing of the relative distances between facial features (by inversion) and the featural information (by blurring) rather than task difficulty. These findings suggest that the FIE might be particularly large for low-pass filtered faces. Although their study did not directly pertain to face inversion, Goffaux and Rossion (2006; Experiments 2 and 3) demonstrated in a composite face paradigm that inversion costs in both accuracy and RT were larger for low-SF faces (< 8 cpf) compared with broad-spectrum and high SF (> 32 cpf) stimuli. Moreover, Nakayama (2003) reported that face discrimination (in a 4-choice task) was most degraded by noise of approximately 12 cpf, with a much broader masking function for inverted than for upright faces. Finally, using an old/new recognition task, Boutet et al. (2003, Experiment 1) found FIEs of comparable magnitude on accuracy for broadband faces, band-pass filtered faces in a range of relatively low SFs (1.25–5 cpf), and intermediate-SF (5–20 cpf) faces, but not for high-SF (20–80 cpf) faces, which did not lead to a significant FIE. These findings, albeit somewhat inconsistent, support the qualitative view of the FIE.

In contrast to the above, Boutet et al. (2003, Experiments 2A and B) found a comparable FIE on accuracy in all SF band-pass filter conditions (1.25–5 cpf; 5–20 cpf; 20–80 cpf) in a sequential matching paradigm and concluded—considering the results of their Experiments 1–3—that SF filtering had little impact on the FIE. Furthermore, Collin, Liu, Troje, McMullen, and Chaudhuri (2004) demonstrated that varying the degree of SF overlap between two faces influences accuracy in a matching task similarly for upright and inverted faces. Finally, using critical band masking—a technique that measures signal thresholds for stimuli to which high-pass (or low-pass) filtered white Gaussian noise is added at different cutoffs—Gaspar, Sekuler, and Bennett (2008) showed that the same narrow SF band (approximately 1.5 octaves wide and centered at about 7 cpf) was used in two 10-choice identification tasks with both upright and inverted faces. These results are consistent with a quantitative view of the FIE. In sum, both orientation dependency and independency of SF use has been reported in the literature.

Here, we re-examined whether the identification of upright and inverted faces is mediated by different SFs using a novel SF variant of the Bubbles technique (Fiset, Blais, Gosselin, & Schyns, 2006; Gosselin & Schyns, 2001; see McCotter, Gosselin, Sowden, & Schyns, 2005, for a distinct attempt at applying Bubbles to SFs). The Bubbles technique has been applied to full-spectrum images (e.g., Gosselin & Schyns, 2001, Experiment 1) as well as band-pass filtered images (e.g., Gosselin & Schyns, 2001, Experiment 2; Schyns et al., 2002). However, it has not previously been used to sample the SF content of stimuli. We randomly varied the availability of SFs on a trial-by-trial basis, thereby testing the contribution of each SF to performance independently. After a sufficient number of trials, we performed multiple linear regressions on the random SF filters and response accuracy or response time (RT).

The SF Bubbles technique allowed us to derive the precise SF tuning curves for upright and inverted face identification. In comparison, traditional high-pass, low-pass, or band-pass filtering techniques (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006) provide only crude estimates of SF tuning. The main strength of the SF Bubbles technique in the present context is that it minimizes the risk that participants adapt to a predictable stimulus manipulation (e.g., low-, band-, or high-pass filtering or critical band masking; see Sowden & Schyns, 2006, for evidence of "channel surfing"), by randomly sampling multiple SFs simultaneously on a trial-by-trial basis.

The present research comprised six experiments. The first experiment was designed to assess the validity of the SF Bubbles technique. Using a plaid (i.e., the sum of two sine wave gratings) detection task, we verified that the SF Bubbles method can uncover precisely the diagnostic SFs. The next three experiments, which constitute the main experiments, examined which SFs are used for the accurate (Experiment 2a and c) and fast (Experiment 2b) identification of upright and inverted faces. The experiments employed two sets of 20 grayscale face photos (10 identities × 2 exemplars per set) that were cropped to an elliptical shape. We decided to focus on inner facial features to bridge the qualitative and quantitative accounts of the FIE as much as possible (e.g., Gaspar, Sekuler, & Bennett, 2008; Goffaux & Rossion, 2006, Experiment 2 and 3; Robbins & McKone, 2003; Sekuler et al., 2004, used faces revealed through an elliptical aperture). To anticipate our main result, we find no difference in SF use between the upright and inverted condition. To rule out that this null result was due to an insensitivity of the SF Bubbles technique to bottom-up or top-down influences on SF tuning, we carried out two more experiments. Experiment 3a re-examined SF tuning in the identification task of Experiment 2a as a function of stimulus size, which is known to influence SF tuning in a bottom-up fashion (e.g., Majaj, Pelli, Kurshan, & Palomares, 2002). In Experiment 3b, task demands were modified (gender or happy/neutral discrimination) to modulate SF tuning in a top-down fashion (e.g., Schyns & Oliva, 1999). In both cases, we revealed subtle differences in SF use, confirming that the SF Bubbles technique is sensitive to bottom-up and top-down induced changes in SF tuning.

Experiment 1

The purpose of the first experiment was to determine whether the SF Bubbles technique can reveal precisely the SFs that convey the information that is diagnostic for the task. We employed a plaid (i.e., the sum of two SFs) detection task: If the SF Bubbles technique works adequately, then we should be able to recover the SFs in the plaid. This experiment is also meant as an illustration of the SF Bubbles technique.

Method

Participants. One male and two female University of Victoria students (22–25 years old; M = 24 years) participated in Experiment 1. All participants had normal or corrected-to-normal vision, and two were naïve to the purpose of the experiment. The third participant was

the first author of the article (Observer 1). Participants gave informed consent approved by the University of Victoria Human Research Ethics Committee.

Apparatus. Experiments 1, 2a, 2b, 3a, and 3b were run on a dual core 2.93 GHz PC at the University of Victoria. Stimuli were displayed on a 22-inch Viewsonic CRT monitor that was calibrated to allow a linear manipulation of luminance. The resulting corrected table contained 154 luminance levels, ranging from 0.3 cd/m^2 to 98.7 cd/m^2. The background luminance was 49.3 cd/m^2. The monitor refresh rate was 85 Hz for all experiments, and the resolution was set to 1024 × 768 pixels (except in Experiment 3a). All experiments were programmed in MATLAB (Natick, MA, USA) using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Procedure. Participants were instructed to perform a plaid detection task. The original plaid comprised a horizontal sine wave grating with an SF of 10 cycles per image (cpi) and a vertical sine wave grating with an SF of 45 cpi (Figure 1) and had a size of 256 × 256 pixels. On "signal present" trials (50% of trials), the SFs of the plaid were randomly sampled (see next section) and displayed embedded in white Gaussian noise. On "signal absent" trials, a white Gaussian noise field of 256 × 256 pixels was displayed. Each trial began with a central fixation cross lasting 435 ms, followed by the stimulus presented for 870 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded by pressing an appropriate key on a computer keyboard. Signal present and signal absent trials occurred in random order. No feedback was provided. Each observer performed eleven 100-trial blocks with breaks between blocks. Participants were seated in a dark room and a chin rest was used to maintain viewing distance at 53 cm; stimuli subtended a visual angle of 10.2˚ × 10.2˚.

Spatial frequency bubbles. All experiments reported in this article revealed SF use by employing the SF Bubbles technique. This section serves two purposes: (a) It describes the SF Bubbles technique in general, and (b) it illustrates the use of the technique in Experiment 1.

On each trial, the SF information of a stimulus was sampled randomly as illustrated in Figure 1. First, the square base stimulus was padded to minimize edge artifacts in the SF domain. It was centered on a uniform gray field of the stimulus' background luminance and twice its size. In Experiment 1, for example, the plaid of size 256 × 256 pixels was padded with a midluminance background of size 512 × 512 pixels. Second, the padded stimulus was

Fourier transformed using functions from the Image Processing Toolbox for MATLAB. The quadrants of the Fourier image were shifted so that low SFs occupied the central region of the complex (i.e., real + imaginary number) amplitude matrix. Third, a random filter was constructed. The construction of this filter involved the following steps: (a) A binary random vector of $2wk$ elements was created, where w was the stimulus width and k a constant that determined the smoothness of the sampling (the higher k, the smoother); k was arbitrarily set to 20 for all the experiments reported in this article. In Experiment 1, the random vector thus had 10,240 elements ($2 \times 256 \times 20$). The vector contained zeros among b ones that were randomly distributed (with repetition); b determined the number of SF bubbles (see below) and was arbitrarily set to 45. (b) To create a smooth filter, the binary vector was convolved with a Gaussian kernel, referred to as an *SF bubble*. The standard deviation of the SF bubble (σ) was arbitrarily set to 1.5 and the maximum to 0.125; all values of the resulting vector above 1 were reduced to 1. The convolution resulted in a "sampling vector" consisting of b randomly located SF bubbles. (c) To ensure that the sampling vector approximately fit the SF sensitivity of the human visual system (see De Valois & De Valois, 1990, for a review), we subjected the smoothed vector to a logarithmic transformation: w elements of the vector were sampled according to the function of

$$f(x) = e^{(x-1)} \frac{\ln(kw)}{(w-1)} + a,$$

with $x = [1{:}w]$ and $a = kw/2$. The constant a was arbitrarily chosen and prevented low and high SFs being sampled less often than intermediate SFs. (d) The resulting w-element filter was then rotated about its origin to create an isotropic random two-dimensional filter of size $w \times w$.

Fourth, filtering was carried out by dot-multiplying the two-dimensional filter with the complex amplitude of the padded base stimulus before subjecting the result to the inverse Fourier transform. We constructed the experimental stimuli by cropping the central $w \times w$ pixel region of the filtered image. White Gaussian noise was added to the SF sampled stimulus to adjust performance. The $w \times w$ noise field was multiplied by $1–c$, with c ranging from 0 to 1 and added to the image multiplied by c. The value of c was increased or decreased on a block-by-block basis by the experimenter (in increments of .02; Experiments 2a–c and 3a) or on a trial-by-trial basis using QUEST (Watson & Pelli, 1983; Experiments 1 and 3b). For example, in Experiment 1, performance was maintained at 75% correct.

To find out which SFs drove the observers' correct and incorrect responses or response times, we performed a multiple linear regression on the random binary vector (see above) and transformations of the observers' correct and incorrect responses or RTs. Here, a multiple linear regression is linearly related to summing all sampling vectors weighted by the transformed responses. Throughout this article, correct and incorrect responses were transformed as follows: Correct responses were given a value of 1–*P*(correct) and incorrect responses a value of –*P*(correct). In Experiment 1, for example, correct responses were assigned a value of 0.25 and incorrect responses a value of -0.75. Similarly, fast responses (RTs shorter than the median RT of the corresponding block) were given a value of 1 and slow responses (those longer than the median RT) were given a value of –1.

The vector of *w* regression coefficients—referred to as *classification vector*—was then transformed into Z-scores for each observer. A group classification vector can be computed by summing the classification vectors of all observers and dividing the resulting vector by $\sqrt{n}$, with *n* equal to the number of observers. A pixel test was used to determine a statistical threshold (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). Note that because of padding, the *x*th element of the classification vector corresponds to *x*/2 cycles per base stimulus width; in this article, all SFs are given either relative to the base stimuli (Experiment 1) or to face width (Experiments 2a–c and 3a and b).

Results and discussion

The first block was considered as practice block and was excluded from the analysis. Results are shown in Figure 2, which plots the Z-scores across all SFs (classification vectors) as well as the significance threshold ($p < .05$; $S_r = 128$; $FWHM = 3.53$; $Z_{crit} = 3.45$; for details, see Chauvin et al., 2005). Individual and group results were very similar (*r*s = .93, .99, and .96); we thus report the group results only. Two significant peaks occurred: the first peak at 10 cpi ($Z_{max} = 10.50$) and the second at 45 cpi ($Z_{max} = 6.87$), with an octave width of 0.42 and 1.39, respectively. The peaks correspond to the two SFs of the plaid, thus demonstrating that the SF Bubbles technique can accurately reveal the SF information that drives the observers' responses.

Experiments 2a–c

Experiments 2a–c were designed to investigate which SFs underlie the identification of upright and inverted faces. For upright faces, several studies demonstrated that a narrow band of intermediate SFs (i.e., centered between 7 and 16 cpf, depending on the size of the face stimuli) is particularly important for recognition (e.g., Costen et al., 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999). For inverted face identification, less is known about the exact SF range used. Previous results indicate that face inversion might be particularly detrimental for low-SF faces (e.g., Collishaw & Hole, 2000; Goffaux & Rossion, 2006), that the SF band used for inverted faces might be much broader than for upright faces (Nakayama, 2003), or that SF tuning might be very similar for upright and inverted face identification (Gaspar, Sekuler, & Bennett, 2008). We tested the independent contribution of each SF to the identification of upright and inverted inner facial features, allowing us to reconstruct the precise SF filters used for effective identification. We used both accuracy (Experiment 2a and c) and RT (Experiment 2b) as measures of effective identification and employed the SF Bubbles technique described above to reveal the SF tuning curves.

Method

Participants. In total, 15 university students (10 women) ages 19–35 years (M = 25.5 years) took part in Experiments 2a–c. Four participants completed both Experiments 2a and b, and 1 of them additionally took part in Experiment 2c. One student participated in Experiment 2a only and one in 2b only. Nine participants took part in Experiment 2c only. All observers had normal or corrected-to-normal vision. Two participants are among the authors of this article, and 13 participants were naïve to the purpose of the study. They received course credit or were paid as compensation.

Apparatus. For Experiments 2a and b, see the corresponding section of Experiment 1. Experiment 2c was run on a dual core 2.2 GHz PC at the Université de Montréal. Stimuli were presented on a 22-inch Hewlett Packard CRT monitor that was calibrated to allow a linear manipulation of luminance. The corrected table contained 117 luminance levels, ranging from 0.3 cd/m^2 to 70.7 cd/m^2; the background luminance was 29.3 cd/m^2. The monitor refresh rate

was 85 Hz, and the resolution was set to 1024 × 768 pixels. The experiment was run in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Stimuli. Two sets of grayscale face photos of 256 × 256 pixels served as base stimuli. Each set comprised two exemplars of 10 faces to make it less likely that observers followed a template-matching strategy. In Set 1 (Experiment 2a and b), the exemplars showed two different expressions (neutral and happy). In Set 2 (Experiment 2c)—a subset of the base faces used by Goffaux and Rossion (2006)—both exemplars had a neutral expression, but the faces were shown from slightly different angles. The main inner facial features (eyes, eyebrows, nose, and mouth) were aligned within each stimulus set using rotation, translation, and scaling. It is important to note that these affine transformations did not alter the shape of facial features or the relative distance between them. The faces were presented behind a midluminance homogeneous field through an elliptical aperture with a horizontal diameter of 158 or 121 pixels and a vertical diameter of 239 or 175 pixels for the first and second stimulus set, respectively (see Figures 3a and 3b). Thereby, only the inner facial features were revealed, and the minor and major axes of the ellipse contained no information relevant to the task at hand. This mode of stimulus presentation was also chosen by Gaspar, Sekuler, and Bennett (2008), Robbins and McKone (2003), and Sekuler et al. (2004). In contrast, Goffaux and Rossion (2006; Experiments 2–4) used different elliptical apertures to reveal the inner facial features of their stimuli. Faces within each stimulus set were equated in mean luminance, contrast, and energy at each SF. Face width subtended a visual angle of 6.5˚ (Set 1) or 5.7˚ (Set 2). The base stimuli were presented upright or inverted (rotated 180˚ in the image plane).

Practice phase. Participants learned to associate the faces with common names (e.g., Mary, John, Peter) from printed grayscale pictures displayed along with names. When the participants were confident that they could identify all faces, the practice session began. Participants were seated in a dark room and a chin rest maintained them at a 53-cm viewing distance from the screen in Experiments 2a and b and at a 45-cm viewing distance for the smaller faces in Experiment 2c. Upright and inverted base stimuli were presented in separate 100-trial blocks, starting with upright faces. Each trial began with the appearance of a central fixation cross on the screen for 435 ms, followed by an upright or an inverted face presented for 435 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded with a key press. Each of the keys (numerals 0 to 9) on a regular computer

keyboard was associated with a particular face name. When participants responded incorrectly, auditory feedback was provided (a brief 3000 Hz pure tone). The first part of the practice session was completed when accuracy for upright faces was above 95% correct for two successive blocks of 100 trials; the second part was completed when the same criterion was reached for inverted faces. On average, participants needed 6.40 or 6.00 practice blocks in the upright condition and 17.00 or 10.50 blocks in the inverted condition with the first or second stimulus set, respectively. Finally, participants performed six additional practice blocks (three with upright and three with inverted faces) in which white Gaussian noise was added to the full-SF spectrum faces in preparation for the experimental blocks.

Experimental phase. In the experimental phase of Experiment 2a, each participant was presented with a total of 2,100 upright and 2,100 inverted face stimuli. Upright and inverted faces were presented in separate 100-trial blocks, starting with upright stimuli and then alternating with inverted stimuli. Accuracy was measured in the same 10-choice identification task as during the practice. The experimental trials differed from the practice trials on the following: (a) The SFs of the base stimuli were randomly sampled (Figure 4; for details, see the *Spatial frequency bubbles* section), (b) no feedback was given, and (c) performance in the upright blocks was maintained between 75% and 85% correct by adjusting the quantity of additive noise block per block. The same amount of noise was used in the following inverted blocks. We chose to equate the quantity of additive noise across conditions instead of accuracy (the latter is what Sekuler et al., 2004, and Gaspar, Sekuler, & Bennett, 2008, did) because in "real-life" upright and inverted faces contain the same amount of information.

The RT version (Experiment 2b) followed Experiment 2a and differed from it only in four respects: (a) Face stimuli remained on the screen until a response was made; (b) participants named aloud the identities of the faces, and a voice key was used to measure response latency; (c) after each trial, the experimenter typed the participant's response using a computer keyboard; and (d) performance was maintained above 90% correct in the upright condition block per block.

Experiment 2c served as control experiment and differed from Experiment 2a only in three respects: (a) The second stimulus set was used (Figure 3b); (b) the number of trials per observer was reduced to 400 per condition; and (c) the σ of the Gaussian filter was increased to 1.8 octaves. The goal of this experiment was to see whether the results obtained with the

first set of stimuli would generalize to another set of faces. Furthermore, we wanted to see whether the results could be replicated with a relatively small number of trials per observer.

Results and discussion

In Experiments 2a–c, the first 100-trial block in each orientation condition was considered as practice and was therefore excluded from the analyses. In Experiment 2a, accuracy across participants was significantly higher for upright (M = 80.52%, SD = 1.35) than for inverted faces (M = 52.43%, SD = 6.32), $t(4) = 8.56$, $p < .01$. In Experiment 2b, where accuracy was adjusted to above 90% for upright faces (M = 92.24%, SD = 2.23), accuracy for inverted faces was again lower (M = 72.12%, SD = 6.12), $t(4) = 6.81$, $p < .01$. Furthermore, RTs on correct trials were significantly shorter in the upright (M = 1479.83 ms, SD = 315.29) than in the inverted condition (M = 1935.01 ms, SD = 410.95), $t(4) = -6.90$, $p < .01$. In Experiment 2c, accuracy was also significantly higher for upright (M = 79.37%, SD = 5.81) than for inverted faces (M = 55.43%, SD = 7.51), $t(9) = 14.87$, $p < .01$. In sum, all versions of the experiment exhibited a clear FIE. Figure 5 shows the mean accuracy (Experiments 2a–c) and RTs (Experiment 2b) for each block.

To reveal the SF ranges that led to accurate (Experiments 2a and c) and fast (Experiment 2b) face identification in the upright and the inverted conditions, we performed multiple linear regressions on the sampling vectors per orientation condition per observer and on the appropriate regressor (see the *Spatial frequency bubbles* section). Because individual and group classification vectors were very similar in Experiments 2a and 2b (average correlations including both orientation conditions of r = .86 in Experiment 2a and r = .77 in Experiment 2b), and because Experiment 2c was based on a relatively small number of trials per observer, we report group results only. The group classification vectors for the upright and inverted conditions and their Z-transformed difference are shown in Figures 6, 7, and 8. The upright group classification vector in Experiment 2a showed a significant SF band of 2.00 octaves and dual peaks, one at 7.14 cpf (Z_{max} = 8.20; $p < .05$; S_r = 128; $FWHM$ = 3.53; Z_{crit} = 3.45) and the other at 12.14 cpf (Z_{max} = 8.10). Similarly, in the inverted condition, a 2.00 octaves wide SF band peaking at 7.14 cpf (Z_{max} = 8.68) and 11.07 cpf (Z_{max}= 7.72) was significant. In Experiment 2b, the group classification vector for the upright condition revealed a significant SF range of 1.94 octaves peaking at 8.57 cpf (Z_{max} = 6.97) and 12.86 cpf (Z_{max} =

5.77). In the inverted condition, an SF range of 1.30 octaves peaking at 12.14 cpf (Z_{max} = 5.06) was significant. In Experiment 2c, the upright classification vector reached significance for an SF band of 1.00 octave with a maximum Z-score at 7.53 cpf (Z_{max} = 4.34; $p < .05$; S_r = 128; *FWHM* = 4.24; Z_{crit} = 3.40). The classification vector for the inverted condition was significant for a SF band of 1.06 octaves, with a peak at 8.06 cpf (Z_{max} = 5.06; $p < .05$; S_r = 128; *FWHM* = 4.24; Z_{crit} = 3.40). None of the difference classification vectors reached statistical significance.

The correlations between classification vectors obtained on accuracy and RTs using the same stimuli (Experiment 2a and b) were very high in both the upright conditions (r = .96) and the inverted conditions (r = .93). For the experiments measuring accuracy but using different stimulus sets (Experiments 2a and c), the correlations were high as well (r = .84 for upright faces and r = .87 for inverted faces). Similarly, we found high correlations between Experiments 2b and c (based on different stimulus sets and different response measures) with r = .87 for upright and r = .95 for inverted faces. Most important, the correlation between the classification vectors for upright and inverted faces was very high in Experiments 2a (r = .97), 2b (r = .95), and 2c (r = .85), and the correlation between the average of the upright and inverted classification vectors of Experiment 2a–c was even higher (r = .98), strongly suggesting that the same SF band was used for identifying upright and inverted faces. To best estimate the center and width of this SF band, we summed all classification vectors for Experiments 2a–c and fitted a Gaussian density function to the logarithm of the average classification vector. The mean of the best fit—9.8 cpf—is our estimate of the center of the SF channel and its *FWHM*—1.9 octaves—is our estimate of the width of the SF channel.

The reliance on this 1.9 octaves wide SF band centered on average on 9.8 cpf appears robust to changes in the measure employed to assess the FIE (accuracy or response time of correct trials), to changes in the face sets, and to changes in procedures (many trials and few subjects in Experiments 2a and 2b; or few trials and relatively many subjects in Experiment 2c). The SF range revealed in the present study is in accordance with the intermediate SFs identified in previous studies for upright faces (e.g., Costen, Parker, & Craw, 1994, 1996; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999). Our results are also consistent with the findings of Gaspar, Sekuler, and Bennett (2008) for both upright and inverted faces, and

suggest that the processing difference does not lie in the extraction of cues at different SFs but rather in how the information within the same SF band is used.

Experiments 3a and b

The results of Experiments 2a–c reveal no difference in SF tuning between the upright and inverted conditions. However, we cannot rule out the possibility of this "null result" being due to a relative insensitivity of the SF Bubbles technique to reveal subtle differences in SF tuning resulting from bottom-up or top-down alterations in the visual strategies employed by observers with upright versus inverted faces (e.g., holistic versus featural processing). We conceived two experiments to test whether the SF Bubbles technique is sensitive enough to reveal subtle differences in SF tuning in situations known to differ solely in bottom-up (Experiment 3a) or top-down (Experiment 3b) SF tuning.

It has been shown that decreasing the size of letters (Chung, Legge, & Tjan, 2002; Majaj et al., 2002) and faces (Loftus & Harley, 2005; Näsänen, 1999) induces a shift in the use of SFs toward lower SFs in a mandatory bottom-up fashion. Experiment 3a examined whether we can reveal such a bottom-up-induced SF tuning change with the SF Bubbles technique. Observers were asked to identify the same faces as in Experiments 2a and b with large versus small stimuli.

Experiment 3b was designed to assess the capacity of the SF Bubbles technique to reveal differences in SF tuning that are due solely to changes in the top-down influence of task demands. More specifically, the SF Bubbles technique was applied to two different tasks previously shown to induce different SF usage patterns (happy/neutral vs. gender discriminations; e.g., Gosselin & Schyns, 2001; Schyns & Oliva, 1999).

Method

Participants. For Experiment 3a, 1 male and 2 female University of Victoria students (ages between 22 and 26 years; M = 23.7 years) were recruited. For Experiment 3b, we recruited 40 students (31 women) between 18 and 42 years of age (M = 19.8 years). All observers had normal or corrected-to-normal vision. Participants were naïve to the purpose of the study, and they received course credits or were paid for participating.

Apparatus. See corresponding section of Experiment 1; for the screen settings in Experiment 3a, see next section.

Stimuli. In Experiment 3a, the base stimuli of Experiment 2a and b were used but their resolution was decreased to 128 × 128 pixels. In the small condition, observers saw the upright face stimuli at a screen resolution of 2,048 × 1,536 pixels and at a distance of 180 cm, resulting in a face width of 0.5˚ of visual angle. In the big condition, they saw them at a screen resolution of 640 × 480 pixels and at a viewing distance of 45 cm, resulting in a face width of 5.9˚ of visual angle.

In Experiment 3b, the same 10 neutral faces (5 males, 5 females) were used as in Experiments 2a, 2b, and 3a. The corresponding 10 happy faces, however, differed from the set previously used in that no teeth were visible (Figure 9). We chose this set to make task difficulty between gender and happy/neutral discriminations more similar; a pilot study showed ceiling effects for happy/neutral discrimination when accuracy for gender was in the 65–75% range. Base stimuli had a resolution of 256 × 256 pixels and face width subtended a visual angle of 6.5˚. They were normalized for a number of low-level visual features and for the main facial feature position as in Experiments 2a, 2b, and 3a.

Procedure. In Experiment 3a, participants performed the same 10-alternative face identification task as in Experiment 2a. Each participant completed 3,300 trials per condition. We adjusted the quantity of additive white Gaussian noise (as described in the *Method* section for Experiment 1) on a block-by-block basis so that performance was approximately 80% correct in the big condition (i.e., the easiest condition). The 100-trial blocks succeeded each other as in Experiment 2a, but this time alternating between big and small rather than between upright and inverted.

Experiment 3b was divided in two parts: Each participant completed six consecutive 100-trial blocks of happy/neutral discrimination and six consecutive 100-trial blocks of gender discrimination. The first 20 participants initially completed the happy/neutral discrimination, followed by the gender discrimination; the last 20 participants completed the tasks in the opposite order. Each trial began with a central fixation cross presented for 412 ms, followed by an upright face presented for 412 ms, and then by a homogeneous mid-gray field that remained on the screen until the observer responded by pressing the appropriate key on a computer keyboard. Keys were counter-balanced across participants. For the initial task, performance was adjusted on a trial-by-trial basis by manipulating the quantity of additive noise using QUEST (Watson & Pelli, 1983). In the second task in Experiment 3b, the same

experimental stimuli were used (i.e., same base stimuli, same sampling vector, and same amount of white Gaussian noise) as in the corresponding trials of the first task in Experiment 3b.

Results and discussion

For each participant, the first 100-trial block of each task was considered as practice and was discarded from data analysis. The analyses were thus performed on 19,200 trials (3,200 trials per size condition × 2 size conditions × 3 participants) and 40,000 trials (500 trials per task × 2 tasks × 40 participants) in Experiments 3a and 3b, respectively. Because individual and group results were very similar for Experiment 3a (with an average correlation of $r = .93$ including both conditions), and because Experiment 3b was based on a large number of observers who each completed a relatively small number of trials, we will report group results only.

Mean accuracy in Experiment 3a was similar in the small ($M = 74.48\%$, $SD = 9.79$) and big conditions ($M = 80.36\%$, $SD = 1.84$), $t(2) = 1.227$, $p > .05$. In the small condition, a range of SFs of octave width 1.8 peaking at 5.00 cpf ($Z_{max} = 12.46$) exceeded the significance threshold ($p < .05$; $S_r = 64$; $FWHM = 3.53$; $Z_{crit} = 3.25$). The significance threshold (Z_{crit}) is slightly lower than in Experiment 2a because of the reduced stimulus resolution (128 × 128 instead of 256 × 256 pixels). In the big condition, an SF range of octave width 2.81 peaking at 8.57 cpf ($Z_{max} = 9.19$) attained significance. This is a replication of the results of Experiments 2a–c obtained with a similar face width. The difference between the group classification vectors of the two conditions reached significance between 3.57 cpf and 5.00 cpf as well as between 9.29 cpf and 17.86 cpf with a maximum at 4.29 cpf ($Z_{max} = 5.27$). Thus, as expected, we observed a clear shift toward lower SFs with smaller stimuli (Figure 10). This shows that the SF Bubbles approach is capable of revealing changes in SF tuning based on the same task and the same stimuli as in Experiment 2a.

In Experiment 3b, mean accuracy across participants was significantly higher for happy/neutral ($M = 81.20\%$, $SD = 5.20$) than for gender discriminations ($M = 65.31\%$, $SD = 5.03$), $t(39) = 16.899$, $p < .001$. The group classification vector results for happy/neutral and gender discriminations are illustrated in Figure 11. We analyzed the results for the different stimulus types separately, that is, happy male, happy female, neutral male, and neutral female

classification vectors for both the happy/neutral and gender discrimination tasks. These eight classification vectors allowed us to compare the use of SF information for the same group of stimuli (e.g., happy females) in the two tasks (gender and happy/neutral) and thus to isolate the top-down effect of task demands.

The group classification vectors for the happy/neutral and gender discriminations revealed different SF patterns for each of the four stimulus types. With happy male faces, the significant portion of the classification vector for the happy/neutral task was shifted into a lower SF range (with peaks at 2.14 cpf, Z_{max} = 6.45, and 5 cpf, Z_{max} = 6.37) than the peak for the gender task (7.86 cpf, Z_{max} = 3.94; $p < .05$; S_r = 128; *FWHM* = 3.53; Z_{crit} = 3.45; for details, see Chauvin et al., 2005). The correlation between the classification vectors was r = -.52. With neutral male faces, the classification vector for the happy/neutral task peaked at 8.57 cpf (Z_{max} = 7.75), and for the gender task at 10 cpf (Z_{max} = 4.10); the correlation was r = .44. With happy female faces, the classification vectors for the two tasks both peaked at 2.86 cpf (Z_{max} = 9.27, happy/neutral; Z_{max} = 11.75, gender) and had a correlation of r = .90. With neutral female faces, two distinct peaks were revealed at 7.86 cpf (Z_{max} = 9.33) for happy/neutral and at 2.14 cpf (Z_{max} = 7.84) for gender discriminations, with a correlation between classification vectors of r = .39. With all four stimulus types, the difference between the classification vectors for happy/neutral and gender reached significance. Furthermore, the results revealed that for both male and female faces, the happy/neutral classification vectors were only weakly correlated (r = -.03 and r = .22, respectively), and that for both happy and neutral faces, the gender classification vectors had a relatively weak correlation as well (r = -.46 and r = -.48). These results are in accordance with the view that the information required for different tasks can reside at different SFs of the same stimulus, and that our visual system is flexibly tuned to extract this information. Overall, the results of Experiment 3b show that SF Bubbles is a technique sensitive to differences in SF tuning that are only due to changes in task demand.

Together the findings of Experiments 3a and b demonstrate that the SF Bubbles approach is capable of revealing subtle differences in SF tuning for complex stimuli induced by a bottom-up factor (Experiment 3a) and by a top-down factor (Experiment 3b). These results also suggest that the null results of Experiment 2a–c are real; thus, the FIE cannot be attributed to qualitative processing differences at the SF level.

General discussion

The goals of the present study were to introduce a new SF probing technique and to uncover the SFs that mediate upright and inverted face identification. It comprised six experiments: The first experiment was designed to assess the validity of the SF Bubbles technique. Using a plaid detection task, we verified that the SF Bubbles method could reveal precisely the two SFs of the plaid. Compared with traditional SF sampling techniques—such as high-pass, low-pass, and band-pass filtering (e.g., Boutet et al., 2003; Goffaux & Rossion, 2006), or critical band masking (e.g., Gaspar, Sekuler, & Bennett, 2008)—SF Bubbles minimizes the risk that observers adapt to a certain SF range during the task by randomly sampling the SF information on a trial-by-trial basis. Another advantage of SF Bubbles in comparison with critical band masking is that SF Bubbles does not assume that SFs are integrated linearly. We could, in principle, evaluate the joint utilization of two or more SF bands provided that we perform computationally taxing second-order analyses, which go beyond the scope of the current article (e.g., Schyns, Bonnar, & Gosselin, 2002).

The next three experiments—the main experiments of the article—examined which SFs are diagnostic for the accurate (Experiments 2a and c) and fast (Experiment 2b) identification of upright and inverted faces. Although accuracy was on average 24% higher and RTs 455 ms shorter with upright faces, thus showing a clear FIE, SF tunings were remarkably similar in both orientation conditions. A single SF band of 1.9 octaves that peaked at 9.8 cpf was used by observers. This result was obtained independently of whether we used the accuracy (Experiments 2a and c) or RT (Experiment 2b) classification vectors. Moreover, this finding appears to be robust to changes in procedures (many trials and few subjects in Experiments 2a and 2b; few trials and relatively many subjects in Experiment 2c) and to changes in face sets. In Experiment 2c, we employed a subset of the base faces used by Goffaux and Rossion (2006) and replicated our findings of Experiment 2a.

To rule out that this null result was due to an insensitivity of the SF Bubbles technique, Experiment 3a re-examined SF tuning in the identification task of Experiment 2a as a function of stimulus size, which is known to influence SF tuning in a bottom-up fashion (Chung et al., 2002; Loftus & Harley, 2005; Majaj et al., 2002; Näsänen, 1999). In Experiment 3b, task demands were modified (gender or happy/neutral discrimination) to modulate SF tuning in a top-down fashion (e.g., Schyns & Oliva, 1999). In both cases, we revealed subtle differences

in SF use, confirming that the SF Bubbles technique is sensitive to bottom-up and top-down induced SF tuning changes. This result suggests that the null results of Experiments 2a–c are real. In addition, the big condition of Experiment 3a, which employed size parameters comparable to those of Experiments 2a–c, closely replicated our previous results.

The SF range revealed in the present study for faces subtending a horizontal visual angle of approximately 6° is consistent with the intermediate SF band (centered between 7 cpf and 16 cpf) identified as optimal in previous experiments for upright faces of visual angles between 2.3° and 9.5° (e.g., Costen et al., 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold et al., 1999; Näsänen, 1999). Our results are also consistent with those of Boutet et al. (2003), who argued that band-pass filtering faces in the low-SF (1.25–5 cpf), medium-SF (5–20 cpf), or high-SF (20–80 cpf) range had little impact on the FIE. Furthermore, our results are in agreement with a recent study on SF use in upright and inverted face identification by Gaspar, Sekuler, and Bennett (2008). They used critical band noise masking to examine SF tuning in two 10-choice identification tasks and reported that SF tuning for upright and inverted faces was mediated by the same SF band (approximately 1.5 octaves wide and centered at about 7 cpf for face stimuli with a width of 2.3° of visual angle). Our study replicates their findings based on accuracy and complements them by including RT analyses, by using a different SF probing technique, and by using different indexes of the FIE. Specifically, they equated accuracy between orientation conditions by manipulating the quantity of signal, and signal threshold was their index of the FIE. In contrast, we used the same quantity of signal and noise for upright and inverted faces, and response accuracy and RT were our indexes of the FIE.

Our face identification results, however, appear to contradict the findings obtained by Collishaw and Hole (2000) and Goffaux and Rossion (2006), which suggest that inversion might be particularly detrimental to faces containing only low SFs.[1] For example, Goffaux and Rossion (2006) reported that the composite face effect (Young et al., 1987) and the whole-part advantage (Tanaka & Farah, 1993)—used as indexes of holistic processing—were more pronounced for low-pass filtered (< 8 cpf) than for high-pass filtered (> 32 cpf) upright faces. In an additional experiment, they found that the composite effect for upright faces in the low-SF condition was also larger than in an intermediate-SF (8–32 cpf) condition. In contrast, no disproportionate composite effect for low SFs was observed when faces were upside-down. The authors concluded that holistic processing is largely supported by low SFs (but see

Cheung, Richler, Palmeri, & Gauthier, 2008, for a re-examination with an extended paradigm). An SF Bubbles experiment is unbiased in comparison with low-pass, high-pass, or band-pass filtering in the sense that on some trials it is equivalent to low-pass filtering, on others to high-pass filtering, and on others to band-pass filtering. On the majority of trials, it is equivalent to band-pass filtering multiple bands simultaneously. At the limit, it contains all possible filtering experiments. If there was a difference in the low-SF range between the identification of upright and inverted faces, this difference should have affected participants' behavior in our experiments at least on the trials in which only low SFs were shown, and we should have seen traces of this in the classification vectors. However, we did not observe such an effect.

How much of the variance between these findings can be explained by different modes of stimulus presentation remains to be investigated. Collishaw and Hole (2002) presented stimuli with both inner and outer facial features (see also Goffaux & Rossion, 2006; Experiment 1). Goffaux and Rossion (2006, Experiments 2–4) showed inner facial features through different elliptical apertures, whereas Gaspar, Sekuler, and Bennett (2008) showed inner facial features through a constant elliptical aperture. We employed the same approach as Gaspar, Sekuler, and Bennett and obtained results consistent with theirs but inconsistent with the studies using a different mode of stimulus presentation.

Another potential explanation for the different findings is that the studies tap into different processes. It is conceivable that holistic processing as indexed by the composite face effect does not correlate with accurate face identification. In fact, Konar, Bennett, and Sekuler (2007, 2008) assessed this possibility by using a composite face task and different identification tasks in a within-subject design. In the composite task, participants were asked to make same-different judgments about the top halves of faces, which were either aligned or misaligned with the bottom halves. In the three identification tasks, participants were asked to either determine whether a target face was present in a lineup, or perform a 10-alternative forced-choice identification with unlimited viewing time, or perform a 4-alternative forced-choice identification with a viewing time of 200 ms (i.e., the viewing duration in the composite task). No correlation was found between the strength of the composite face effect and accuracy in any of these face identification tasks. If there is a qualitative difference between the processing of upright and inverted faces—possibly in SF use—that is neither

correlated with accuracy nor with RT in face identification, it will not be revealed in our classification vectors. Thus, even though we did not find differences in SF use for the identification of upright and inverted faces, our results are not necessarily inconsistent with a qualitative view of the FIE.

It also remains possible that the same SFs are used, but in a different fashion. For example, in the upright condition, observers might use information at 9.8 cpf to encode the distance between eye and eyebrow; in the inverted condition, they might use information at the same SF to encode local features of the eye. Accordingly, Boutet et al. (2003) showed that intermediate SFs are optimal for both configural and featural modifications (but see Goffaux, Hault, Michel, Vuong, & Rossion, 2005, for a dissociation between low- and high-SF information and the extraction of configural and featural cues). Although the present study provides a precise estimate of the SF tuning curves and demonstrates that the performance drop with inversion is not due to a shift to less informative SFs, it remains an avenue for future research to shed light on how exactly the revealed SF information is used at other processing stages.

The most parsimonious explanation for the present findings is provided by the quantitative account of the FIE (Sekuler et al., 2004). According to this view, the FIE can be explained in terms of a decrease in the sensitivity of the same process. Sensitivity can be broken down into *calculation efficiency*, its deterministic component, and *internal noise*, its stochastic component. Gaspar, Bennett, and Sekuler (2008) recently found evidence that face inversion leads to a reduction in calculation efficiency but does not alter internal noise. We interact much more frequently with upright than with inverted faces, and it is thus plausible that a neural mechanism similar to that thought to mediate *perceptual learning* is responsible for the FIE. For instance, Kobatake, Wang, and Tanaka (1998) showed that monkeys trained on a set of novel stimuli have more inferotemporal neurons responsive to features of these stimuli than untrained monkeys. The more such selective neurons, the greater is the sensitivity. Perhaps more fusiform gyrus neurons are responsive to upright than to inverted facial features (see also Perrett, Oram, & Ashbridge, 1998, for a review) within the critical SF band characterized in this article.

Conclusion

The present study introduced a novel SF sampling technique that was applied to investigate SF tuning in upright and inverted face identification. The results show that the same SFs (dependent on stimulus size) were used for the accurate and fast identification of upright and inverted inner facial features, and performance was higher in the upright condition. The findings place an additional constraint on theories of qualitative processing differences and are consistent with predictions of the quantitative account of the FIE.

Acknowledgments

We thank Bruno Rossion and Valérie Goffaux for providing the face images of Set 2.

Footnote

[1] Nakayama (2003) reported a broader masking function for inverted than for upright faces, which is also in disagreement with our SF results. To our knowledge, this study has only been published as an abstract, and we do not have sufficient details to discuss it any further.

References

Ahumada, A., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America, 49*, 1751–1756.

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*, 281–316.

Barton, J. J., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology, 92*, 527–549.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35*, 1089–1105.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics, 65*, 1078–1093.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*, 659–667.

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1327–1336.

Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42*, 2137–2152.

Collin, C. A., Liu, C. H., Troje, N. F., McMullen, P. A., & Chaudhuri, A. (2004). Face recognition is affected by similarity in spatial frequency range to a greater degree than within-category object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 975–987.

Collishaw, S. M., & Hole, J. C. (2000). Featural and configurational processes in the recognition of faces of different familiarity. *Perception, 29*, 893–909.

Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception*, *23*, 129–146.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics, 58*, 602–612.

Davies, G., Ellis, H., & Shepherd, J. (1977). Cue saliency in faces as assessed by the ‘‘photofit’’ technique. *Perception, 6*, 263–269.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. New York, NY: Oxford University Press.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2(1),* i-i, doi: 10.1167/2.1.i.

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,* 628–634.

Fiset, D., Blais, C., Gosselin, F., & Schyns, P. (2006). Effective frequency tuning of three face categorization tasks [Abstract]. *Journal of Vision*, *6*(6), 273a.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception, 29*, 159–170.

Gaspar, C. M., Bennett, P. J., & Sekuler, A. B. (2008). The effects of face inversion and contrast-reversal on efficiency and internal noise. *Vision Research, 48*, 1084–1095.

Gaspar, C. M., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*, 2817–2826.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception, 34*, 77–86.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial"—holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 1023–1039.

Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*, 3537–3560.

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28*, 167–207.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Gosselin, F., & Schyns, P. G. (2005). Bubbles: A user's guide. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building Object Categories in Developmental Time* (pp. 91–106). Hillsdale, NJ: Erlbaum.

Haig, N. D. (1985). How faces differ—a new comparative technique. *Perception, 14*, 601–615.

Haig, N. D. (1986). Exploring recognition with interchanged facial features. *Perception, 15*, 235–247.

Hochberg, J., & Galper, R. R. (1967). Recognition of faces: An explanatory study. *Psychological Science, 9*, 619–620.

Kobatake, E., Wang, G., & Tanaka, K. (1998). Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology, 80*, 324–330.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2007). The composite face effect is not correlated with face identification accuracy [Abstract]. *Journal of Vision*, *7*(9), 501a.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2008). The composite face effect is still not correlated with face identification accuracy [Abstract]. *Journal of Vision*, *8*(6), 891a.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51*(A), 449–473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53*(A), 513–536.

Leder, H., Candrian, G., Huber, O., & Bruce, V. (2001). Configural features in the context of upright and inverted faces. *Perception, 30*, 73–83.

Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology, 97*, 19–29.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature, 410*, 890.

Loftus, G. R., & Harley, E. M. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review*, *12*, 43–65.

Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, *42*, 1165–1184.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.

McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. G. (2005). The use of visual information in natural scenes. *Visual Cognition, 12*, 938–953.

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin and Review*, *8*, 454–469.

Nakayama, K. (2003). Face specific processing: Role of local features in an affine metric [Abstract]. *Journal of Vision*, *3*(9), 91a.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, *39*, 3824–3833.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition, 67*, 111–145.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception, 17*, 43–63.

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition, 47*, 25–57.

Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B: Biological Sciences, 271*, S448–S450.

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition, 88*, 79–107.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103,* 34–79.

Rock, I. (1974). The perception of disoriented figures. *Scientific American*, *230*, 78–85.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128,* 274–289.

Rossion, B., & Boremanse, A. (2008). Nonlinear relationship between holistic processing of individual faces and picture-plane rotation: evidence from the face composite illusion. *Journal of Vision, 8,* 1–13.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*, 62–74.

Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: An integrative review of the role of spatial frequencies. *Psychological Research, 70*, 273–292.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402–409.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*, 391–396.

Sergent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75*, 221–242.

Shepherd, J., Davies, G., & Ellis, H. (1981). Studies of cue saliency. In G. Davies, H. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.

Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences, 10*, 538–545.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *46*(A), 225–245.

Tanaka, J. W., & Sengco, J. (1997). Features and their configuration in face recognition. *Memory and Cognition, 25*, 583–592.

Thompson, P. (1980). Margaret Thatcher: A new illusion. *Perception, 9*, 483–484.

Valentine, T. (1988). Upside-down faces: a review of the effect of inversion upon face recognition. *British Journal of Psychology, 79*, 471–491.

Walker-Smith, G. J., Gale, A. G., & Findlay, J. M. (1977). Eye movement strategies involved in face perception. *Perception, 6*, 313–326.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33,* 113–120.

Williams, C. C., & Henderson, J. M. (2007). The face inversion effect is not a consequence of aberrant eye movements. *Memory and Cognition, 35*, 1977–1985.

Yarmey, A. D. (1971). Recognition memory for familiar "public" faces: Effects of orientation and delay. *Psychonomic Science*, *24*, 286–288.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron, 44*, 889–898.

Author note

Verena Willenbockel, Daniel Fiset, James Tanaka, and Daniel Bub are at the Department of Psychology at the University of Victoria. Alan Chauvin is at the Laboratoire URECA of the Université Charles-de-Gaulle Lille III. Caroline Blais, Martin Arguin, and Frédéric Gosselin are at the Département de Psychologie of the Université de Montréal.

Portions of this research were submitted in thesis form to the Faculty of Graduate Studies at the University of Victoria. The research was supported by a graduate scholarship to Verena Willenbockel from the German Academic Exchange Service (DAAD), by a scholarship from the James S. McDonnell Foundation (Perceptual Expertise Network) and a postdoctoral scholarship from the Fonds Québécois de Recherche en Nature et Technologies (FQRNT) to Daniel Fiset, by a FQRNT graduate scholarship to Caroline Blais, by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) to Frédéric Gosselin and Martin Arguin and to James Tanaka, and by grants from FQRNT to Martin Arguin, Frédéric Gosselin, and Pierre Jolicoeur.

Figure captions

Figure 1. Illustration of the SF Bubbles technique. (1) Padding of the base stimulus with a uniform gray background. (2) Fast Fourier transform (FFT) of the padded stimulus. (3) Construction of a random SF filter: (a) creation of a binary random vector of length *2wk* (*w* = image width; *k* = 20) consisting of 45 ones among zeros; (b) convolution of the random vector with a Gaussian kernel (an "SF bubble"), resulting in a random sampling vector; (c) log-scaling of sampling vector (see text for details), resulting in a one-dimensional (1D) filter; (d) construction of a 2D filter by rotating the sampling vector about its origin. (4) SF filtering by dot-multiplying the 2D filter with the padded stimulus' complex FFT amplitudes. (5) Inverse Fourier transform (IFFT) and (6) cropping of the output to create the SF sampled experimental stimulus.

Figure 2. Individual and group classification vectors obtained in the plaid detection task (Experiment 1). The SF Bubbles technique revealed significant peaks at 10 cycles per image (cpi) and at 45 cpi (Z_{crit} = 3.45, $p < .05$), thus accurately showing the diagnostic SFs (i.e., the SFs of the plaid).

Figure 3. a) Base stimuli used in Experiments 2a, 2b, and 3a. Faces of ten identities × two expressions (neutral, happy) were shown. b) Base stimuli used in Experiment 2c (from the base set of Goffaux & Rossion, 2006). Faces displayed ten neutral identities × two slightly different viewpoints.

Figure 4. Three sample stimuli after filtering with the SF Bubbles technique. Fourier energy averaged across orientations is plotted as a function of SF in cycles per image (cpi).

Figure 5. Mean accuracy across observers over the 20 blocks of Experiments 2a (a) and 2b (b), as well as over the three blocks of Experiment 2c (d). Mean response times (RTs) on correct trials over the 20 blocks of Experiment 2b are shown in (c). Error bars give the standard errors. Both accuracy and RTs show a clear FIE.

Figure 6. Accuracy group classification vectors obtained for upright and inverted face identification in Experiment 2a. The two faces at the bottom show the SFs (cpf = cycles per face) that reached statistical significance (Z_{crit} = 3.45, $p < .05$) in both orientation conditions. There was no significant difference between the upright and inverted classification vectors.

Figure 7. Response time group classification vectors obtained for upright and inverted face identification in Experiment 2b. The two faces at the bottom show the SFs (cpf = cycles per face) that reached statistical significance (Z_{crit} = 3.45, p < .05) in both orientation conditions. There was no significant difference between the two classification vectors.

Figure 8. Accuracy group classification vectors obtained for upright and inverted face identification in Experiment 2c with stimuli from Goffaux and Rossion (2006). The two faces at the bottom show the SFs (cpf = cycles per face) that reached statistical significance (Z_{crit} = 3.40, p < .05) in both orientation conditions. There was no significant difference between the two classification vectors.

Figure 9. Base stimuli used in Experiment 3b. Faces displayed 10 identities × 2 expressions (neutral, happy). The neutral faces were the same as those used in Experiments 2a, 2b, and 3a, but the happy faces differed from those previously used in that no teeth were visible.

Figure 10. Accuracy group classification vectors for upright face identification obtained in the *big* and *small* conditions of Experiment 3a. The two faces at the bottom show the SFs (cpf = cycles per face) that reached statistical significance (Z_{crit} = 3.25, p < .05) in both size conditions. The difference between classification vectors exceeded the significance threshold.

Figure 11. Accuracy group classification vectors obtained for two different tasks (happy/neutral and gender discriminations) performed on the same face set (Experiment 3b). For all stimulus types, the difference between the classification vectors for the two tasks exceeded the significance threshold (Z_{crit} = 3.45, p < .05).

Figures

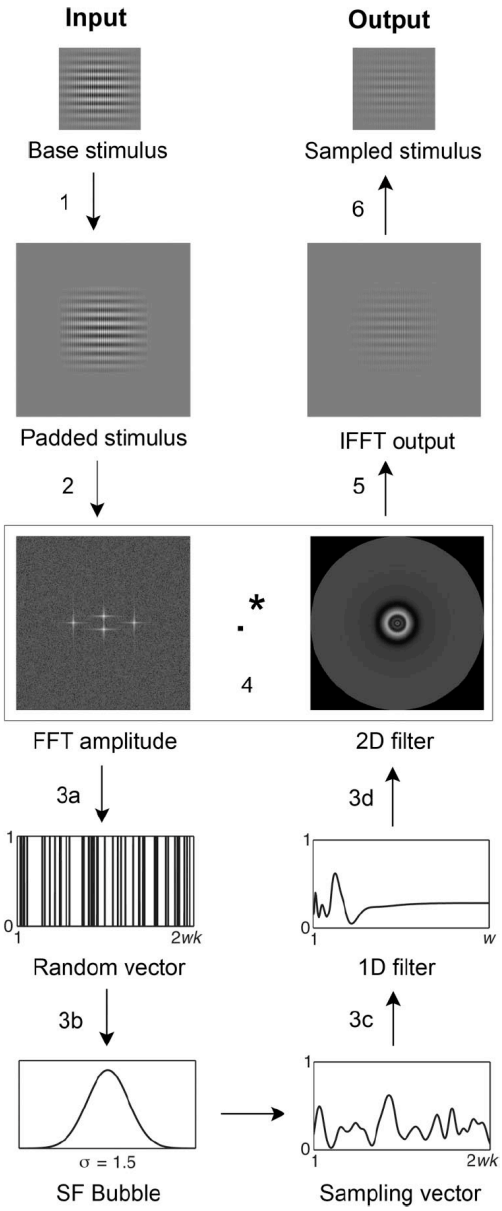


Figure 1: Illustration of the SF Bubbles technique

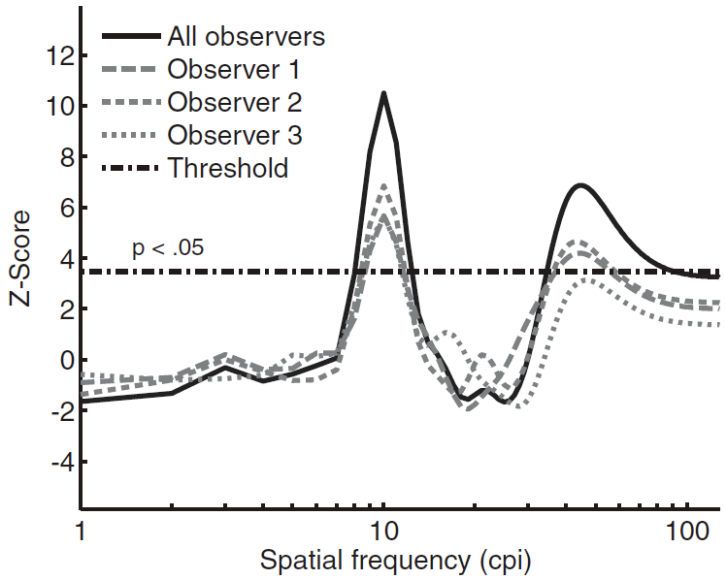
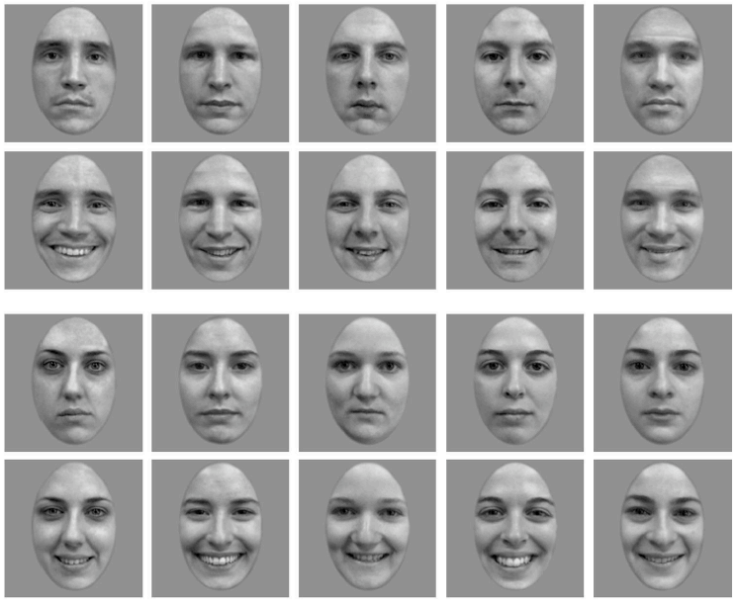


Figure 2: Results of Experiment 1

a)



b)

Figure 3: Base face images used in Experiments 2a–c and 3a

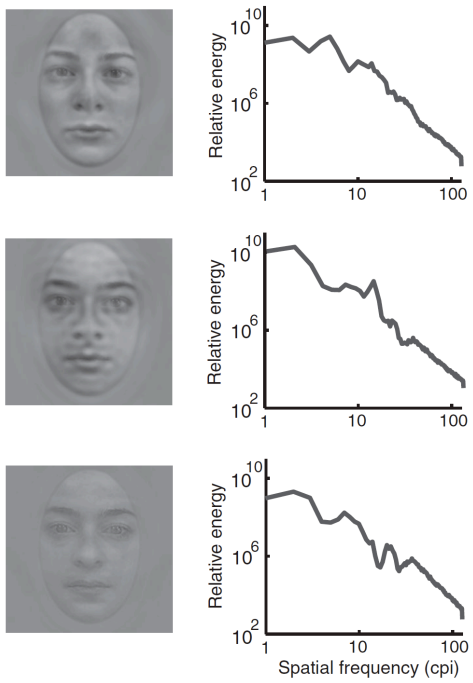


Figure 4: Sample image filtered with SF Bubbles

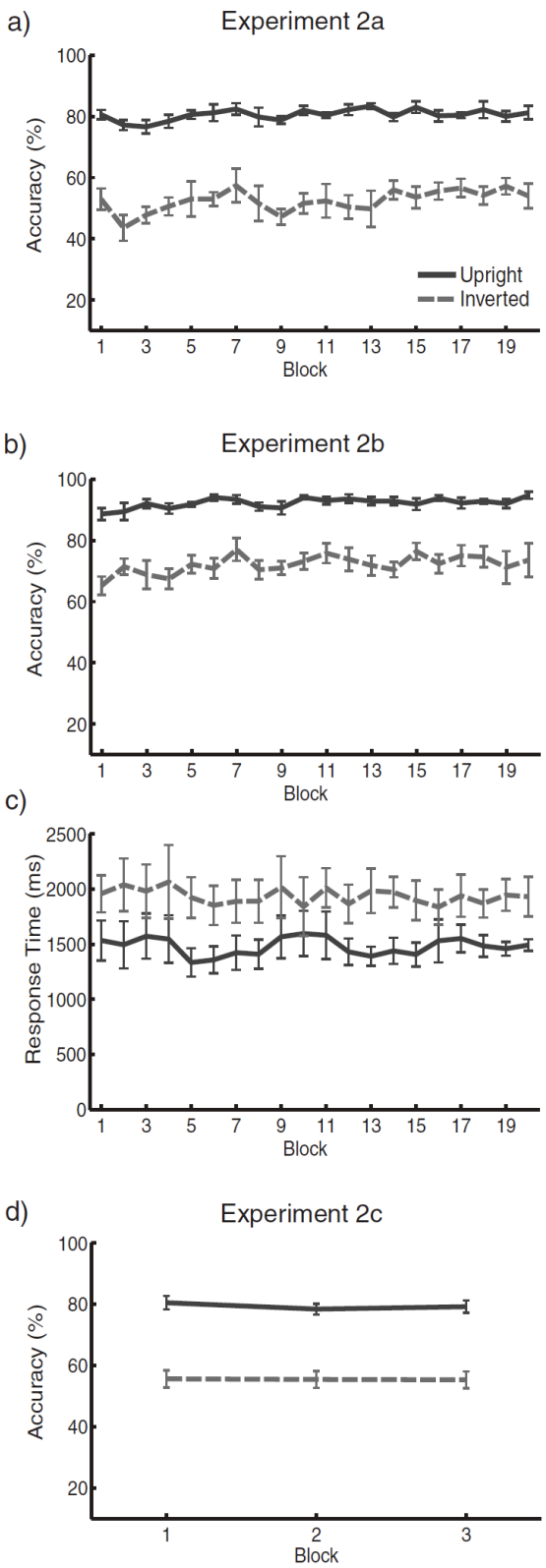


Figure 5: Results of Experiments 2a–c (accuracy and RT)

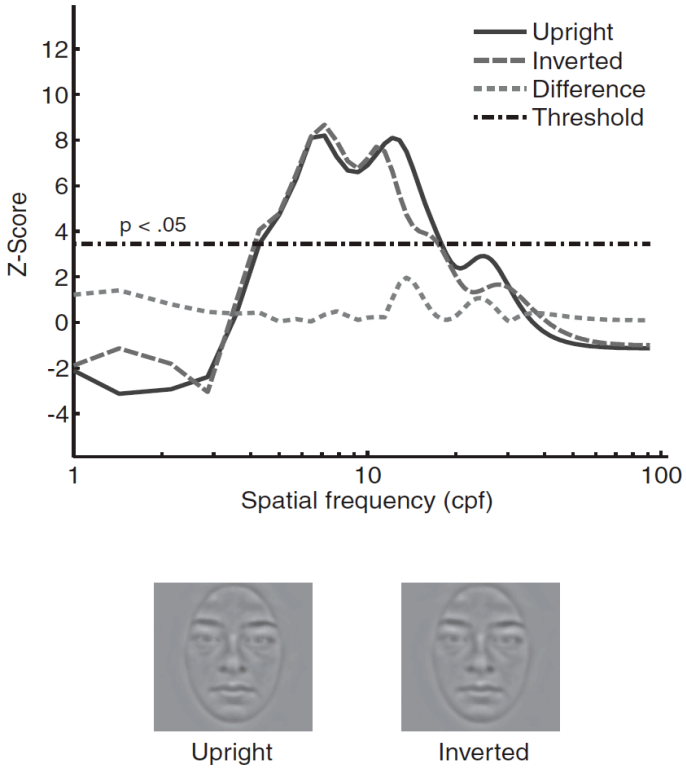


Figure 6: Results of Experiment 2a

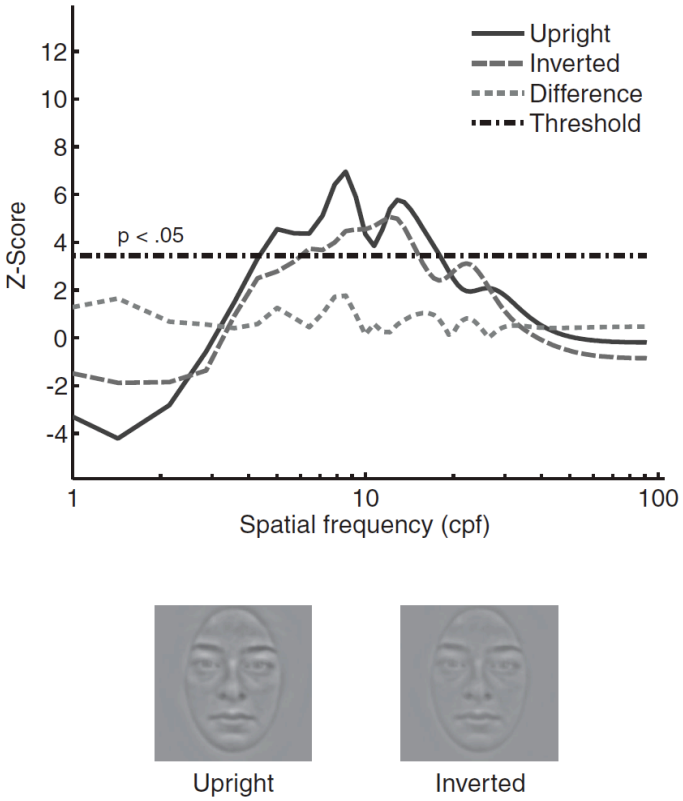


Figure 7: Results of Experiment 2b

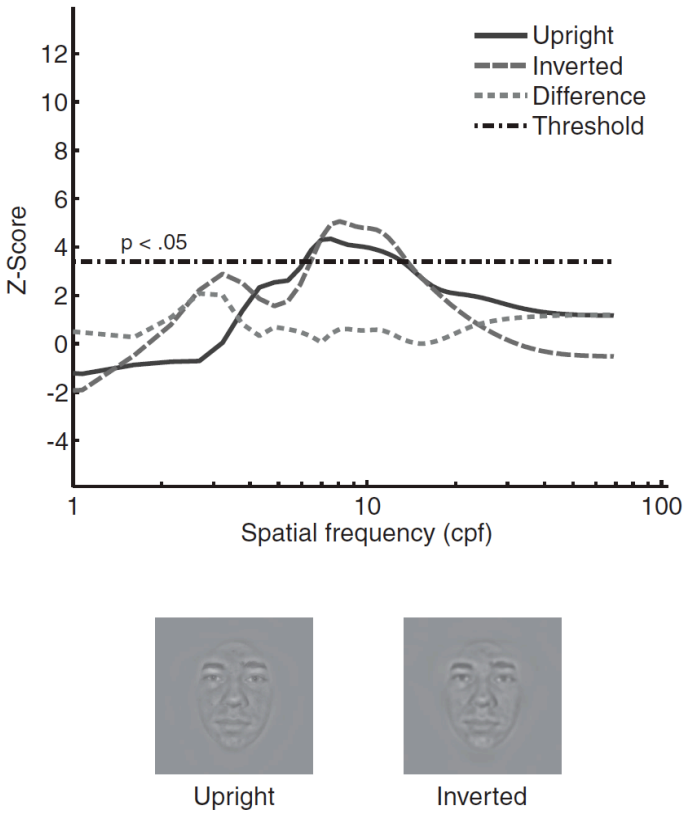


Figure 8: Results of Experiment 2c

Figure 9: Base face images used in Experiment 3b

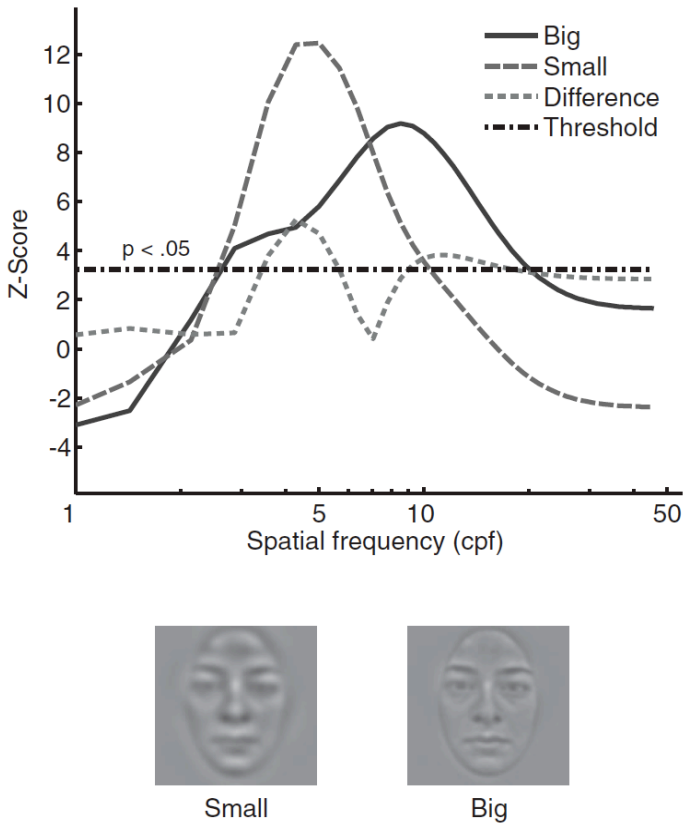


Figure 10: Results of Experiment 3a

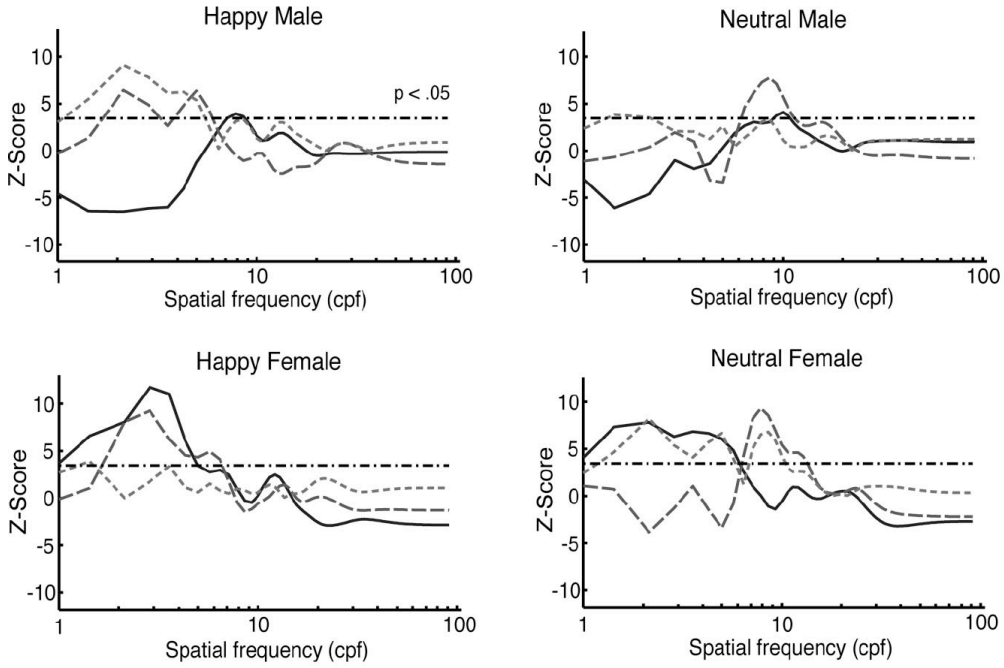


Figure 11: Results of Experiment 3b

Transitional text # 2

Chapter 3 has introduced and illustrated the SF Bubbles technique and shown that the method is sensitive to subtle changes in SF tuning. Moreover, the chapter served as an introduction to SF tuning during face perception. Our results showed that a mid-SF band is particularly important for identifying consciously perceived faces accurately and rapidly, irrespective of face orientation.

In Chapter 4, the focus shifts to the informational correlates of consciousness during face perception. Here, we investigated SF tuning during conscious and non-conscious face-gender perception by combining the SF Bubbles technique with a masked priming paradigm.

Chapter 4: Third article

Article accepted for publication:

Willenbockel, V., Lepore, F., Bacon, B. A., & Gosselin, F. (in press). The "informational correlates" of conscious and non-conscious face-gender perception. *Journal of Vision.*

The "informational correlates" of conscious and non-conscious face-gender perception

Verena Willenbockel[1,2], Franco Lepore[1,2], Benoit A. Bacon[1,3], & Frédéric Gosselin[1,2]

[1]Centre de Recherche en Neuropsychologie et Cognition (CERNEC), Montréal, QC, Canada
[2]Département de Psychologie, Université de Montréal, Montréal, QC, Canada
[3]Department of Psychology, Bishop's University, Sherbrooke, QC, Canada

Abstract

We used a face-gender repetition priming paradigm to precisely map the spatial frequencies (SFs) that influence observers' responses under different prime awareness conditions. A "visible prime" condition was set up by presenting the stimulus sequence mask-blank-prime-blank-mask-target, and an "invisible prime" condition by reversing the order of the masks and the blanks (see also Dehaene et al., 2001). The prime faces (~4.6º × 3.1º) were randomly filtered trial-by-trial with the SF Bubbles technique (Willenbockel et al., 2010a). Classification vectors, derived by summing the SF filters from each trial weighted by observers' transformed response times, revealed that SFs around 12 cycles per face width modulated responses in both prime awareness conditions. The significant SFs closely matched those optimal for accurate performance in a direct face-gender classification paradigm. Surprisingly, the significant SFs facilitated observers' responses in the visible prime condition, whereas they *slowed* responses in the invisible prime condition. Our findings suggest that SF tuning per se remains robust under different prime awareness conditions but that diagnostic visual cues might be utilized in a qualitatively different fashion as a function of awareness.

Keywords: consciousness, face perception, priming, spatial frequency

Introduction

Only a fraction of the visual input that impinges on our retinas actually enters our awareness, i.e., is perceived consciously. Numerous studies have shown, however, that some stimuli of which we are not aware can still impact our neural activity and behavior, in a "non-conscious" manner (e.g., see Kouider & Dehaene, 2007, for a review). The question remains as to how the visual processes that underlie conscious perception differ from those underlying non-conscious perception.

Many effective demonstrations of non-conscious perception in neurotypical individuals have involved priming effects: the processing of a visible target stimulus is influenced by a related preceding ("prime") stimulus, even when participants are not aware of the prime (e.g., Abrams, Klinger, & Greenwald, 2002; De Gardelle & Kouider, 2010; Dehaene et al., 1998, 2001; Finkbeiner & Palermo, 2009; Naccache, Blandin, & Dehaene, 2002; Naccache & Dehaene, 2001). Typically, a direct measure (e.g., prime detection or discrimination) is employed to demonstrate null sensitivity for the prime. An indirect measure (i.e., the effects of the prime on target perception) establishes that the prime information was nonetheless encoded and processed. Many studies reported facilitatory priming effects for non-consciously perceived primes that were identical or congruent to the target (see Kouider & Dehaene, 2007, for a review). However, "negative compatibility effects" in which primes that are congruent with the target inhibit observers' responses have also been reported (see, e.g., Eimer & Schlaghecken, 2003, for a review).

A common way of rendering a prime "invisible" is to present it only briefly and to have it immediately followed by a masking stimulus (Breitmeyer & Öğmen, 2006; Kouider & Dehaene, 2007). Masking is thought to decrease the bottom-up stimulus strength and consequently to help suppress the stimulus from awareness (specifically, to render it "subliminal"; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). An elegant masking paradigm was used by Dehaene et al. (2001) to investigate the neural correlates of word perception under different awareness conditions. In the invisible, non-conscious condition, participants were presented with a stimuli sequence that consisted of a blank screen, followed by a first mask, a word, a second mask, and then another blank. In the visible condition, the masks and the blanks were switched—so that the word was immediately surrounded by

blanks. Using these minimally different conditions in combination with functional magnetic resonance imaging (fMRI) and event-related potential (ERP) recordings, Dehaene and colleagues demonstrated that activation to invisible stimuli was much less than activation to visible stimuli in several brain areas, and that some areas only significantly responded to visible stimuli. In a second experiment, they demonstrated that the invisible words, used as primes, led to repetition priming (see also Kouider, Dehaene, Jobert, & Le Bihan, 2007).

Likewise, many other studies have employed masking paradigms to contrast brain activation between minimally different conscious and non-conscious conditions, which has revealed several neural markers of consciousness (see Dehaene & Changeux, 2011, for a recent review). Surprisingly few studies, however, have investigated possible distinctions in the *visual information* that leads to conscious vs. non-conscious priming effects. Since typically only a subset of the visual input reaches our awareness while a considerable amount of information remains non-consciously processed (e.g., Dehaene & Changeux, 2011), it is conceivable that different aspects of a complex visual stimulus lead to priming effects under different awareness conditions. For example, it is possible that different spatial frequencies (SFs) of the words used by Dehaene et al. (2001) were processed as a function of awareness.

It is well established that the human visual system processes input with multiple channels, each tuned to specific SFs (see De Valois & De Valois, 1990, for a review). Low SFs represent coarse information (e.g., the blurred shape of a face) whereas high SFs represent precise, detailed information (e.g., the fine wrinkles in a face). It has been proposed that different SF ranges are processed at different speeds via distinct neuroanatomical pathways (Livingstone & Hubel, 1988), play different functional roles (e.g., Bar, 2003; Bullier, 2001), and possibly interact differently with awareness (De Gardelle & Kouider, 2010; Khalid, Finkbeiner, König, & Ansorge, in press).

To investigate putative interactions between SF processing and awareness, De Gardelle and Kouider (2010) employed a masked priming paradigm with foveally presented hybrid faces as primes (i.e., the low SFs of one face image were combined with the high SFs of another; Schyns & Oliva, 1994). The primes were displayed as briefly as 43 ms and as long as 300 ms to create four visibility conditions. The observers were asked to perform a fame judgment task on visible target faces. Based on the priming effects observed for famous faces, the authors reported two main results: First, they found comparable non-conscious priming

effects for low-SF (<12 cycles per face width, cpf) and high-SF information (>12 cpf). Second, they observed that the magnitude of the priming effects increased with prime duration for high-SF and full-bandwidth stimuli, but not for low-SF stimuli. However, in creating the hybrids, the continuous SF spectrum of the face images was divided into "halves" by using an arbitrary cutoff of 12 cpf, which falls into the SF band that is diagnostic for recognizing famous faces (Butler, Blais, Gosselin, Bub, & Fiset, 2010). Therefore, it remains possible that qualitative differences (i.e., in terms of optimal SF or SF bandwidth) were present but not revealed given the filtering method and task employed.

A recent study on non-conscious face-gender priming including low- and high-pass filtered primes provided support for qualitative differences (Khalid et al., in press). In two experiments, the authors measured face-gender priming effects using peripherally presented low-pass or high-pass filtered primes (similar in SF content to those by De Gardelle & Kouider, 2010). The authors found converging evidence that non-conscious priming occurs with low- but not with high-SF primes.

The aim of the present study was to map the SFs that modulate observers' responses in different prime awareness conditions with a higher SF resolution than previous work and independent from any cutoff frequencies. The key aspect of the present demonstration is that we randomly sampled the SFs of the prime faces on a trial-by-trial basis using the SF Bubbles technique (Willenbockel et al., 2010a). Sampling SFs over the whole spectrum allowed us to correlate participants' response times (RTs) with the information that was made available to them and reveal the exact SF tuning curves for the task at hand. Observers were asked to perform face-gender judgments—a natural two-choice task, for which both low and high SFs have been found to be useful (Schyns & Oliva, 1999) and for which non-conscious priming has been demonstrated (Finkbeiner & Palermo, 2009; Khalid et al., in press). To minimize the difference between our visible prime and invisible[1] prime conditions we employed a masking paradigm adapted from Dehaene and colleagues (2001) for use with "SF bubblized" face primes. Using this approach, we tested whether the same information is diagnostic for face-gender perception as a function of awareness, and whether consciously and non-consciously processed diagnostic cues lead to the same behavioral effects.

Methods

Participants

Twelve adults (7 women; 21–37 years old; *Mdn* = 25.50 years) took part in Experiment 1. Eighteen adults (9 women; 19–30 years old; *Mdn* = 22.00 years) took part in Experiment 2. Four participants completed both experiments. All participants were recruited at the Université de Montréal, reported to have normal or corrected-to-normal vision, and provided written informed consent. The experiments were conducted in accordance with the Declaration of Helsinki and were approved by the CÉRFAS (Comité d'éthique de la recherche de la faculté des arts et des sciences) of the Université de Montréal.

Materials

Twenty grayscale photographs of faces from Schyns and Oliva (1999) were used as base stimuli (see Figure 1 for an example). The images (256 × 256 pixels) depicted five male and five female faces (width = ~3.1º, height = ~4.6º), each showing a happy and a neutral expression. The position of the main facial features, hairstyle, orientation, and lighting were normalized, and the faces were equated in mean luminance and contrast [root mean square (RMS) contrast = 0.43] using the SHINE toolbox (Willenbockel et al., 2010b). The targets were constructed by reducing the RMS contrast of the face images to 0.32, and the primes were created by randomly SF filtering the images using the SF Bubbles technique (see Figure 1 for three examples and Willenbockel et al., 2010a, for a detailed description and an illustration of the filtering procedure). On each trial, a given base image was first padded with a uniform gray background and then transformed into the frequency domain using a fast Fourier transform. The amplitude spectrum of the transformed image was multiplied element-wise with a "random" filter that was constructed in the following way: A vector consisting of randomly distributed binary elements (10,195 zeros and 45 ones) was convolved with a Gaussian kernel (an "SF bubble"; σ = 1.8). As a result, a smooth sampling vector was obtained. To account for the finding that the human visual system is more sensitive to low than to high SFs (e.g., see De Valois & De Valois, 1990, for a review), the sampling vector was transformed using a logarithmic function (see Willenbockel et al., 2010a, for details). The log-transformed, smoothed sampling vector was then "rotated" about its origin to obtain a 2D filter. After multiplying this filter element-wise with the base image's amplitude spectrum, we

back-transformed the result into the image domain via an inverse fast Fourier transform. The filtered image contained a random subset of the base image's SF information. For the analysis, we essentially used the log-transformed, smoothed sampling vector, henceforth referred to as SF filter.

On average the primes had a mean RMS contrast of 0.16 (*SD* = 0.04). Masks were random noise textures of 256 × 256 pixels generated on each trial (Figure 2)[2]. They had a mean RMS contrast of 0.52 (*SD* = 0.03) and subtended a visual angle of 5.8º × 5.8º. All stimuli were displayed on a 40.3 cd/m^2 background using a calibrated CRT monitor. The experiment was programmed in MATLAB with the Psychophysics toolbox (Brainard, 1997; Pelli, 1997).

Procedure

We conducted two experiments, which both consisted of a practice phase, visibility test ("pretest"), testing phase, and another visibility test ("posttest"). The experiments differed mainly in the number of conditions and the number of trials per condition in the testing phase. Experiment 1, which was the main experiment, was designed to map SF tuning as a function of prime awareness. It required a large number of trials per condition; therefore, we minimized the number of conditions (visible prime, invisible prime). Experiment 2, which served as a control experiment, included an additional prime-absent condition, but relatively few trials per condition.

During all phases, the participants were seated in a dark room, and a chin rest was used to maintain a viewing distance of one meter from the screen. The practice phase, which was identical for both experiments, started after the participants had seen all base images. Each practice trial consisted of the presentation of a central fixation cross (500 ms), followed by a mask (50 ms), a uniform gray field ("blank"; 50 ms), an SF sampled face (33–142 ms), another blank (33 ms), another mask (17 ms), and phase noise created from the average of all target faces (until a response was made; see Figure 2). The duration of the SF sampled face image was adjusted trial-by-trial using QUEST (Watson & Pelli, 1983) to maintain performance at 90% correct. Participants were instructed to accurately identify the gender of the faces by pressing labeled keys, counterbalanced across observers, on a regular computer keyboard. Auditory feedback in the form of a brief 3000-Hz tone was provided when an incorrect

response was made. Each observer performed a minimum of three 100-trial blocks. The practice was completed once the face-stimulus duration decreased to 50 ms or less. In Experiment 1, participants completed a total of 17,400 practice trials (M = 1,450 trials per subject, min = 300 trials, max = 3,100 trials, 95% confidence interval [1067, 1958]), and in Experiment 2 they completed 19,000 trials (M = 1,911 trials per subject, min = 300 trials, max = 4,900 trials, 95% confidence interval [1483, 2508]) (see Table I).

The pretest included a *visible* face and an *invisible* face condition (Figure 2b). The visible face condition was identical to the mask-blank-face-blank-mask-noise sequence of the practice phase, except that the face was presented at a fixed duration of 50 ms (see also, e.g., Finkbeiner & Palermo, 2009; Khalid et al., in press); this resulted in a face-mask stimulus onset asynchrony (SOA) of 83 ms. In the invisible face condition, the masks and the blanks were switched (see also Dehaene et al., 2001) so that the masks immediately surrounded the face image (blank-mask-face-mask-blank-noise; face-mask SOA = 50 ms). The two visibility conditions were interleaved in random order. Observers were asked to accurately identify the gender of the randomly SF filtered faces. In Experiment 1, observers performed the pretest with both original-contrast (80 trials) and contrast-reduced faces (320 trials) that were randomly intermixed. It turned out that contrast reduction was not necessary to obtain chance-level performance in the invisible face condition; thus, original-contrast faces were used in the remainder of the study, and only those results will be reported. In Experiment 2, observers completed 200 pretest trials (with original-contrast faces).

The testing phase followed the same basic procedure as the pretest except that (a) a target full-spectrum face replaced the phase noise at the end of each trial, (b) a control condition was added to Experiment 2, and (c) the focus was on RT rather than accuracy. Specifically, in the testing phase of both experiments, the last image in the stimulus sequence was always an unfiltered face that remained on the screen until the observer made a response. The target face was identical to the SF filtered face (i.e., the prime) on prime-present trials. In Experiment 1, there was always a prime in the sequence (invisible prime, visible prime). In Experiment 2, the prime faces were replaced by a blank on 50% of the trials, resulting in four conditions (invisible blank, invisible prime, visible blank, visible prime). In both experiments, observers were instructed to pay attention to the whole stimulus sequence and to identify the gender of the full-spectrum target face as accurately and as quickly as possible. The testing

phase of Experiment 1 consisted of 25 blocks of 80 trials, for a total of 1,000 trials per condition per observer; that of Experiment 2 consisted of 10 blocks of 80 trials (200 per condition per observer).

In order to reassess the visibility of the primes after the testing phase, each participant completed a 200-trial posttest that followed the same procedure as the pretest and was identical for both experiments.

Analysis and results

SF usage for direct, accurate face-gender discrimination

First, we analyzed the data from the practice phases of Experiments 1 and 2. This allowed us to see which SFs are optimal for direct, accurate face-gender categorizations and to compare SF tuning for the two groups of participants. We summed the SF filters from each practice trial weighted by the observers' transformed accuracies. The accuracies were transformed as follows: Correct responses were given a value of *P*(incorrect), which denotes the probability of observing an incorrect response—i.e., here .1—and incorrect responses a value of -*P*(correct)—i.e., here -.9. Henceforth, we will call the result of this weighted sum a *classification vector*. One such classification vector was computed per block per participant. One participant classification vector was calculated for each observer by summing all respective block classification vectors. Finally, we derived one group classification vector for Experiment 1 and another for Experiment 2 by summing the appropriate participant classification vectors and transforming the results into Z-scores (see Willenbockel et al., 2010a, for details). Statistical significance was evaluated by applying the Cluster test (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005): Given the clusters greater than an arbitrary Z-score threshold, $Z_{arbitrary}$, the test gives a cluster size, k_{crit}, above which the specified p-value is satisfied (k_{crit} = 4.81 pixels, $p < .05$, two-tailed, $Z_{arbitrary} = \pm 2.3$, $S_r = 128$, $FWHM = 4.24$).

The Z-transformed group classification vector for Experiment 1 showed a positively significant SF range ($k_{max} = 58$) with a local maximum at 3.19 cpf ($Z_{local\ max} = 3.98$) and a global maximum at 11.16 cpf ($Z_{max} = 12.46$). Similarly, the group classification vector for Experiment 2 showed a positively significant SF range ($k_{max} = 53$) with a local maximum at 3.19 cpf ($Z_{local\ max} = 4.47$) and a global maximum at 10.63 cpf ($Z_{max} = 13.74$) (Figure 3). The

two classification vectors were highly correlated (r = .99), suggesting that SF tuning was very consistent across the two groups of participants. The peak SFs for individual participant classification vectors are given in Table I.

Prime visibility

Next, we analyzed the pre- and posttest data of both experiments to verify prime visibility (Figure 4). Pretest accuracy in the visible face condition was significantly higher than chance in both Experiment 1 (M = 83.14%, SE = 2.63%), $t(11) = 12.60$, $p < .001$, and Experiment 2 (M = 84.33%, SE = 1.24%), $t(17) = 27.76$, $p < .001$. In the invisible face condition, pretest accuracy was neither significantly different from chance in Experiment 1 (M = 49.83%, SE = 1.00%), $t(11) = -.17$, $p = .87$, nor in Experiment 2 (M = 51.22%, SE = 1.32%), $t(17) = .92$, $p = .37$.

Similarly, posttest accuracy in the visible face condition was significantly above chance in Experiment 1 (M = 81.25%, SE = 2.14%), $t(11) = 14.58$, $p < .001$, and in Experiment 2 (M = 79.44%, SE = 1.66%), $t(17) = 17.75$, $p < .001$. Posttest accuracy in the invisible face condition did not attain statistical significance, neither in Experiment 1 (M = 52.67%, SE = 1.59%), $t(11) = 1.68$, $p = .12$, nor in Experiment 2 (M = 48.89%, SE = .86%), $t(17) = -1.29$, $p = .21$. These pre- and posttest results for both experiments also held up when using *d'* instead of accuracy. Overall, the visibility tests showed that both groups of participants could reliably identify the gender of the faces in the visible face condition but not in the invisible face condition. As can be seen in Figure 4, however, one participant performed above chance in the posttest of Experiment 1 (65% correct). Interestingly, it is the only participant who completed the posttest on a different day than the pretest and the main experiment.

Which SFs prime?

To examine which SFs within the primes modulate observers' responses to the target, we analyzed the data from the testing phase of Experiment 1 (specifically, block 5 to 25 for each participant; 20,160 trials in total). Trials with incorrect face-gender discrimination responses were excluded. Classification vectors were derived by summing the SF filters of the prime stimuli weighted by the observers' transformed RTs. The RTs were transformed as follows: $(-\log(RT+1) - \text{mean}_{-\log(RT+1)})) / \text{std}_{-\log(RT+1)}$. One classification vector was computed

per condition and per block. One participant classification vector per condition was then calculated for each observer by summing all respective block classification vectors. Group classification vectors were derived for each condition by summing the respective participant classification vectors and transforming the results into Z-scores. To explore the difference in SF tuning between the conditions, we subtracted the group classification vectors for the two awareness conditions and transformed the result into Z-scores. The Cluster test was again used to evaluate statistical significance.

Figure 5 depicts the Z-scored classification vectors for the two prime visibility conditions. The classification vector for the visible prime condition showed positive Z-scores across nearly the whole SF spectrum, with a significant peak at 12.22 cpf (Z_{max} = 3.70; k_{max} = 21). The classification vector for the invisible prime condition showed positive Z-scores for a range of low and high SFs, but without any significant peaks. Strikingly, for mid-SFs, Z-scores became negative, with a significant dip at 11.69 cpf (Z_{min} = -2.84; k_{max} = 5). The difference classification vector (visible prime vs. invisible prime) peaked significantly at 11.69 cpf (Z_{max} = 4.59; k_{max} = 11). We also computed the classification vectors without the outlier participant's data (see the posttest results for Experiment 1) and the difference between our visibility conditions remained significant for basically the same SFs (the peak was at 12.22 cpf; Z_{max} = 4.29; k_{max} = 10).

In sum, the correlation was maximal for the same mid-SFs in both visibility conditions; however, these SFs were linked with fast responses in the visible prime condition, whereas they were linked with *slow* responses in the invisible prime condition.

Priming effects

The sign reversal for the significant Z-scores in Experiment 1 raises the question of the nature of the overall priming effects in the two visibility conditions. To assess the direction and magnitude of priming, we analyzed the RTs from correct trials in the testing phase of Experiment 2. This was done using a repeated-measures ANOVA with the factors prime visibility (invisible prime, visible prime) and presence (prime absent, prime present). The results showed that there was no main effect of prime visibility, $F(1,17) = .75$, $p = .40$, but a significant main effect of presence, $F(1,17) = 83.94$, $p < .001$. The latter reflected faster responses when the primes were present than when they were absent, i.e., facilitatory priming.

There was also a significant visibility × presence interaction, $F(1,17) = 19.18$, $p < .001$. The interaction was driven by a larger facilitatory priming effect in the visible prime (28 ms; 95% confidence interval [20, 34]) than in the invisible prime (9 ms; 95% confidence interval [3, 14]) condition. Contrasts for the interaction term revealed that the priming effects were significant in both visibility conditions [visible: $F(1,17) = 54.32$, $p < .001$; invisible: $F(1,17) = 10.03$, $p < .01$]. No significant effects were found in accuracy; see Table II for mean accuracy and RTs in Experiments 1 and 2, and Table III for RT priming effects for each observer in Experiment 2.

Discussion

We examined which SFs modulate observers' responses as a function of prime awareness. A face-gender repetition priming paradigm adapted from Dehaene and colleagues (2001) allowed us to set up visible and invisible prime conditions that differed solely in the timing of mask onset. By combining this paradigm with the SF Bubbles technique (Willenbockel et al., 2010a), we were able to map in an unbiased manner which SFs influenced observers' RTs. To our knowledge, the present study is one of the first to examine SF tuning as a function of awareness at such a fine SF resolution (see also Willenbockel, Lepore, Nguyen, Bouthillier, & Gosselin, 2012). Our main results show that the same SFs affected RTs in both prime visibility conditions but, surprisingly, in opposite ways.

With regard to SF tuning, we found that information around 12 cpf significantly influenced observers' RTs (given face stimuli subtending a horizontal visual angle of ~3 degrees). This finding closely resembles SF tuning results obtained in other face perception tasks. For instance, an SF band centered between 7 and 16 cpf was found to be optimal for the identification of faces subtending visual angles between 2.3 and 9.5 degrees (e.g., Costen, Parker, & Craw, 1994, 1996; Gaspar, Sekuler, & Bennett, 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999; Willenbockel et al., 2010a). Moreover, SFs around 11 cpf were maximally correlated with observers' accuracy in the direct face-gender discrimination task of the practice phases of Experiments 1 and 2. This suggests that SF tuning per se is robust across face identification, direct face-gender discrimination, and face-gender priming, even under different prime awareness conditions. It also indicates that the masks we used to manipulate prime visibility did not alter SF tuning.

The present SF results extend the findings of De Gardelle and Kouider (2010) and Khalid et al. (in press) by revealing the precise "informational correlates" of conscious and non-conscious priming. Both previous studies demonstrated non-conscious face priming effects for low SFs. Additionally, De Gardelle and Kouider, but not Khalid and colleagues, observed non-conscious priming for high SFs. However, the filtering methods employed in those studies are limited in that they rely on a cutoff frequency for low- and high-pass filtering (3 cycles per degree, which corresponded to approximately 12 cpf). SF Bubbles has the advantage over low-, high-, and band-pass filtering that it is unbiased and can reveal subtle differences in peak SFs or bandwidths (see also Thurman & Grossman, 2011, for a comparison of SF Bubbles with band-pass filtering). This way, potential pitfalls related to the largely arbitrary choice of cutoff frequencies can be avoided. For instance, if our experiment was re-run with primes that were low- and high-pass filtered at 12 cpf, one would expect to see low- and high-SF priming independently of awareness; however, if it was re-run with primes that were low- and high-pass filtered at 17 cpf, one would expect to see low- and high-SF priming in the visible prime condition but only low-SF priming in the invisible prime condition. In fact, as we have shown, the same SFs were maximally correlated with observers' responses in both visibility conditions, with slightly greater absolute values in the visible prime condition.

A surprising aspect of our results is that the significant SFs influenced RTs in opposite directions in the two awareness conditions. Whereas in the visible prime condition, the significant SFs led to fast responses, they led to *slow* responses in the invisible prime condition. Such a reversal was seen neither in De Gardelle and Kouider's (2010) nor in Khalid et al.'s (in press) RT priming effects. This could be due to methodological differences, such as the choice of cutoff frequencies—possibly, the reversal would have been present for a mid-SF band-pass condition containing, e.g., SFs between 8 and 16 cpf.

To shed light on the nature of the overall priming effects in our paradigm, we ran a control experiment (Experiment 2). It allowed us to compare the RTs from prime-present trials with those from prime-absent trials, separately for the two prime awareness conditions. Priming studies often measure congruence effects (e.g., same-gender vs. different-gender prime-target trials in Khalid et al., in press). However, the purpose of Experiment 2 was to uncover how the difference we observed in SF tuning (in Experiment 1) translates into overall

RT effects, while introducing as few changes with regard to Experiment 1 as possible. Therefore, the prime-present vs. prime-absent contrast seemed more appropriate: First, in Experiment 1, prime and target were always identical. Adding a different-gender condition at this stage would have introduced the potential confound that different gender implies different identities and different photos. Second, a different-gender condition could potentially have led to a change in the observers' strategy, since prime information would not always have been useful (e.g., participants might have paid less attention to the primes overall compared to Experiment 1).

The results of Experiment 2 revealed *facilitatory* priming effects of 9 ms and 28 ms for the invisible and visible prime conditions, respectively. By themselves, these findings appear consistent with previous face-gender priming results (e.g., Finkbeiner & Palermo, 2009; Goshen-Gottstein & Ganel, 2000; Henson et al., 2003; Khalid et al., in press). For instance, Finkbeiner and Palermo (2009) observed faster responses to target faces displaying the same gender as the invisible (masked) prime faces than to targets of opposite gender (congruence effects of 10 ms, 9 ms, and 8 ms for different SOA conditions in Experiment 3). Khalid and colleagues found congruence effects of 4–6 ms for invisible low-SF primes. The studies by Goshen-Gottstein and Ganel (2000) and Henson and colleagues (2003) revealed long-lag repetition priming effects (i.e., shorter RTs to repeated versus unrepeated faces, with repetition lags of ~10 min) of approximately 20–30 ms for consciously perceived familiar and unfamiliar faces. Furthermore, several studies using different tasks demonstrated that priming effects are larger when the primes are consciously perceived than when they are rendered non-conscious (e.g., De Gardelle & Kouider, 2010; Kouider et al., 2007). Thus, the results of Experiment 2 closely replicate classic priming effects, both in terms of the direction of the priming (i.e., facilitatory) and magnitude.

How can these facilitatory priming effects observed in Experiment 2 be reconciled with the SF results obtained in Experiment 1? All "blind" statistical tests for classification images—including the Cluster test (Chauvin et al., 2005) that we employed—assume relatively focal signals. Accordingly, several SFs that show positive Z-scores could have facilitated observers' responses despite being outside of the significant SF band in both classification vectors. In the visible prime condition, the band of SFs that attained statistical significance would also have facilitated observers' responses, resulting in an enhanced facilitatory effect. In the invisible

prime condition, in contrast, the same band of SFs would have somewhat hindered observers' responses, resulting overall in a relatively small facilitatory effect.

This does not explain, however, why the influences of the diagnostic SFs are reversed as a function of awareness. This reversal could be related to a number of previous results demonstrating opposite priming influences in contexts of different awareness conditions (e.g., Eimer & Schlaghecken, 2002; Frings & Wentura, 2005; Klapp & Hinkley, 2002; for reviews see Eimer & Schlaghecken, 2003; Sumner, 2007). For example, Eimer and Schlaghecken (2002) found negative compatibility effects when the primes were not consciously perceived but positive compatibility effects when the primes did reach the observers' awareness. The negative compatibility effect was initially demonstrated using a left/right forced choice task with arrows as stimuli (Eimer & Schlaghecken, 1998, 2002) but has now been replicated in other paradigms, for instance, with emotional faces (Bennett, Lleras, Oriet, & Enns, 2007). Possibly, the results in our invisible prime condition reflect signs of a transition from positive to negative priming, with most SFs facilitating responses (hence the overall positive priming effect) but the SF band most systematically facilitating responses in the visible condition inhibiting responses (hence the local negative priming effect). It is possible that if we had measured gender congruence effects in our control task instead of the prime-present vs. prime-absent contrast, we would have observed a negative priming effect.

It is still debated which factors determine the direction of priming effects (e.g., see Sumner, 2007, for a review). It has been suggested that there is a causal link between prime awareness and the direction of the compatibility effects (e.g., Eimer & Schlaghecken, 2002; Frings & Wentura, 2005). However, there are also results speaking against this possibility (e.g., Verleger, Jaśkowski, Aydemir, Van der Lubbe, & Groen, 2004; see Sumner, 2007, for a review). Alternatively, it has been suggested that interactions between the mask and the prime might play a role (e.g., see Sumner, 2007, for a review). It would be conceivable that the increase in RTs observed in our study is related to interactions between information at ~12 cpf in the invisible prime face and the immediately following mask; this interaction might not be present in the visible prime condition due to the inserted blank. In any case, more work will be needed to clarify the links between the direction of priming, masking, and awareness.

Several discussions in the field of non-conscious perception have emphasized the importance of revealing qualitative differences in the effects obtained with visible and

invisible stimuli. In fact, reliably demonstrating non-conscious perception has posed many challenges (e.g., for reviews see Holender, 1986; Kouider & Dehaene, 2007), and revealing differences of a qualitative nature has been suggested to be the most convincing way to show that the conscious/non-conscious distinction is meaningful (Cheesman & Merikle, 1986). Here, we examined the possibility that qualitative differences reside in the visual information that is encoded and processed consciously vs. non-consciously. Results showing that, for instance, low SFs are processed non-consciously whereas high SFs are not available during non-conscious processing would have strongly supported the qualitative views (see also Khalid et al., in press). However, the present results did not reveal differences in SF tuning per se. This could mean that the same underlying process(es) played a role in both awareness conditions (e.g., Holender & Duscherer, 2004; Perruchet & Vinter, 2002). It could also be that the composite RT measure that we employed does not contain enough signal to reveal subtle SF tuning differences in foveal vision. In a number of other recent studies results have been reported that were interpreted as qualitative differences between conscious and non-conscious perception (e.g., Barbot & Kouider, 2012; Eimer & Schlaghecken, 2002; Frings & Wentura, 2005; Snodgrass & Shevrin, 2006); both Willenbockel et al. (2012) and Khalid et al. (in press) found differences in SF tuning as a function of awareness in face perception. We think that two interesting avenues for future research would be to map SF tuning with a high resolution in peripheral vision (see Khalid et al., in press) and to further investigate the reversal of priming influences we observed specifically for the SFs that are diagnostic for the task at hand.

Acknowledgments

We thank Josiane Jauniaux for help with data collection for Experiment 2 and Quoc Vuong and two anonymous reviewers for their helpful comments and suggestions.

Footnotes

[1] By "invisible" we mean "gender-invisible" (see also, e.g., Finkbeiner & Palermo, 2009; Khalid et al., in press).

[2] MATLAB code to generate the masks:

```
mask=abs(imresize(randn(30,30),[256,256]));
mask(mask>1)=1;
mask=(mask-.5)*.85+.5;
```

References

Abrams, R. L., Klinger, M. R., & Greenwald, A. G. (2002). Subliminal words activate semantic categories (not automated motor responses). *Psychonomic Bulletin and Review, 9,* 100–106.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*, 600–609.

Barbot, A., & Kouider, S. (2012). Longer is not better: nonconscious overstimulation reverses priming influences under interocular suppression. *Attention, Perception, & Psychophysics, 74,* 174–184.

Bennett, J. D., Lleras, A., Oriet, C., & Enns, J. T. (2007). A negative compatibility effect in priming of emotional faces. *Psychonomic Bulletin and Review, 14*, 908–912.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Breitmeyer, B. G., & Öğmen, H. (2006). *Visual masking: time slices through conscious and unconscious vision*. Oxford: Oxford University Press.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36*, 96–107.

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics, 72,* 1444–1449.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, *5*, 659–667.

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology, 40*, 343–367.

Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception, 23*, 129–146.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics, 58*, 602–612.

De Gardelle, V., & Kouider, S. (2010). How spatial frequencies and visual awareness interact during face processing. *Psychological Science, 21*, 58–66.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*, 200–227.

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*, 204–211.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., Van de Moortele, P. F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature, 395*, 597–600.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision.* New York, NY: Oxford University Press.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activation: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1737–1747.

Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: Evidence from masked priming. *Psychonomic Bulletin and Review, 9*, 514–520.

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology, 64*, 7–26.

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious

processing: A comparison of face and nonface stimuli. *Psychological Science, 20*, 42–51.

Frings, C., & Wentura, D. (2005). Negative priming with masked distractor-only prime trials: awareness moderates negative priming. *Experimental Psychology, 52*, 131–139.

Gaspar, C., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*, 2817–2826.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*, 3537–3560.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1198–1214.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13*, 793–805.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences, 9*, 1–66.

Holender, D., & Duscherer, K. (2004). Unconscious perception: The need for a paradigm shift. *Perception & Psychophysics, 66*, 872–881.

Khalid, S., Finkbeiner, M., König, P., & Ansorge, U. (in press). Subcortical human face processing? Evidence from masked priming. *Journal of Experimental Psychology: Human Perception and Performance.*

Klapp, S. T., & Hinkley, L. B. (2002). The negative compatibility effect: Unconscious inhibition influences reaction time and response selection. *Journal of Experimental Psychology: General, 131,* 255–269.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362*, 857–875.

Kouider, S., Dehaene, S., Jobert, A., & Le Bihan, D. (2007). Cerebral bases of subliminal and supraliminal priming during reading. *Cerebral Cortex, 17*, 2019–2029.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science, 240*, 740–749.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science, 13*, 416–424.

Naccache, L., & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition, 80,* 215–229.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*, 3824–3833.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25*, 297–330.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5*, 195–200.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: Replicable and qualitatively different unconscious perceptual effects. *Cognition, 101*, 43–79.

Sumner, P. (2007). Negative and positive masked-priming – implications for motor inhibition. *Advances in Cognitive Psychology, 3*, 317–326.

Thurman, S. M., & Grossman, E. D. (2011). Diagnostic spatial frequencies and human efficiency for discriminating actions. *Attention, Perception, & Psychophysics, 73,* 572–580.

Verleger, R., Jaśkowski, P., Aydemir, A., Van der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and nonconscious processing? On inverse priming induced by masked arrows. *Journal of Experimental Psychology: General*, *133,* 494–515.

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010a). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010b). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*, 671–684.

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, 3:237*, doi: 10.3389/fpsyg.2012.00237.

Table captions

Table I. Spatial frequency (SF) tuning results for individual observers in the practice phases of Experiments 1 and 2. The table gives the number of trials, peak SF in cycles per face width (cpf), maximum Z-score, cluster size, and significance as evaluated using the Cluster test (Chauvin et al., 2005).

Table II. Face-gender judgment accuracy and response times (RTs) in the testing phases of Experiments 1 and 2.

Table III. Individual priming effects (response times from prime absent trials minus response times from prime present trials) for the two visibility conditions in the testing phase of Experiment 2.

Figure captions

Figure 1. Sample stimuli. The top left panel shows one of the base face images and a plot of its spatial frequency (SF) content (energy plotted as a function of SF in cycles per image, cpi). The other panels show sample results after filtering the base face image with the SF Bubbles technique on three hypothetical trials.

Figure 2. Illustration of the experimental paradigm used in Experiments 1 and 2. A) Stimuli sequence for the practice phase. The duration of the face image was adjusted trial-by-trial to maintain accuracy at 90%. B) Stimuli sequence for the pre- and post visibility tests. The stimuli sequence at the top corresponds to the *visible face* condition and the stimuli sequence at the bottom to the *invisible face* condition. The only difference between the two conditions is the temporal order of the masks and blanks. C) Visible and invisible prime sequences used in the testing phases of both experiments, which included a full-spectrum target face at the end of each trial. Note that in Experiment 2, the prime face was replaced by a blank on 50% of the trials.

Figure 3. Spatial frequency tuning for direct, accurate face-gender discrimination. The graph depicts the group classification vectors derived from the practice phases of Experiment 1 (red; the pink area shows the 95% confidence interval) and Experiment 2 (green; the light green area shows the 95% confidence interval). Stars mark the significant segments of the classification vectors.

Figure 4. Pre- and posttest results from A) Experiment 1 and B) Experiment 2 for the visible and invisible face conditions. Data is depicted by box plots, with red horizontal lines indicating medians, edges of the blue boxes indicating 25^{th} and 75^{th} percentiles, black whiskers extending to the extremes not considered outliers, and red crosses marking outliers.

Figure 5. Spatial frequency tuning for conscious and non-conscious face-gender priming. The graph shows the group classification vectors derived from the testing phase of Experiment 1 for the visible prime condition (blue), the invisible prime condition (red), and the normalized difference between the two (green). The light areas show the respective 95% confidence intervals. Stars mark the significant segments of the classification vectors.

Tables

Table I: Spatial frequency results for individual observers in the practice phase

Experiment	Participant	Number of trials	Peak SF	Z_{max}	Cluster size *k*	Significance
1	1	1600	7.97	3.99	20	*
	2	2900	14.34	6.32	49	*
	3	1500	13.28	6.29	29	*
	4	800	10.63	3.22	8	*
	5	300	9.56	3.19	11	*
	6	3100	10.63	5.96	24	*
	7	600	9.56	2.79	6	*
	8	2000	12.22	3.51	12	*
	9	1000	13.28	2.78	10	*
	10	1100	11.69	3.23	7	*
	11	1500	11.69	3.76	23	*
	12	1000	7.97	3.47	9	*
2	1	1200	17.53	3.03	15	*
	2	3000	11.16	4.97	23	*
	3	2800	10.63	5.86	39	*
	4	2200	7.97	3.17	25	*
	5	1200	14.88	5.26	25	*
	6	1300	8.50	2.82	4	n.s.
	7	2400	11.69	2.32	2	n.s.
	8	3200	15.94	4.60	37	*
	9	4900	9.03	8.14	27	*
	10	1500	8.50	4.25	11	*
	11	1400	10.63	2.80	7	*
	12	1900	8.50	4.13	17	*
	13	700	11.16	4.10	14	*
	14	900	14.88	4.26	27	*
	15	700	12.21	3.67	17	*
	16	2100	13.28	3.90	20	*
	17	300	4.78	1.39	0	n.s.
	18	2700	17.53	4.42	104	*

* significant using the Cluster test (k_{crit} = 4.81 pixels, $p < .05$, two-tailed, $Z_{arbitrary} = \pm 2.3$, S_r = 128, *FWHM* = 4.24; Chauvin et al., 2005); n.s. = not significant

Table II: Face-gender judgment accuracy and response times in the testing phase

Exp.	Condition	Accuracy (% correct)				RT (ms)			
		Mean	**Min**	**Max**	**95% confidence interval**	**Mean**	**Min**	**Max**	**95% confidence interval**
1	Visible prime	96.09	93.21	98.21	[95.34, 96.83]	500.84	439.22	814.84	[465.68, 620.37]
	Invisible prime	95.70	91.19	98.45	[94.36, 96.65]	521.52	454.27	842.44	[485.55, 618.44]
2	Visible prime	94.25	87.50	99.50	[92.72, 95.33]	578.46	527.60	664.78	[565.16, 595.73]
	Visible blank	94.19	87.00	99.00	[92.53, 95.58]	606.19	564.42	703.36	[593.62, 627.86]
	Invisible prime	94.47	91.50	99.50	[93.49, 95.77]	589.35	538.55	689.55	[574.65, 609.25]
	Invisible blank	94.56	89.00	99.50	[93.22, 95.92]	597.93	549.95	719.75	[581.80, 623.88]

Table III: Priming effects for individual observers in Experiment 2

Participant	Priming effect (ms)	
	Visible prime	Invisible prime
1	12	14
2	-11	21
3	27	13
4	29	16
5	27	8
6	20	17
7	45	-2
8	6	4
9	8	0
10	39	30
11	37	-16
12	43	-13
13	29	6
14	28	6
15	26	18
16	42	11
17	51	6
18	40	16

Figures

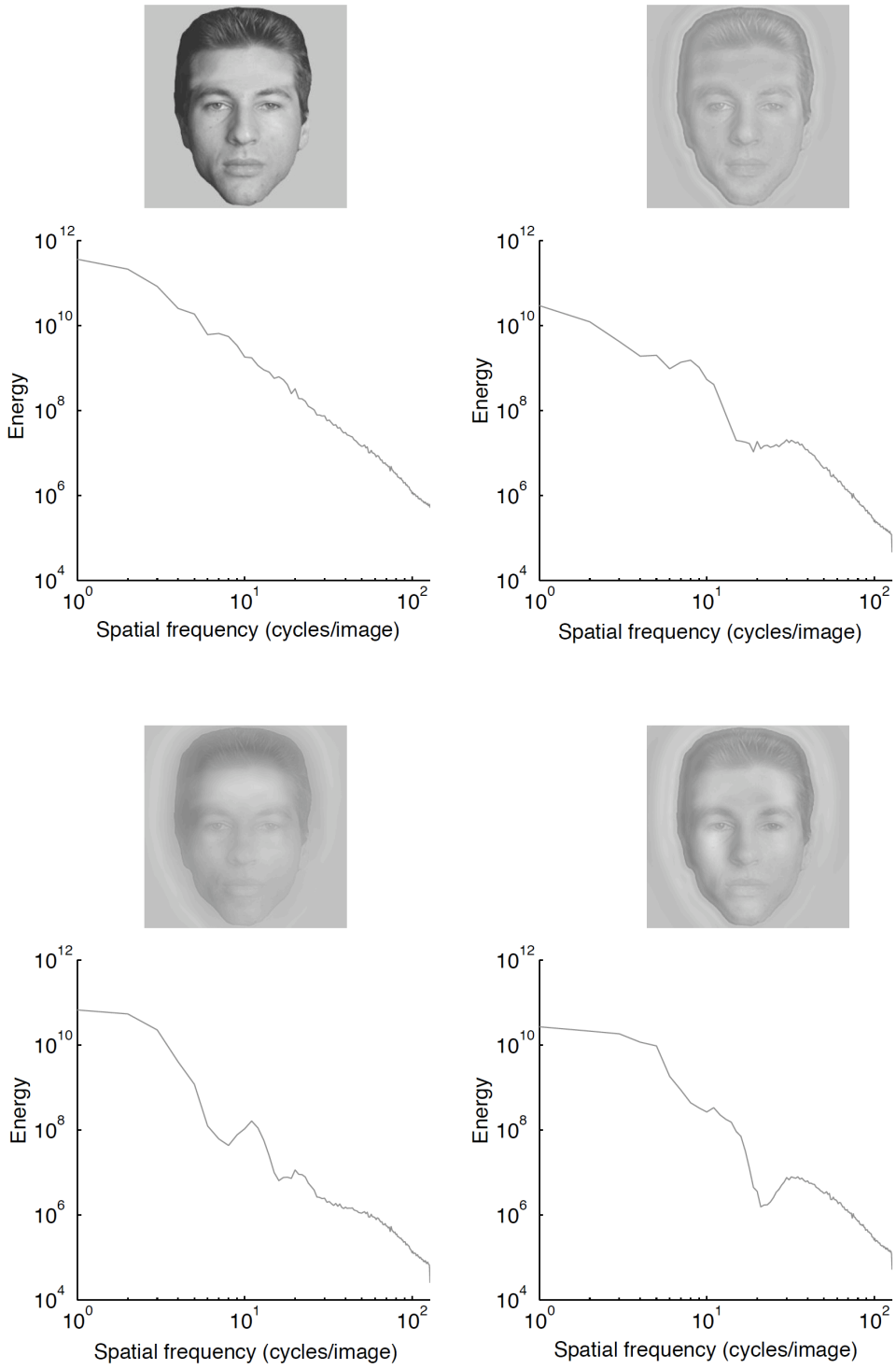


Figure 1: Sample stimuli and spectral content

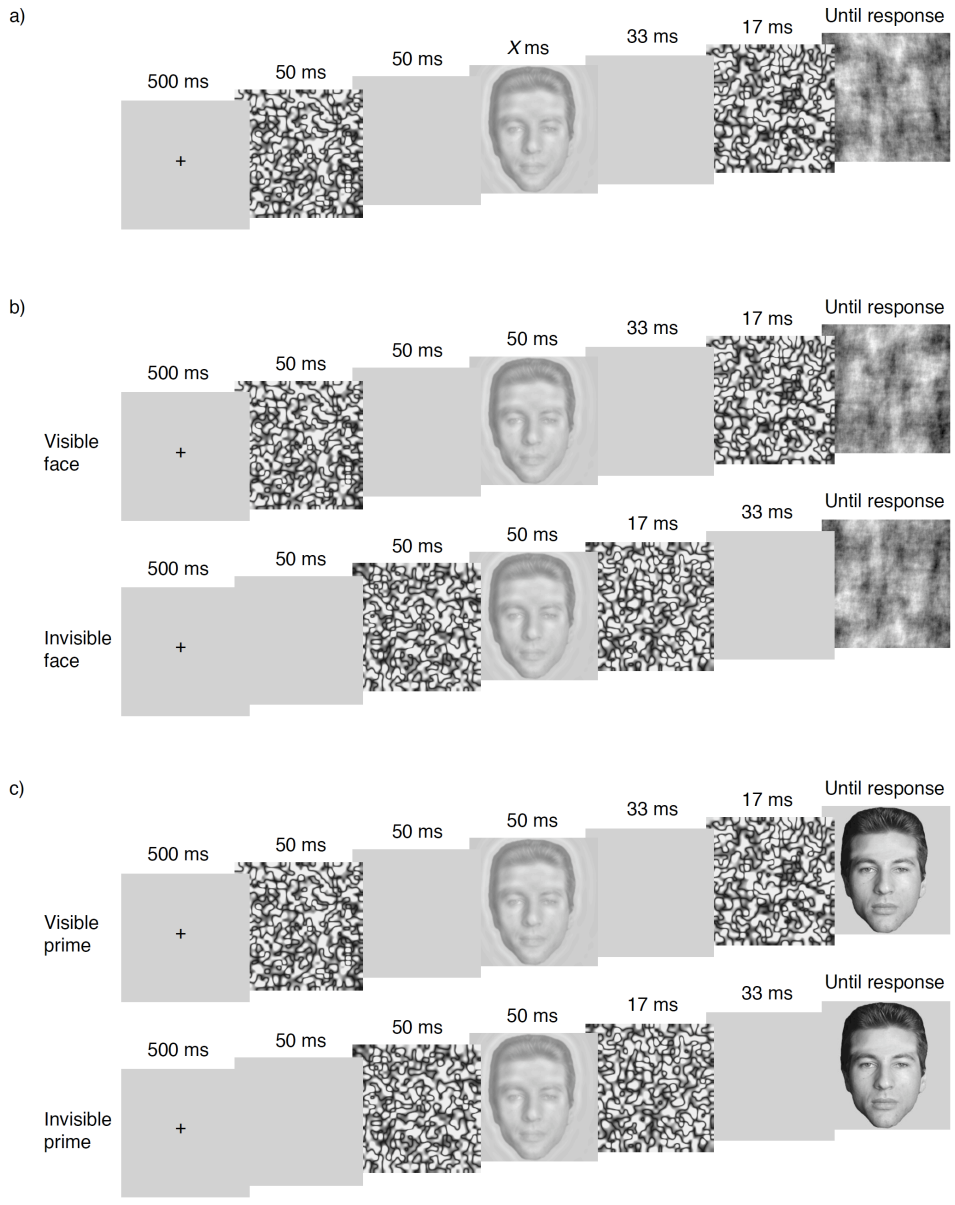


Figure 2: Experimental paradigm

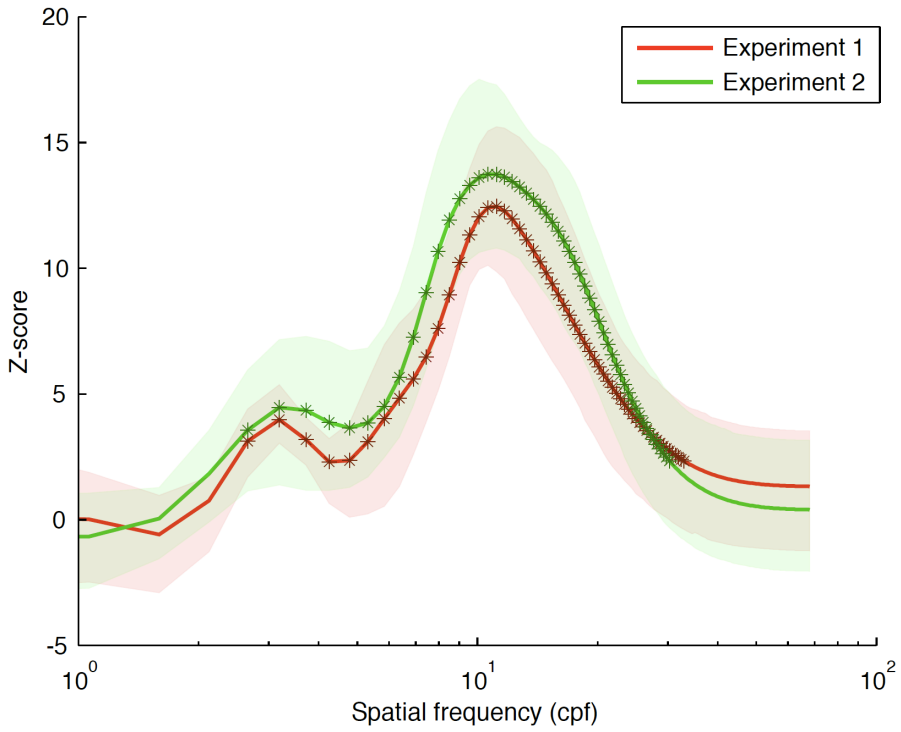


Figure 3: Spatial frequency tuning for accurate face-gender discrimination

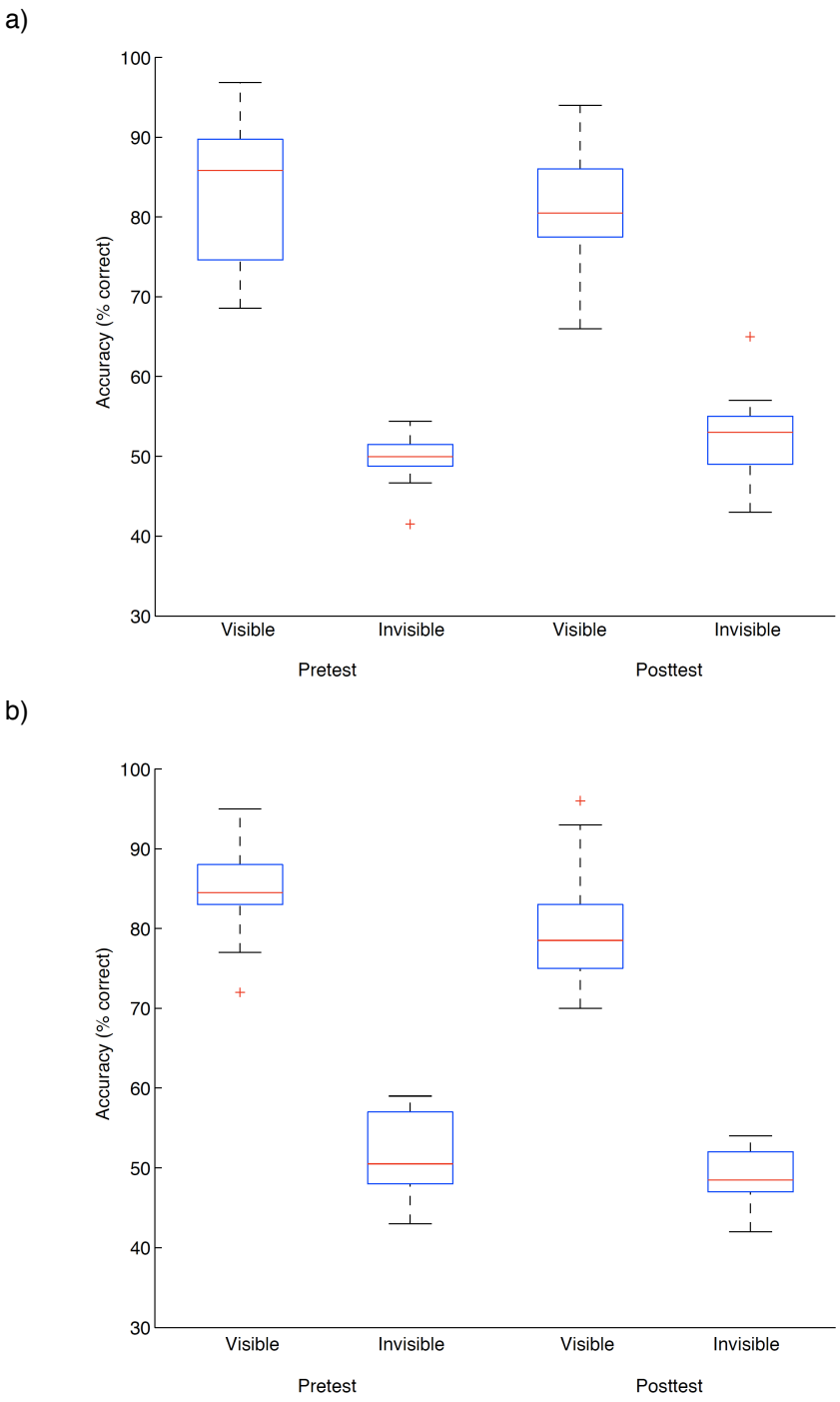


Figure 4: Visibility test results

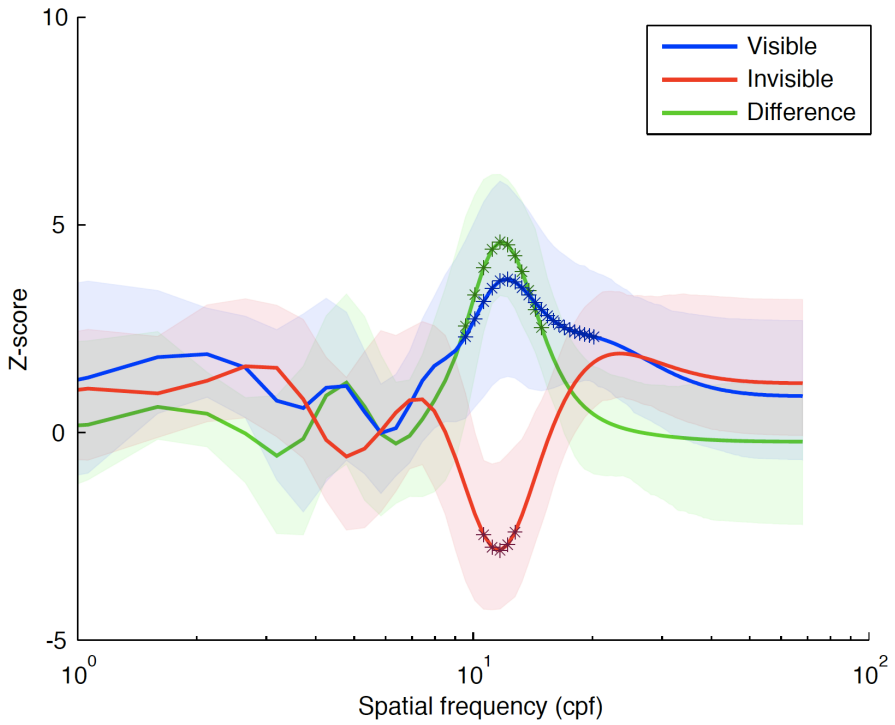


Figure 5: Spatial frequency tuning for conscious and non-conscious priming

Transitional text # 3

In Chapter 4, we mapped the informational correlates of consciousness in the context of face perception using behavioral measures. We did not find any changes in SF tuning as a function of awareness. However, it remains possible that qualitative differences can be revealed when examining SF tuning at a higher temporal and spatial resolution, i.e., for individual brain regions or at certain processing stages. This was tested in Chapter 5, again using the SF Bubbles technique. We "zoomed in" by exploiting the high spatial and temporal resolution of intracranial ERPs to map SF tuning for specific brain regions over time. We had the opportunity to record from brain regions associated with the processing of emotional facial expressions (the insula and the amygdala) in patients undergoing epilepsy monitoring.

Chapter 5: Fourth article

Article published:

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, 3:237*. doi: 10.3389/fpsyg.2012.00237.

Spatial frequency tuning during the conscious and non-conscious perception of emotional facial expressions—an intracranial ERP study

Verena Willenbockel[1], Franco Lepore[1], Dang Khoa Nguyen[2], Alain Bouthillier[2], & Frédéric Gosselin[1]

[1]CERNEC, Département de Psychologie, Université de Montréal, Montréal, QC, Canada
[2]Centre Hospitalier de l'Université de Montréal, Hôpital Notre-Dame, Montréal, QC, Canada

Abstract

Previous studies have shown that complex visual stimuli, such as emotional facial expressions, can influence brain activity independently of the observers' awareness. Little is known yet, however, about the "informational correlates" of consciousness—i.e., which low-level information correlates with brain activation during conscious vs. non-conscious perception. Here, we investigated this question in the spatial frequency (SF) domain. We examined which SFs in disgusted and fearful faces modulate activation in the insula and amygdala over time and as a function of awareness, using a combination of intracranial event-related potentials (ERPs), SF Bubbles (Willenbockel et al., 2010a), and Continuous Flash Suppression (CFS; Tsuchiya & Koch, 2005). Patients implanted with electrodes for epilepsy monitoring viewed face photographs (13° × 7°) that were randomly SF filtered on a trial-by-trial basis. In the conscious condition, the faces were visible; in the non-conscious condition, they were rendered invisible using CFS. The data were analyzed by performing multiple linear regressions on the SF filters from each trial and the transformed ERP amplitudes across time. The resulting classification images suggest that many SFs are involved in the conscious and non-conscious perception of emotional expressions, with SFs between 6 and 10 cycles per face width being particularly important early on. The results also revealed qualitative differences between the awareness conditions for both regions. Non-conscious processing relied on low

SFs more and was faster than conscious processing. Overall, our findings are consistent with the idea that different pathways are employed for the processing of emotional stimuli under different degrees of awareness. The present study represents a first step to mapping how SF information "flows" through the emotion-processing network with a high temporal resolution and to shedding light on the informational correlates of consciousness in general.

Keywords: consciousness, emotional facial expressions, spatial frequency

Introduction

The look on someone's face can speak volumes. Emotional facial expressions convey a wealth of information, such as cues about a person's state of mind or warning signs of potentially threatening situations (e.g., reflected by fear) or materials (e.g., reflected by disgust). Human faces and brains are thought to have co-evolved to be efficient transmitters and decoders of emotional signals, respectively (Schyns, Petro, & Smith, 2007, 2009; Smith, Cottrell, Gosselin, & Schyns, 2005). Moreover, it has been claimed that emotional information from a face can be extracted without the observer's awareness (see Tamietto & De Gelder, 2010, for a review). Numerous studies have shown that face stimuli rendered "invisible" using techniques such as backward masking (e.g., Smith, in press), binocular rivalry (e.g., Williams, Morris, McGlone, Abbott, & Mattingley, 2004), or Continuous Flash Suppression (e.g., Jiang & He, 2006; Jiang et al., 2009; Tsuchiya & Koch, 2005) can be processed sufficiently for the healthy brain to distinguish neutral from emotional expressions, including fear, disgust, and happiness. Differential brain responses to both invisible and visible facial expressions have been measured, for instance, using functional magnetic resonance imaging (fMRI) (e.g., Jiang & He, 2006; Williams et al., 2004) and surface event-related potentials (ERPs) (e.g., Jiang et al., 2009; Smith, in press). Thus, it is now widely thought that facial expressions can influence neural activity and behavior independently of awareness, and that they constitute a stimulus class well suited for investigating differences between conscious and non-conscious perception in the human brain.

One fundamental question, which is the focus of the present article, concerns which "low-level" aspects of facial-expression signals modulate brain responses as a function of awareness. Faces are complex stimuli that contain information at various spatial frequencies (SFs). Broadly speaking, low SFs represent the coarse information in an image (e.g., luminance blobs), whereas high SFs represent the fine-grained information (e.g., fine wrinkles in a face). It is well known that the visual system filters any retinal input with multiple quasi-linear band-pass filters, each tuned to a specific range of SFs (see De Valois & De Valois, 1990, for a review). The contribution of different SFs to the perception of facial expressions has been investigated in a number of fMRI (Morawetz, Baudewig, Treue, & Dechent, 2011; Vuilleumier, Armony, Driver, & Dolan, 2003; Winston, Vuilleumier, & Dolan, 2003) and

surface ERP (Holmes, Winston, & Eimer, 2005; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Schyns et al., 2007, 2009; Vlamings, Goffaux, & Kemner, 2009) studies. However, the studies led to mixed findings and were limited in several respects. For instance, the low temporal resolution of fMRI and the low spatial resolution of surface ERPs did not allow for conclusions to be drawn about the precise temporal dynamics of SF processing in specific brain regions. Moreover, the SF filtering methods that were employed (low-pass, high-pass, or band-pass filtering) provided only a crude estimate of SF tuning. Also, the studies were restricted to consciously perceived face stimuli. Therefore, not much is known yet about the "informational correlates" of consciousness in this context—i.e., precisely which SFs are correlated with localized brain signals during the conscious vs. non-conscious perception of emotional expressions.

The aim of the present study was to examine which SFs are correlated with brain signals in specific regions of the emotion-processing network under different awareness conditions. We had the opportunity to record intracranial ERPs from the insula and, to a lesser extent, from the amygdala of patients undergoing monitoring for medically intractable epilepsy. The insula and amygdala have previously been associated with the processing of disgust and fear, respectively (e.g., Adolphs, Tranel, Damasio, & Damasio, 1994, 1995; Phillips et al., 1997, 1998, 2004; Krolak-Salmon et al., 2003; Krolak-Salmon, Hénaff, Vighetto, Bertrand, & Mauguière, 2004; but for evidence that the insula also responds to fear, see, e.g., Morris et al., 1998, and for evidence that the amygdala also responds to disgust, see Fitzgerald, Angstadt, Jelsone, Nathan, & Phan, 2006; Van der Gaag, Minderaa, & Keysers, 2007; Winston, O'Doherty, & Dolan, 2003). Here, we traced which SFs in disgusted and fearful faces modulate activation in these two interconnected brain structures over time. Our study employed a novel combination of Continuous Flash Suppression (Tsuchiya & Koch, 2005), intracranial recordings, and the SF Bubbles technique (Willenbockel et al., 2010a). We will elaborate on the three methods in the following paragraphs and briefly review some of their applications in previous studies.

CFS is a powerful method to render visual stimuli invisible (Tsuchiya & Koch, 2005). One of its main strengths is that it allows for suppressing stimuli from awareness for a long duration (i.e., up to several seconds). A second strength of CFS is that the onset of the suppression can be precisely timed. CFS involves presenting a static image to one of the

observer's eyes, while dynamic high-contrast noise (e.g., Mondrian patterns flashed at 10 Hz; Tsuchiya & Koch, 2005) is presented to the other eye. As a result of this dichoptic stimulation, typically only the noise is consciously perceived; the static stimulus is suppressed from the observer's awareness but nevertheless processed in the brain (e.g., Jiang & He, 2006; Jiang et al., 2009; Tsuchiya & Koch, 2005). Using CFS and fMRI, Jiang and He (2006) found that suppressed fearful compared with scrambled faces elicited significant activation in the fusiform face area, superior temporal sulcus, and the bilateral amygdalae. The amygdalae were also more activated by fearful than by neutral faces, independently of awareness. Using CFS in combination with surface ERPs, Jiang and colleagues (2009) observed significant amplitude differences to suppressed fearful vs. scrambled faces beginning at 140 ms and to suppressed fearful vs. neutral faces starting at 220 ms after stimulus onset. Overall, combining CFS with fMRI, which has a high spatial resolution, or with surface ERPs, which have a high temporal resolution, has provided important insights into the "where" or "when" of non-conscious facial expression processing—but not both aspects simultaneously. In the present study, we combined CFS with intracranial recordings, which combine some of the advantages of fMRI and surface ERPs.

It has been argued that intracranial recordings currently provide the best combination of high temporal *and* high spatial resolution, plus large anatomical field-of-view and wide frequency bandwidth (Tsuchiya, Kawasaki, Oya, Howard, & Adolphs, 2008). A number of previous intracranial ERP studies with patients undergoing epilepsy monitoring investigated the temporal dynamics of conscious emotional facial expression processing in the insula and amygdala (Krolak-Salmon et al., 2003, 2004; Pourtois, Spinelli, Seeck, & Vuilleumier, 2010). Krolak-Salmon and colleagues (2003) found amplitude differences to disgusted vs. neutral, fearful, and happy expressions in the ventral anterior insula. This "disgust effect" started at approximately 300 ms post stimulus onset when observers were engaged in an expression task and approximately 100 ms later when they performed a face-gender task. In a similar study, a "fear effect" was observed in the amygdala, starting at 200 ms in an expression task and later (after 600 ms) in a face-gender task (Krolak-Salmon et al., 2004). Pourtois and colleagues (2010) observed earlier amplitude differences to fearful vs. neutral faces in the amygdala, starting at 140 ms post stimulus onset. This early effect was not affected by attention but an attentional modulation of emotional responses occurred at longer latencies (after 700 ms).

Intracranial ERPs were also used to study amygdala activation to masked emotional words (Naccache et al., 2005). Differences between invisible threatening and neutral words were found after 800 ms post stimulus onset. In the current study, we combined intracranial recordings with CFS to investigate the temporal dynamics of non-conscious emotional expression processing in these brain regions. Furthermore, we went beyond previous studies by examining precisely which SFs in fearful and disgusted faces modulate brain signals over time by combining intracranial recordings with the SF Bubbles technique (Willenbockel et al., 2010a).

The Bubbles method (Gosselin & Schyns, 2001) is a classification image technique that can be used to reveal which stimulus information modulates observers' behavioral (e.g., Adolphs et al., 2005; Smith et al., 2005) or brain (Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin, 2003; Smith, Gosselin, & Schyns, 2004, 2007; Schyns et al., 2007, 2009) responses. SF Bubbles (Willenbockel et al., 2010a) is a variant of the technique that can be employed to examine which information in the SF domain correlates with observers' responses. SF Bubbles involves randomly sampling the energy of visual stimuli at different SFs on a trial-by-trial basis and then performing a multiple linear regression on the information samples and the response measure of interest to precisely reveal the SF tuning curves for a given task. For example, Willenbockel and colleagues (2010a) used the technique to compare the SF tuning of upright and inverted face identification, and Thurman and Grossman (2011) employed it to investigate SF tuning for discriminating videos of human actions. In the latter study, the results obtained with SF Bubbles were directly compared with those from a more traditional band-pass filtering approach. The results from both methods were consistent but the authors stressed that SF Bubbles offers several advantages. Specifically, SF Bubbles allows for deriving SF tuning curves—spanning the whole SF spectrum—at a much higher resolution and based on a smaller number of trials. A second strength of the method is that randomly sampling multiple SFs simultaneously on a trial-by-trial basis minimizes the risk that participants adapt to a predictable stimulus manipulation (e.g., band-, low-, or high-pass filtering or critical band masking; see Sowden & Schyns, 2006, for evidence of "channel surfing"). Moreover, SF Bubbles is unbiased in that no cutoff frequencies have to be chosen—a parameter that differs considerably between previous

experiments using traditional filtering methods (for examples from the emotion-processing literature see, e.g., Morawetz et al., 2011; Vlamings et al., 2009; Vuilleumier et al., 2003).

The combination of SF Bubbles with intracranial recordings employed in the current study allowed us to map the SF tuning of the insula and amygdala over time. In one condition, we used CFS to render SF filtered disgusted and fearful faces invisible (i.e., dynamic Mondrian patterns were presented to one eye while an "SF bubblized" emotional face was presented to the other eye). In the other condition, the filtered faces were visible (i.e., an "SF bubblized" face was presented to both eyes). Overall, this study represents a unique opportunity to shed light on the neural processing dynamics for ecologically important visual information as a function of awareness.

Material and methods

Participants

Three patients with medically intractable epilepsy gave their written informed consent and participated in this experiment. The patients were undergoing epilepsy monitoring at the Hôpital Notre-Dame, Montréal, to guide neurosurgical treatment. For this purpose, they had electrodes implanted under a clinical protocol; the electrode locations were chosen solely based on medical considerations. Our study was approved by the CHUM (Centre Hospitalier de l'Université de Montréal) ethics committee and took place at the hospital approximately 6–10 days after the electrode implantation. The participants were naïve to the awareness aspect of the study until the debriefing after the experiment. All of them had normal or corrected-to-normal vision; further participant information is summarized in Table I.

Anatomical location of the electrodes of interest

All patients had depth electrodes (Ad-Tech Medical Instrument Corporation, Racine, WI, USA) implanted in the insula, and additional grid, strip, or depth electrodes in other regions. One of the patients had a depth electrode implanted in the amygdala. The implantation schemes are described in detail in a previous article (Surbeck et al., 2011). Patient 1 underwent an open microdissection of the Sylvian fissure (Type I implantation). In the anterior, medial, and posterior insula each, she had a Spencer depth electrode with a diameter of 1.1 mm, which featured four contacts along its length. The contacts were of 2.3 mm in length and spaced 5 mm apart from center to center. Two contacts per electrode ended

up in the insular cortex. In the amygdala, she also had a depth electrode with four contacts (1.1 mm diameter, 2.3 mm length, 10 mm spacing). Patient 2 underwent the combined Yale-Grenoble stereotactic implantation (Type II). In the anterior and posterior insula each, he was implanted with a 10-contact Spencer depth electrode (1.1 mm diameter, 2.3 mm length, 10 mm spacing). Patient 3 underwent a Type I implantation with a new hybrid operculo-insular electrode (Ad-Tech Medical Instrument Corporation, Wisconsin, USA), among other regular subdural electrodes. The hybrid electrode combined the design of a depth and a subdural strip electrode. The depth component featuring two contacts was implanted into the insular cortex. The length of that segment was 10 mm and the diameter 1.1 mm. The length of each contact was 2.4 mm. Further information can be found in an article by Bouthillier, Surbeck, Weil, Tayah, and Nguyen (2012).

High-resolution MRIs with 1-mm-thick slices were obtained after the implantation to determine the exact position of the electrodes (Figure 1). A 3D representation of the electrodes with respect to the patient's brain was generated using GridView software (Stellate Systems Inc., Montreal, QC, Canada; see also Wang, Agarwal, Nguyen, Domocos, & Gotman, 2005). In the analyses presented here we included two contacts per electrode implanted either in the anterior insula (Participants 1–3) or in the amygdala (Participant 1).

Electrophysiological recording and stimulus display

Intracranial EEG was recorded at 2 kHz using a Stellate Harmonie system (Stellate Systems Inc., Montreal, QC, Canada). Either a subdural parietal contact (Participant 1), a subdural temporal contact (Participant 2), or the mastoids (Participant 3) served as a reference. This heterogeneity was not a concern to us because we were interested in the correlations between random SF filters and trial-by-trial voltage variations, which are robust to reference changes. The timing of the stimulus onsets was determined based on the recording of digital trigger signals by the Stellate eAmp using the eAMP Trigger Interface. A dual core 2.19 GHz PC (AMD Athlon 64 X2 4200+) and a 17" LCD display (VE700, ViewSonic, CA, USA) were used for presenting the stimuli. The gamma parameter was set to 1, to linearize the relationship between the RGB values and corresponding luminance values. The refresh rate was 60 Hz and the resolution 1024 × 768 pixels. The luminance range in the green channel was diminished to match the red channel, which typically has a lower maximum luminance

(min = 0.4 cd/m^2, max = 33.3 cd/m^2). All stimuli were shown on a gray background (13.57 cd/m^2) using the Psychophysics toolbox (Brainard, 1997; Pelli, 1997) for MATLAB (Mathworks, Natick, MA, USA).

Stimuli

Twelve grayscale face photographs (256 × 256 pixels) from the STOIC database (Roy, Roy, Éthier-Majcher, Fortin, Belin, & Gosselin, submitted) served as base stimuli. The photographs depicted three male and three female faces, each with a disgusted and a fearful expression (Figure 2). The faces were cropped to exclude non-facial cues, and they were equated in mean luminance and contrast (root mean square (RMS) contrast of 0.2) using the SHINE toolbox (Willenbockel et al., 2010b).

The SFs of the base images (see Supplementary Figure 1 for a plot of the spectral content of the base faces) were randomly sampled trial-by-trial using the SF Bubbles technique (Willenbockel et al., 2010a). In brief, the to-be-filtered base image was padded with a uniform gray background and then subjected to a fast Fourier transform. The amplitude spectrum of the padded image was multiplied element-wise with a filter constructed in the following way: A vector consisting of randomly distributed binary elements (45 ones among 10195 zeros) was convolved with a Gaussian kernel, referred to as an "SF bubble" ($\sigma = 1.8$). This yielded a smoothed sampling vector. The sampling vector was subjected to a logarithmic transformation to take into account the fact that the human visual system is more sensitive to low than to high SFs (e.g., De Valois & De Valois, 1990). To obtain a two-dimensional filter, the log-transformed, smoothed sampling vector was then "rotated" about its origin. After multiplying the two-dimensional filter element-wise with the amplitude spectrum of the base image, the result was back-transformed into the image domain via an inverse fast Fourier transform. The "SF bubblized" image contained a random subset of the base image's SF content (see Figure 3 for sample stimuli; for an illustration of the filtering procedure, see Figure 1 in Willenbockel et al., 2010a).

The contrast level of the SF sampled stimuli was kept constant across experimental conditions but was adjusted for each participant so he/she reported being able to recognize the facial expressions in the *visible face* condition (see the Procedure section) but did not detect the faces in the *invisible face* condition. For Participants 1 and 2, this resulted in a mean RMS

contrast of 0.019 and for Participant 3 of 0.024. To be able to display stimuli with low contrast, we used Floyd-Steinberg dithering (Floyd & Steinberg, 1976), which enhances the luminance resolution (see also Allard & Faubert, 2008). The face stimuli subtended visual angles of approximately 7.1° horizontally and 12.8° vertically.

The high-contrast noise used for CFS (Tsuchiya & Koch, 2005) consisted of random elliptical Mondrian patterns (Figure 4; see also, e.g., Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009). The mean RMS contrast of the Mondrians was 0.80 (*SD* = 0.11). The noise fields were of 256 × 256 pixels and subtended horizontal and vertical visual angles of approximately 10.6° and 13.7°, respectively (see Supplementary Figure 1 for a plot of the spectral content of the Mondrians).

Target stimuli consisted of face/Mondrian composites (Figure 4). A composite was constructed by multiplying the pixel values of an SF sampled face image (RMS contrast = 0.04) element-wise with those of a Mondrian noise field and then adjusting the contrast so it matched the Mondrians. For each target trial, five Mondrian/face composites were constructed using the same face image but different Mondrian patterns.

Procedure

The participants took part in the experiment while sitting comfortably in their dimly lit hospital room. All stimuli appeared at the center of the computer screen and were viewed from a distance of 56 cm through red-green anaglyph glasses. The glasses allowed us to simultaneously present distinct information to each eye of the participant (i.e., one eye with information in red and the other with information in green). Each trial began with a fixation cross presented for 500–900 ms (to both eyes), followed by a blank screen for 500–900 ms. Then, a face stimulus was displayed for 500 ms in one of three conditions: the *invisible face* condition, the *visible face* condition, or the *target* condition (Figure 4).

In the invisible face condition, we employed CFS to suppress the face stimulus from awareness. The static SF sampled face image was presented to one eye (by showing it in the red layer of the RGB image) while the other eye was presented with suppression noise (i.e., Mondrians were presented in the green layer). The Mondrians changed at a rate of 10 Hz (see also Tsuchiya & Koch, 2005). As a result, only the dynamic Mondrians were consciously perceived. In the visible face condition, both eyes were presented with the same SF filtered

face by displaying it in both the red and green layers. On target trials, a face was shown to one eye (i.e., in the red layer) while a Mondrian/face composite was shown to the other eye (i.e., in the green layer) at 10 Hz.

The participants were instructed to look at all images carefully and to press the space bar on a regular computer keyboard if they perceived Mondrian patterns and a face together on a given trial. The detection task allowed us to see if participants were paying attention to the stimuli and to evaluate for each CFS trial whether the faces were successfully suppressed. The interstimulus interval was adjusted for each participant to ensure that he/she had enough time for the keypress (see also Jiang & He, 2006).

One experimental session typically consisted of five 105-trial blocks (plus one practice block in the first session), with breaks in between. After each session, the red and green lenses were swapped. 45.7% of the trials were invisible face trials, 45.7% were visible face trials, and 8.6% were target trials. The different trial types were randomly intermixed within each block. We recorded four sessions per participant (with a maximum of two sessions per day, depending on the patient's willingness for research participation and on clinical constraints). In total, Participant 1 completed 20 blocks, Participant 2, 19 blocks, and Participant 3, 21 blocks.

Analysis

The intracranial EEG data from all contacts of interest for each participant were segmented from 200 ms before stimulus onset until 1500 ms after stimulus onset and baseline corrected using Brain Vision Analyzer 2.0.1 (Brain Products GmbH, Munich, Germany). The following analyses were carried out with custom MATLAB programs. Target trials and all other trials on which a keypress was made were excluded from the SF analysis. Table I provides the exact number of trials included in the analyses for each participant.

To trace which SFs modulate the EEG amplitudes recorded from the insula or amygdala over time, we ran multiple linear regressions on the SF filters from each trial and the transformed EEG amplitudes within time bins of 20 ms (separately for each participant, brain region, condition, and session). EEG amplitudes within a given time bin were transformed as follows: First, we averaged the recorded EEG amplitudes within the time bin and across the two contacts of interest from each electrode. Then we performed a median split across trials:

we set the amplitude from a given trial to 1 if it was greater than or equal to the median of all trials or to -1 if it fell below the median. This way, the impact of any abnormal amplitudes (e.g., due to epileptic spikes) was minimized without having to rely on a subjective trial rejection criterion. We then summed the filters from all trials weighted by the transformed amplitudes, which, here, is equivalent to a multiple linear regression. This was done separately for each of the 85 bins between 200 ms before stimulus onset and 1500 ms after stimulus onset.

The vectors of regression coefficients obtained for each time bin were stored in a time segment × SF sampling points matrix and smoothed using Gaussian kernels with a standard deviation of 4.0 time bins and 300 sampling points. The result was transformed into Z-scores—henceforth called classification images (CIs). We focus here on the overall CIs for each brain structure (insula and amygdala) and for each awareness condition to maximize the signal-to-noise ratio. Separately for the visible and invisible face conditions, we summed the CIs across sessions and divided the result by the square root of the number of sessions (i.e., √4). We then summed the resulting CIs across emotional expressions and divided by the square root of the number of expressions (i.e., √2). In addition, to compute the insula CIs, we summed the respective CIs across participants and divided by the square root of the number of participants (i.e., √3). Statistical significance was evaluated using the Pixel test from the Stat4Ci toolbox ($p < .05$, $S_r = 870400$, $FWHM = 99.91$, $Z_{crit} = \pm 3.78$; Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005).

Results

Behavioral results

The detection task served two purposes: (a) to ensure that participants stayed alert during the experiment, and (b) to check on each CFS trial whether the face broke through the suppression noise. The percentage of correctly detected targets for the three participants was very high ($M = 97.07\%$, $SD = 1.13\%$), suggesting that the participants paid attention to the stimuli. The percentage of detected non-targets was small ($M = 0.30\%$, $SD = 0.43\%$), which confirmed that the faces were successfully suppressed from awareness in the invisible face condition.

Spatial frequency results

Figure 5 depicts the significant pixels (regardless of polarity) for each SF and time bin, up to 1.5 seconds after stimulus onset for the overall insula and amygdala CIs (see Supplementary Figure 2 for the raw, non-thresholded, CIs). The purple pixels correspond to the visible face condition, the green pixels to the invisible face condition, and the black pixels indicate overlaps between the conditions. We will focus on the SFs that reached significance during stimulus presentation (0–500 ms). Note, however, that for both regions and visibility conditions, we found multiple other low-, mid-, and high-SF clusters to be significant after the offset of the stimulus.

Figure 5a shows the results for the insula (Participants 1–3). In the visible face condition, SFs around 8.75 cycles per face width (cpf) reached significance at approximately 340 ms after stimulus onset. In the invisible face condition, SFs around 9.40 cpf became significant at approximately 140 ms, followed by very low SFs around 2.27 cpf. The latter attained significance at approximately 200 ms and again at 420 ms. The significant pixels of the two visibility conditions overlap for SFs around 9.04 cpf between 340 and 400 ms.

Figure 5b displays the results for the amygdala (Participant 1). In the visible face condition, SFs around 6.48 cpf attained significance at approximately 240 ms. In the invisible face condition, SFs around 5.51 cpf became significant at about 140 ms. Then, at approximately 260 ms, very low SFs (1.95 cpf) reached significance.

The line graphs (summation plots) on top of the CIs depict the number of significant pixels for each SF, collapsed across time. For both the insula and amygdala, they show quite clearly that processing in the invisible face condition relied on low SFs more than processing in the visible face condition. Likewise, the graphs on the right of the CIs show the number of significant pixels for each time bin, collapsed across SFs. For both regions, they indicate that significant correlations between SFs and brain signals occurred earlier for the invisible than for the visible face condition.

Discussion

The aim of the present study was to shed light on the "informational correlates" of consciousness in the context of emotional facial expression perception. Specifically, we examined which SFs in consciously and non-consciously perceived stimuli are correlated with

brain signals in two key structures of the emotion-processing network—the insula and the amygdala. We employed a novel combination of three techniques: intracranial recordings in awake human participants, SF Bubbles (Willenbockel et al., 2010a), and CFS (Tsuchiya & Koch, 2005). To our knowledge, this is the first study that mapped the time course of SF tuning for specific regions of the emotion-processing network and as a function of awareness. In the following, we will put our findings into context by focusing, in turn, on a) emotional expression perception and awareness, b) awareness and SF processing (during face perception in general), and c) SF processing and emotional expression perception. We will then briefly discuss our findings in light of theories on the neural pathways involved in emotion perception.

Disgusted and fearful faces were used as stimuli because previous work has shown that the insula and amygdala are implicated in the processing of these facial expressions. In particular, numerous studies led to the conclusion that the anterior insula is important for the processing of disgust (e.g., Phillips et al., 1997, 1998; Krolak-Salmon et al., 2003), and the amygdala for the processing of fear (e.g., Adolphs et al., 1994, 1995; Krolak-Salmon et al., 2004; Morris et al., 1996; Phillips et al., 1998); taken together, other studies indicated that these brain regions respond to both disgusted and fearful faces (e.g., Fitzgerald et al., 2006; Morris et al., 1998; Van der Gaag et al., 2007; Winston et al., 2003; see also Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003). Our results replicate these findings. However, the emotion-specificity of the responses in these regions goes beyond the scope of this article.

Previous work has also shown that both disgusted and fearful expressions can be perceived independently of awareness (e.g., Smith, in press). Using various methods to render stimuli invisible, neuroimaging studies demonstrated that the amygdala is involved in the non-conscious processing of emotional faces (e.g., Jiang & He, 2006; Morris, Öhman, & Dolan, 1999; Pasley, Mayes, & Schultz, 2004; Whalen et al., 1998; Williams et al., 2004, 2006; but see Phillips et al., 2004). Scarce studies found support for an involvement of the insula in the non-conscious processing of emotional stimuli (e.g., Sabatini et al., 2009; but see Anderson et al., 2003, and Phillips et al., 2004, for results that speak against automatic facial expression processing in the insula). The present results indicate that both structures play a role in perceiving emotional faces, consciously and non-consciously.

In our visible face condition, the first significant correlations between stimulus information and brain signals occurred at approximately 340 ms and 240 ms after stimulus onset in the insula and amygdala, respectively. In our invisible face condition, they were present as early as 140 ms in both regions. Moreover, in both visibility conditions, we found significant correlations at long latencies, up to 1500 ms after stimulus onset. These temporal dynamics appear largely consistent with the results from previous intracranial ERP studies on conscious emotional facial expression perception, although a direct comparison is difficult due to important methodological differences. In line with our finding that the response of the insula occurred later than that of the amygdala, previous results revealed emotional effects as early as 300 ms post stimulus onset in the insula (Krolak-Salmon et al., 2003), and as early as 140 ms (Pourtois et al., 2010) or 200 ms (Krolak-Salmon et al., 2004) in the amygdala.

Furthermore, long-latency effects were present in previous intracranial ERP data as well. For instance, Krolak-Salmon and colleagues (2004) observed differential responses to fear vs. neutral or happy faces until 1100 ms after stimulus onset in the amygdala of one patient. Pourtois and colleagues (2010) found late emotional effects in the amygdala that were modulated by attention, starting at approximately 700 ms after stimulus onset and lasting more than 300 ms. Finally, such late effects were seen in response to invisible emotional words in the amygdala (after 800 ms after stimulus onset; Naccache et al., 2005), suggesting that considerable time is needed for extracting emotional meaning. Naccache and colleagues (2005) speculated that top-down influences might amplify non-conscious amygdala activation in this context, without making information accessible to conscious report. Possibly, the late significant correlations with low-level information that we found also reflect feedback or top-down influences that amplify certain aspects of the stimuli later on.

Our CIs show complex patterns of SF tuning over time, for both the insula and the amygdala. A comparison of the CIs between awareness conditions revealed that invisible face processing relied on very low SFs (< 3 cpf) more than visible face processing, especially within the first 600 ms after stimulus onset. The idea that SF processing and awareness interact during face perception has come up repeatedly in the literature but has, as far as we know, only been investigated in one published study (De Gardelle & Kouider, 2010). The authors employed a masked priming paradigm with hybrid prime stimuli—composed of the low SFs of one face and the high SFs of another face—and a fame judgment task. Using

behavioral measures, they discovered that both low SFs (< 12 cpf) and high SFs (> 12 cpf) could be processed without awareness. The influence of high SFs correlated with prime visibility (i.e., prime duration), whereas the influence of low SFs did not. De Gardelle and Kouider's results are consistent with ours inasmuch as we also found a broad range of SFs to be processed non-consciously. The qualitative differences that we observed between our awareness conditions, however, were not seen in their data. This discrepancy could be due to several methodological differences between the studies.

Whereas not much work has been done on SF processing and awareness, several studies have looked at SF processing during the conscious perception of fearful faces. The majority of studies imply that low SFs are particularly important for the perception of fear (Pourtois et al., 2005; Vlamings et al., 2009; Vuilleumier et al., 2003; Winston et al., 2003; but see Holmes et al., 2005; Morawetz et al., 2011). For instance, in an fMRI study, Vuilleumier and colleagues (2003) observed larger amygdala responses to fearful than to neutral faces when stimuli were unfiltered or low-pass filtered (< 6 cpf), but not when they were high-pass filtered (> 24 cpf). In a recent surface ERP study (Vlamings et al., 2009), it was found that fearful relative to neutral faces elicited a larger P1 component (i.e., a positive deflection around 100 ms post stimulus onset) and a larger N170 (i.e., a negative deflection around 170 ms), also only for low-pass (≤ 12 cpf), not for high-pass (≥ 36 cpf), filtered faces. These findings are in line with our amygdala CI: many pixels reached significance for SFs under 6 cpf but very few attained significance for SFs above 24 cpf (see the summation plot in Figure 5b (top)). However, as discussed above, we did not find any significant SFs for latencies below 200 ms in our visible face condition, suggesting that the early emotional effects observed using surface ERPs (Vlamings et al., 2009; see also Pourtois et al., 2005) are probably not driven by the amygdala or insula.

The SF tuning patterns we found raise the question about the underlying neural mechanisms of SF processing as a function of awareness. Specifically, through which pathways does SF information arrive at the insula and amygdala? Currently two theories are discussed in the emotion-processing literature, namely the subcortical pathway hypothesis (for recent reviews see Tamietto & De Gelder, 2010; De Gelder, Van Honk, & Tamietto, 2011) and the multiple waves model (Pessoa & Adolphs, 2010, 2011). According to the former, low-SF information from emotional stimuli is conveyed quickly and automatically via

a subcortical route through the superior colliculus and the pulvinar nucleus of the thalamus to the amygdala, whereas high SFs are processed more slowly along a cortical route. The multiple waves model, in contrast, suggests that emotional information is processed in parallel by multiple cortical pathways, without reliance on a direct subcortical route to the amygdala. Our study was not designed to test these theories; however, our results appear to be consistent with the multiple waves model, while they challenge the subcortical pathway hypothesis in at least two ways. The first hurdle for the subcortical pathway hypothesis is that the early low-SF clusters revealed to be significant in the invisible face condition are *not* present in the visible face condition. The second hurdle is that the latencies we found in both awareness conditions (140 ms in the invisible face condition, and 340 ms or 240 ms in the visible face condition for the insula and amygdala, respectively) do not appear faster than cortical visual processing (see Pessoa & Adolphs, 2010, for a review). More work will be needed to directly test these two theories.

One limitation of the current study is that since we recorded brain signals from patients with epilepsy, we cannot be entirely sure that our data are representative of the healthy population. For Participant 1, epileptic spikes were found in the insula; we therefore recomputed our CIs without her data for the insular contacts. However, we did not find any changes in the main results (see Supplementary Figure 3). The structures we recorded from in all participants were structurally normal-appearing on high-resolution MRI. Thus, we think it is reasonable to assume that the results we report here can be generalized. Intracranial recordings from volunteers with epilepsy have previously been used in several studies (e.g., Naccache et al., 2005; Pourtois et al., 2010; Krolak-Salmon et al., 2003, 2004; Oya, Kawasaki, Howard, & Adolphs, 2002; Tsuchiya et al., 2008) because they bear a number of advantages—specifically, a millisecond temporal resolution combined with a high spatial resolution—and are thus considered to provide an important window into the workings of the human brain.

A second drawback is that in creating the two awareness conditions, we introduced differences in physical stimulation. In the visible face condition, a static face was presented to both eyes, whereas in the invisible face condition, dynamic high-contrast noise replaced the face presented to one eye. This has the disadvantage that we do not know to what extent and how the flashing of the noise patterns influenced our SF results (see Yang & Blake, 2012). We chose suppression noise that was used in several previous studies and found to be very

effective (i.e., high-contrast Mondrian patterns; e.g., Jiang & He, 2006; Jiang et al., 2009; Tsuchiya & Koch, 2005). The spectral energy of our Mondrians was highly correlated with that of our base faces (see Supplementary Figure 1; the correlation between the average across faces and the average across noise patterns was $r = .95$). Our Mondrians consisted of elliptical elements (see also Tsuchiya et al., 2009; Adams, Gray, Garner, & Graf, 2010) and thus contained energy at all orientations. It is not yet known what the optimal suppression noise would be, and basically any method used to render visual stimuli invisible for normal-sighted observers introduce differences in stimulation. Therefore, this problem is difficult to overcome (see, e.g., the review by Tamietto & De Gelder, 2010). We used CFS because it results in longer suppression than other techniques, such as backward masking or binocular rivalry. Also, it has the advantage that the suppression can be precisely timed.

Investigating the informational correlates of consciousness from several angles—i.e., with different awareness-manipulating techniques and paradigms (e.g., Faivre, Berthet, & Kouider, 2012; for a review see Kim & Blake, 2005)—might currently be the best approach to overcome the limitations of the present study. For example, one could combine the SF Bubbles technique with masked priming to examine which SFs of visible vs. invisible primes influence observers' responses to a visible target. It might also be a good idea to use SF Bubbles together with a crowding paradigm, which has recently been emphasized as a more ecologically valid approach than masking or CFS (Faivre et al., 2012). Furthermore, it might be insightful to combine SF Bubbles with an attentional blink (e.g., Raymond, Shapiro, & Arnell, 1992) paradigm, where physical stimulation remains constant but stimuli can be rendered invisible by diverting the observers' attention. This could represent a promising avenue for contrasting SF tuning between conscious and *preconscious* processing (see Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). The present study is meant as a first step towards gathering converging evidence about the informational correlates of consciousness.

Conclusion

Using state-of-the art techniques, we mapped the SF tuning of the insula and amygdala as a function of awareness. Our results are consistent with the idea that a wide range of SFs plays a role in the conscious and non-conscious perception of emotional facial expressions,

with SFs between 6 and 10 cpf appearing particularly important early on (for faces subtending approximately 7°). That being said, qualitative differences in SF tuning were observed between our awareness conditions—particularly in the early processing of very low SFs—that are consistent with the idea that different neural pathways are employed for conveying visual information to the amygdala and insula under different degrees of awareness. The present study paves the way for future work that investigates the temporal dynamics of SF processing in specific structures of the emotion-processing network and for elucidating the informational correlates of consciousness in general.

Acknowledgments

We thank the technicians at the CERNEC and the Hôpital Notre-Dame for help with data collection and the participants for their time and effort. We also thank the two reviewers for their helpful comments. This research was supported by a graduate student scholarship from the Fonds Québécois de la Recherche sur la Nature et les Technologies (FQRNT) to Verena Willenbockel, as well as by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grants awarded to Frédéric Gosselin and to Franco Lepore.

References

Adams, W. J., Gray, K. L. H., Garner, M., & Graf, E. W. (2010). High-level face adaptation without awareness. *Psychological Science, 21*, 205–210.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433,* 68–72.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372,* 669–672.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *Journal of Neuroscience, 15*, 5879–5891.

Allard, R., & Faubert, J. (2008). The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution. *Behavior Research Methods, 40,* 735–743.

Anderson, A. K., Christoff, K., Panitz, D., De Rosa, E., & Gabrieli, J. D. E. (2003). Neural correlates of the automatic processing of threat facial signals. *Journal of Neuroscience, 23,* 5627–5633.

Bouthillier, A., Surbeck, W., Weil, A. G., Tayah, T., & Nguyen, D. K. (2012). The hybrid operculo-insular electrode: A new electrode for intracranial investigation of peri-sylvian/insular refractory epilepsy. *Neurosurgery, 70,* 1574–1580.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 5,* 659–667.

De Gardelle, V., & Kouider, S. (2010). How spatial frequencies and visual awareness interact during face processing. *Psychological Science, 21*, 58–66.

De Gelder, B., Van Honk, J., & Tamietto, M. (2011). Emotion in the brain: of low roads, high roads and roads less travelled. *Nature Reviews Neuroscience, 12,* 425.

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*, 204–211.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision.* New York, NY: Oxford University Press.

Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: a comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Consciousness Research, 3*, 129.

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat: Amygdala reactivity across multiple expressions of facial affect. *NeuroImage, 30*, 1441–1448.

Floyd, R. W., & Steinberg, L. (1976). An adaptive algorithm for spatial greyscale. *Proceedings of the Society for Information Display, 17*, 75–77.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Holmes, A., Winston, J. S., & Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Cognitive Brain Research, 25,* 508–520.

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-information processing. *Current Biology, 16,* 2023–2029.

Jiang, Y., 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,* 1171–1177.

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*, 381–388.

Krolak-Salmon, P., Hénaff, M.-A., Isnard, J., Tallon-Baudry, C., Guénot, M., Vighetto, A., Bertrand, O., & Mauguière, F. (2003). An attention modulated response to disgust in human ventral anterior insula. *Annals of Neurology, 53,* 446–453.

Krolak-Salmon, P., Hénaff, M.-A., Vighetto, A., Bertrand, O., & Mauguière, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: A depth electrode ERP study in human. *Neuron, 42,* 665–676.

Morawetz, C., Baudewig, J., Treue, S., & Dechent, P. (2011). Effects of spatial frequency and location of fearful faces on human amygdala activity. *Brain Research, 1371*, 87–99.

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain, 121*, 47–57.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383*, 812–815.

Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences, 96*, 1680–1685.

Naccache, L., Gaillard, R., Adam, C., Hasboun, D., Clémenceau, S., Baulac, M., Dehaene, S., & Cohen, L. (2005). A direct intracranial record of emotions evoked by subliminal words. *Proceedings of the National Academy of Sciences, 102,* 7713–7717.

Oya, H., Kawasaki, H., Howard III, M. A., & Adolphs, R. (2002). Electrophysiological responses in the human amygdala discriminate emotion categories of complex visual stimuli. *Journal of Neuroscience, 22,* 9502–9512.

Pasley, B. N., Mayes, L. C., & Schultz, R. T. (2004). Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron, 42*, 163–172.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

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,* 773–782.

Pessoa, L., & Adolphs, R. (2011). Emotion and the brain: multiple roads are better than one. *Nature Reviews Neuroscience, 12*, 425–426.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389,* 495–498.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., Williams, S. C. R., Bullmore, E. T., Brammer, M., & Gray, J. A. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London B, Biological Sciences, 265,* 1809–1817.

Phillips, M. L., Williams, L. M., Heining, M., Herba, C. M., Russell, T., Andrew, C., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Morgan, M., Young, A. W., & Gray, J. A. (2004). Differential neural responses to overt and covert presentations of facial expressions of fear and disgust. *NeuroImage, 21,* 1484–1496.

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping, 26,* 65–79.

Pourtois, G., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010). Temporal precedence of emotion over attention modulations in the lateral amygdala: Intracranial ERP evidence

from a patient with temporal lobe epilepsy. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 83–93.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*, 849–860.

Sabatini, E., Penna, S. D., Franciotti, R., Ferretti, A., Zoccolotti, P., Rossini, P. M., Romani, G. L., & Gainotti, G. (2009). Brain structures activated by overt and covert emotional visual stimuli. *Brain Research Bulletin, 79,* 258–264.

Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *NeuroReport, 14,* 1665–1669.

Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Current Biology, 17,* 1580–1585.

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: behavioral and brain evidence. *PLoS ONE, 4,* e5625.

Smith, M. L. (in press). Rapid processing of emotional expressions without conscious awareness. *Cerebral Cortex.*

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16,* 184–189.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *NeuroImage, 37*, 974–984.

Smith, M. L., Gosselin, F., & Schyns, P. G. (2004). Receptive fields for flexible face categorizations. *Psychological Science, 15,* 753–761.

Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences, 10*, 538–545.

Surbeck, W., Bouthillier, A., Weil, A. G., Crevier, L., Carmant, L., Lortie, A., Major, P., & Nguyen, D. K. (2011). The combination of subdural and depth electrodes for intracranial EEG investigation of suspected insular (perisylvian) epilepsy. *Epilepsia, 52,* 458–466.

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience, 11,* 697–709.

Thurman, S. M., & Grossman, E. D. (2011). Diagnostic spatial frequencies and human efficiency for discriminating actions. *Attention, Perception, & Psychophysics, 73,* 572–580.

Tsuchiya, N., Kawasaki, H., Oya, H., Howard III, M. A., & Adolphs, R. (2008). Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PLoS ONE, 3,* e3892.

Tsuchiya, N., & Koch, C. (2005). Continuous Flash Suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096–1101.

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,* 1224–1225.

Van der Gaag, C., Minderaa, R. B., & Keysers, C. (2007). The BOLD signal in the amygdala does not differentiate between dynamic facial expressions. *Social Cognitive and Affective Neuroscience, 2*, 93–103.

Vlamings, P. H. J. M., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *Journal of Vision, 9,* 1–13.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6,* 624–631.

Wang, Y., Agarwal, R., Nguyen, D., Domocos, V., & Gotman, J. (2005). Intracranial electrode visualization in invasive pre-surgical evaluation for epilepsy. *Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, September 1–4, 2005.*

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,* 411–418.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010a). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010b). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*, 671–684.

Williams, L. M., Das, P., Liddell, B. J., Kemp, A. H., Rennie, C. J., & Gordon, E. (2006). Mode of functional connectivity in amygdala pathways dissociates level of awareness for signals of fear. *Journal of Neuroscience, 26,* 9264–9271.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *Journal of Neuroscience, 24,* 2898–2904.

Winston, J. S., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage, 20*, 84–97.

Winston, J. S., Vuilleumier, P., & Dolan, R. J. (2003). Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. *Current Biology, 13,* 1824–1829.

Yang, E., & Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of Vision, 12,* 1–14.

Figure captions

Figure 1. Panels show the locations of the electrode contacts of interest for each participant based on post-implantation MRIs (A = anterior; P = posterior; L = left; R = right).

Figure 2. Base face images with disgusted and fearful expressions used in the experiment.

Figure 3. Example of a face image filtered with the SF Bubbles technique on six hypothetical trials.

Figure 4. Illustration of the paradigm. In the visible face condition, a stationary SF filtered face image was shown to both eyes simultaneously (i.e., in the red and green layers). In the invisible face condition, an SF filtered face was shown to one eye (i.e., in the red layer) while dynamic noise patterns were presented at 10 Hz to the other eye (i.e., in the green layer). Target stimuli consisted of a stationary SF filtered stimulus presented to one eye and a combination of a face and noise patterns to the other eye. As a result, participants typically perceived the face image on visible face trials, only the dynamic noise on invisible face trials, and both face and noise on target trials. (Note that the contrast and brightness of the images was slightly modified in the figure to improve readability.)

Figure 5. Classification images for a) the insula and b) the amygdala. The classification images show the significant pixels for the invisible face condition (green) and the visible face condition (purple) for each spatial frequency (in cpf = cycles per face) and time segment between 200 ms before stimulus onset and 1500 ms after stimulus onset. Black regions indicate the overlap between the awareness conditions. The line graphs (summation plots) show the number of significant pixels across time for each spatial frequency (top) or across spatial frequencies for each time segment (right).

Tables

Table I: Participant information

ID	Gender	Age (years)	Handedness	Seizure focus	Hemisphere recorded from	Number of trials
1	female	39	ambidextrous	Frontal operculum Temporal operculum Insula Superior temporal gyrus	Left	1920
2	male	34	ambidextrous	Hippocampus	Left	1823
3	male	35	right	Inferior frontal gyrus	Right	2007

Figures

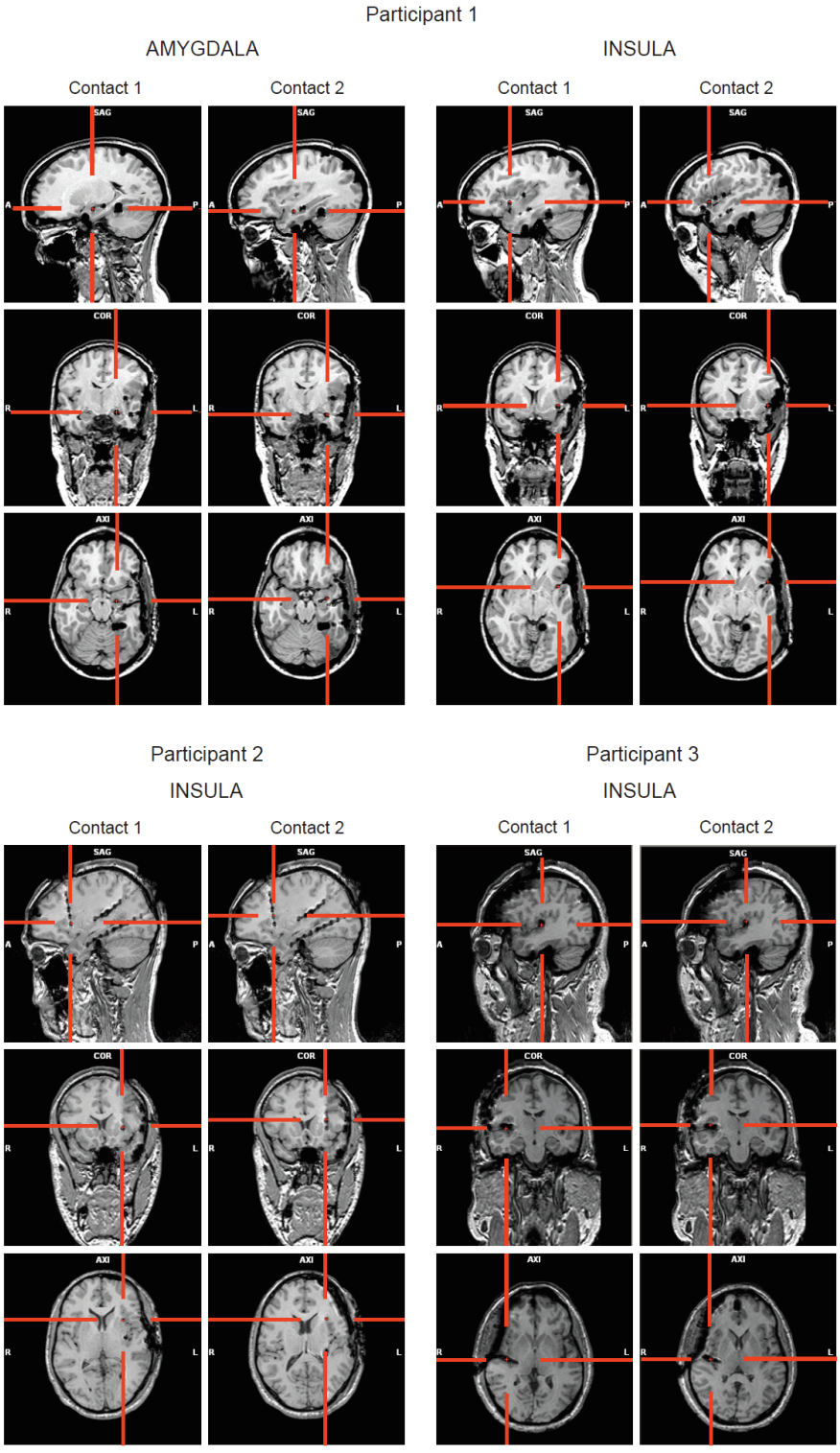


Figure 1: Electrode locations

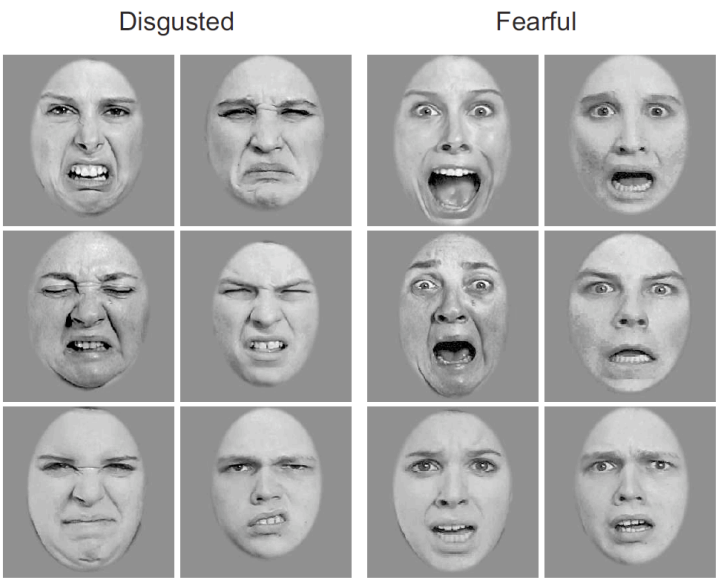


Figure 2: Base face images

Figure 3: Sample face image filtered with SF Bubbles

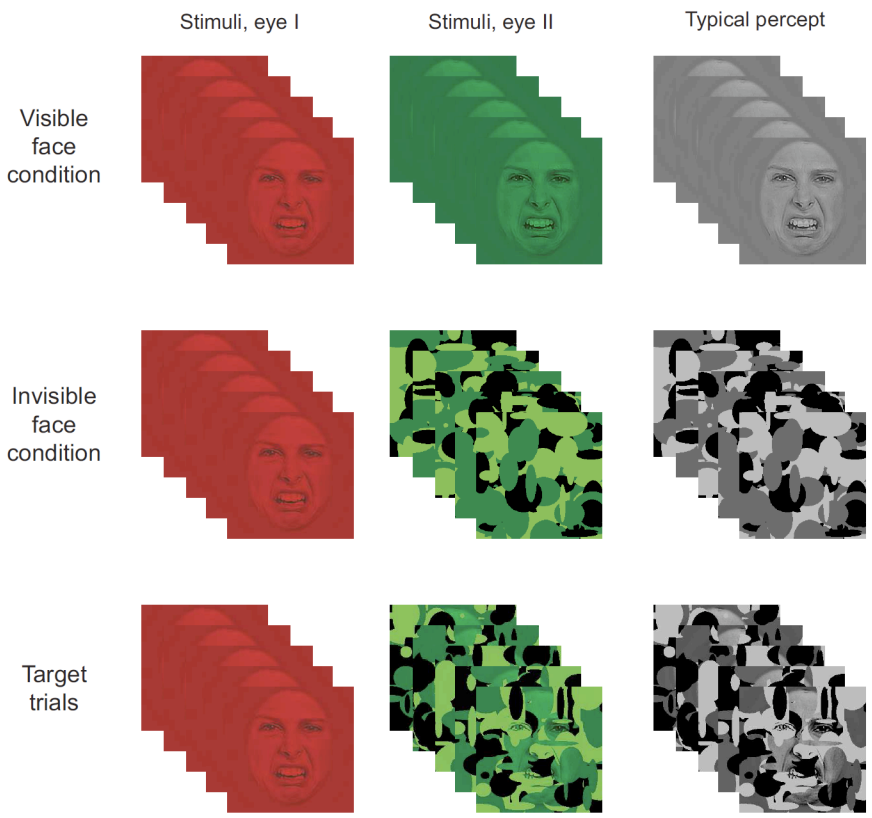


Figure 4: Illustration of the paradigm

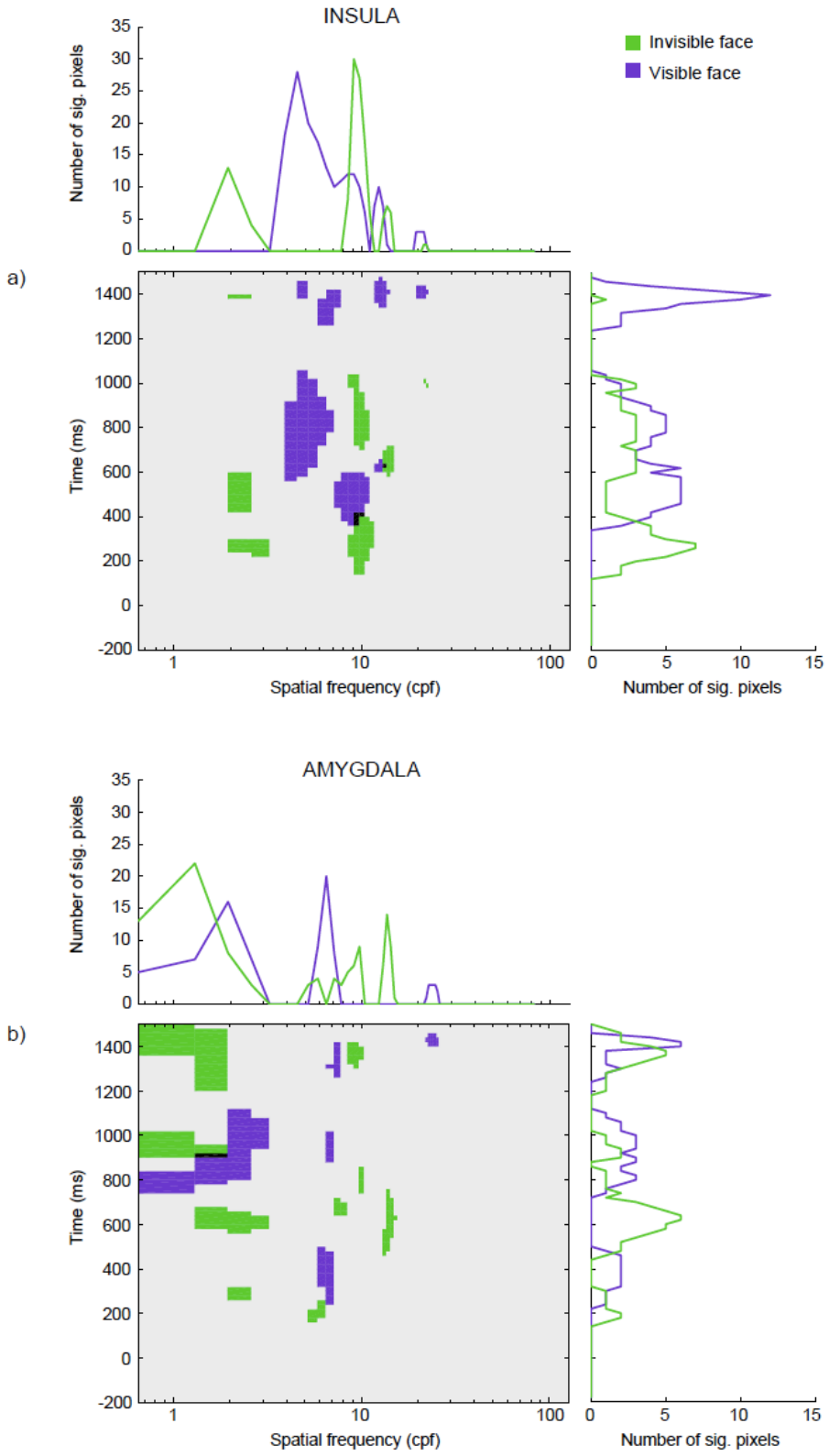


Figure 5: Classification images for the insula and amygdala

Chapter 6: General discussion

A lot of research has been devoted to unraveling the brain mechanisms underlying our ability to quickly and effortlessly perceive the visual world around us. Thanks to new neuroimaging techniques, it is now possible to get high-resolution glimpses into the workings of the human brain. These insights have much advanced our knowledge of these mechanisms. However, many open questions still remain: For instance, what are the mechanisms underlying our remarkable face perception skills? What causes the FIE? How exactly do bottom-up and top-down processes interact? What information is processed consciously vs. non-consciously in the brain? And ultimately, why does it feel a certain way when we see something, like the color blue vs. the color red? These are just a few examples of "puzzle pieces" that have not yet been found or fallen into place.

The research here was motivated by the questions raised above. The studies included in this thesis focus on low-level visual information processing in the context of face perception and examine which factors modulate this processing. The research took us from investigating SF tuning for face identification as a function of orientation to shed light on the FIE (Chapter 3) to SF tuning for face-gender perception as a function of awareness (Chapter 4) on to the conscious and non-conscious perception of SF information in emotional facial expressions (Chapter 5). Taken together, the pieces that this thesis contributes to the big puzzle consist of shedding light on the mechanisms underlying face perception from an information-processing perspective and, most importantly, on what we call the *informational correlates of consciousness*—which visual information modulates observers' responses during conscious vs. non-conscious perception.

Revealing the material correlates of consciousness is regarded by many as helpful for coming up with a final theory of consciousness (e.g., Koch, 2004). Several philosophers and researchers defend the view that once the so-called easy problems (e.g., What characterizes consciousness?) are solved, the hard problem (i.e., What causes consciousness?) will also be solved. Previous work has concentrated on a different material correlate of consciousness, the NCC, by contrasting *brain activation patterns* between conscious and non-conscious conditions. Here, we contrasted *SF tuning* between conscious and non-conscious conditions to learn more about the contents of consciousness per se. Looking at the informational correlates of consciousness is a novel contribution of this thesis. In a way, the informational correlates of

consciousness and the NCC can be considered as complementary material correlates of consciousness.

In Chapters 4 and 5, we examined the informational correlates of consciousness from two different perspectives: We used different response measures (RTs and intracranial ERP amplitudes) to target different processing levels—behavioral and neural. Moreover, we employed different paradigms to, ideally, collect converging evidence with different awareness-manipulating techniques. As far as we know, our studies are the first to investigate with a high SF resolution which information in complex, socially relevant stimuli modulates observers' responses as a function of awareness. Moreover, a novel aspect of our work is that we traced the precise SF tuning over time in the insula and amygdala, two important structures of the emotion-processing network.

Before we turn to our results, however, it is important to mention that this thesis also makes methodological contributions to visual perception research that appear to have been well received in the scientific community. The SHINE toolbox and the SF Bubbles technique introduced in Chapters 2 and 3 not only have been useful for studying the informational correlates of consciousness, they have also found use in several other studies by different labs to shed light on research questions from various domains (e.g., Arnold & Lipp, 2011; Butler, Blais, Gosselin, Bub, & Fiset, 2010; Herzmann, Willenbockel, Tanaka, & Curran, 2011; Nemrodov & Itier, 2011; Tanaka, Meixner, & Kantner, 2011; Thurman & Grossman, 2011; Vogel, Monesson, & Scott, 2012; Willenbockel, Fiset, & Tanaka, 2011; Williams, Willenbockel, & Gauthier, 20009; Wyatte, Curran, & O'Reilly, 2012; Zeidman, Mullally, Schwarzkopf, & Maguire, 2012). In the following section on the methodological implications, I will discuss the two methods in turn.

Methodological contributions

The SHINE toolbox

One fundamental aspect to ensure when conducting an experiment on visual perception is that one has good stimuli. This typically entails making sure that the stimulus set does not introduce any confounding variables. For example, when one is interested in studying what distinguishes face and object perception at the neural level, one might want to ensure that the images displaying objects and faces cannot be distinguished based on the overall brightness

alone. The first methodological contribution of this thesis targets this fundamental issue. In Chapter 2 (Willenbockel et al., 2010b), we introduced the SHINE toolbox, which provides easy-to-use MATLAB functions for the control of low-level image properties including luminance, contrast, and SF. Previous research has shown that these properties can impact brain responses (e.g., certain ERP components such as the P1; see, e.g., Luck, 2005), and it is thus easy to see that controlling them is important. SHINE is a computationally straightforward but powerful and flexible tool that could be applied for image preprocessing in a variety of experiments.

The studies we conducted provide some examples: SHINE was used in Chapters 3–5 of this thesis to match low-level properties across face stimuli. In those studies, the same stimuli were used across experimental conditions (e.g., the same face images in the upright and inverted face conditions in Chapter 3 or in the visible and invisible face conditions in Chapters 4 and 5), and SHINE thus merely served to ensure that the within stimulus set variation in low-level properties was minimized. More powerful applications of the toolbox can be found in other studies that are not included in this thesis (because they would have broadened its scope even more). There, we employed SHINE functions to equate the low-level properties of images of different categories (e.g., Herzmann et al., 2011; Williams et al., 2009) or to parametrically vary the low-level content of the stimuli in different conditions (e.g., Willenbockel et al., 2011).

In the latter study, for instance, we used SHINE to study the relative contribution of lightness and facial morphology to race perception. SHINE was employed to create five different lightness levels ranging from the original lightness distribution of Caucasian faces to the lightness distribution of African-American faces. These distributions were then applied to both Caucasian and African-American faces and morphs thereof (5 morphing levels; see Figure 1 below), which were presented upright or inverted. We found that observers' race categorization responses were mainly affected by morphing levels; however, when the faces were upside-down or when morphological cues were ambiguous, the influence of lightness increased and the influence of morphological information decreased. This approach thus revealed that the default mechanism for race categorization relies mostly on morphological information but lightness also plays a role, especially when morphological features are difficult to extract.

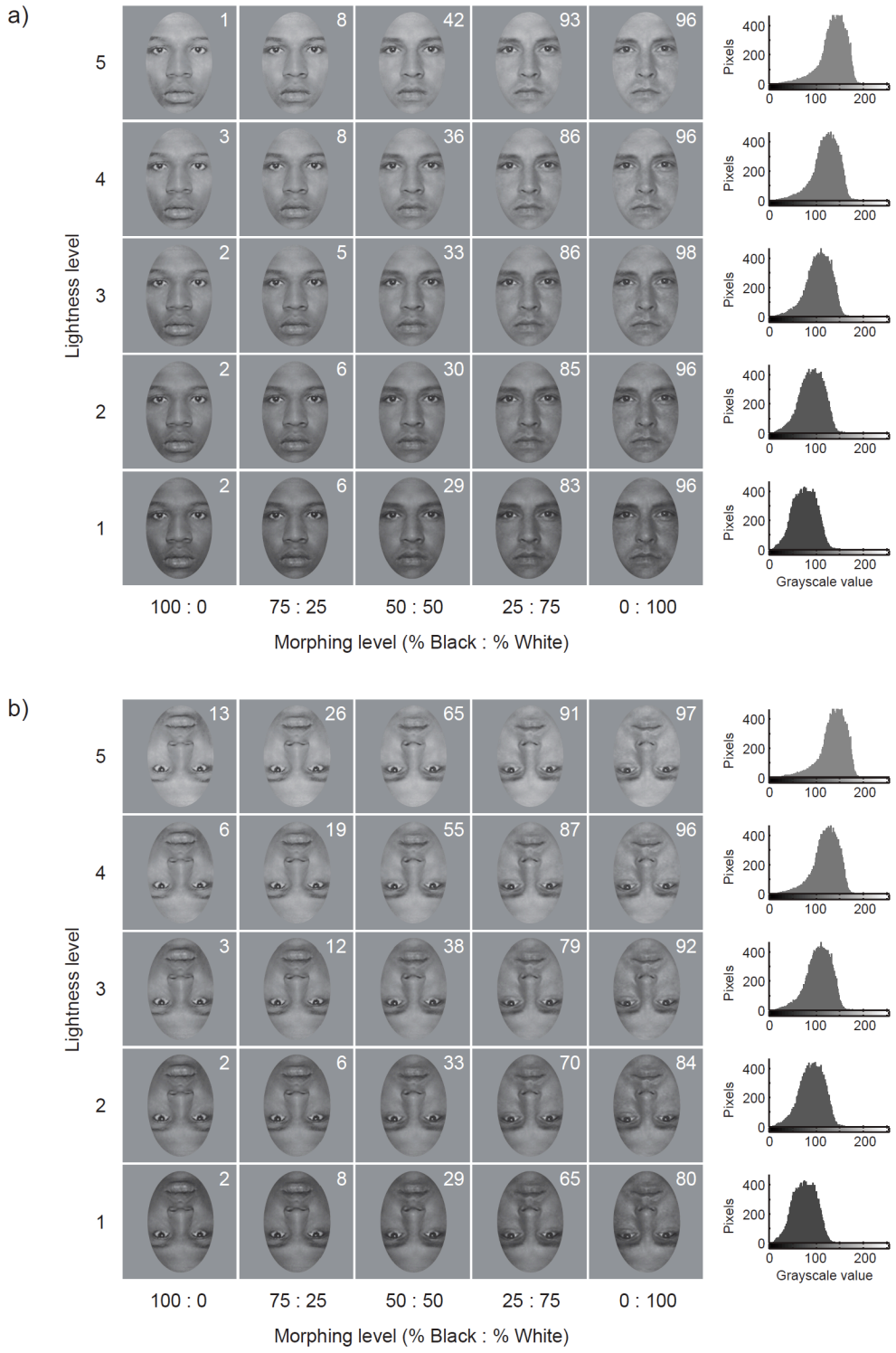


Figure 1: Illustration of the stimuli used by Willenbockel, Fiset, and Tanaka (2011). The white numbers indicate the percentage of "White" categorizations for the a) upright and b) inverted condition.

Besides in our own work, SHINE has found use in several behavioral, ERP, and fMRI studies from other labs to ensure that the high-level effects of interest are not undermined by low-level confounds (e.g., Arnold & Lipp, 2011; Butler et al., 2010; Nemrodov & Itier, 2011; Tanaka et al., 2011; Vogel et al., 2012; Wyatte et al., 2012; Zeidman et al., 2012). The number of citations (16 so far) and emails regarding SHINE I have received suggest that there is a considerable demand for such a toolbox in the field.

The SF Bubbles technique

The second methodological contribution was made in Chapter 3 (Willenbockel et al., 2010a), in which we developed and validated the SF Bubbles technique. SF Bubbles belongs to the family of classification image techniques (Eckstein & Ahumada, 2002) and can be used to derive the precise SF tuning curves for a given task. The strengths of SF Bubbles compared to more traditional SF sampling techniques (e.g., low-pass, high-pass, or band-pass filtering or critical band masking) include that it provides high-resolution SF tuning curves for a given task with a relatively small number of trials (e.g., Thurman & Grossman, 2011), minimizes the risk that observers adapt to certain SF ranges (by randomly sampling SFs on a trial-by-trial basis), and that it is unbiased in that no cutoff frequencies have to be chosen. Moreover, SF Bubbles can reveal multiple significant SF bands simultaneously. These features make SF Bubbles especially suitable for research that is of rather exploratory nature (i.e., where no good a priori justifications for certain cutoff frequencies exist), as was the case with the informational correlates of consciousness.

SF Bubbles has been applied with different response measures and in different contexts. In the studies presented here, it was employed to investigate which SFs correlate with observers' behavioral (Chapter 3 and 4) or brain (i.e., intracranial ERP; Chapter 5) responses during face perception, and was shown to be sensitive to both bottom-up and top-down induced changes in SF tuning. SF Bubbles has also been used in other recent behavioral studies to reveal the precise SF tuning curves for various tasks, such as discriminating videos of human actions (Thurman & Grossman, 2011), visual word recognition (Tadros et al., in press), and music reading (Hammal, Gosselin, Peretz, & Hébert, submitted). Thurman and Grossman directly compared the results obtained with SF Bubbles and those obtained with a more traditional SF sampling technique and emphasized the above-mentioned advantages of

SF Bubbles. Tadros and colleagues introduced a variation of SF Bubbles in which performance was adjusted by increasing or decreasing the number of Bubbles, whereas in our studies, we kept the number of Bubbles constant; both variants were found to work well.

SF Bubbles is rooted in the same idea as previous variants of the Bubbles technique developed by Gosselin and Schyns (2001), but differs from it in important ways. The underlying idea is that if diagnostic information for the task at hand is revealed to the observer on a given trial, the probability of the observer responding correctly and/or quickly (and showing the associated brain responses) should increase. The different variants of the Bubbles technique differ in the domain in which information is sampled. Whereas Gosselin and Schyns (2001) randomly varied the x,y-coordinates of Gaussian "bubble-like" windows placed on the image (in the image/spatial domain), we varied the center SFs of Gaussian filters that were applied after Fourier transforming the image (in the SF domain). The SF variant has the major advantage over previous Bubbles variants that it is more universally applicable (for an exception see the variant by McCotter, Gosselin, Sowden, & Schyns, 2005, described below). The variants by Gosselin and Schyns (2001) require carefully aligned stimuli that depict comparable features in the image domain. SF Bubbles, in contrast, can be used to compare diagnostic information between stimulus categories that look very different in the image domain (e.g., one could easily compare the optimal SFs for identifying faces and letters), because SFs are the building blocks of any image. SF Bubbles is thus complementary to previous Bubbles variants and more versatile.

As indicated, there is one previous Bubbles variant that can also be universally applied and might at first glance appear similar to SF Bubbles. It was developed by McCotter et al. (2005) to investigate which information in scenes is used for basic-level categorizations. In that variant of the technique, the Fourier coefficients of the phase spectrum are randomly sampled. Whereas sampled Fourier coefficients keep their original phase, the phase of non-sampled coefficients is substituted with that of white noise. The method differs, however, from SF Bubbles in that it selectively alters the phase in the Fourier coefficients while normalizing (whitening) the amplitude of each component frequency; SF Bubbles does not alter the phase spectrum of images but the amplitude spectrum. Therefore, although the techniques both operate in the Fourier domain, they are complementary.

In general, classification image techniques including Bubbles are powerful methods because they lead to a result in any case (assuming that the experiment is properly set up and a sufficient number of trials is completed): Even if no differences are found between conditions, such as in our experiments on upright vs. inverted face identification (Chapter 3), one still determines the effective stimulus as a product of the experiment (see Figures 6–8 in Chapter 3 for examples). Effective stimuli contain only the information that is relevant for the task. It is conceivable to use the effective stimuli in subsequent experiments to increase the signal-to-noise ratio compared to using stimuli that contain both diagnostic and non-diagnostic information. For instance, one could carry out an SF Bubbles experiment to "carve out" the information that is important for the task of interest, and then use the results to generate optimized stimuli for studies in which only a relatively small number of trials can be obtained, e.g., in neuroimaging studies with patients.

Empirical and theoretical implications

As indicated at the beginning, this thesis provides several somewhat spread-out puzzle pieces: From empirical and theoretical standpoints, it enhances our knowledge of face perception, SF processing, and visual consciousness. The focus, however, was on the latter.

SF tuning underlying conscious face perception

Faces convey various cues that are helpful for social interactions, such as a person's gender, identity, and emotional state. Humans are typically very good at extracting relevant facial information "at a glance". To shed light on observers' face perception strategies, researchers have begun to map which specific facial cues are used for different tasks. In the spatial domain, previous variants of the Bubbles technique and other classification image techniques have been very helpful in this endeavor (e.g., Butler et al., 2010; Gosselin & Schyns, 2001; Sekuler, Gaspar, Gold, & Bennett, 2004; Smith, Cottrell, Gosselin, & Schyns, 2005). In the SF domain, this was investigated using methods such as hybrid faces, low-, high-, and band-pass filtering, and critical band masking (e.g., Schyns & Oliva, 1999; Goffaux & Rossion, 2006; Näsänen, 1999; Gaspar, Sekuler, & Bennett, 2008). However, as mentioned above, these methods do not reveal which individual SFs modulate observers' responses. Here, we exploited the advantages of SF Bubbles to investigate, with a high resolution and low experimental bias, the SF tuning for different face perception tasks.

In the main experiments of Chapter 3, we examined which SFs underlie the identification of upright vs. inverted faces to see whether SF tuning differences contribute to the FIE. This was of particular interest because an influential theory posits that upright faces are processed holistically, presumably based on relatively low SFs, whereas inverted faces are processed featurally, possibly using higher SFs (see, e.g., Goffaux & Rossion, 2006; Rossion, 2008). Our results, however, revealed that the same mid-SFs underlie the fast and accurate identification of both upright and inverted faces (based on inner facial features). The SFs we identified correspond well with the results obtained in previous studies on upright face identification that used various filtering approaches (e.g., Costen, Parker, & Craw, 1994, 1996; Gaspar et al., 2008; Gold, Bennett, & Sekuler, 1999; Näsänen, 1999; Willenbockel et al., 2010a). Our findings are also in accordance with the SF results on both upright and inverted face identification reported in a paper that was submitted and published while ours was under review (Gaspar et al., 2008). The results of Gaspar and colleagues' study and ours, based on different methods, clearly show that the same SFs are used to identify inner facial features irrespective of face orientation, and thus constrain where putative qualitative differences underlying the FIE might reside.

In Chapter 4, we examined the SFs influencing face-gender categorization responses in a direct classification task and an indirect priming paradigm. We found that mid-SFs similar to the ones revealed in Chapter 3 for face identification were significant for both accurate direct and fast indirect face-gender classifications. Thus, for some aspects of face perception that rely on static facial cues, SF tuning appears to be rather robust.

Although our results highlight the importance of mid-SFs for extracting facial identity and gender information, we also found some support for flexible SF usage. In Chapter 3, we carried out two additional experiments to investigate whether bottom-up and top-down induced changes in SF tuning could be revealed using SF Bubbles. Consistent with previous work (e.g., Näsänen, 1999; see also Sowden & Schyns, 2006), we found that stimulus size affects SF tuning to some degree (it should be noted, however, that the change is relatively small so that it is still feasible to report results in cycles per object or image). We also observed that SF tuning is affected by task demands. This finding is in accordance with earlier results based on hybrid faces obtained by Schyns and Oliva (1999) and supports the flexible SF usage hypothesis.

The informational correlates of consciousness

The main goal of this thesis was to provide the first steps towards shedding light on the informational correlates of consciousness. Faces constitute a suitable stimulus class for this endeavor since numerous studies have shown that facial information can influence observers' responses in the absence of awareness (e.g., Finkbeiner & Palermo, 2009; Jiang & He, 2006; Jiang et al., 2009; Kouider et al., 2009; Williams et al., 2004). Moreover, as outlined in the previous section, we already have a benchmark of SF tuning for conscious face perception (e.g., see Morrison & Schyns, 2001, and Ruiz-Soler & Beltran, 2006, for reviews). The idea to investigate whether SF processing during face perception differs as a function of awareness is not entirely new. Recently researchers have begun to investigate this with low- vs. high-SF faces in two different priming paradigms (De Gardelle & Kouider, 2010; Khalid et al., in press). However, whereas De Gardelle and Kouider observed that both low and high SFs of unseen primes influenced observers' RTs, Khalid and colleagues found non-conscious influences from low-SF but not high-SF primes. Here, we wanted to know in more detail—i.e., with a higher SF and brain resolution—which SFs are correlated with observers' responses during conscious vs. non-conscious perception.

To this end, we investigated the informational correlates of consciousness at two levels—behavioral and neural. This was accomplished by using the SF Bubbles technique with different response measures. In Chapter 4 (Willenbockel et al., submitted), we employed a masked priming paradigm and examined which SFs in the primes influence observers' behavioral responses when judging the gender of full-spectrum target faces. Here, we found that the same mid-SFs were correlated with observers' RTs under different prime awareness conditions; however, surprisingly the significant SFs influenced RTs in opposite ways. In Chapter 5 (Willenbockel et al., 2012), we used a novel combination of SF Bubbles, CFS, and intracranial ERP recordings to investigate the SF tuning of the insula and amygdala during the perception of emotional faces with a high temporal resolution. Here, we found differences in SF tuning as a function of awareness at several SFs and time points. Overall, the intracranial ERP results suggest that non-conscious processing relies on low SFs more and is faster than conscious processing. Altogether, our two studies thus provided us with quite different puzzle pieces about the informational correlates of consciousness.

How can these results be reconciled? One possibility is that the difference simply arises from the different response measures used. It has been argued that RTs, being a composite measure of many processes, are less sensitive than local brain signals, such as intracranial ERPs (Naccache et al., 2005; Sternberg, 2001). RTs may have reflected the sum of priming effects occurring at various stages in the visual system (Dehaene et al., 2004) and may have not been sensitive to subtle SF tuning differences that occurred at only some of these stages. An alternative possibility is that other methodological differences gave rise to the different findings, including differences between tasks (face-gender vs. detection task) and the awareness-manipulating techniques that were used. We employed masking and CFS, which are widely used techniques to render visual stimuli invisible (e.g., Breitmeyer & Öğmen, 2006; Faivre, Berthet, & Kouider, 2012; Kouider & Dehaene, 2007; Tsuchiya & Koch, 2005). However, only recently have researchers begun to systematically compare the fate of stimuli rendered invisible using these methods (e.g., Faivre et al., 2012), and open questions remain regarding the underlying mechanisms. The difference between the findings of our studies suggests that the interaction between SF processing and awareness is influenced by a number of factors, and that SF tuning differences are subtle and thus best traceable with measures that have a high spatial and high temporal resolution.

Interestingly, neither the results from our behavioral study nor from our intracranial ERP study provide support for the hypothesis that low SFs can be processed quickly independently of awareness, whereas high SFs are extracted more slowly and are unavailable during non-conscious processing (see also De Gardelle & Kouider, 2010). This hypothesis was derived from theories suggesting that low SFs quickly provide us with the gist of the visual input before high SFs serve to flesh out the details (e.g., Bar, 2003; Bullier 2001a, b) together with theories proposing that non-conscious processing precedes conscious processing (e.g., Dehaene & Naccache, 2001; Lamme, 2003; Zeki, 2003). Support also came from the proposal of a direct subcortical route to the amygdala that quickly and largely automatically conveys low SFs (e.g., see Tamietto & De Gelder, 2010, for a recent review). In our classification images, however, we observed significant early low-SF clusters for the non-conscious condition only and significant high-SF clusters for both awareness conditions. It remains possible though that this hypothesis holds for other brain regions, such as the ones attributed key roles in the models by Bar and Bullier.

Our behavioral results do fit, however, with the findings obtained by De Gardelle and Kouider (2010) who observed that both SFs below 12 cpf and SFs above 12 cpf could influence observers' responses non-consciously. In accordance with this, our results from Chapter 4 revealed that SFs around 12 cpf are significant irrespective of awareness. De Gardelle and Kouider interpreted their findings in terms of the flexible SF usage hypothesis, according to which high SFs can be extracted flexibly and thus early on during non-conscious processing. Our intracranial ERP results also appear to support the flexible SF usage hypothesis over the coarse-to-fine hypothesis: Our classification images indicate that, at least for the insula and the amygdala, SF tuning does not strictly proceed from low to high over time but that SF tuning patterns are rather complex.

Our findings from both studies also support the growing body of results suggesting that qualitative differences exist between conscious and non-conscious processing (e.g., Barbot & Kouider, 2012; Eimer & Schlaghecken, 2002; Frings & Wentura, 2005; Snodgrass & Shevrin, 2006). In Chapter 4, the significant SFs influenced RTs in opposite ways as a function of awareness, which we interpreted as a qualitative difference, and in Chapter 5, we found qualitative differences in SF tuning over time between our awareness conditions. More work will be needed to shed light on the exact neural mechanisms that give rise to these differences.

Overall, our studies provided the first insights into the precise informational correlates of consciousness in the SF domain from behavioral and neural perspectives. We have to conclude, however, that more puzzle pieces are needed before we can clearly see the big picture of how stimulus information, the contents of consciousness, and specific brain activation patterns are linked.

Future directions

Each of the four articles included in this thesis led to some interesting ideas for future research, which I will briefly summarize below.

First, regarding SHINE, one might want to expand the toolbox and also run some studies examining the effects of "SHINEing" on observers' responses. SHINE in its current form has shown to be useful for controlling low-level image properties in grayscale images, which it was designed to do. However, many researchers nowadays turn to color images in order to use stimuli that look as realistic as possible. As mentioned in Chapter 2, there are

ways to apply the current, grayscale version of SHINE to images in appropriate color spaces. However, currently this requires the user to have some background knowledge of color image processing in MATLAB. Thus, to broaden the scope of the toolbox, it might be useful to implement a color SHINE version.

Besides this, one could take the SHINE project into an empirical direction. One might want to systematically investigate how different modes of equalization and the related degradations of perceived image quality affect observers' behavioral and brain responses (e.g., the P1 or N170 components). This might be useful with regard to questions like: When does it suffice to equate mean luminance and contrast and when should the exact histogram matching be used? When does it suffice to equate the rotational average of the Fourier amplitude spectra and when should the spectra be exactly equated? Although the answers to these questions will probably differ to some degree for different stimulus sets, this line of research might be helpful for establishing some guidelines about when to use which degree of equalization.

Second, with regard to our face inversion study (Chapter 3), one might want to re-examine SF tuning under more ecologically valid conditions. In our study and in previous work in which only the inner facial features were presented (e.g., Gaspar et al., 2008), no support was found for qualitative differences between the processing of upright and inverted faces. In fact, our SF tuning curves were remarkably similar for both face orientations. However, it is possible that the story does not end there. As mentioned in the discussion of Chapter 3, qualitative differences in SF use could perhaps be uncovered using stimuli that show full faces (i.e., including face contours). This idea was supported by pilot data we acquired when re-running Experiment 2c with face images from the set of Goffaux and Rossion (2006) that retained face-contour information (see Figure 2 below).

Surprisingly, however, the difference trend we observed did still not support the holistic face processing account: We observed a shift towards lower SFs for inverted compared to upright faces, whereas the holistic account of the FIE would have predicted that lower SFs are used for upright, holistically processed faces than for inverted ones. This preliminary finding sparked our interest, and we are currently running a series of follow-up experiments with a new face set to examine more closely the SF tuning for identifying faces without vs. with contour information as a function of face orientation (for preliminary results, see Fiset et al., 2010).

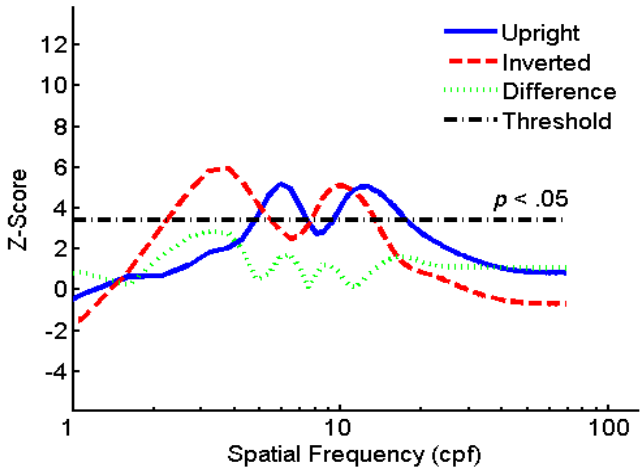


Figure 2: Group classification vectors from 14 observers obtained when re-running Experiment 2c of Chapter 3 (Willenbockel et al. 2010a) with faces retaining contour information. Significant peaks were found at 3.49 cpf and 10.22 cpf for inverted faces and at 5.91 cpf and 12.37 cpf for upright faces.

Third, for our masked priming study in Chapter 4, two avenues for future research come to mind: One consists of further investigating the reversal of priming influences we observed as a function of awareness. From previous research it appears that non-conscious priming is not always negative (i.e., inhibiting responses) but that negative priming has more often been found in non-conscious than conscious conditions (e.g., Sumner, 2007). The reasons for the reversal of priming influences are still unclear. The second avenue for future research consists of testing whether we can replicate the findings by Khalid et al. (in press) using SF Bubbles. To this end, one might want to present the face images in the periphery as Khalid and colleagues did. Perhaps this could reconcile the findings obtained by Khalid et al. (in press), De Gardelle and Kouider (2010), and our study: The behavioral studies presenting stimuli in foveal vision did not observe qualitative differences as a function of awareness, whereas the one with peripheral stimuli did. In general, it would be interesting to look at the interaction between awareness and SF processing as a function of eccentricity because it is well known that SF tuning changes as a function of eccentricity (i.e., shifts towards lower

SFs). Intuitively, it appears plausible that there are mechanisms in peripheral vision to non-consciously monitor our surroundings (e.g., to quickly detect approaching sources of danger).

Fourth, regarding our intracranial ERP experiment, a promising line of future research would be to replicate our results with more participants and other awareness-manipulating techniques, and eventually map out the SF tuning over time for several other brain regions as well. It would be nice to know how information "flows" through the emotion-processing network and to more directly test the subcortical pathway hypothesis (e.g., Johnson, 2005; Tamietto & De Gelder, 2010) vs. the multiple waves model (Pessoa & Adolphs, 2010). It would also be interesting to map SF tuning for the areas that were attributed key roles in the models on object recognition (e.g., Bar, 2003; Bullier, 2001a, b).

In general, to further our understanding of the informational correlates of consciousness, future research might want to continue along the lines of the present studies and map SF tuning as a function of awareness for different tasks using behavioral and neuroimaging measures. Combining the SF Bubbles technique with crowding to render stimuli invisible might be a way to investigate visual information use as a function of awareness in a more realistic setting. One might also want to combine it with paradigms that render stimuli invisible by diverting the observer's attention (i.e., render stimuli "preconscious"; Dehaene et al., 2006), such as the attentional blink (Raymond, Shapiro, & Arnell, 1992). Having a detailed picture of what information is processed under different awareness conditions—also in terms of other low-level attributes, such as color—could help to disentangle which neural pathways are specifically involved in conscious perception. Knowing both the informational and neuronal correlates might perhaps one day take us to a bridge across the explanatory gap if such a bridge exists.

Conclusion

In this thesis, two methods are introduced—the SHINE toolbox for image preprocessing and the SF Bubbles technique for obtaining high-resolution SF tuning curves for a given task. We used these methods to investigate whether SF tuning during face perception is affected by a number of factors—specifically, picture-plane inversion, viewing distance, task demands, and, most importantly, level of awareness. We did not find SF tuning changes as a function of face orientation but as a function of viewing distance and task demands. With

regard to different levels of awareness, we observed that the same SFs were correlated with RTs during conscious and non-conscious face-gender priming in a masking paradigm. SF tuning differences were revealed, however, when examining the conscious vs. non-conscious perception of emotional faces with the high spatial and temporal resolution of intracranial ERPs in a CFS paradigm. The studies presented in this thesis provide important insights into how the building blocks of complex visual stimuli are used by the human visual system to consciously and non-consciously perceive a class of stimuli that is highly significant for social interactions. I believe that shedding light onto the informational correlates of consciousness, in addition to shedding light onto the NCC, provides fundamental pieces towards understanding the big puzzle of consciousness.

General references

Arnold, D. H., & Lipp, O. V. (2011). Discrepant integration times for upright and inverted faces. *Perception, 40*, 989–999.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*, 600–609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617–629.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Science, 103*, 449–454.

Barbot, A., & Kouider, S. (2012). Longer is not better: nonconscious overstimulation reverses priming influences under interocular suppression. *Attention, Perception, & Psychophysics, 74,* 174–184.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9*, 46–52.

Bredfeldt, C. E., & Ringach, D. L. (2002). Dynamics of spatial frequency tuning in macaque V1. *Journal of Neuroscience, 22*, 1976–1984.

Breitmeyer, B. G., & Öğmen, H. (2006). *Visual masking: time slices through conscious and unconscious vision.* Oxford: Oxford University Press.

Briggs, F., & Usrey, W. M. (2011). Corticogeniculate feedback and visual processing in the primate. *Journal of Physiology, 589*, 33–40.

Bullier, J. (2001a). Feedback connections and conscious vision. *Trends in Cognitive Sciences, 5*, 369–370.

Bullier, J. (2001b). Integrated model of visual processing. *Brain Research Reviews, 36*, 96–107.

Butler, S., Blais, C., Gosselin, F., Bub, D., & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics, 72*, 1444–1449.

Byrne, A. (2001). Intentionalism defended. *The Philosophical Review, 110*, 199–240.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197*, 551–566.

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*, 200–219.

Chalmers, D. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.

Costen, N. P., Parker, D. M., & Craw, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception, 23*, 129–146.

Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics, 58*, 602–612.

Crick, F., & Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex, 8*, 97–107.

Crick, F., & Koch, C. (2003). A framework for consciousness. *Nature Neuroscience, 6*, 119–126.

De Gardelle, V., & Kouider, S. (2010). How spatial frequencies and visual awareness interact during face processing. *Psychological Science, 21*, 58–66.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*, 545–549.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision.* New York, NY: Oxford University Press.

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70*, 200–227.

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*, 204–211.

Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J.-B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words: behavioral and neuroimaging evidence. *Psychological Science, 15*, 307–313.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*, 1–37.

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J.-F., Poline, J.-B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, *4*, 752–758.

Dehaene, S., Naccache, L., Le Clec'H, G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., Van de Moortele, P.-F., & Le Bihan, D. (1998). Imaging unconscious semantic priming. *Nature, 395*, 597–600.

Dennett, D. (1996). *The Intentional Stance*. Cambridge, MA: MIT Press.

Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2*, i.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination*. New York, NY: Basic Books.

Eger, E., Schyns, P. G., & Kleinschmidt, A. (2004). Scale invariant adaptation in fusiform face-responsive regions. *NeuroImage, 22,* 232–242.

Eimer, M., & Schlaghecken, F. (2002). Links between conscious awareness and response inhibition: Evidence from masked priming. *Psychonomic Bulletin and Review, 9*, 514–520.

Faivre, N., Berthet, V., & Kouider, S. (2012). Nonconscious influences from emotional faces: a comparison of visual crowding, masking, and continuous flash suppression. *Frontiers in Consciousness Research, 3*, 129.

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience, 8*, 1380–1385.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1–47.

Ferrera, V. P., Nealey, T. A., & Maunsell, J. H. R. (1994). Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LGN pathways. *Journal of Neuroscience, 14*, 2080–2088.

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious processing: a comparison of face and nonface stimuli. *Psychological Science, 20*, 42–51.

Fiset, D., Willenbockel, V., Bourdon, M., Arguin, M., & Gosselin, F. (2010). The role of contour information in the spatial frequency tuning of upright and inverted face identification [Abstract]. *Journal of Vision, 10(7)*, 646a.

Frazor, R. A., Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2004). Visual cortex neurons of monkeys and cats: temporal dynamics of the spatial frequency response function. *Journal of Neurophysiology, 91*, 2607–2627.

Frings, C., & Wentura, D. (2005). Negative priming with masked distractor-only prime trials: awareness moderates negative priming. *Experimental Psychology, 52*, 131–139.

Gaspar, C., Sekuler, A. B., & Bennett, P. J. (2008). Spatial frequency tuning of upright and inverted face identification. *Vision Research, 48*, 2817–2826.

Goffaux, V., Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cerebral Cortex, 21*, 467–476.

Goffaux, V., & Rossion, B. (2006). Faces are "spatial" – holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1023–1039.

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass filtered letters and faces by human and ideal observers. *Vision Research, 39*, 3537–3560.

Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271.

Halit, H., De Haan, M., Schyns, P. G., & Johnson, M. H. (2006). Is high-spatial frequency information used in the early stages of face detection? *Brain Research, 1117*, 154–161.

Hammal, Z., Gosselin, F., Peretz, I., & Hébert, S. (submitted). Music to my eyes: Spatial frequencies mediating music reading.

Hannula, D. E., Simons, D. J., & Cohen, N. J. (2005). Imaging implicit perception: promise and pitfalls. *Nature Reviews Neuroscience, 6*, 247–255.

Haynes, J.-D., & Rees, G. (2005). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7*, 523–534.

Hegdé, J. (2008). Time course of visual perception: coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*, 405–439.

Henriksson, L., Nurminen, L., Hyvärinen, A., & Vanni, S. (2008). Spatial frequency tuning in human retinotopic visual areas. *Journal of Vision, 8*, 1–13.

Herzmann, G., Willenbockel, V., Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition for own-race and other-race faces. *Neuropsychologia, 49*, 3103–3115.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron, 36*, 791–804.

Holender, D., & Duscherer, K. (2004). Unconscious perception: The need for a paradigm shift. *Perception & Psychophysics, 66*, 872–881.

Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience, 8*, 197–230.

Iidaka, T., Yamashita, K., Kashikura, K., & Yonekura, Y. (2004). Spatial frequency of visual image modulates neural responses in the temporo-occipital lobe. An investigation with event-related fMRI. *Cognitive Brain Research, 18*, 196–204.

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: dissociating subsystems for facial-information processing. *Current Biology, 16,* 2023–2029.

Jiang, Y., 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,* 1171–1177.

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience, 6,* 766–774.

Kandel, E. R., Schwartz, J. H., & Jessell, T. M. (2000). *Principles of Neural Science*, 4th ed. New York, NY: McGraw-Hill.

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*, 4302–4311.

Khalid, S., Finkbeiner, M., König, P., & Ansorge, U. (in press). Subcortical human face processing? Evidence from masked priming. *Journal of Experimental Psychology: Human Perception and Performance.*

Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: rendering the visible 'invisible'. *Trends in Cognitive Sciences, 9*, 381–388.

Koch, C. (2004). *The quest for consciousness: a neurobiological approach*. Englewood, CO: Roberts & Company.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. *Trends in Cognitive Sciences, 11*, 16–22.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: a critical review of visual masking. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 362*, 857–875.

Kouider, S. (2009). Neurobiological theories of consciousness. In W. Banks (Ed.), *Encyclopedia of Consciousness* (pp. 87–100). Oxford: Elsevier Press.

Kouider, S., Eger, E., Dolan, R. J., & Henson, R. N. (2009). Activity in face-responsive brain regions is modulated by invisible, attended faces: evidence from masked priming. *Cerebral Cortex, 19*, 13–23.

Lamme, V. A. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7,* 12–18.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*, 354–361.

Lewis, C. I. (1929). Mind and the world order: outline of a theory of knowledge. Dover reprint, 1956.

Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT press.

Ludwig, K. (2007). The Mind-Body Problem: An overview. In S. P. Stich, & T. A. Warfield (Eds.), *The Blackwell Guide to Philosophy of Mind* (pp. 1–46), Malden, MA: Blackwell Publishing Ltd.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: Freeman.

Martin, P. R., & Solomon, S. G. (2011). Information processing in the primate visual system. *Journal of Physiology, 589*, 29–31.

Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., & Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Proceedings of the National Academy of Sciences USA, 99*, 1645–1650.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex, 9*, 431–444.

McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. G. (2005). The use of visual information in natural scenes. *Visual Cognition, 12*, 938–953.

Merigan, W. H., & Maunsell, J. H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience, 16*, 369–402.

Mihaylova, M., Stomonyakov, V., & Vassilev, A. (1999). Peripheral and central delay in processing high spatial frequencies: reaction time and VEP latency studies. *Vision Research, 39*, 699–705.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Morrison, D. J., & Schyns, P. G. (2001). Usage of spatial scales for the categorization of faces, objects, and scenes. *Psychonomic Bulletin and Review*, *8*, 454–469.

Musselwhite, M. J., & Jeffreys, D. A. (1985). The influence of spatial frequency on the reaction times and evoked potentials recorded to grating pattern stimuli. *Vision Research, 25*, 1545–1555.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science, 13*, 416–424.

Naccache, L., Gaillard, R., Adam, C., Hasboun, D., Clémenceau, S., Baulac, M., Dehaene, S., & Cohen, L. (2005). A direct intracranial record of emotions evoked by subliminal words. *Proceedings of the National Academy of Sciences, 102,* 7713–7717.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*, 435–450.

Näsänen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research, 39*, 3824–3833.

Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience, 10*, 360–372.

Nemrodov, D., & Itier, R. (2011). The role of eyes in early face processing: A rapid adaptation study of the inversion effect. *British Journal of Psychology, 102*, 783–798.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology, 34,* 72–107.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607–609.

Özgen, E., Sowden, P., Schyns, P., & Daoutis, C. (2005). Top-down modulation of spatial frequency processing in scene perception. *Visual Cognition, 12*, 925–937.

Parker, D. M., & Dutch, S. (1987). Perceptual latency and spatial frequency. *Vision Research, 27*, 1279–1283.

Parker, D. M., Lishman, J. R., & Hughes, J. (1992). Temporal integration of spatially filtered visual images. *Perception, 21*, 147–160.

Parker, D. M., Lishman, J. R., & Hughes, J. (1997). Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception, 26*, 1169–1180.

Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain Sciences, 25*, 297–330.

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*, 773–783.

Peyrin, C., Mermillod, M., Chokron, S., & Marendaz, C. (2006). Effect of temporal constraints on hemispheric asymmetries during spatial frequency processing. *Brain and Cognition, 62*, 214–220.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18*, 849–860.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience, 3*, 261–270.

Rodieck, R. W. (1998). *The first steps in seeing.* Sunderland, MA: Sinauer Associates.

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128,* 274–289.

Rotshtein, P., Vuilleumier, P., Winston, J., Driver, J., & Dolan, R. (2007). Distinct and convergent visual processing of high and low spatial frequency information in faces. *Cerebral Cortex, 17*, 2713–2724.

Ruiz-Soler, M., & Beltran, F. S. (2006). Face perception: an integrative review of the role of spatial frequencies. *Psychological Research, 70*, 273–292.

Sahraie, A., Weiskrantz, L., Barbur, J. L., Simmons, A., Williams, S. C. R., & Brammer, M. J. (1997). Pattern of neuronal activity associated with conscious and unconscious processing of visual signals. *Proceedings of the National Academy of Sciences USA, 94*, 9406–9411.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science, 5*, 195–200.

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of facial expressions of emotion co-evolved with their efficient decoding in the brain: behavioral and brain evidence. *PLoS ONE, 4,* e5625.

Searle, J. R. (1998). How to study consciousness scientifically. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 353,* 1935–1942.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, *14*, 391–396.

Simpson, W. A., & McFadden, S. M. (2005). Spatial frequency channels derived from individual differences. *Vision Research, 45*, 2723–2727.

Smith, M. L. (2012). Rapid processing of emotional expressions without conscious awareness. *Cerebral Cortex, 22,* 1748–1760.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science, 16*, 184–189.

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: replicable and qualitatively different unconscious perceptual effects. *Cognition, 101*, 43–79.

Sowden, P. T., & Schyns, P. G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences, 10*, 538–545.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica, 106*, 147–246.

Sumner, P. (2007). Negative and positive masked-priming – implications for motor inhibition. *Advances in Cognitive Psychology, 3*, 317–326.

Tadros, K., Dupuis-Roy, N., Fiset, D., Arguin, M., & Gosselin, F. (in press). Reading laterally: The cerebral hemispheric use of spatial frequencies in visual word recognition. *Journal of Vision.*

Tamietto, M., & De Gelder, B. (2010). Neural bases of the non-conscious perception of emotional signals. *Nature Reviews Neuroscience, 11,* 697–709.

Tanaka, J. W., Meixner, T. L., & Kantner, J. (2011). Exploring the perceptual spaces of faces, cars and birds in children and adults. *Developmental Science, 14*, 762–768.

Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited: magnocellular and parvocellular contributions to conscious and nonconscious vision. *Psychological Science, 22*, 934–942.

Thurman, S. M., & Grossman, E. D. (2011). Diagnostic spatial frequencies and human efficiency for discriminating actions. *Attention, Perception, & Psychophysics, 73,* 572–580.

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: an update. *Annals of the New York Academy of Sciences, 1124*, 239–261.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8,* 1096–1101.

Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586), Cambridge, MA: MIT Press.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science, 255*, 419–423.

Van Gaal, S., & Lamme, V. A. (2011). Unconscious high-level information processing: implication for neurobiological theories of consciousness. *Neuroscientist, 18*, 287–301.

Verleger, R., Jaśkowski, P., Aydemir, A., Van der Lubbe, R. H. J., & Groen, M. (2004). Qualitative differences between conscious and nonconscious processing? On inverse

priming induced by masked arrows. *Journal of Experimental Psychology: General*, *133*, 494–515.

Vlamings, P. H., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *Journal of Vision, 9*, 1–13.

Vogel, M., Monesson, A., & Scott, L. S. (2012). Building biases in infancy: the influence of race on face and voice emotion matching. *Developmental Science, 15*, 359–372.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6,* 624–631.

Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America A, 4*, 2006–2021.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010a). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Willenbockel, V., Fiset, D., & Tanaka, J. W. (2011). Relative influences of lightness and facial morphology on perceived race. *Perception, 40*, 621–624.

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, 3:237*. doi: 10.3389/fpsyg.2012.00237.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010b). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42,* 671–684.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *Journal of Neuroscience, 24*, 2898–2904.

Williams, N. R., Willenbockel, V., & Gauthier, I. (2009). Sensitivity to spatial frequency and orientation content is not specific to face perception. *Vision Research, 49*, 2353–2362.

Wilson, H. R., & Wilkinson, F. (1997). Evolving concepts of spatial channels in vision: from independence to nonlinear interactions. *Perception, 26*, 939–960.

Wyatte, D., Curran, T., & O'Reilly, R. (2012). The limits of feedforward vision: recurrent processing promotes robust object recognition when objects are degraded. *Journal of Cognitive Neuroscience, 24,* 2248–2261.

Zeidman, P., Mullally, S. L., Schwarzkopf, D. S., Maguire, E. A. (2012). Exploring the parahippocampal cortex response to high and low spatial frequency spaces. *NeuroReport, 23*, 503–507.

Zeki, S. (2003). The disunity of consciousness. *Trends in Cognitive Sciences, 7,* 214–218.

Appendix

Supplementary materials

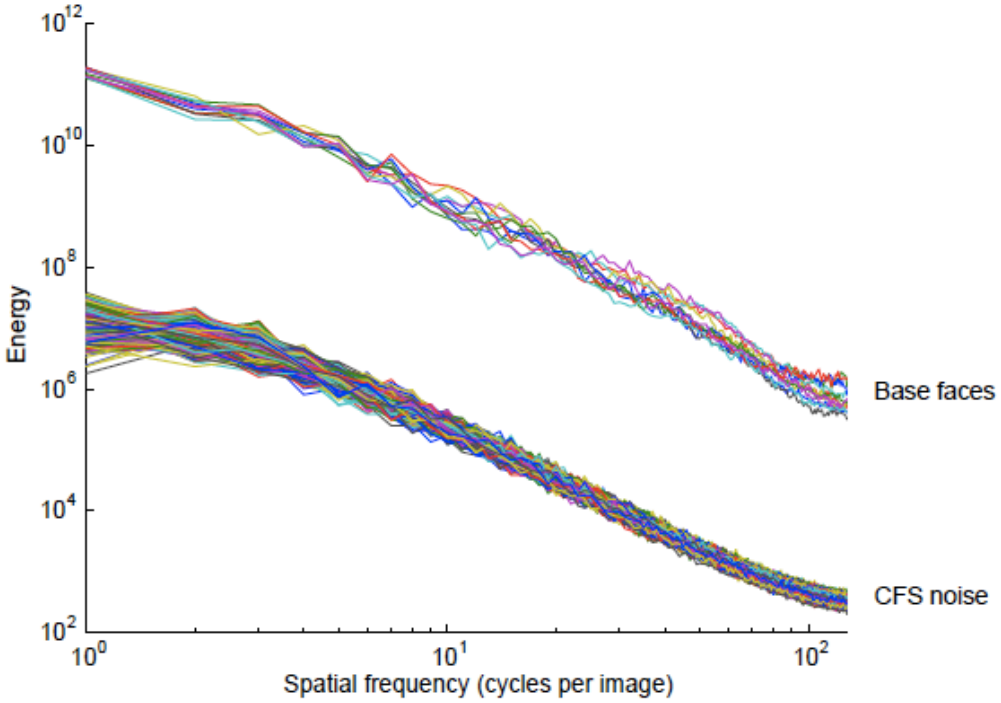
for

Spatial frequency tuning during the conscious and non-conscious perception of emotional facial expressions—an intracranial ERP study

Verena Willenbockel[1], Franco Lepore[1], Dang Khoa Nguyen[2], Alain Bouthillier[2], Frédéric Gosselin[1]

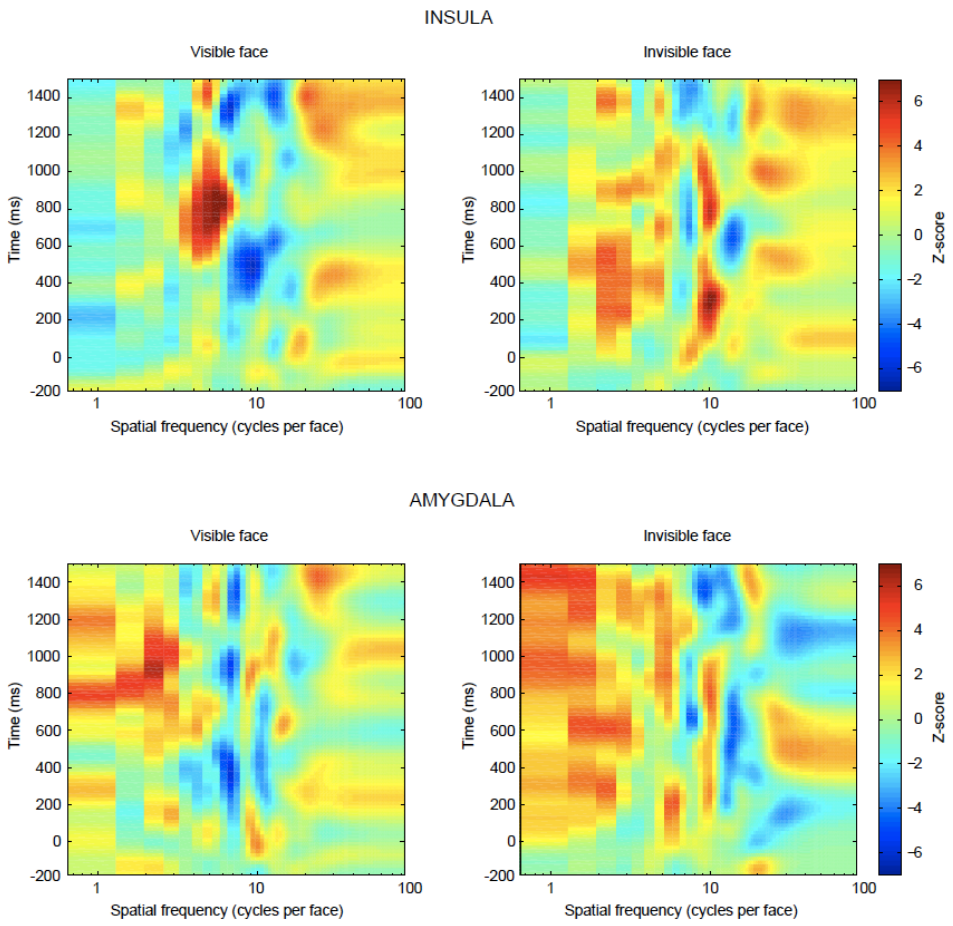
[1]CERNEC, Département de Psychologie, Université de Montréal, Montréal, QC, Canada
[2]Centre Hospitalier de l'Université de Montréal, Hôpital Notre-Dame, Montréal, QC, Canada

Supplementary Figures 1–3



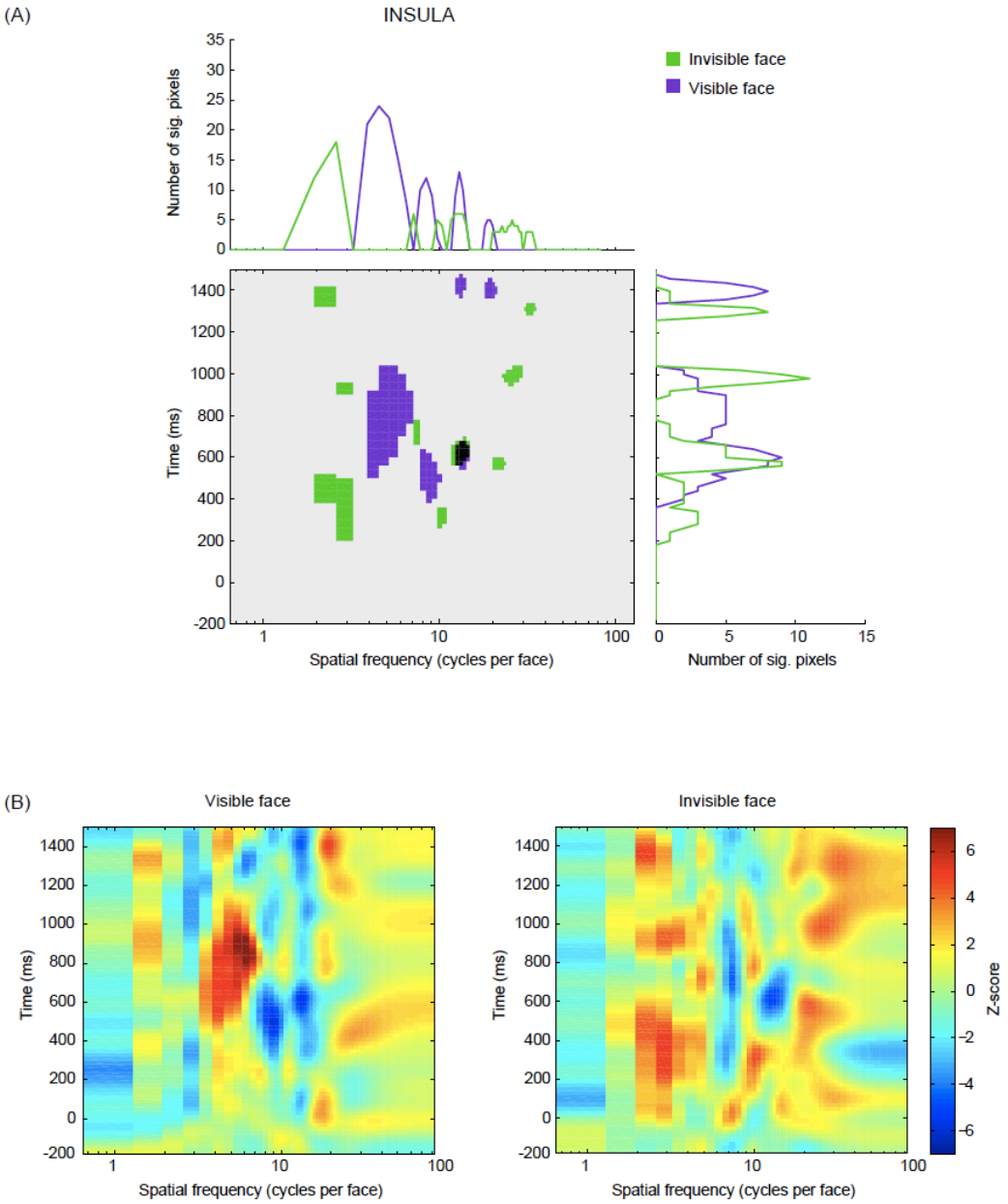
Supplementary Figure 1

Rotational average of the power spectra of the 12 base faces and the suppression noise (540 Mondrian patterns).



Supplementary Figure 2

Raw classification images for the two regions (top: insula; bottom: amygdala) and awareness conditions (left: visible face; right: invisible face).



Supplementary Figure 3

Insula classification images (A: thresholded; B: raw) computed from the data of Participants 2 and 3 only.

Curriculum Vitae (articles)

Willenbockel, V., Lepore, F., Bacon, B., & Gosselin, F. (in press). The "informational correlates" of conscious and non-conscious face-gender perception. *Journal of Vision.*

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, 3:237*. doi: 10.3389/fpsyg.2012.00237.

Herzmann, G., **Willenbockel, V.**, Tanaka, J. W., & Curran, T. (2011). The neural correlates of memory encoding and recognition of own-race and other-race faces. *Neuropsychologia, 49,* 3103–3115.

Willenbockel, V., Fiset, D., & Tanaka, J. W. (2011). Relative influences of lightness and facial morphology on perceived race. *Perception, 40,* 621–624.

Frey, H.-P., Wirz, K., **Willenbockel, V.**, Betz, T., Schreiber, C., Troscianko, T., & König, P. (2011). Beyond correlation: Do color features influence attention in rainforests? *Frontiers in Human Neuroscience, 5:36*. doi: 10.3389/fnhum.2011.00036

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,* 671–684.

Willenbockel, V., Fiset, D., Chauvin, A., Blais, C., Arguin, M., Tanaka, J. W., Bub, D. N., & Gosselin, F. (2010). Does face inversion change spatial frequency tuning? *Journal of Experimental Psychology: Human Perception and Performance, 36,* 122–135.

Williams, N. R., **Willenbockel, V.**, & Gauthier, I. (2009). Sensitivity to spatial frequency and orientation content is not specific to face perception. *Vision Research, 49,* 2353–2362.

Curby, K., **Willenbockel, V.**, Tanaka, J. W., & Schultz, R. (2009). Face processing in autism: Insights from the perceptual expertise framework. In D. Bub, M. J. Tarr, & I. Gauthier (Eds.),

Perceptual expertise: Bridging brain and behavior (pp. 139–166). New York: Oxford University Press.