UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Emotions and the temporal perception of events

Antonio Schettino

Promotor: Prof. Dr. Gilles Pourtois

Proefschrift ingediend tot het behalen van de academische graad
van Doctor in de Psychologie

2012

ACKNOWLEDGEMENTS

There's a lot of boring stuff to read here, I'd better get the acknowledgements done ASAP[1].
I apologize in advance if I forget someone, but if you know me you already expect my memory to fail miserably. If you don't know me, well... you've just found out why you're not mentioned here.

First and foremost, I would like to thank my advisor, Prof. Gilles Pourtois. Throughout these last four years you haven't ceased to amaze me. You gave me countless opportunities to discuss my unripe ideas, and transformed them into something more digestible. You are an invaluable source of inspiration, a volcano of ideas, and a hardcore party animal. Every single conference we attended together was memorable (as much as alcohol allowed it to be), and I truly hope there will be more in the future. Thanks for fulfilling my dream of doing a high-level PhD abroad, I hope you don't regret having chosen me.

I am also greatly indebted to Prof. Tom Loeys. Tom, you were not officially my co-supervisor, but I can safely say that you were my "ghost co-supervisor": without your kind and patient help I would have never been able to extract so much valuable information from my data. I hope we will still collaborate in the future, I have much to learn from you.

Thanks to the members of my Guidance Committee: Rudi De Raedt, Adriaan Spruyt, and Monica Dhar. I greatly appreciated your constructive comments, they steered me to the right direction and prevented me to waste precious time making silly mistakes.

Thanks to the PanLab: Ernst, Sven, Chris, An, Anouk, Jun, Manon, Marie-Anne, Vale, Ine, Igor, Jasmina, Jonas, Lynn, Thomas, Sofie, Anamaria, Leen, all the newcomers whose names I don't know (yet), and the special guests (Nath, Anders, Mandy, Elliot, Menno).

---

[1] That's not true. I'm taking my time.

Kristien and Naomi, sharing the office with you was a blast. We laughed, complained, consumed large quantities of food (especially Naomi), and helped each other along the way. I would do it all over again.

I shared funny moments with many other people, too: RoelJan, Kim, Evelien (you can’t hide your pet rabbit forever), Lies (I’ll come and visit you in Perth, but please get rid of the golden silk orb-weavers in your garden), Kristina & Bas, Helen, Marijke, Zanzo, Julia, Bariş, my friends in Geneva (Christian, Géraldine, Alison, Leonie, Aline, Ben). Jeff, Neeltje, Toon, Emma, Tom: thanks for the “Manic Munchies”. Thanks to my past and present housemates (Fred, Arne, Veerle, Tanguy, Alyssa, Tristan, Jinzhi), Alessia (my fashion stylist and good friend), the Antwerpen boys & gals (Jeroen, Dirk, Ellen, Astrid, Frank).

A special mention goes to: (i) my metal friend Tzom: you are to music as Neal DeGrasse Tyson is to astrophysics; (ii) Niclas: since you started sharing your knowledge with me, I no longer need IMDB.com; (iii) Jonathan and Maud: the future will bring us more gin tonic, and we will enjoy every drop of it.

I would also like to thank the people at Hot Yoga Gent, especially Mareike and Karel. Writing this thesis could have been much harder without your yoga classes, you managed to keep my anxiety level in the “normal” range.

Good Italian friends are not forgotten: Neein & Stefano (we shared amazing trips, and I couldn’t ask for a better company... let’s do it over and over again!), Ari (I’ll soon find a comfortable situation to use your birthday present... I’ll send pictures), Fabio (whose art I have proudly carved in my left arm), Adri (I don’t play hard to get, I’m really busy!), Mike, Silvia, Sara, Barattolo, Cliv, Palloncino, Wallace.
Matti, your courage inspired me to move abroad and fulfill my dreams. I can’t thank you enough for that.

Special thanks to Bruno. Unsurprisingly, this is not the first time you have a special mention in my acknowledgements (remember my master degree thesis, *marmottone ninja*?). Of all the people I have met, of all the wonderful friends I have, you are the one who was, is, and always will be present. I can count on you, and I hope you know you can count on me... no matter how distant we live.

I would also like to thank a whole country, if I may. Japan, your sushi, sashimi & okonomiyaki opened my stomach; your gorgeous girls opened my eyes; your zen gardens opened my Anahata... in short, you gave me what I had been looking for throughout my entire life: peace. I will treasure it forever, or die trying.

I would also like to thank my mom, dad, and brother. Some say “Distance is disaffection”, and I couldn’t disagree more. During my stay in Belgium, I realized how deep my love for you is, and the best part is that I am not afraid to show it anymore.

And you... you. Describing all the multifarious ways in which you make me feel good is the greatest of all challenges. Indeed I had to fight for you, but it was worth every tear... and now I know that pure love does exist.

I love you, Inez.

Antonio
September 2012

There is never [...] anoetic sentience. The organism in perception is in one way or another in a state of expectancy about the environment.

Jerome S. Bruner & Leo Postman, 1949

TABLE OF CONTENTS

CHAPTER 1

General introduction

“Feeling is not free of thought, nor is thought free of feeling” (Zajonc, 1980). Consistent with this famous quote, mounting evidence suggests that affective evaluations are not simply by-products of our mental life, but instead they appear to be essential and play a causal role in a broad range of psychological phenomena, including attitudes formation (Ito & Cacioppo, 2001), prejudice (Forgas & Fiedler, 1996), moral judgment (Greene & Haidt, 2002), personality (Watson, 2000), and psychopathology (Davidson, 2000). In this context, it is not surprising to see that what was in the past known as “cold cognition” (i.e., devoid of any affective connotation) has recently been shown to be profoundly shaped by affective processes (Dolan, 2002; Mesulam, 1998; Schwarz, 2000), including in the domain of memory (Bower, 1981; Dolcos & Denkova, 2008), decision-making (Damasio, 1996), or selective attention (Vuilleumier, 2005; Vuilleumier & Huang, 2009). The goal of this thesis was to highlight and better characterize the extent to which affective processes could dynamically influence early visual processes and hence guide the content of visual cognition.

In this general introduction, I will first situate the complex construct of emotion in psychophysiology and psychology, and delineate a number of key mental processes classically involved in the rapid and efficient processing of basic emotional stimuli or complex visual scenes. Next, I will review evidence for the interplay between emotion and perceptual brain processes via attention mechanisms, focusing on the classical experimental paradigms commonly used in the literature to disclose these puzzling cognition-emotion interaction effects. In addition, I will describe an alternative task (temporal order judgment) better suited to gain insight into the putative exogenous capture of attention by motivationally salient stimuli, and which I have extensively used in the first empirical part of this dissertation (see Chapter 2). As will appear more clearly here below, however, an intrinsic limitation of these paradigms is that effects of emotion on visual perception are essentially deemed “reactive” and “bottom-up”. The underlying assumption is that motivationally significant (in particular threat-related) stimuli exert “automatic” gating effects on early visual perception, these effects being typically measured following the onset of the full-blown emotional stimulus. However, emotion processes may also dynamically and proactively change perception *before* stimulus onset, presumably via top-down expectation and associative brain mechanisms (Bar, 2009b; Gosselin & Schyns, 2003; Niemi & Näätänen, 1981). Accordingly, I will move on to “proactive” effects of emotion on vision, focusing on predictive models of visual perception (Friston, 2005; Friston & Kiebel, 2009) and accumulator models of perceptual decision making (Gold & Shadlen, 2007; Heekeren, Marrett, & Ungerleider, 2008). I will then present a new experimental paradigm based on the progressive revelation of the stimulus content, which is suited to timely delineate how expectations (here primarily related to emotion) may bias “online” visual scene recognition processes. The validation of this new experimental paradigm forms the core of the second section of the experimental part of this dissertation (Chapters 3-5). Finally, I will briefly describe the methods used in this thesis, namely the specific statistical analyses used to model the behavioral data collected in several samples of adult healthy participants (Chapters 2-5), as well as the ERP topographic mapping analyses employed in Chapters 3-4 to delineate the rapidly evolving spatio-temporal dynamics of visual perceptual processes, and their modulation by top-down expectations regarding the emotional vs. non-emotional content of the retinal input.

1. Emotion as motivational drive

1.1. The Evaluative Space Model (ESM)

Many theoretical models of emotion have been proposed in the literature, including discrete or basic emotion theories (assuming a limited number of innate, hard-wired affect reactions; Ekman, 1992; Izard, 1993), neo-constructivist theories (emotions are primarily described in relation to valence and arousal dimensions, as categorized by each person’s memory and cultural concepts; Barrett, 2006a; Russell, 2003, 2009), and appraisal theories (emotion is a dynamic process based on an individual's subjective appraisal of significant events, which in turn determines action; Fontaine, Scherer, Roesch, & Ellsworth, 2007; Frijda, 1987; Scherer, 2009). Throughout this dissertation, emotion will be referred to as a general response of the body meant to act in reaction to stimuli that either threaten or sustain life (Damasio, 1994; Frijda, 1987; Lang, 1985) via the concurrent activation of either a defensive or appetitive motivational system (Bradley, 2009; Cacioppo & Gardner, 1999; Lang & Bradley, 2010). The *defensive* system is activated in threatening contexts, and mainly elicits “fight-or-flight” behavior such as withdrawal, escape, and attack (Cannon, 1932; Lang & Bradley, 2010). On the other hand, the *appetitive* system promotes survival (and, phylogenetically, the continuation of the species) in non-dangerous situations through search for nourishment and opportunities for procreation (Bradley, 2009). These two systems allow the organism to constantly appraise stimuli or situations on the basis of common motivational parameters, such as valence and arousal (Cacioppo & Berntson, 1994; Cacioppo & Gardner, 1999; Lang, 1995; Öhman, Hamm, & Hugdahl, 2000). *Valence* refers to the categorization of stimuli in terms of the degree of pleasantness or unpleasantness (Barrett, 2006b; Russell, 2003): the appetitive system prompts positive affect, whereas negative affect originates from the defensive system (Lang & Bradley, 2010). *Arousal* reflects the intensity of activation of these motivational systems (Russell, 1980), and can be conceptualized as the extent to which a stimulus elicits an individual sensation of energy, an “experience of feeling active” (Barrett, Mesquita, Ochsner, & Gross, 2007; Gerber, et al., 2008).

The ability to categorize and discriminate stimuli in the environment in terms of valence and arousal is fundamental for the organism’s survival (Brosch, Pourtois, & Sander, 2010), because it sparks off fast behavioral responses (approach or withdrawal) that could turn out to be crucial, particularly in situations of danger (Bradley, 2009; Cacioppo & Gardner, 1999; Lang & Bradley, 2010). Moreover, valence and arousal are relevant concepts from a practical perspective, because the visual stimuli usually employed in research on the interplay between vision and emotion are usually categorized along these two fundamental and independent dimensions (including the complex scenes used in the second part of this doctoral thesis, i.e., IAPS database; Lang, Bradley, & Cuthbert, 2008). However, this bipolar structural organization of the affect system may not be sufficient to explain the immense complexity and richness of all the possible emotional repertoire (Cacioppo, Gardner, & Berntson, 1997; Fontaine, et al., 2007; Norris, Gollan, Berntson, & Cacioppo, 2010; Scherer, 2009). Theories that emphasize the categorical or dimensional organization of affective experience have usually based their observations on (self-)reported feeling states (Ekman, Friesen, & Ellsworth, 1972; Russell & Carroll, 1999; Watson, Wiese, Vaidya, & Tellegen, 1999), often overlooking the fact that many cognitive processes occur outside of awareness and therefore are not prone to introspection (Norris, et al., 2010; Wilson & Bar-Anan, 2008). Moreover, although behavioral output may perhaps be operationalized through a single dimension with two opposite anchors (approach vs. withdrawal), this does not necessarily imply that the underlying structure of the affective system follows such a relatively simple organization (Cacioppo & Gardner, 1999; Cacioppo, et al., 1997; Norris, et al., 2010). To overcome these limitations, the Evaluative Space Model (*ESM*; Cacioppo & Berntson, 1994, 1999; Cacioppo & Gardner, 1999; Cacioppo, et al., 1997; Norris, et al., 2010) postulates that the affective system receives inputs from (at least) two specialized evaluative channels that process information in parallel: one processes threat-related (negative) inputs, while the other one simultaneously analyzes appetitive (positive) information (Cacioppo & Gardner, 1999; Cacioppo, Gardner, & Berntson, 1999; Lang, Bradley, & Cuthbert, 1990). Given these two (partially) segregated evaluative channels, the range of possible behaviors of the affective system is larger (Cacioppo & Gardner, 1999). The main advantage of this approach over other theories of emotion (e.g., Barrett, 2006a; Ekman, 1992; Russell, 1980) is that positivity and negativity serve as driving motivational forces that concur to directly shape behavior: separate (but synergistic) evaluative channels may explain how individuals enjoy both the benefits of exploration and the self-preservative advantages of a predisposition to avoid threatening situations (Cacioppo & Gardner, 1999; Cacioppo, et al., 1997, 1999). This architecture of the affective system further enables to model and predict different “activation functions” or motivational drives during the online evaluation of (and subsequent response to) emotion-laden stimuli, as described here below.

1.2. Positivity offset

Positivity offset refers to the fact that, when inputs to the affective system are low and minimal, the output of positivity is higher than negativity or, in other words, positive affect outweighs negative affect (Ito, Cacioppo, & Lang, 1998; Norris, et al., 2010). As a direct consequence, at low levels of activation (e.g., in non-threatening or goal-conducive situations) the general motivation to approach is usually stronger than the motivation to avoid or protect. From an evolutionary perspective, this activation function enables organisms to approach neutral or unfamiliar stimuli as well as explore novel environments, with the ultimate goal to find new or additional sources of nourishment and protection, eventually promoting mating and reproduction (Cacioppo & Gardner, 1999).

Several lines of research provide evidence for positivity offset effects under specific circumstances. For instance, participants tend to have a positive first impression of unknown individuals (*positivity bias*; Klar & Giladi, 1997; Sears, 1983). Moreover, they are usually too optimistic (Sharot, Riccardi, Raio, & Phelps, 2007), and hence tend to overestimate the number of positive events they will experience in the near future (Pulford & Colman, 1996), as well as attribute lower (health) risks to themselves than others (Hoorens & Buunk, 1993). Also, when individuals are requested to evaluate objects or other people, positive words are generally used more frequently and easily than negative words (*Pollyanna hypothesis*; Boucher & Osgood, 1969).

1.3. Negativity bias

Negativity bias reflects heightened sensitivity to negative information: due to their enhanced motivational value (it is more difficult to reverse the consequences of an injury than those of a missed opportunity), aversive stimuli usually elicit stronger bodily responses compared to neutral or appetitive ones (Cacioppo & Gardner, 1999; Cacioppo & Ito, 1999; Ito, Larsen, Smith, & Cacioppo, 1998). Thus, negative events usually have a greater impact on individuals and longer lasting effects than positively valenced experiences (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Sheldon, Ryan, & Reis, 1996). A large number of studies have extensively demonstrated that negative (particularly threat-related) stimuli receive prioritized access to awareness by biasing perceptual and attentional processes (Bradley, 2009; Compton, 2003; Öhman, Flykt, & Esteves, 2001; Pourtois, Schettino, & Vuilleumier, 2012; Vuilleumier, 2005). In the next section, I will provide a general presentation of cognitive processes (and brain mechanisms) through which (negative) emotion can profoundly bias visual perception and (selective) attention.

2. How emotion colors visual perception: the role of attention

2.1. General mechanisms of attention control

Due to the limited processing capacity of the sensory systems, which cannot simultaneously process multiple objects at a given moment in time (Broadbent, 1957; Marois & Ivanoff, 2005; Neisser, 1967; Posner, 1980), attention selection mechanisms operate as strong and efficient filters during stimulus processing in order to focus on the most relevant information in an overwhelmingly rich environment (Driver, 2001; Kastner & Ungerleider, 2000). Which stimulus eventually “wins” this competition and receives priority for further processing (as well as access to awareness) is determined by both *bottom-up*, stimulus-driven factors that reflect changes in salient perceptual properties, and *top-down* factors, such as prior knowledge, expectations and current goals (Corbetta & Shulman, 2002; Egeth & Yantis, 1997; Posner, 1980; Serences, et al., 2005; Theeuwes, 1994; Yantis, 2000). More specifically, task-irrelevant stimuli or features (if they carry novelty or perceptual salience, such as motion, contrast, or luminance) may draw attention in a bottom-up, exogenous way, in the sense that attention is reflexively and automatically captured by these events (or features), even though current goals may actually concern other properties (Kastner & Ungerleider, 2000; Theeuwes, 1994, 2005; Wolfe & Horowitz, 2004). Alternatively, the processing of a perceptually weak or hidden stimulus may be prioritized via top-down, *endogenous* attention mechanisms, when it matches with current (exploration) goals, expectations, or intentions (Bacon & Egeth, 1994; Folk, Remington, & Johnston, 1992; Vogt, De Houwer, Moors, Van Damme, & Crombez, 2010). Importantly, exogenous and endogenous attention mechanisms are not mutually exclusive, but they may show reciprocal interaction effects at various levels along the processing hierarchy (Corbetta & Shulman, 2002; Hopfinger & West, 2006; McMains & Kastner, 2011). Exogenous attention, although reflexive, rapid and automatic (Mulckhuyse & Theeuwes, 2010; Posner, Snyder, & Davidson, 1980), can in fact be modulated by top-down attentional factors, related to expectations and task demands (Folk, et al., 1992). Symmetrically, endogenous attention, which is usually assumed to depend on effortful and voluntary control mechanisms (Kastner & Ungerleider, 2000), can be influenced by both conscious (Vogt, et al., 2010) and unconscious (Moskowitz, 2002) activation of goals. This dynamic interplay between exogenous and endogenous attention control mechanisms is further supported by a number of neuroimaging studies showing largely overlapping brain networks for these two attention control mechanisms, particularly involving frontoparietal areas (Corbetta & Shulman, 2002; Kim, et al., 1999; Mayer, Dorflinger, Rao, & Seidenberg, 2004; McMains & Kastner, 2011; Peelen, Heslenfeld, & Theeuwes, 2004; Rosen, et al., 1999). Even though they show overlapping activations, they can also be dissociated at the neural level. Endogenous attention control is typically mediated by the intraparietal sulcus and frontal eye fields, while exogenous orienting of the attentional focus is controlled by more ventral frontal and parietal regions (with a right hemispheric preference), i.e., the ventral frontal cortex and the temporoparietal junction (Brosch, Pourtois, Sander, & Vuilleumier, 2011; Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002; Peelen, et al., 2004).

Most of the evidence for the existence of top-down (endogenous) and bottom-up (exogenous) attention control mechanisms reviewed here above has been obtained with stimuli or situations devoid of emotion. However, a large body of behavioral and neuroimaging studies have shown that the rapid and efficient selection of sensory information for further perceptual analysis is not only determined by physical (bottom-up) or goal-related (top-down) factors, but also by the emotional or motivational significance of the stimulus (or situation) for the individual (Bradley, 2009; Compton, 2003; Öhman, 2001; Pourtois, et al., 2012; Storbeck & Clore, 2007; Vuilleumier, 2005; Yiend, 2010).

2.2. Emotional attention

Recent models emphasize that emotion could influence (visual) perception via brain mechanisms that do not overlap with those involved in the exogenous or endogenous control of attention (Pourtois, et al., 2012; Todd, Cunningham, Anderson, & Thompson, 2012; Vuilleumier, 2005). Hence, besides these two fundamental mechanisms, attention control could also be initiated by a third category of stimuli or events, namely emotional stimuli. In this view, effects of emotion on attention can be dissociated from “pure” exogenous or endogenous factors (Brosch, et al., 2011; Keil, Moratti, Sabatinelli, Bradley, & Lang, 2005). More specifically, emotion is thought to influence attentional control processes via specific neural mechanisms and brain pathways (Pourtois, et al., 2012). First, emotional stimuli usually lead to an enhanced sensory processing (and hence neural responses in the visual cortex) compared to neutral stimuli, similarly to classical attention gain control mechanism (Desimone & Duncan, 1995; Hillyard, 1985; Kastner & Ungerleider, 2000; Luck, Woodman, & Vogel, 2000). This effect includes signal-to-noise ratio changes in category-selective regions (e.g., fusiform gyrus in response to emotional vs. neutral faces; Vuilleumier, Armony, Driver, & Dolan, 2001) as well as in earlier visual areas, including V1 (Damaraju, Huang, Barrett, & Pessoa, 2009; Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000; Pourtois, Grandjean, Sander, & Vuilleumier, 2004). Second, emotional stimuli can influence, early on following stimulus onset, activation of limbic brain areas implicated in the processing of the relevance or affective significance of these stimuli (LeDoux, 1996), such as the amygdala or the orbitofrontal cortex (Kawasaki, et al., 2001; Luo, Holroyd, Jones, Hendler, & Blair, 2007; Luo, et al., 2010; Pourtois, Spinelli, Seeck, & Vuilleumier, 2010b). Thanks to the reciprocal anatomical connections between these limbic structures and the occipital/temporal visual cortex (Amaral, Behniea, & Kelly, 2003; Catani, Jones, Donato, & ffytche, 2003; Gschwind, Pourtois, Schwartz, de Ville, & Vuilleumier, 2012; Vuilleumier, Richardson, Armony, Driver, & Dolan, 2004), these early effects following stimulus onset taking place in the amygdala and orbitofrontal cortex can in turn “gate” emotional stimulus processing in distant regions of the visual cortex at a later latency (Keil, et al., 2009; Pourtois, Spinelli, Seeck, & Vuilleumier, 2010a; Sabatinelli, Lang, Bradley, Costa, & Keil, 2009). Hence, given this specific neural pathway, effects of emotion on attention appear to occur in parallel to other attentional gating effects mediated by the classical dorsal frontoparietal networks involved in the endogenous and exogenous control of attention (Corbetta & Shulman, 2002; Peelen, et al., 2004).

At the behavioral level, several experimental paradigms have been used in order to substantiate how emotional -- particularly threatening -- stimuli may bias early visual perceptual processes more readily than neutral stimuli, via the putative mediation of enhanced attention control mechanisms (Anderson, 2005; Öhman, 2001; Öhman, Lundqvist, & Esteves, 2001; Phelps, 2006). However, as will appear more clearly here below, each of these paradigms does not provide a direct, unique or pure measure of early attentional capture by emotion, but multiple (attention) processes are presumably operating in parallel such that it remains difficult to tease apart a specific contribution of emotion on visual perception by means of these specific tasks (Horstmann, 2007; West, Anderson, & Pratt, 2009). To overcome these problems, after a brief overview of the main experimental paradigms most frequently used in the literature to investigate emotional attention effects on perception, I will outline an alternative experimental paradigm (namely the temporal order judgment task) that appears better suited to explore early emotion-dependent orienting effects of attention and influences on (temporal) perception. This experimental paradigm has been extensively used in the first empirical section of this thesis (Chapter 2) in order to corroborate the assumption of early, reflexive effects of (negative) emotion on (temporal) perception.

2.3. Emotional Stroop task

In the classical Stroop task, participants have to name as quickly as possible the color of the ink in which words are printed. The (semantic) meaning of the word itself is thus irrelevant for the task at hand, but it is “automatically” processed nonetheless and hence it slows down color naming, particularly when the meaning and the ink of the word mismatch

(Macleod, 1991; Stroop, 1935). Similarly, in the *emotional* Stroop task, naming the color or simply counting emotion-laden (particularly negative) words is typically slower compared to naming neutral words, even though the word meaning is task-irrelevant and should be ignored (Algom, Chajut, & Lev, 2004; Phaf & Kan, 2007). For example, response latencies to naming the color of the ink of the word *cancer* are generally slower than naming the color of the ink of the word *flower*. These results are interpreted as reflecting enhanced attentional orienting towards the content of these motivationally relevant stimuli (Yiend, 2010). This effect has been observed in non-preselected adult healthy participants (Pratto & John, 1991; Wentura, Rothermund, & Bak, 2000), but is usually exacerbated in subclinical anxiety (Fox, 1993; Salters-Pedneault, Gentes, & Roemer, 2007) as well as in specific psychopathological conditions, including generalized anxiety disorder (Mogg, Mathews, & Weinman, 1989), phobias (Watts, McKenna, Sharrock, & Trezise, 1986), post-traumatic stress disorder (Cisler, et al., 2011), obsessive-compulsive disorder (Moritz, et al., 2004), and depression (Epp, Dobson, Dozois, & Frewen, 2012; Gotlib & McCann, 1984).

Despite its broad use in the literature and value to reveal effects of emotion on attention, however, the emotional Stroop task suffers from a number of theoretical and methodological limitations, which overall complicate the interpretation of the results. First, it is unclear whether the emotional Stroop task actually measures a standard Stroop effect or something else (Algom, et al., 2004). Specifically, in the classical Stroop effect the contrast is between congruent (e.g., the word *yellow* presented in yellow ink) and incongruent trials (e.g., the word *red* presented in green ink). Thus, a clear one-to-one relationship between the color of the ink and the word meaning is always present, creating interference when these two mismatch. Such a clear contrast is not evident for the stimuli used in emotional Stroop tasks: following the example outlined above, no direct relationship can be inferred between the word *cancer* and the color of the ink. Therefore, standard incongruency effects cannot be calculated with the emotional Stroop task (Algom, et al., 2004). Moreover, the reported delayed response latencies for negative vs. neutral stimuli may very well stem from post-attentional processes (Algom, et al., 2004; Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007): when the task is administered to sub-clinical or clinical (e.g., anxious) populations, threat-related words might inherently increase the current negative affect state of participants, thus slowing down reaction times at the response level (MacLeod, Mathews, & Tata, 1986). Alternatively, responses might be delayed as a consequence of participants’ active effort in avoiding (or diverting attention away from) threat-related information (Cisler & Koster, 2010; De Ruiter & Brosschot, 1994; Koster, Verschuere, Crombez, & Van Damme, 2005). Also, participants might develop a generalized sensitivity for negative information only because they are presented with an increasing number of negative words throughout several experimental blocks, thus leading to an enhanced perception of threat (Bar-Haim, et al., 2007; McKenna & Sharma, 2004; Phaf & Kan, 2007). In light of these limitations, other experimental paradigms have been devised and used in the literature in order to highlight interaction effects between attentional and emotional processes. At any rate, it is also clear that the emotional Stroop task cannot be used to make strong, unequivocal inferences regarding a rapid bias during attention allocation towards motivationally relevant stimuli.

2.4. Attentional blink task

In the attentional blink (*AB*) paradigm, a series of stimuli (usually letters or words, sometimes faces) is rapidly presented on the screen at a rate of ~10 stimuli per second (rapid serial visual presentation; *RSVP*), and participants have to identify one or more (usually two) pre-defined targets based on their diagnostic color (Broadbent & Broadbent, 1987; Raymond, Shapiro, & Arnell, 1992). Although any single target (presented for only 100 ms) can be accurately reported, the second of two consecutive targets often fails to be reported when the temporal gap between the onset of the first (T1) and second (T2) stimulus is kept short, i.e., 200-500 ms (Reeves & Sperling, 1986; Shapiro, Arnell, & Raymond, 1997). Interestingly, when emotional stimuli -- e.g., words (Anderson, 2005; Ihssen & Keil, 2009; Keil & Ihssen, 2004; Tibboel, Van Bockstaele, & De Houwer, 2011) or pictures (Most, Chun, Widders, & Zald, 2005) -- are embedded in a RSVP, two different emotional attention effects have been observed: (i) AB is reduced when T2 carries a (negative) emotional meaning or is highly arousing (Anderson, 2005; Keil & Ihssen, 2004; Keil, Ihssen, & Heim, 2006; Schwabe, et al., 2011); (ii) AB is prolonged if T1 is highly arousing (Anderson, 2005; Arnell, Killman, & Fijavz, 2007; Maratos, 2011; Mathewson, Arnell, & Mansfield, 2008).

Several theoretical accounts have been proposed to explain the classical AB effect (notably, the debate is still ongoing; for a detailed review, see Dux & Marois, 2009) that can also be borrowed to interpret changes in the profile of the AB created by emotional stimuli. One of the most influential theories, the *interference account* (Shapiro, Raymond, & Arnell, 1994), postulates that the perceptual representation of each and every stimulus of the RSVP sequence is compared with selection templates stored in working memory, and the stimuli that most closely match with these templates are registered and assigned a weight based on the available space. Due to the limited capacity of working memory, the processing of T2 is impaired because the system is still busy storing, processing, and retaining T1, thereby leading to an AB effect (see also Chun & Potter, 1995; Giesbrecht & Di Lollo, 1998). Similarly, because of the intrinsic motivational relevance of (negative) emotional stimuli (e.g., arousing or taboo words; Mathewson, et al., 2008; Tibboel, De Houwer, & Crombez, 2008), attending to an emotion-laden T1 reduces the attentional resources required to encode a non-emotional T2 in working memory, thereby impairing performance (Arnell, et al., 2007; Mathewson, et al., 2008). Likewise, the processing of arousing T2 could still be enhanced (following neutral T1) because these stimuli “automatically” attract attention given their intrinsic motivational relevance, eventually resulting in a diminished AB (Ihssen & Keil, 2009; Keil & Ihssen, 2004; Schwabe, et al., 2011). Alternative accounts posit that the AB mostly stems from an inefficient allocation of temporal attention (a consequence of the RSVP), which can somehow be restored “paradoxically”, when central attentional resources are occupied by a concurrent, distracting secondary task (see Olivers & Nieuwenhuis, 2006).

The AB paradigm has proven to be extremely effective and valuable to highlight and decompose key features of attention selection and their likely modulation by emotion processes. However, it does not enable to pinpoint whether an early attentional capture by emotion occurs during the processing of the successive (verbal) stimuli or not. Multiple parallel attention and memory processes are likely contributing to the AB effect, including the active inhibition of physical features of T1 that impairs the subsequent identification of T2 (Raymond, et al., 1992), the switching between the semantic category of T1 and T2 (Di Lollo, Kawahara, Ghorashi, & Enns, 2005), the prolonged attentional re-engagement towards T2 (Nieuwenstein, Potter, & Theeuwes, 2009), and the overall working memory capacity (Chun & Potter, 1995; Shapiro, et al., 1994). Therefore, it remains difficult, based on the use of the AB exclusively, to decide whether emotional (unlike neutral) stimuli reflexively attract attention and in turn impact on visual perception.

2.5. *Visual search task*

In *visual search* tasks, participants are instructed to look for a specific target embedded in an array of perceptually similar distracters. Attention selection and search can be guided either by automatic processes in the case of singletons (e.g., a red “5” among several blue “5”s) or by more controlled processes in the case of a conjunction of features (e.g., a red “5” among several blue “5”s, red “6”s, and blue “6”s) (Treisman & Gelade, 1980; Wolfe &

Horowitz, 2004). Interestingly, emotional (particularly threatening) faces have been found to capture attention quicker and more effectively relative to neutral faces, serving as evidence for the “face in the crowd” effect (Eastwood, Smilek, & Merikle, 2003; Fox, Russo, Bowles, & Dutton, 2001; Hansen & Hansen, 1988; Lucas & Vuilleumier, 2008; Öhman, Flykt, et al., 2001). However, these findings have been disputed by many authors given the equivocal nature of the effect (see Wolfe & Horowitz, 2004 for a review). Among the well-known limitations, with visual search tasks it remains somewhat difficult to disentangle the impact of emotion or semantic property of the stimulus from other low-level, perceptual or physical features that can potentially bias visual attention very rapidly, independently from a deep processing of the stimulus at the emotional or semantic level (e.g., mouth or eyes; Calvo & Nummenmaa, 2011; Fox & Damjanovic, 2006; Horstmann & Bauland, 2006; VanRullen, 2006). Moreover, this paradigm alike does not exclusively titrate a genuine early spatial orienting effect of attention towards the location of motivationally significant objects. In fact, according to some authors in vision research (e.g., Horstmann, 2007; Wolfe, 2001; but see Joseph, Chun, & Nakayama, 1997), efficient visual search occurs when a target stimulus is detected regardless of the number of distracter items. If an angry face is detected within 500 ms when embedded in an array of happy faces regardless of the number of these distracters, it means that search for this target is efficient, and some specific features of the angry face (not present, or at least not to the same extent, in happy faces) may actively capture attention. Conversely, if the detection latency of a happy face among angry faces is slower as a function of the increasing number of these distracters, search is non-efficient and attention is assumed to be deployed serially on each and every item in the array until the target is detected (Wolfe, 2001; Wolfe & Horowitz, 2004). Many results in literature support a serial search hypothesis for emotional relative to neutral face stimuli (e.g., Eastwood, Smilek, & Merikle, 2001; Fox, et al., 2000; Öhman, Lundqvist, et al., 2001). One reason that could potentially account for this outcome is the fact that emotional faces are usually rather complex stimuli (even when schematic faces are used), whose identification is the result of a composite process requiring the rapid identification of distinct visual features (such as a specific curvature of the mouth or the first order configuration regarding the position of the eyes relative to the mouth; see Maurer, Le Grand, & Mondloch, 2002). Because early attention orienting has been found to be driven mostly by changes in basic perceptual features (i.e., color, orientation, size, luminance, and spatial frequency; Egeth & Yantis, 1997; Theeuwes, 1994; Wolfe & Horowitz, 2004), it is unclear, based on visual search tasks alone, whether the negative emotional face expression *per se* can capture attention early on following stimulus onset, or other (confounding) low-level perceptual variables can account for this effect (Horstmann, 2007).

2.6. (Spatial) cueing and dot probe tasks

In dot probe tasks (Fox, et al., 2001; MacLeod, et al., 1986; Mogg & Bradley, 1999; Posner, 1980), two cues (e.g., faces or pictures) are simultaneously presented on the left and right side or fixation (or above vs. below fixation), and participants have to either detect or discriminate the content of a subsequent target stimulus replacing one of the cues after a short time interval. Usually the cueing is non-predictive, such that the target appears regardless of the position of the preceding cues. Cueing tasks are a variant of the dot probe tasks, where only one cue is typically used. Importantly, one cue is emotionally significant (e.g., threatening) and the other one is neutral. Participants are usually better (as reflected by faster reaction times or higher accuracy) at processing the (non-emotional) target stimulus if it appears at the location previously occupied by the emotional cue (valid trials) (Bocanegra & Zeelenberg, 2009, 2011b; Brosch, et al., 2011; Fox, et al., 2001; Phelps, Ling, & Carrasco, 2006; Pourtois, et al., 2004). Sometimes, in particular for high anxious or dysphoric participants, a disengagement effect rather than a facilitated orienting towards emotional stimuli is seen with (exogenous) cueing tasks, corresponding to a prolonged latency to process the target for invalid compared to valid or neutral trials. This observation is consistent with the idea of a holding rather than a capture effect (i.e., attention is held on the spatial location of an emotion-laden stimulus; Fox, et al., 2001; Koster, Crombez, Verschuere, Van Damme, & Wiersema, 2006). Whether being expressed by faster RTs for valid trials or slower RTs for invalid trials, the underlying logic is shared in these two cases and one usually concludes that attention is “automatically” captured by the task-irrelevant, motivationally salient stimuli (Bradley, 2009), an effect which in turn leads to an enhanced processing of the subsequent (task-relevant) target (Fox, et al., 2001).

Despite the value of these cueing paradigms to study effects of emotion on the rapid deployment or reorienting of attention, alternative interpretations of these findings are also proposed, which do not postulate any direct mediation by emotional processes. Most of the time, results are solely based on RTs (or discrimination accuracy) corresponding to the processing of the non-emotional target as opposed to the emotional cue (which is typically task-irrelevant). Therefore, only indirect inferences on how (spatial) attention is differentially allocated during the presentation of the emotional cue can be made (Fox, et al., 2001; West, et al., 2009). Presumably, RTs likely reflect the end-product of many distinct but interactive attentional processes (e.g., orienting, disengagement, shifting, and re-engagement), and in these conditions it remains difficult to know which of them is selectively modulated by affective processes (Posner, 1980).

2.7. Temporal Order Judgment

So far, I have reviewed four classical experimental paradigms used in the literature in order to gain insight into effects of emotion on attention control processes. Although each of them rests on specific assumptions and provides useful hints on the interplay between emotion and attention processes, they all have specific limitations. In particular, none of them appears suited to measure and explore the early and initial differential allocation of attention towards emotional (as opposed to neutral) stimuli exclusively (Horstmann, 2007; West, et al., 2009). By contrast, temporal order judgment (TOJ) tasks provide a direct, sensitive and accurate measure of attentional capture (Jaskowski, 1993; Stelmach, Herdman, & Mcneil, 1994; Titchener, 1908). Surprisingly, very few studies (Fecica & Stolz, 2008; West, et al., 2009) have actually capitalized on the strengths of this specific experimental paradigm in order to corroborate the assumption of an early differential orienting effect triggered by emotion on attention control processes and, in turn, low-level visual perception. Therefore, one of the goals of my dissertation was to explore whether (negative) emotional stimuli had the propensity to capture “reflexively” attention and hence bias TOJs.

In a typical TOJ task, attention is oriented (using symbolic cues) either to the left or the right side of fixation, and participants have to judge which of two subsequent and competing visual stimuli, displayed on the left and right at various stimulus onset asynchronies (SOAs), was presented first. Because attention “accelerates” sensory processing (Desimone & Duncan, 1995; Serences & Yantis, 2006), stimuli appearing on the attended location are processed faster and, as a consequence, perceived as appearing first (an effect termed *visual prior entry*; Schneider & Bavelier, 2003; Spence & Parise, 2010; Spence, Shore, & Klein, 2001). Due to their intrinsic motivational relevance, (negative) emotional facial expressions (i.e., angry or fearful faces) could capture attention in an automatic fashion (Esteves, Dimberg, & Öhman, 1994; Lang, Bradley, & Cuthbert, 1997; Vuilleumier, et al., 2001; Vuilleumier & Pourtois, 2007) and, as a consequence, a prior entry effect for this class of stimuli could be observed. More specifically, emotion (similarly to attention) is thought to “accelerate” sensory processing (see also Droit-Volet & Meck, 2007). This prediction was at the core of the experimental work reported in the first empirical part of this thesis (Chapter 2).

In a series of five successive experiments, I sought to investigate whether negative emotional facial expressions (i.e., angry or fearful faces), when competing for attention selection with neutral faces, readily capture attention control processes in a genuine bottom-up way or not, as revealed by a prior entry effect for these former face stimuli during a standard TOJ. I used standard face stimuli previously validated in the literature (Ekman & Friesen, 1976) and that were consistently shown to bias, in a bottom-up fashion, attention control brain processes in previous studies based on different tasks (Esteves, et al., 1994; Pourtois, et al., 2012; Vuilleumier, 2005; Vuilleumier, et al., 2001). To ascertain that the emotional content of the stimuli was the main factor influencing perceptual judgments, on each trial I presented pairs of faces that were always visually dissimilar, both in terms of identity and facial expression. The idea behind this experimental manipulation was that visual dissimilarity between the two faces was randomized across trials, thereby making it unlikely that dissimilarity alone could strategically be used by participants to perform the task (Duncan & Humphreys, 1989). In addition, at the end of each experiment I asked participants to rate the emotional intensity of the face stimuli, to provide evidence that the (negative and neutral) facial expressions were recognized “explicitly” in accordance with the normative ratings (Ekman & Friesen, 1976). The five experiments reported in Chapter 2 were aimed at: (i) corroborating and extending previous results (Fecica & Stolz, 2008; West, et al., 2009) showing prior entry effects for emotional relative to neutral faces, indicating early attentional capture for this class of stimuli; (ii) comparing the magnitude of the prior entry effect for fearful vs. angry faces, given that early attention allocation could likely be influenced by the perceived threat value conveyed by the face (Anderson, Christoff, Panitz, De Rosa, & Gabrieli, 2003; Cristinzio, N'Diaye, Seeck, Vuilleumier, & Sander, 2010); (iii) disentangling engagement vs. disengagement effects during the “automatic” orienting of attention towards negative emotional stimuli (Koster, Crombez, Van Damme, Verschuere, & De Houwer, 2004; Salemink, van den Hout, & Kindt, 2007) by means of a variant of the of the TOJ task where the processing of the offset of the two faces was emphasized, as opposed to their onset (see Chapter 2 for details).

3. Expectation-driven influences on visual perception: the role of emotion

A vast majority of studies and models reviewed above has primarily focused on “magical” (in the sense of “automatic”, “irrepressible”, “compelling”) bottom-up influences of emotion on (visual) perception, with the key assumption that specific attention processes somehow reflexively mediate these influences. However, the evidence for such an assumption is rather mixed (see Pourtois, et al., 2012, for a recent review). As outlined above, attention is essentially multifaceted and it comprises not only bottom-up guidance mechanisms, but also top-down factors, including prior knowledge, expectations and current goals, which also substantially contribute to selection and allocation processes during (early) stimulus processing (Bar, 2009b; Summerfield & Egner, 2009). In the second part of my dissertation, I concentrated on these elements and designed studies aimed at addressing the question whether emotion could reliably bias expectation processes generated “proactively” during visual perception, with a focus on emotional scene perception (see Chapters 3-5). Before I turn to the short presentation of these studies (and the link to emotion), I will first review existing models which have dealt with proactive effects of expectations on visual object (and scene) recognition, when these stimuli were primarily “neutral” (i.e., devoid of any motivational or emotional value for the participant). In this framework, a key assumption has been made regarding the contribution of diagnostic low spatial frequency information during the proactive generation of guesses or predictions regarding the actual content of the visual stimulus.

3.1. Beyond bottom-up hierarchical visual processing during object and scene recognition

Visual object recognition has classically been conceived as a serial process in which visual stimuli, detected by the eye’s retina and conveyed through the lateral geniculate nuclei, are processed in the primary visual cortex (V1) and, from there, undergo further processing in hierarchically higher cortical regions located primarily in the infero-temporal cortex (Felleman & Van Essen, 1991; Marr, 1982; Palmer, 1999). During visual processing, two main bottom-up pathways have been identified (Biederman, 1987; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982): (i) a dorsal pathway (from V1 to posterior parietal regions, such as motion-sensitive V5/MT) concerned with spatial properties of vision (answering the question “where?”); (ii) a ventral pathway (from V1 to color-sensitive V4 within the infero-temporal cortex) dedicated to the identification of visual objects (answering the question “what?”) (Farivar, 2009; Goodale & Milner, 1992). These hierarchical models, however, mainly emphasize the predominant role of *bottom-up* processes leading to visual object (or scene) recognition, while they typically do not emphasize *top-down* modulatory influences on these bottom-up processing streams, including attention (Blair, Watson, Walshe, & Maj, 2009; Treisman & Kanwisher, 1998), task-relevance of specific visual features (Egner & Hirsch, 2005), the prior probability of the occurrence of an event (Summerfield & Egner, 2009), image storage and retrieval in working memory (Ranganath, Cohen, Dam, & D'Esposito, 2004), the proactive use of contextual information (Bar, 2004; Oliva & Torralba, 2007), and the online monitoring of the outcome of a decision (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuiss, 2004; Ullsperger, Volz, & von Cramon, 2004). Evidence had accumulated during the last decade to show that each of these top-down sources could substantially bias sensory processing either in the “what” or “where” brain pathway, sometimes at very early latencies following stimulus onset (Thorpe, Fize, & Marlot, 1996), and in early visual areas, including V1 (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Smith & Muckli, 2010). Thus, the rapid and “effortful” bottom-up processing of visual input in low-level cortical areas can be assisted by the use of abstract information extracted in frontal or anterior temporal brain regions during stimulus processing (or even prior to its onset), which eventually constrains the number and magnitude of computations made in these early visual areas (Bar, et al., 2006; Ganis, Schendan, & Kosslyn, 2007).

3.2. Predictive coding models of visual perception

Recent theoretical accounts try to better model the role of expectations in visual object (or scene) recognition (Bar, 2009b; Kersten, Mamassian, & Yuille, 2004; Summerfield & Egner, 2009; Summerfield & Koechlin, 2008; Yuille & Kersten, 2006). In contrast to a phenomenology mainly based on reactive brain events time-locked to the bottom-up processing of specific features, these models highlight the proactive nature of human visual perception: the analysis of the surrounding environment is not only based on incoming perceptual input, but it is linked to previously encountered, familiar information (Bar, 2007; Kourtzi & Connor, 2011). In other words, while the visual input travels through feedforward routes from lower to higher level cortical regions, high level areas are not silent, but they quickly extract and evaluate the gist of the scene based on a coarse analysis of the stimulus, eventually integrating the outputs of bottom-up “hierarchies” and top-down “reverse hierarchies” to achieve recognition (Hochstein & Ahissar, 2002).

Recently, many theoretical efforts have been made in the literature in order to best describe how the visual input may quickly propagate through successive stages of a hierarchically organized cortical architecture to eventually be processed and reach (visual) awareness (Mumford, 1992; Ullman, 1995). According to these *predictive coding* models (Enns & Lleras, 2008; Friston, 2003, 2005; Grossberg, 2009; Rao & Ballard, 1999; Spratling, 2008), predictions based on prior experience are projected from higher level to lower lever brain areas, with the aim to facilitate recognition by limiting the number of candidate representations of an object that need to be considered (see also Bar, 2003). When a discrepancy between such top-down predictions and the visual input is detected, an error signal is generated and propagated back to higher-level regions, where a new prediction, updated by the error signal, is produced (Friston, 2005; Rao & Ballard, 1999). Expected information and bottom-up sensory input are thus dynamically and iteratively compared in recursive, interacting loops of activity, such that sensory input is interpreted at each cortical stage within the context of prior expectations until the visual system is able to decide on a single (or unequivocal) perceptual interpretation (Clark, 2012; Friston, 2002; Summerfield & Koechlin, 2008). From a quantitative standpoint, these expectations translate into empirical priors for a Bayesian inference process that dynamically takes place at each hierarchical level of the visual system (Friston, 2005), following the assumption that visual perception is essentially a probabilistic, inferential process (Kersten, et al., 2004; Yuille & Kersten, 2006).

3.3. Prediction-based visual perception through magnocellular vision

Interestingly, asymmetries in speed of processing and visual pathways between low spatial frequency (LSF) and high spatial frequency (HSF) information could provide a mechanistic account to explain predictive coding effects during early stages of visual object recognition. Several pioneering studies have already ascertained the differential contribution of LSF and HSF information during the processing of faces (Goffaux, et al., 2011; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Winston, Vuilleumier, & Dolan, 2003) as well as complex visual scenes (Oliva & Schyns, 1997; Schyns & Oliva, 1994; Torralba & Oliva, 2003). For example, visual perception relies on LSF information when presentation time is brief (30 ms), but when presentation time is prolonged (150 ms) HSF information is prioritized (Schyns & Oliva, 1994). In a recent study (Peyrin, et al., 2010), the authors combined functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs) to characterize the brain correlates of the prioritization of LSF over HSF information during the processing of visual input (Bullier, 2001; Hegdé, 2008). For each trial, two images were displayed for 100 ms in rapid succession, with either an LSF image followed by an HSF image (LtH sequence) or vice versa (HtL sequence), and participants had to judge whether the two scenes were from the same category (indoor, beach or city) or not. FMRI results revealed selective increases during LtH sequences in a broad network comprising occipital, frontal and temporo-parietal areas. The analysis of the electrophysiological data further specified that higher order areas in frontal and temporo-parietal regions responded more to LSF stimuli when presented first, corresponding to a “coarse-to-fine” decomposition of the stimulus, whereas occipital visual cortex responded more to HSF presented after LSF. Thus, the combination of different imaging methods revealed that the quick processing of LSF in higher order (frontal) brain regions influenced scene recognition by enhancing the processing of diagnostic HSF information in early visual cortex (Peyrin, et al., 2010).

As indirectly suggested by this study (Peyrin, et al., 2010), the differential sensory processing of LSF and HSF information could very well underlie reciprocal interaction effects between (bottom-up) sensory processing and (top-down) predictions (Bullier, 2001; Hegdé, 2008). Specifically, the coarse analysis of the visual input based on fast magnocellular vision might actually serve to generate, rapidly following stimulus onset, a restricted number of guesses or predictions about its actual content, as recently put forward by Moshe Bar (Bar, 2003, 2004, 2007, 2009b). According to this dominant model (see Figure 1), an impoverished (blurry) version of the input image, mainly composed of LSF information, is rapidly projected (via a fast dorsal magnocellular route) from early visual cortex to orbitofrontal (OFC) and parahippocampal (PHC) areas (see also Dolan, et al., 1997; George, et al., 1999). In the OFC, this coarse visual information would be sufficient to activate the most likely interpretations of the main *objects* in the scene, which are then “transferred” to the infero-temporal cortex for further verification (see also Gamond, et al., 2011). At the same time, in the PHC, this blurred image activates the most probable (experience-driven) guesses about the *context* in which the main objects are embedded, which are subsequently projected to the infero-temporal cortex for further sensory processing, including category-selective responses in the fusiform gyrus and posterior parahippocampal gyrus. The integration of the representations of the specific context and the candidate interpretations of the target object, gradually refined with details conveyed by HSF information, would eventually occur in the infero-temporal cortex, thus leading to overt recognition (Bar, 2004, 2007). A pioneering study using magnetoencephalography (MEG) and fMRI confirmed an early activation in the OFC, beginning at approximately 80 ms after stimulus onset, during the presentation of familiar objects that were presented for 63 ms and interposed between two masks (Bar, et al., 2006). Crucially, activity in this frontal region preceded the activation of brain areas in the fusiform gyrus classically involved in object recognition, lending support to the idea that the OFC might generate online predictions that were subsequently projected back to hierarchically lower areas, with the aim to limit the number of possible candidates for recognition (Bar, 2003).

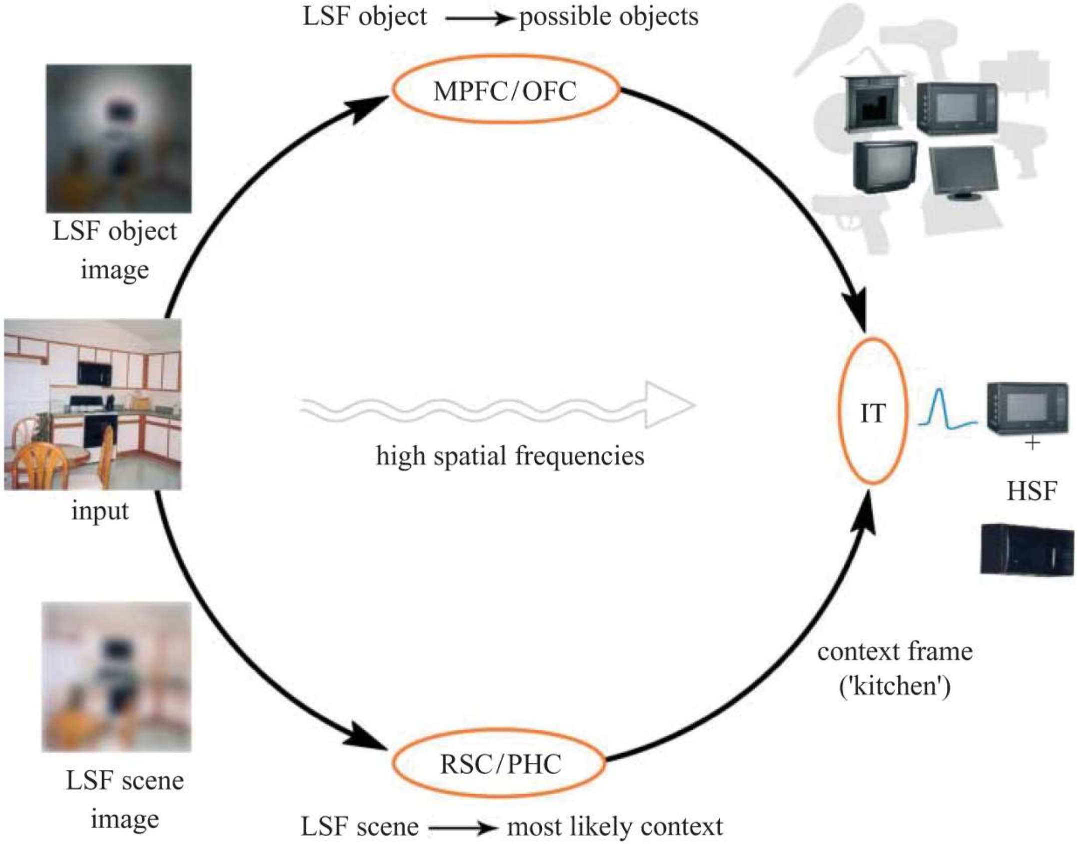


Figure 1. Moshe Bar’s model for object and scene recognition (from Bar, 2009). In parallel to the bottom-up, serial analysis of the stimulus content based on the extraction of HSF information (corresponding to the fine details of the stimulus, travelling via parvocellular pathways), LSF information (corresponding to the overall shape of the stimulus), processed via magnocellular pathways, rapidly reaches specific regions located in the medial frontal and ventral temporal areas. This coarse information serves to generate an initial prediction (“early guess”) about (i) the main object content in the scene (in MPFC and OFC), and (ii) the context in which it is likely embedded (in RSC and PHC). These predictions are later validated and refined in the IT during the processing of diagnostic HSF information. MPFC: medial prefrontal cortex; OFC: orbitofrontal cortex; RSC: retrosplenial complex; PHC: parahippocampal cortex; IT: inferior temporal cortex.

3.4. Computational modeling of visual perception guided by predictions

From a computational perspective, expectations may actually lower the threshold that needs to be overcome to make a decision in favor of one option, similar to what accumulator models of decision making propose (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). A higher probability of encountering a particular object embedded in a given context, in fact, could bias the visual system towards a more rapid and accurate accumulation of evidence in favor of a specific response, thereby resolving more quickly the conflict between similar options generated by the ambiguous visual input (Heekeren, et al., 2008; Melloni, Schwiedrzik, Muller, Rodriguez, & Singer, 2011; Philiastides, Auksztulewicz, Heekeren, & Blankenburg, 2011). According to these computational models, when a choice has to be made between two competing alternatives, two separate evidence accumulators collect sensory evidence for each response. When the amount of evidence in favor of one option reaches the response threshold, the decision corresponding to that accumulator is elicited (Figure 2) (Gold & Shadlen, 2007; Heekeren, et al., 2008). Crucially, this accumulation of evidence obeys to non-linear principles, and the competition between the two alternatives results in a progressive decay of perceptual evidence over time for the more unlikely alternative (Usher & McClelland, 2001). Importantly, the evidence accumulation process is stochastic. These random, moment-to-moment fluctuations in the amount of evidence supporting each response alternative may ultimately explain the variability in reaction times as well as the production of errors during perceptual decision making (Brown & Heathcote, 2008; Niemi & Näätänen, 1981). In addition to the accumulation of evidence process *per se*, two additional random factors need to be taken into account in these models: the start points and the drift rates. The *start points* refer to the different initial amount of evidence in favor of each response (potentially influenced by prior knowledge and expectations), whereas the *drift rates* parameterize the (fluctuating) average speed of evidence accumulation for each response (Gold & Shadlen, 2007; Usher & McClelland, 2001). This framework has initially been applied to model neurophysiological data collected in nonhuman primates (Hanes & Schall, 1996; Kim & Shadlen, 1999; Shadlen & Newsome, 2001), and later it has proven highly effective in describing human performance during recognition memory (Ratcliff, 1978), economic decisions (Sanfey, Loewenstein, McClure, & Cohen, 2006), as well as semantic and lexical tasks (Ratcliff & McKoon, 1982; Ratcliff, Thapar, Gomez, & McKoon, 2004).

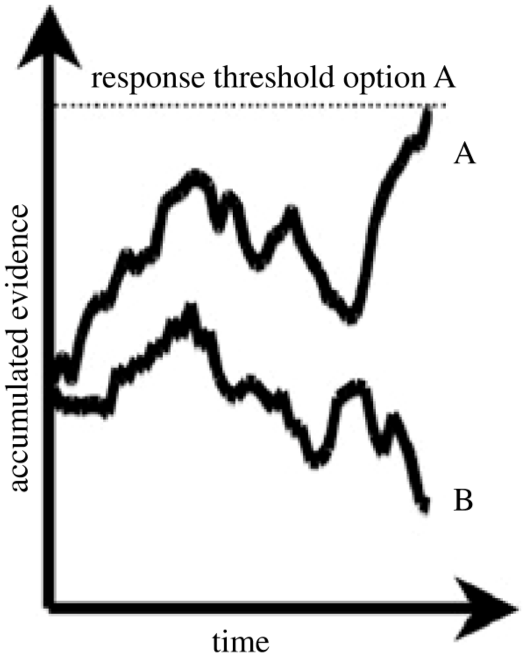


Figure 2. Usher & McClelland's model for perceptual decision making between two alternatives (adapted from Brown & Heathcote, 2008). When a choice has to be made between two competing alternatives, two separate evidence accumulators (A and B in the figure) collect sensory evidence for each response. Response A is chosen when the evidence in its favor surpasses the threshold (dotted horizontal line).

An increasing number of behavioral and neuroimaging studies have attempted to better characterize the processes and identify the neural correlates of perceptual evidence accumulation during a wide range of perceptual decision making tasks (Heekeren, Marrett, Bandettini, & Ungerleider, 2004; Melloni, et al., 2011; Philiastides, et al., 2011; Philiastides & Sajda, 2007)[1]. In one of these earlier studies, Heekeren and colleagues (2004) used fMRI to investigate perceptual decision making using a face-house categorization task. Participants were presented with images that were either difficult (i.e., perceptual noise was added in order to make the recognition more demanding) or easy, and the task was to identify whether the compound object was a face or a house. Results showed greater BOLD response in the fusiform face area (FFA) during trials showing clear images of faces compared to their degraded (noisy) version, in accordance with the notion that this region preferentially responds to face stimuli than other visual objects or categories (Kanwisher, McDermott, & Chun, 1997; Kanwisher & Yovel, 2006). A similar pattern was observed in the parahippocampal place area (PPA) during the processing of houses, confirming the pivotal role of this region in the perception and analysis of buildings and other objects (or scenes) that can serve as landmarks (Aguirre, Zarahn, & D'Esposito, 1998; Epstein & Kanwisher, 1998). This result is also consistent with the idea that, in these category-selective infero-

---

[1] These tasks include, for instance, the identification of an object identity based on the processing of fragmented outlines (e.g., Torfs, Panis, & Wagemans, 2010), decreased perceptual noise (e.g., James, Humphrey, Gati, Menon, & Goodale, 2000), or increased high spatial frequency information (e.g., Pourtois, Dan, et al., 2005).

temporal brain areas, the evidence for the two object categories accrues differently as a function of the signal-to-noise ratio of the stimuli: the processing of degraded objects took longer to reach the response threshold and, as a consequence, reaction times increased. More interestingly, however, the left posterior dorsolateral prefrontal cortex (DLPFC) responded more to clear compared to degraded stimuli, and its activity correlated with the difference between the output signals of FFA and PPA. Thus, the left DLPFC could serve as an “integration hub”, where sensory information coming from low-level areas is analyzed and compared in order to compute a perceptual decision (see also Heekeren, et al., 2008).

Other fMRI studies using paradigms based on a progressive and gradual revelation of the stimulus content have also shed light on brain systems involved in the online generation of predictions during perceptual decision making (Carlson, Grol, & Verstraten, 2006; James, et al., 2000; Ploran, et al., 2007; Ploran, Tremel, Nelson, & Wheeler, 2011; Wheeler, Petersen, Nelson, Ploran, & Velanova, 2008). Among them, Ploran, et al. (2007) devised a task in which masked picture stimuli were gradually revealed over eight discrete steps in a maximum revelation time of 14 seconds, and participants had to press a button as soon as they could identify the picture with a reasonable degree of confidence. Different brain areas showed distinct patterns of neural response. In several posterior occipital regions, activity increased monotonically as a function of the amount of visual information aggregated by the visual system (sensory processors). Conversely, in inferior temporal, frontal and parietal regions (including the bilateral fusiform gyrus and the DLPFC), a gradual increase in activity peaking at the time of recognition was compatible with the assumption of a genuine accumulation of evidence process taking place in these regions, that was presumably necessary to recognize the identity of the target object (accumulators). Finally, in several regions of medial frontal cortex -- including the dorsal anterior cingulate cortex (dACC) and the anterior insula -- activity remained near baseline until the time of recognition when a substantial increase in the BOLD signal occurred, suggesting their selective involvement in decision-related, as opposed to evidence accumulation processes (Ploran, et al., 2007; see also Ploran, et al., 2011).

The aforementioned brain imaging studies extended earlier knowledge on mechanisms of perceptual decision making, providing evidence for the idea that accumulation and moment-of-recognition processes could be subtended by non-overlapping brain networks (Heekeren, et al., 2004; Ploran, et al., 2007; Ploran, et al., 2011). However, the progressive stimulus revelation procedure used in these studies were adapted to be compatible with the sluggish temporal resolution of fMRI, and only a compound measure of brain activity across many events during a given trial could eventually be modeled with this haemodynamic imaging technique. EEG (and local field potentials), on the other hand, offers the advantage to track brain activity using a millisecond temporal resolution (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). A recent study (Philiastides & Sajda, 2006) used single-trial EEG analysis to identify the neural correlates of decision making during a face-car discrimination task similar to the face-house task described above. Two distinct ERP components were found to maximally discriminate between faces and cars: (i) an early negative component, occurring between 130 and 200 ms after stimulus onset (consistent with the well-known face-selective N170 component; Bentin, Allison, Puce, Perez, & McCarthy, 1996; Rossion, Joyce, Cottrell, & Tarr, 2003); (ii) a late component, peaking between 300 and 450 ms post-stimulus onset. Interestingly, the N170 was not affected by the amount of noise in the stimulus, whereas the subsequent ERP component correlated with task difficulty (i.e., signal-to-noise ratio) and change in response time. These findings were interpreted as an indication of accumulation processes that were dependent on the strength of perceptual evidence provided by the environment (Philiastides & Sajda, 2006). In another recent study (Jemel, Schuller, & Goffaux, 2010), the authors characterized the spatiotemporal dynamics of electrical brain activity prior to recognition of famous faces, in order to investigate whether face recognition processes (taking place in the infero-temporal cortex) showed a categorical as opposed to a more gradual effect. Famous and unfamiliar faces were progressively revealed by parametrically decreasing, in a stepwise fashion, the amount of perceptual noise in subsequent image frames (see also James, et al., 2000, for a similar procedure applied to complex visual scenes), and participants were required to perform familiarity judgments after each image level. Electrophysiological results revealed both gradual and categorical processes at work during overt recognition of familiar faces: whereas early and mid-latency ERP components (i.e., N170 and N250) showed an abrupt activity change at the moment of recognition, the N400 as well as the late positive component (LPC) exhibited an incremental increase in amplitude when approaching to actual time of recognition. These results are consistent with the fMRI results described above (Carlson, et al., 2006; James, et al., 2000; Ploran, et al., 2007), in that they suggest two distinct patterns of activation (gradual evidence accumulation vs. categorical, all-or-nothing activity) likely subtended by non-overlapping neural networks.

3.5. *Emotion-based predictions during visual perception*

Predictive coding and accumulator models of visual object or scene recognition provide versatile functional accounts to describe and characterize proactive processes at work prior to overt recognition, when these processes operate primarily on “neutral” content (i.e., devoid of any emotional or motivational value). Therefore, one of the central research questions addressed in this dissertation was to explore whether, when encountering emotion-laden (as opposed to neutral) scenes, systematic changes in the extent and speed of proactive guesses during sensory processing may be revealed. As reviewed here above (see section 2), emotion can have profound influences during (early) stimulus processing, and hence it can bias visual perception, presumably because dedicated attention control processes are “automatically” recruited upon encounter of motivationally significant events or stimuli (Bocanegra & Zeelenberg, 2009; Öhman, Flykt, et al., 2001; Phelps, et al., 2006; Vuilleumier, 2005). Whereas these effects inform about the propensity of emotion to reactively influence and guide perception (once the stimulus has been processed, at least to some extent), the question arises as to whether emotion can also dynamically and proactively alter the generation of predictions during early sensory processing (i.e., before the actual content of the object or scene is fully disclosed), eventually influencing visual perception.

In Chapter 3, I explored the electrophysiological correlates of upstream processes leading to visual scene recognition with the use of a novel experimental paradigm called *progressive unfolding task*. In this paradigm, the content of either emotional (i.e., pleasant and unpleasant) or neutral scenes is progressively revealed, using a six-step procedure, until participants were able to perform an animacy judgment task. Each new step progressively added a constant amount of LSF and HSF information, starting from a initial blurry and undistinguishable stimulus up to a fully resolved and meaningful picture. While participants performed this task, high density EEG (128 channels) was recorded in order to gain insight into the exact spatio-temporal dynamics associated with the online generation of predictions during sensory processing regarding the most likely interpretation of the retinal input. Critically, the actual moment of recognition was identified for each and every trial and, based on this specific time point, I looked “backward” at electrical brain activity generated prior to overt recognition, with the aim to track distinctive neural events prior to recognition presumably corresponding to the online generation of predictions.

In Chapter 4, I modified this progressive unfolding task in an attempt to better constrain the generation and use of specific predictions regarding the identity of the

scenes gradually revealed. More specifically, in Chapter 3 no specific constrain or prior information was given to participants regarding the likely identity of the visual scenes progressively revealed during the task. As a consequence, a bias towards the preferential processing of perceptual cues (hence with little control over the actual content and number of predictions generated “online” by the participants) was probably inadvertently introduced. To overcome this limitation, in Chapter 4 I introduced a standard memory manipulation (Goldman-Rakic, 1990; Rugg, et al., 1998). Specifically, on each trial, participants were first required to encode and store in short-term memory the content of a complex visual scene (carrying either an emotional or a neutral meaning), before the content of the same or a new scene was progressively revealed (similarly to the procedure employed in Chapter 3) and a delayed match-to-sample task was required. Again, high density EEG was recorded throughout the experiment, allowing me to investigate the precise spatio-temporal dynamics involved in the generation of predictions during scene perception, as well as its systematic influence by the emotional content of the scene.

Finally, in Chapter 5 I ran additional experiments in order to verify that the results obtained in Chapters 3 and 4 could not solely be explained by uncontrolled variations along non-emotional factors, with a focus on visual complexity and familiarity. *Picture complexity* is commonly defined as the extent to which an object can be segregated from its background (Bradley, Hamby, Low, & Lang, 2007). Since figure-ground segregation (and more generally perceptual grouping) is a fundamental and early perceptual process during visual perception (Albright & Stoner, 2002; Craft, Schutze, Niebur, & von der Heydt, 2007; Wagemans, et al., 2012), scenes having a clear figure-ground segregation might artificially lead to faster and easier recognition than more complex, perceptually “noisy” scenes, especially during impoverished viewing conditions (see Chapters 3-4). Accordingly, in Chapter 5 I aimed at verifying whether emotional and neutral scenes shared similar picture complexity or not, and how this variable could eventually influence, on a trial-by-trial basis, overt recognition during the progressive unfolding task described in Chapter 4. Likewise, the degree of *familiarity*, defined here as the frequency of experience associated with a given stimulus content (Libkuman, Otam, Kern, Viger, & Novak, 2007), also likely contributes to the speed and accuracy of visual scene recognition. Previous research has already reported that novelty can engage the same cardiovascular systems as valence and arousal (Mendes, Blascovich, Hunter, Lickel, & Jost, 2007). Furthermore, novel scenes elicit startle reflex similarly to emotional pictures (Bradley, Lang, & Cuthbert, 1993). Moreover, neuroimaging studies have found enhanced activity to novel stimuli in several affective brain areas, such as the amygdala (Schwartz, Wright, Shin, Kagan, & Rauch, 2003), OFC and dACC (Weierich, Wright, Negreira, Dickerson, & Barrett, 2010), providing further evidence for the assumption that novelty provides a significant stimulus dimension that can interact with (or even mimic) specific emotional effects (see also Ranganath & Rainer, 2003). Therefore, in Chapter 5 I also looked at the possible confounding effect of familiarity during recognition of emotional vs. neutral scenes, in particular when proactive guesses were made by the participants during the resolution of the complex progressive unfolding task.

4. Methodological considerations

4.1. Temporal order judgment and point of subjective simultaneity

In Chapter 2, following standard practice, I calculated each participant’s point of subjective simultaneity (*PSS*) as a reliable estimate of prior entry, which is indicative of early attentional capture towards one of the two face stimuli in the pair. The PSS indicates the time interval needed for the participant to perceive the two stimuli as arriving simultaneously or, in other words, the SOA between two stimuli at which observers report maximal uncertainty (Schneider & Bavelier, 2003; Shore, Spence, & Klein, 2001; Spence & Parise, 2010). To obtain the PSS, I first calculated the proportion of “right first” responses for each condition and SOA. I then applied the inverse of the standard normal distribution function to these scores (*probit* analysis; Finney, 1964). The transformed *z*-scores of the proportion of “right first” responses were finally included in a linear regression to derive the PSS, calculated from the slope and intercept of the best-fitted line of the *z*-scores (PSS = -slope/intercept). To account for the correlation of measurements, each participant had its own intercept and slope with estimated random effects from a bivariate zero mean normal distribution.

4.2. Progressive unfolding task and mixed proportional odds model

In Chapters 3-5, I used a complex data analysis to assess whether emotion reliably influenced, at the behavioral level, accumulation of evidence processes leading to visual scene recognition. Because the six steps used to progressively reveal the content of the scenes were not independent with one another, the data were first expressed as the cumulative percentage of correctly recognized pictures, from the first, blurred image until the sixth, intact scene. This resulted in an inverted s-shaped psychometric function, computed separately for each emotion condition (i.e., pleasant, neutral, and unpleasant scenes). Next, to verify whether these psychometric curves shifted as a function of the emotional content of the scenes (thereby indicating slower or faster accumulation of evidence for pleasant or unpleasant compared to neutral scenes), I used a *proportional odds model* (Agresti, 2007). This regression analysis for ordinal dependent variables (here applied on the single-trial data), which estimates correct recognition at each image level, enabled me to model the cumulative probability up to and including recognition from image level k (k = 1, ?, 5). Specifically, if one defines Y (Y=1,...,6) as the level of recognition at which a correct response is given, the cumulative probability of Y at the k-th level of recognition is $\Pr(Y \leq k)$ (note that, since all participants eventually give a correct response at the end of the image sequence, $\Pr(Y \leq 6) = 100\%$). Because it is mathematically more convenient to model the odds (i.e., the ratio of the probability that recognition will occur to the probability that it will not happen), it follows that:

$$\frac{\Pr(Y \leq k)}{1 - \Pr(Y \leq k)}$$

The derived odds ratio expresses how much the odds of recognition from image level k or earlier is increased (if larger than 1) or decreased (if smaller than 1) across the different experimental conditions, and thus it provides a single number capturing the shift in psychometric curve. Of note, this model has to take into account the fact that several observations are collected from different participants, and measurements from the same participants are expected to be correlated (De Boeck, 2008). Therefore, to account for dependencies of trials within the same participant, a multi-level version of the proportional odds model was used, in which subject-specific random effects were assumed.

4.3. ERP topographic mapping analyses

In Chapters 3-4, high density EEG was recorded by means of 128 (active) electrodes evenly covering the scalp surface in agreement with the five percent electrode system (Oostenveld & Praamstra, 2001). To fully characterize differences in electrical brain activity between experimental conditions while minimizing the number of priors, I opted for the use of reference-free topographical analyses (Lehmann & Skrandies, 1980; Michel & Murray, 2012; Michel, Seeck, & Landis, 1999; Michel, et al., 2001; Murray, Brunet, & Michel, 2008; Pourtois, Delplanque, Michel, & Vuilleumier, 2008) that depart from more conventional peak analyses in ERP research (Picton, et al., 2000). This data-driven clustering method allows to summarize and identify, within the fast evolving ERP signal, the dominant spatial configurations of the global electric field measured for all 128 channels concurrently (i.e., topographical maps), which is time-locked and phase-locked to the onset of the stimulus. This data analysis method heavily rests on the assumption that different topographic maps necessarily denote the activation of different brain generators, as previously demonstrated based on simulations and mathematical reconstructions (Brandeis & Lehmann, 1986; Lehmann & Skrandies, 1980; Michel, et al., 1999). This data-driven clustering method provides at least three advantages compared to standard peak measurements (Picton, et al., 2000), which assume many priors regarding where and when (ERP) events of interest should take place following stimulus onset (Makeig, Debener, Onton, & Delorme, 2004). First, the amplitude and actual morphology of the so-called ERP components heavily depend upon the reference montage that is used by the experimenter (see Murray, et al., 2008 for direct evidence). This has important implications for the interpretation of ERP components as well as their comparison across studies, because different reference montages (e.g., average reference, tip of the nose, or linked mastoids) are often used, leading to profound changes in the actual expression (e.g., amplitude) of these ERP components. By contrast, the resulting geometry or spatial configuration of the electric field is reference-free, in the sense that it does not vary depending on local (or global) amplitude variations (Michel & Murray, 2012; Pourtois, et al., 2008). Second, the *a priori* selection of peaks and/or troughs in the ERP waveform typically overlooks low signal amplitude information, albeit potentially indicative of important neural events, given the intrinsic oscillatory nature of the recorded ERP signal (Murray, et al., 2008). Reference-free topographic analyses, by comparison, allow to cluster dominant topographies embedded in the continuous (grand average) ERP signal, regardless of (local or global) changes in amplitude. As a result, not only the heuristic, but also the overall interpretability of the observed ERP effects is substantially increased with this analysis, compared to standard peak measurements (Pourtois, et al., 2008). As pointed out above, because changes in the configuration of the active cerebral sources necessarily translate into changes in the scalp topographies (though the reverse inference is not necessarily true from a mathematical standpoint, an element that further complicates the well-known inverse problem in EEG/ERP research; see Scherg & Voncramon, 1986), data analysis techniques that capitalize on these spatial properties allow to draw (indirect) inferences about the underlying configuration of neural generators at a given moment in time, and how it can vary depending on the actual experimental condition (Michel & Murray, 2012). Third, using ERP topographic mapping methods, one can more easily perform source imaging than with a standard peak analysis -- given that the dominant maps are reflecting different configurations of the underlying brain generators --, and hence gain insight into the putative neural networks involved at specific latency (following stimulus onset) and/or conditions (Michel & Murray, 2012; Michel, et al., 2004).

4.4. Source localization

Once the dominant scalp topographies embedded in the ERP data set have been identified by the analysis outlined above (see also Chapters 3-4), I used a specific source localization algorithm in order to gain insight into the likely location and distribution of the intracranial generators underlying these dominant topographical maps. For this purpose, I used standardized low resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002), a standard linear distributed model for the mathematical reconstruction and statistical testing of the (possible) cortical intracerebral generators. Even though it remains mathematically impossible to solve the inverse problem (i.e., activity recorded on the scalp can be virtually generated by an infinite number of dipoles; Luck, 2005; Plonsey, 1963), the localization error of this algorithm can be minimized and some simulations even showed that, on average, a 1-voxel resolution (keeping in mind that voxels/nodes here are 6 mm isometric) could be obtained, which is remarkably better than other inverse solution methods, including single (regional) dipoles (Grech, et al., 2008; Pascual-Marqui, 1999). In short, this (non-parametric) distributed inverse solution is based on the neurophysiological assumption of coherent coactivation of highly synchronized adjacent cortical areas (Silva, Amitai, & Connors, 1991), and is therefore suitable to compute multiple simultaneously active sources without any *a priori* assumption on the number and position of the underlying cortical generators.

CHAPTER 2

No prior entry for negative emotional facial expressions: evidence from temporal order judgments[1]

Previous research has shown that threat-related faces, due to their intrinsic motivational relevance, capture attention more readily than neutral faces. Here we used a standard temporal order judgment (TOJ) task to assess whether negative (either angry or fearful) emotional faces, when competing with neutral faces for attention selection, may lead to a prior entry effect and hence are perceived as appearing first, especially when uncertainty is high regarding the order of the two onsets. However, we failed to find evidence for this conjecture across five different experiments, despite the fact that participants were invariably influenced by asynchronies in the respective onsets (or offsets) of the two competing faces in the pair, and could reliably identify the emotion in the faces. Importantly, by systematically varying task demands across experiments, we could rule out confounds related to suboptimal stimulus presentation, inappropriate task demands, or the presence of a confounding disengagement effect away from threat. These findings challenge the notion of an early automatic capture of attention by (negative) emotion. We argue that non-emotional, motion-related cues can probably be used implicitly by participants during the TOJ task to decide which of the two competing faces appeared first. In this model, the lack of systematic bias of attention by emotion is imputed to the primacy of a non-emotional cue to resolve the TOJ task, which in turn prevents negative emotion to exert an early bottom-up influence on the guidance of spatial and temporal attention.

---

[1] Schettino, A., Loeys, T., & Pourtois, G. Manuscript in preparation.

1. Introduction

Results obtained from a variety of experimental paradigms suggest that, under specific circumstances, negative emotional stimuli may receive prioritized access to awareness by biasing perceptual and attentional processes (Bradley, 2009; Compton, 2003; Öhman, Flykt, et al., 2001; Pourtois, et al., 2012; Vuilleumier, 2005). In variants of the Stroop task, for instance, naming the color of a word is slower when the stimulus carries a negative emotional meaning, even though this semantic feature is task-irrelevant (Bar-Haim, et al., 2007; Phaf & Kan, 2007; Yiend, 2010). Similarly, in visual search tasks participants are usually faster at detecting negative emotional targets embedded in an array of neutral distracters (Eastwood, et al., 2001; Öhman, Flykt, et al., 2001; Olatunji, Ciesielski, Armstrong, & Zald, 2011). Furthermore, the well-known deficit in perceiving the second of two targets presented rapidly one after another among a stream of distracter items (*attentional blink*; Broadbent & Broadbent, 1987; Shapiro, et al., 1997) is reduced if the second target carries a negative emotional meaning (Keil, et al., 2006; Maratos, 2011; Schwabe, et al., 2011), or prolonged if the first target is (highly) arousing (Anderson, 2005; Anderson & Phelps, 2001; Arnell, et al., 2007; Mathewson, et al., 2008). Finally, studies using spatial cueing tasks have shown that emotion-laden stimuli facilitate the processing of (non-emotional) targets subsequently presented at the same location, consistent with the assumption of a rapid orienting of attention towards these (task-irrelevant) stimuli, as opposed to neutral ones (Armony & Dolan, 2002; Bocanegra & Zeelenberg, 2009, 2011b; Fox, et al., 2001; Mogg & Bradley, 1999; Most & Wang, 2011; Phelps, et al., 2006; Pourtois, et al., 2004).

Taken together, these findings suggest that motivationally relevant stimuli (including negative facial expressions) can exert a strong modulatory influence on attentional control processes. However, still little is known about how these stimuli are *initially* prioritized by dedicated attentional control systems, mainly because the initial attentional deployment (or orienting) was not directly measured in these earlier studies. Visual search, spatial cueing, or attentional blink tasks, in fact, require participants to quickly engage, disengage and reallocate attention concurrently towards different competing stimuli in order to resolve the task. Therefore, these paradigms are not suited to titrate changes in the initial allocation of attention towards emotional vs. neutral stimuli (Horstmann, 2007; West, et al., 2009). By contrast, temporal order judgment (TOJ) tasks provide a more direct, sensitive and accurate measure of attentional capture (Jaskowski, 1993; Stelmach & Herdman, 1991; Stelmach, et al., 1994; Titchener, 1908). In a typical TOJ task, attention is oriented either to the left or the right side of fixation, and participants have to judge which of two competing stimuli, displayed on the left and right at various stimulus onset asynchronies (SOAs), was presented first. Because attention accelerates sensory processing (Desimone & Duncan, 1995; Serences & Yantis, 2006), the stimulus appearing on the attended location is processed faster and, as a consequence, its onset is perceived as taking place first (*visual prior entry*; Schneider & Bavelier, 2003; Spence & Parise, 2010; Spence, et al., 2001). In this study, we capitalized on these well-established functional properties to assess whether intrinsic negative emotional stimuli might lead to a genuine prior entry effect or not, when they compete for attention selection with neutral visual stimuli. We opted for the use of fearful and angry facial expressions, given their enhanced intrinsic biological relevance (Anderson, et al., 2003; Whalen, et al., 1998). Moreover, converging evidence from earlier studies shows that, across various task settings and experimental conditions, fearful and angry facial expressions have the propensity to reliably bias attention control and early perceptual processes (Esteves, et al., 1994; Lang, et al., 1997; Vuilleumier, et al., 2001; Vuilleumier & Pourtois, 2007).

To date, two earlier studies used TOJ tasks to assess whether (negative) emotional faces could lead to a prior entry effect when competing with neutral faces or not. In their study, Fecica & Stolz (2008) presented schematic neutral, happy or angry faces -- separated by SOAs of 0, 17, 34, or 100 ms -- on the left and right side of fixation, and asked participants to decide which of the two competing face stimuli appeared first. Results showed that, in conditions of high uncertainty (i.e., at short as opposed to long SOAs), happy and angry faces were consistently perceived as appearing first compared to neutral faces. Moreover and unexpectedly, a stronger prior entry effect was observed for happy relative to angry faces. This latter result is at variance with the well-known *negativity bias* for threatening stimuli (Baumeister, et al., 2001; Eastwood, et al., 2001; Fox, et al., 2001; Öhman, Lundqvist, et al., 2001), and might ultimately be explained, at least in part, by the use of a small number of stimuli (i.e., three schematic faces repeated many times across trials) which may have introduced systematic attentional biases based on the (fast) processing of specific low-level perceptual features (Mak-Fan, Thompson, & Green, 2011; Wolfe & Horowitz, 2004).

West, et al. (2009) conducted several experiments using the TOJ task to investigate whether motivationally significant stimuli could be prioritized over neutral ones. First, they reported a prior entry effect for schematic upright (neutral) faces when competing with inverted schematic faces, providing evidence for a bias in the early allocation of attention towards these biologically relevant stimuli. Moreover, they found that this initial attentional deployment was influenced by the emotional content of the faces (i.e., schematic angry faces were prioritized over neutral faces), and was further enhanced by the use of realistic photographs of angry faces. However, in this study alike, a limited number of face stimuli was used (i.e., four angry and four neutral face identities). Therefore, it remains unclear whether the negative emotional facial expression *per se*, or rather uncontrolled perceptual factors, led to a differential early allocation of attention towards these emotion-laden stimuli.

In our study we used a large set of realistic photographs of faces (extensively validated in the literature) and assessed whether negative emotional facial expressions could lead to a prior entry effect when competing with neutral faces. Importantly, to overcome any low-level perceptual confound, on each and every trial we presented participants with a pair of faces (with a variable SOA between the two onsets) that were always visually dissimilar, both in terms of identity and facial expression (i.e., either neutral or emotional). The rationale of this manipulation is that, across trials, visual dissimilarity between the two competing faces is always present and variable -- and thus uninformative -- and, accordingly, it cannot implicitly be used by participants as a distinctive visual cue to decide which of the two faces appeared first (Duncan & Humphreys, 1989; see also Pourtois, et al., 2004; Vuilleumier, et al., 2001). In these conditions, presumably, only the differential emotional content of the face would influence perceptual judgments. Furthermore, to verify that the emotional facial expressions were recognized as such, at the end of the experiment we asked participants to rate the emotion intensity of each and every face stimulus used during the main TOJ task. The goal of our study was threefold. First, we evaluated if negative emotional faces were processed faster than neutral faces (see Pourtois, et al., 2012), thereby showing prior entry consistent with the assumption of early attentional capture. Second, we compared the size of the prior entry effect for fearful (indirect threat) vs. angry faces (direct threat), bearing in mind that the capture of attention by negative emotion could very well depend on the perceived relevance of the threat displayed in the face (Cristinzio, et al., 2010; Sander, Grandjean, Kaiser, Wehrle, & Scherer, 2007; Whalen, et al., 2004). Finally, considering previous studies showing that attentional biases for threat-related information might result from a difficulty to disengage attention away from threat (i.e., holding effect rather than capture; see (Fox, et al., 2001; Koster, et al., 2006; Salemink, et al., 2007)), we devised a variant of the TOJ task that was suited to disentangle possible engagement from disengagement effects driven by the emotional content of the faces. In this version of the task, the onset of the two faces in the pair was identical but their offset was manipulated as a function of a variable SOA. Participants were asked to judge which of the two faces *disappeared* first. We hypothesized that, if the enhanced attention allocation towards negative emotional expressions was due to a difficulty to disengage attention away from these emotion-laden stimuli, in conditions of high uncertainty (i.e., at short SOAs), participants would perceive negative emotional faces as disappearing the latest on the screen.

2. Experiment 1

2.1. Participants

Thirty-seven psychology students of Ghent University participated in the study in exchange of course credits. All volunteers were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders. The data of five participants were excluded from subsequent analyses due to abnormal psychometric functions in at least one experimental condition (Perez, et al., 2009; Spence, et al., 2001; Weiß & Scharlau, 2012), suggesting that (in this specific condition) their performance was not influenced by the main SOA manipulation (see below). Thus, the final sample consisted of 32 participants (27 women, mean age 19 years, range 18-22).

The study was approved by the local ethics committee of the Faculty of Psychology and Educational Sciences, Ghent University. All volunteers gave informed written consent prior to their participation.

2.2. Stimuli

We used pairs of grayscale photographs of ten different individuals (four women) selected from the standardized Ekman database (Ekman & Friesen, 1976). In order to remove most of the external facial features (e.g., hair and ears) and to standardize the spatial layout occupied by each face, each stimulus was enclosed in an oval frame encompassing 8.86° x 7.63° of visual angle (Figure 1; for a similar procedure, see also Pourtois, et al., 2004). Means and standard deviations of pixel luminance were extracted using ImageJ (v1.44; http://rsb.info.nih.gov/ij/), and apparent contrast, defined as the standard deviation divided by the mean, was calculated for each and every face stimulus. Independent samples t-tests revealed that neutral and fearful faces did not differ with regard to apparent contrast [$t(18) = -0.65$, $p = .523$].

2.3. *Procedure*

The experiment was conducted in a small, dimly lit room on a PC connected to a 19" CRT monitor (refresh rate: 100 Hz) running E-Prime 2.0. (http://www.pstnet.com/products/e-prime/). Viewing distance was held constant at 60 cm throughout the experimental session, with head motions restrained by a chinrest. After filling out the informed consent, participants were presented with either the onset or the offset task (see below), consisting of four blocks each (90 trials per block). Both tasks, whose presentation was counterbalanced across participants, were preceded by verbal task instructions, as well as practice blocks containing 10 trials with happy and neutral faces.

Trials of the *onset* task were structured as follows (Figure 1, left panel). A central black cross (0.96° x 0.96°) was displayed for 1000 ms on a white background. Participants were instructed to maintain fixation on this cross. Afterwards, the first face (8.86° x 7.63°) appeared in one of two placeholders located on the left or right side of fixation. After a variable time interval (SOAs: 100, 30, or 10 ms), the second face appeared on the opposite side. Both stimuli were equidistant from fixation (distance between the center of the cross and the center of the face: 10.29°). Both faces remained on the screen for 100 ms before being replaced in synchrony by a uniform mask until response. The task was to indicate, as fast and accurately as possible, the location (either left or right) of the stimulus that was perceived as appearing first (i.e., two-alternative forced choice task), using numbers 2 or 8 on the numeric pad of a standard AZERTY keyboard. In order to avoid any stimulus-response compatibility effects (Schneider & Bavelier, 2003), we opted for the use of specific response buttons whose (vertical) alignment was orthogonal to the stimuli appearing on the screen along a horizontal axis. Response buttons were counterbalanced across participants. Importantly, each face pair always consisted of two different identities, resulting in a total number of 90 face pairs per condition. In 50% of the trials, one face conveyed a fearful expression, while the other one displayed a neutral expression. Each emotion expression appeared equally often to the left or right of the central fixation cross. As control conditions, either two neutral faces (25% of the trials) or two fearful faces (25% of the trials) were displayed on screen. Thus, three stimulus pair conditions were presented in random order: fearful face-neutral face (FearNeut), fearful face-fearful face (FearFear), neutral face-neutral face (NeutNeut).

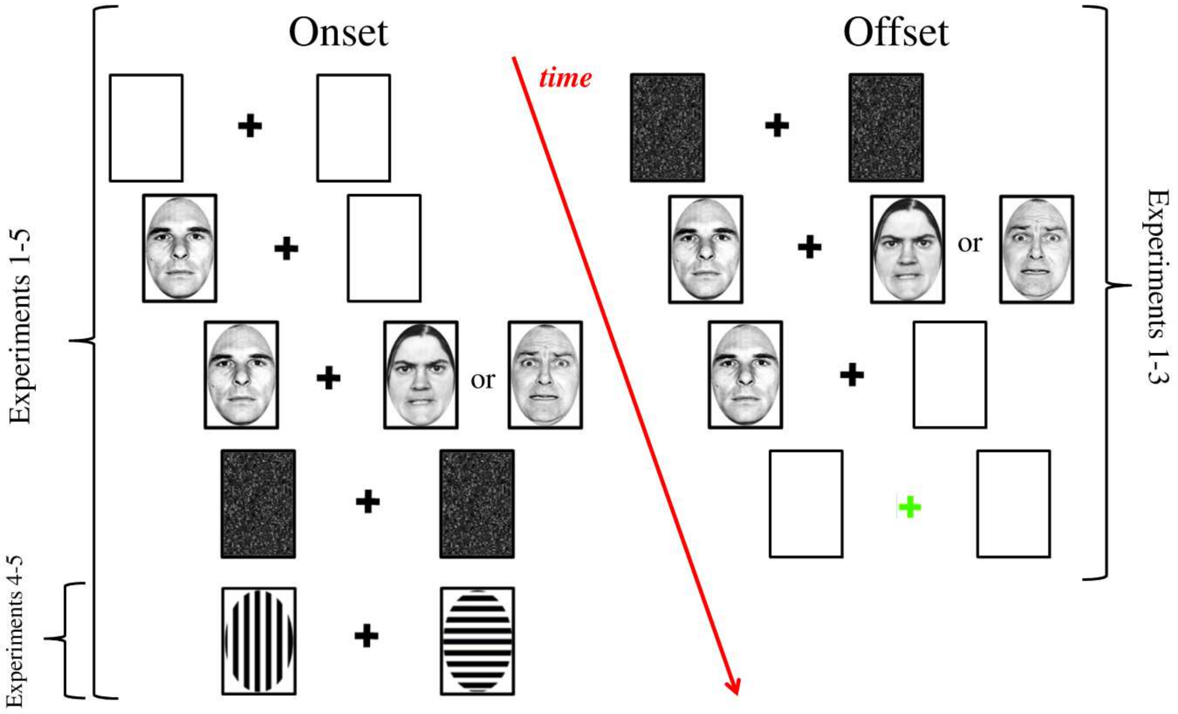


Figure 1. Stimuli and procedure used in Experiments 1-5. In the onset task (left panel), participants were presented with two placeholders on either side of fixation. After 1000 ms, one of the two face stimuli in the pair appeared either in the left or right box for a given stimulus onset asynchrony (SOA; 10, 30, or 100 ms), immediately followed by the second face stimulus. The stimulus pair stayed on screen for an additional 100 ms before being masked in synchrony until participants could decide which stimulus appeared first (left or right in Experiments 1-2; emotional or neutral in Experiments 3-5). In the offset task (right panel), each trial began with both masks in the placeholders for 500 ms, replaced by facial stimuli. After 1000 ms, one of the two faces disappeared for the duration of the SOAs. The remaining face stayed on screen for an additional 100 ms, followed by a green fixation cross indicating that a response had to be made. Participants had to report which stimulus disappeared first.

The *offset* task (Figure 1, right panel) was similar to the onset task described above, but here each trial began with a 500ms display including the fixation cross and both masks on the left and right side. Each mask was then replaced by a face (stimulus pairs: FearNeut, FearFear, NeutNeut) for 1000 ms, after which one of the two stimuli disappeared for the duration of the SOAs (100, 30, or 10 ms). The remaining face stayed on screen for an additional 100 ms, and then the black fixation cross turned green, signaling that a response had to be made. Participants had to perform a similar two-alternative forced choice task and judge the location of the face stimulus that was perceived as disappearing first, using numbers 2 or 8 on the numeric pad.

Finally, to verify that the emotional content of the faces selected in our study was perceived in line with the normative ratings, at the end of the experiment we asked participants to rate the amount of fear conveyed by each neutral and fearful face. A standard

9-point Likert scale was used for this purpose, with anchor 1 corresponding to “not afraid” and anchor 9 to “extremely afraid”.

2.4. Questionnaires

At the end of the experiment (also valid for Experiments 2-5), participants were asked to fill out two questionnaires, in order to assess whether specific affective or personality traits might be related to task performance. Levels of trait anxiety were measured using the Dutch version of the State-Trait Anxiety Inventory, trait characteristics (Van der Ploeg, Defares, & Spielberger, 1979). Participants also completed the Need For Affect Scale (Maio & Esses, 2001), which provides an estimate of participants’ general motivation to either approach or avoid emotion-inducing situations. The results confirmed normal scores of trait anxiety and Need For Affect (Table 1). However, no significant correlation was found between these scores and the behavioral results obtained across the five experiments. Therefore, the potential modulatory role of these two personality factors on the prioritized allocation of attention towards negative emotional stimuli is not discussed further.

Table 1. Mean values and standard deviations (in parenthesis) of the scores obtained for each questionnaire (and relative subscales) administered at the end of the experiment.

Questionnaire	Scores				
	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
STAI-T	41.91 *(10.08)*	40.00 *(7.11)*	40.90 *(10.43)*	40.18 *(9.08)*	43.56 *(11.68)*
NFAS	3.95 *(0.47)*	3.92 *(0.41)*	3.86 *(0.46)*	4.04 *(0.38)*	4.06 *(0.54)*
Approach	4.72 *(0.85)*	4.78 *(0.71)*	4.95 (0.49)	4.86 *(0.68)*	4.80 *(0.70)*
Avoidance	3.18 *(0.73)*	3.06 *(0.63)*	2.77 *(0.71)*	3.22 *(0.90)*	3.32 *(1.03)*

Note. STAI-T: State-Trait Anxiety Inventory, trait version; NFAS: Need for Affect Scale. STAI-T scores range from 20 to 80. NFAS scores were obtained using a 7-points Likert scale.

2.5. Data analysis

Accuracy was expressed as the proportion of “right first” responses. Positive SOAs refer to cases when the first stimulus was presented on the right hemifield, whereas negative SOAs indicate that the first stimulus was presented on the left side (see Figure 2A). The effect of prior entry was assessed by calculating each participant’s point of subjective simultaneity (*PSS*). This parameter indicates the time interval needed for the participant to perceive the two stimuli as arriving simultaneously or, in other words, an estimate of the SOA at which participants would be likely to make each response equally often (Perez, et al., 2009; Shore, et al., 2001; Sinnett, Juncadella, Rafal, Azanon, & Soto-Faraco, 2007; Spence & Parise, 2010; Spence, et al., 2001). To compute the PSS, transformed *z*-scores of the proportion of “right first” responses were first obtained by applying the inverse of the standard normal distribution function to the raw proportion scores (*probit* analysis; Finney, 1964). This transformation enabled us to perform a linear regression on the transformed data to derive the PSS, calculated from the slope and intercept of the best-fitted line of the *z*-scores (PSS = -slope/intercept)[2]. To account for the correlation of measurements within the same subject, we used a mixed probit regression model, where each participant had his/her own intercept and slope with estimated random effects from a bivariate zero-mean normal distribution. If a PSS value was falling outside the SOA range (i.e., > +100 or < -100 ms), the data of this participant were excluded from further analyses (for a similar procedure, see Spence, et al., 2001; Weiß & Scharlau, 2012). Based on previous research (West, et al., 2009), we hypothesized a prior entry effect (i.e., PSS significantly different from zero, as assessed by one-sample t-tests) for fearful compared to neutral faces in the FearNeut condition, whereas no difference ought to be observed in the two control conditions (i.e., FearFear and NeutNeut)[3]

Additionally, reaction times (RTs) were analyzed by means of repeated measures analysis of variance (ANOVA), with factors *SOA* (6 levels: -100, -30, -10, +10, +30, +100) and *stimulus pair condition* (3 levels: fear-neutral, fear-fear, neutral-neutral). We predicted that the SOAs for which uncertainty was high (i.e., ±30 and ±10 ms) would be associated with slower RTs, compared to longer SOAs (i.e., ±100 ms).

---

[2] We also computed and analyzed the so-called “just noticeable difference” (*JND*). JND corresponds to the slope of the best-fitted line of the z-scores (0.675/slope). This metric reflects the smallest temporal interval between two stimuli needed for an observer to correctly judge which stimulus had been presented first on 75% of the trials, since ±0.675 represents the 75% and 25% point on the cumulative normal distribution (Perez, et al., 2009; Sinnett, et al., 2007; Spence & Parise, 2010; Vatakis & Spence, 2006). However, from a theoretical standpoint, the effects of spatial attention on JND in a TOJ task are still unclear (Shore & Spence, 2005; Sinnett, et al., 2007). In addition, our analyses performed on the JND obtained for each of the five experiments did not reveal any valuable (compared to the PSS) information regarding a potential differential prior entry effects for emotional relative to neutral faces. Therefore, the results obtained for the JND are not reported.

[3] Several studies point either to a possible advantage of the right hemisphere in attention selection and processing (Holländer, Corballis, & Hamm, 2005; Verleger, Smigasiewicz, & Moller, 2011; Verleger, et al., 2009), or a disadvantage of the left hemisphere in these processes (Hellige, 1983; Hellige, Cox, & Litvac, 1979). Moreover, earlier work suggested that the right hemisphere could preferentially be engaged in the processing of emotion-laden stimuli (Gainotti, 1972, 2012; Mammucari, et al., 1988; Wittling & Roschmann, 1993). Accordingly, in all the experiments reported here, we also assessed whether any enhanced prior entry effect could be observed when the first (emotional or neutral) face in the pair was presented in the left vs. right hemifield relative to fixation. However, we did not find any effect of the side of presentation during the TOJ task, a finding which is also consistent with the study by Fecica & Stolz (2008).

Whenever Mauchly’s test indicated that the assumption of sphericity had been violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates. The alpha level for all statistical analyses was set at $p < 0.05$.

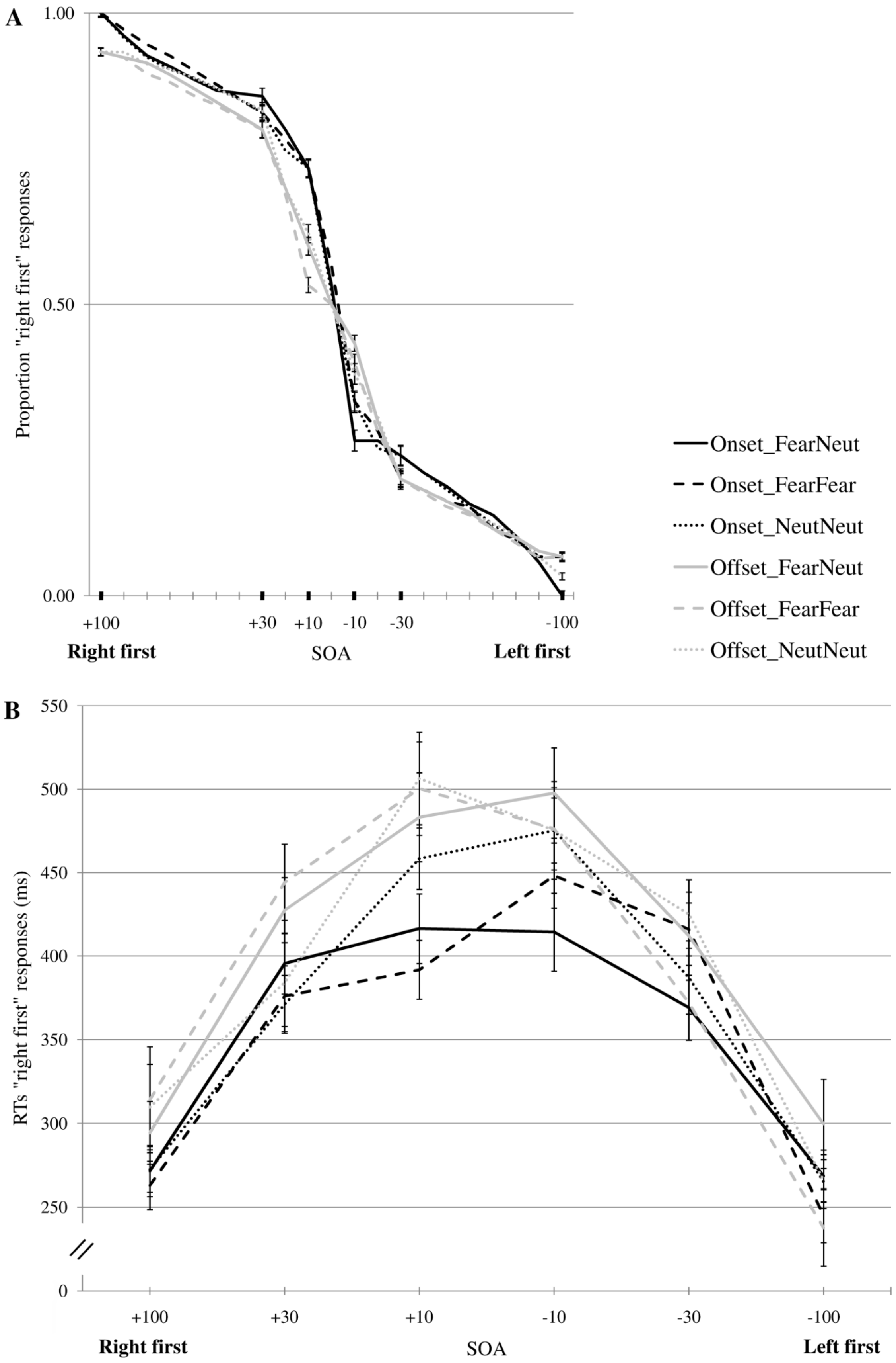
2.6. *Results*

Trials whose RTs were slower than three standard deviations from the mean (calculated for each condition and SOA separately across participants) were removed from the analysis (M = 1.12%, SD = 0.73).

Figure 2A shows the proportion of “right first” responses for each task (onset and offset) and condition (FearNeut, FearFear, NeutNeut). A clear inverted S-shaped psychometric function was obtained for each condition, providing evidence that the main experimental manipulation (i.e., SOA) was successful. Thus, participants perceived the onsets (or offsets) of the two stimuli in accordance with their respective occurrences. More specifically, participants’ TOJs were more uncertain (i.e., the proportion of “right first” responses was close to chance) at short (i.e., ±30 and ±10 ms) compared to long (i.e., ±100 ms) SOAs. The analysis of RTs corroborated this assumption. Participants were slower at short relative to long SOAs, as confirmed by a significant main effect of SOA [onset: $F(5, 155) = 52.88$, $p < .001$, $\eta_p^2 = .630$; offset: $F(5, 155) = 46.49$, $p < .001$, $\eta_p^2 = .600$], as well as its significant quadratic trend [onset: $F(1, 31) = 98.42$, $p < .001$, $\eta_p^2 = .760$; offset: $F(1, 31) = 99.17$, $p < .001$, $\eta_p^2 = .762$]. The SOA x condition interaction was not significant [onset: $F(10, 310) = 0.36$, $p = .964$; offset: $F(10, 310) = 1.46$, $p = .153$], indicating a comparable response speed across the three experimental conditions (Figure 2B).

The PSS values for each condition are reported in Table 2, separately for offset and onset tasks. In the onset task, for none of the three conditions did the one-sample t-test reach significance [FearNeut: $t(31) = 1.15$, $p = .260$; FearFear: $t(31) = 1.37$, $p = .180$; NeutNeut: $t(31) = 1.82$, $p = .079$]. Similar results were obtained for the offset task [FearNeut: $t(31) = 0.72$, $p = .479$; FearFear: $t(31) = -0.20$, $p = .847$; NeutNeut: $t(31) = -0.15$, $p = .884$], indicating no reliable prior entry effect for fearful compared to neutral faces (Figure 2C).

Importantly, results of the post-experiment ratings unequivocally confirmed that fearful faces were perceived as more fearful compared to neutral faces [$t(31) = 28.10$, $p < .001$] (Figure 2D).


A
1.00
0.50
0.00
Proportion "right first" responses
+100
+30
+10
-10
-30
-100
Right first
SOA
Left first
Onset_FearNeut
Onset_FearFear
Onset_NeutNeut
Offset_FearNeut
Offset_FearFear
Offset_NeutNeut
B
550
500
450
400
350
300
250
0
RTs "right first" responses (ms)
+100
+30
+10
-10
-30
-100
Right first
SOA
Left first

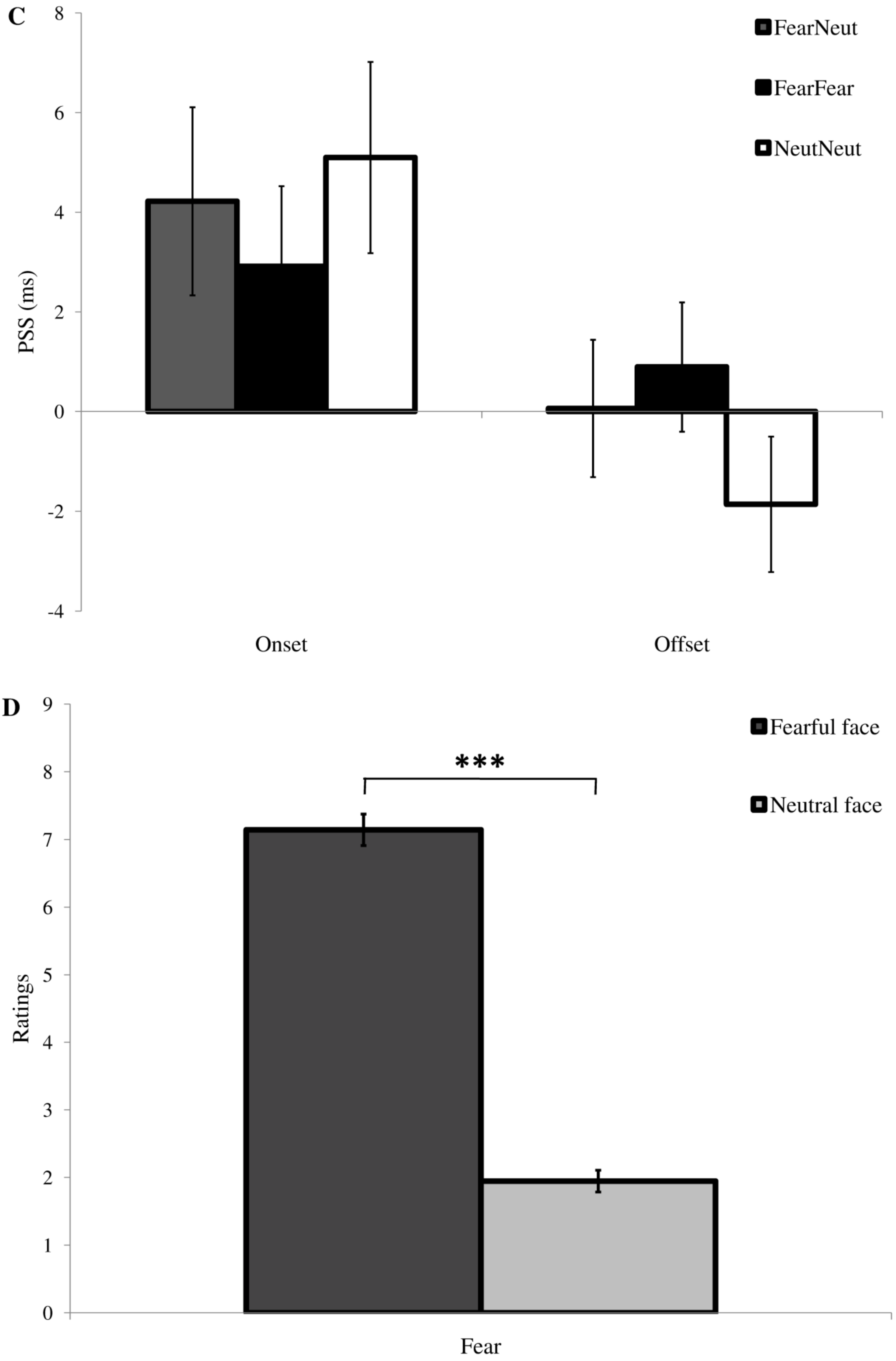


Figure 2. Results of Experiment 1. (A) The average proportion of “right first” responses, separately for onset (black lines) and offset (grey lines) tasks, displayed for each condition as a function of SOA. Positive SOAs

indicate that the first stimulus actually appeared on the right hemifield, whereas negative SOAs refer to first stimuli appearing on the left. The different conditions are: fearful-neutral (FearNeut, solid lines), fearful-fearful (FearFear, dashed lines), neutral-neutral (NeutNeut, dotted lines). The horizontal line corresponds to the 50% response mark (chance level), that is when participants responded "left" or "right" equally often. Significant visual prior entry effects (indicating attentional capture for one of the two stimuli in the pair) would be visualized as horizontal shifts of the point of maximum uncertainty across the 50% response mark. (B) The average RTs of "right first" responses, separately for onset (black lines) and offset (grey lines) tasks, for FearNeut (solid lines), FearFear (dashed lines) and NeutNeut (dotted lines) conditions as a function of SOA. Slower RTs were observed at short relative to long SOAs, indicating increased response uncertainty. However, no difference in this inverted U-shape was evidenced across conditions. (C) PSS values (in ms) for onset (left) and offset (right) tasks, separately for FearNeut (grey bar), FearFear (black bar) and NeutNeut (white bar) conditions. Positive values indicate prior entry for the left stimulus in the pair, while negative values correspond to prior entry for the right stimulus. None of these values was significantly different from zero, indicating no prior entry for any of the experimental conditions. (D) Mean fear ratings collected at the end of the main experiment, separately for fearful (dark grey bar) and neutral (light grey bar) faces. Fearful faces were consistently rated as more fearful than neutral faces. *** $p < .001$. Vertical bars correspond to standard errors of the means.

Table 2. Median values and standard deviations (in parenthesis) of the PSS scores obtained in Experiment 1 for each condition, separately for onset and offset tasks.

Task	Condition	PSS
Onset	FearNeut	4.22 *(14.70)*
	FearFear	2.92 *(15.56)*
	NeutNeut	5.10 *(15.35)*
Offset	FearNeut	0.06 *(18.21)*
	FearFear	0.89 *(21.35)*
	NeutNeut	-1.86 *(21.71)*

Note. Positive values reflect processing prioritization (i.e., prior entry) for the left stimulus in the pair, whereas negative values refer to prior entry for the right stimulus in the pair.

2.7. Discussion

In Experiment 1, participants were presented with pairs of fearful and neutral faces, and were instructed to report whether the first stimulus appeared (or disappeared) on the left or right visual hemifield. We hypothesized that fearful faces, because of their enhanced intrinsic motivational salience compared to neutral faces, could rapidly capture exogenous attention and, accordingly, bias TOJs (as reflected by PSS values being significantly different from zero in the FearNeut condition). However, we failed to observe such pattern of results. Importantly, these null findings could not easily be accounted for by mere task difficulty or abnormal temporal perception and/or attentional allocation throughout the visual field, since most of the participants could correctly identify the first onset in the pair (as evidenced by the presence of a clear inverted S-shaped psychometric function observed for each experimental condition; see Figure 2A). Moreover, post-experiment ratings confirmed that fearful faces were clearly recognized as such compared to neutral faces (Figure 2D), ruling out the possibility that the fearful faces selected in this experiment displayed weak or undifferentiated negative emotional expressions. Moreover, the results obtained for the offset task did not show any significant or differential effect for the FearNeut condition, suggesting that neither a rapid orienting (onset task) nor a prolonged disengagement effect (offset task) was evidenced for fearful faces in this TOJ task.

Although fearful faces were previously shown to influence early attention selection processes (see Pourtois, et al., 2012, for a recent review), the lack of a reliable prior entry for fearful faces might be explained by the fact that the threat displayed in these faces is indirect in essence, thereby affecting the motivational significance to a lower extent than angry faces (Sander, et al., 2007; Whalen, et al., 2004). Moreover, earlier studies already reported prior entry effects for (schematic) angry faces (Fecica & Stolz, 2008; West, et al., 2009). Therefore, in Experiment 2, we used angry and fearful faces concurrently in order to assess whether any prior entry effect for negative emotional facial expressions might be specific to angry faces or not. Furthermore, we substantially reduced the size of the face stimuli compared to Experiment 1, as well as their eccentricity relative to fixation. We reasoned that the use of large face stimuli (i.e., subtending 8.86° x 7.63° of visual angle) shown in the far periphery (i.e., 10.29° from fixation) may have favored the use of low-level features to perform the TOJ task in Experiment 1. By comparison, West and colleagues (2009) presented schematic or human faces in squared boxes subtending 3.80° x 3.60° of visual angle at a lower horizontal eccentricity (3.15° from fixation). Accordingly, in Experiment 2, our stimulus parameters were closely matched to those used previously by West, et al. (2009).

3. Experiment 2

3.1. Participants

Forty healthy psychology students participated in the study in exchange of course credits. None of them had participated to Experiment 1. All volunteers gave informed written consent prior to their participation. The data of two participants were excluded from further analyses due to an abnormal inverted S-shaped psychometric function in at least one experimental condition (similarly to Experiment 1). Thus, the final sample contained 38 participants (32 women, mean age 18 years, range 17-22).

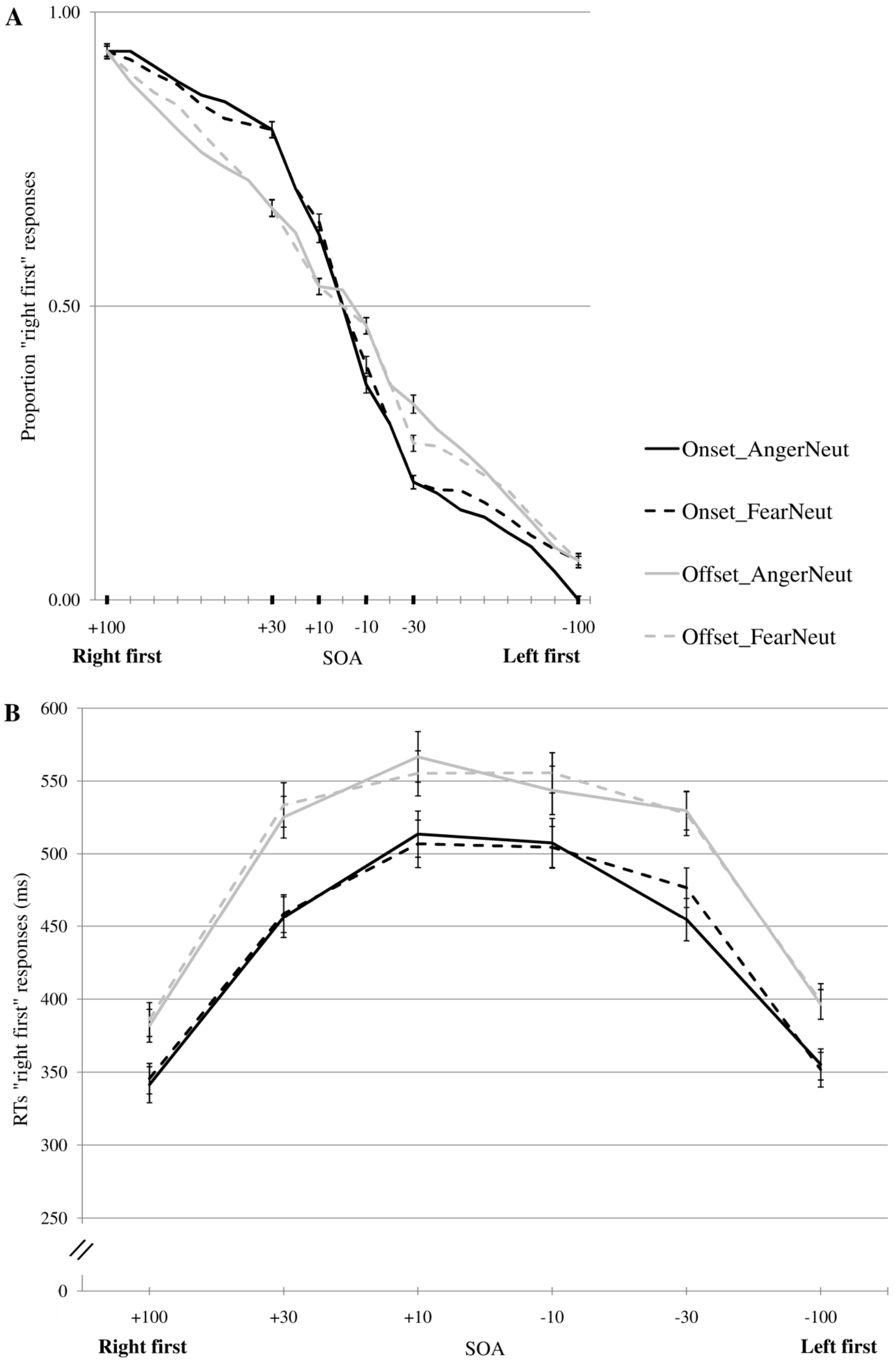
3.2. *Stimuli*

Fearful and neutral faces were identical to the ones used in Experiment 1. However, they were now enclosed in a smaller oval frame, spanning 4.77° x 3.06° of visual angle. In addition, 10 faces displaying an angry expression were selected from the same standardized Ekman series (Ekman & Friesen, 1976). Apparent contrast was also calculated for angry faces, and independent samples t-tests revealed no significant difference between neutral and angry faces [$t(18) = -0.99$, $p = .334$], as well as fearful and angry faces [$t(18) = -0.16$, $p = .877$].

3.3. *Procedure*

The procedure and design of the onset and offset tasks were similar to Experiment 1. However, here the facial stimuli were presented closer to fixation (distance between the center of the fixation cross and the center of the face: 6.68°) compared to Experiment 1. The stimulus pair conditions were angry face-neutral face (AngerNeut) and fearful face-neutral face (FearNeut). No additional condition (i.e., AngerAnger, FearFear or NeutNeut) was included, in order to avoid an excessively high number of trials and long testing session likely causing drops or lapses in attention. Note that the use of the AngerNeut and/or FearNeut conditions alone is sufficient to establish whether any reliable prior entry (for either angry or fearful faces) was present or not (West, et al., 2009).

Ratings of perceived anger and fear conveyed by each face stimulus were collected at the end of the main TOJ task by means of 9-point Likert scales, ranging from 1 (“not afraid/angry”) to 9 (“extremely afraid/angry”). Additionally, participants were asked to provide ratings of perceived brightness for each emotional and neutral face (from 1, “very dark”, to 9, “very bright”), to further corroborate the lack of clear difference in this low-level visual property across the three emotion categories (i.e., neutral, angry, and fearful).


A
1.00
0.50
0.00
Proportion "right first" responses
+100
+30
+10
-10
-30
-100
Right first
SOA
Left first
Onset_AngerNeut
Onset_FearNeut
Offset_AngerNeut
Offset_FearNeut
B
600
550
500
450
400
350
300
250
0
RTs "right first" responses (ms)
+100
+30
+10
-10
-30
-100
Right first
SOA
Left first

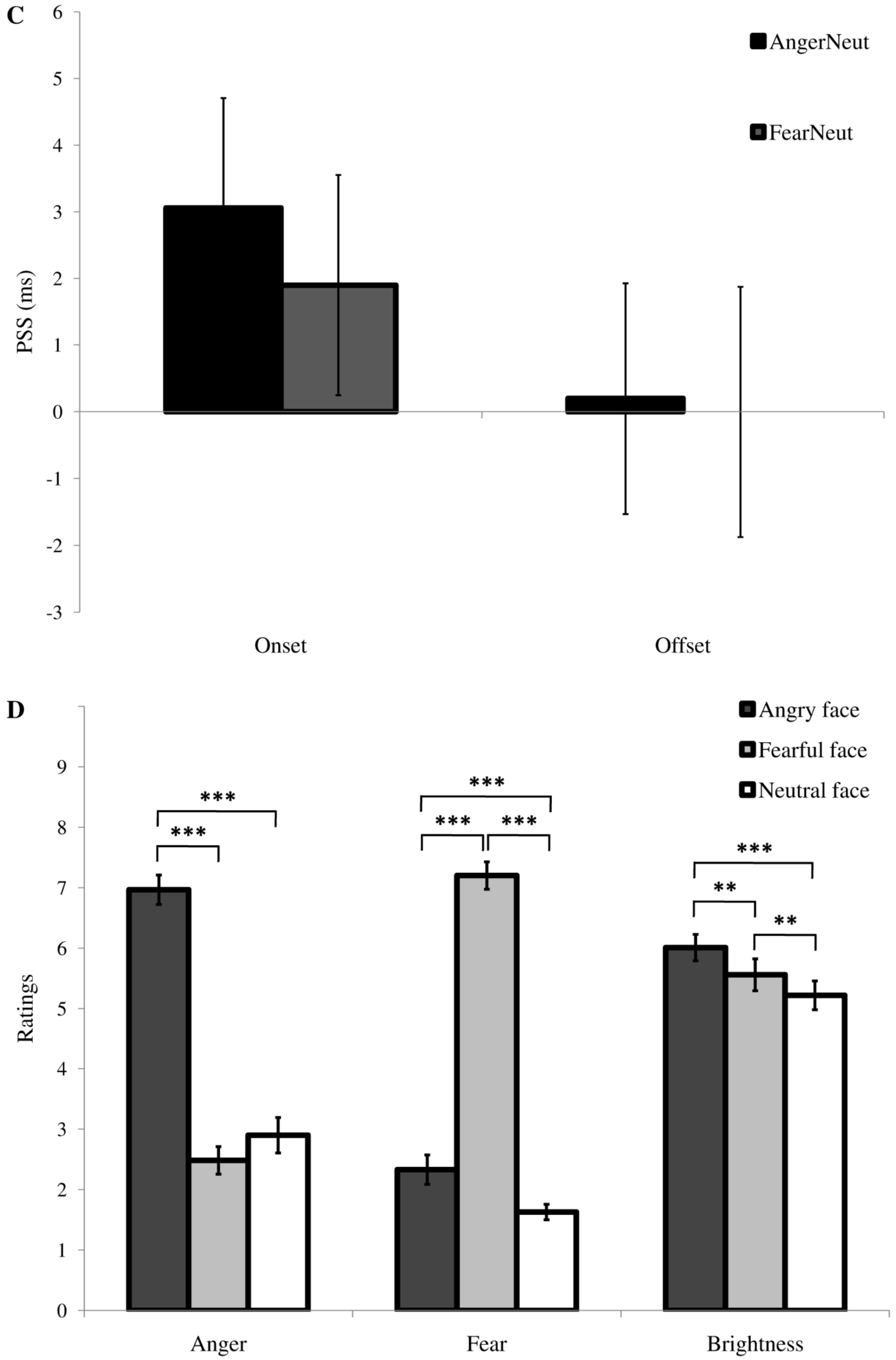


Figure 3. Results of Experiment 2. (A) The average proportion of “right first” responses, separately for onset and offset tasks, for each condition as a function of SOA. (B) The average RTs of “right first” responses,

separately for onset and offset tasks, for AngerNeut and FearNeut conditions. Participants were more uncertain at short compared to long SOAs, but no reliable differences were observed across conditions. (C) PSS values for onset and offset tasks, separately for AngerNeut and FearNeut stimulus pairs. No prior entry was found for any of the experimental conditions. (D) Mean anger (left), fear (middle) and brightness (right) ratings, separately for angry (dark grey bar), fearful (light grey bar) and neutral (white bar) faces. Emotion ratings were consistent with the *a priori* stimulus categorization. Of note, the statistically different ratings in perceived brightness did not influence behavior, as confirmed by a lack of prior entry effect for angry faces. ** $p < .01$; *** $p < .001$. Vertical bars correspond to standard errors of the means.

3.4. Results

Following standard practice, trials whose RTs were slower than three standard deviations from the mean were discarded (M = 0.98%, SD = 0.66).

Behavioral results showed that, as reported in Figure 3A, the distribution of the proportion of “right first” responses was consistent with the results obtained in Experiment 1: responses were close to chance level at short compared to long SOAs, both in the onset and offset tasks. Accordingly, RTs were slower at short relative to long SOAs. A 6 (SOA) x 2 (condition: AngerNeut, FearNeut) repeated measures ANOVA confirmed a significant main effect of SOA [onset: $F(5, 185) = 63.70$, $p < .001$, $\eta_p^2 = .633$; offset: $F(5, 185) = 43.62$, $p < .001$, $\eta_p^2 = .541$], as well as a significant quadratic trend [onset: $F(1, 37) = 103.03$, $p < .001$, $\eta_p^2 = .736$; offset: $F(1, 37) = 66.82$, $p < .001$, $\eta_p^2 = .644$] (see Figure 3B). In addition, in the offset task exclusively, a marginally significant SOA x condition interaction was found [$F(5, 185) = 2.47$, $p = .055$, $\eta_p^2 = .063$]. However, post-hoc t-tests failed to evidence meaningful differences across conditions when corrected for multiple comparisons (ps > .05).

Table 3 shows the PSS values for each condition, separately for offset and onset tasks. None of the one-sample t-tests were significantly different from zero, both in the onset [AngerNeut: $t(37) = 0.99$, $p = .327$; FearNeut: $t(37) = 0.74$, $p = .466$] and offset tasks [AngerNeut: $t(37) = 0.04$, $p = .969$; FearNeut: $t(37) = 0.37$, $p = .718$] (see also Figure 3C). Thus, no prior entry for negative emotional facial expressions (either fear or anger) was evidenced.

Post-experiment ratings confirmed that fearful faces were perceived as more fearful compared to neutral [$t(37) = 34.02$, $p < .001$] and angry faces [$t(37) = 29.60$, $p < .001$]. Similarly, angry faces were rated higher along the anger intensity dimension compared to neutral [$t(37) = 33.15$, $p < .001$] and fearful faces [$t(37) = 25.97$, $p < .001$] (Figure 3D). Thus, participants correctly perceived the respective emotion content and intensity displayed by the selected face stimuli. Results further revealed higher perceived brightness for emotional relative to neutral faces [anger vs. neutral: $t(37) = 4.73$, $p < .001$; fear vs. neutral: $t(37) =$

2.97, $p < .001$]. Note that, despite these subjective differences in brightness, no prior entry effect for either angry or fearful faces was found compared to neutral faces.

Table 3. Median values and standard deviations (in parenthesis) of the PSS scores obtained in Experiment 2 for each condition, separately for onset and offset tasks.

Task	Condition	PSS
Onset	AngerNeut	3.06 *(20.33)*
	FearNeut	1.90 *(20.39)*
Offset	AngerNeut	0.20 *(21.34)*
	FearNeut	0.00 *(23.12)*

3.5. Discussion

Results of Experiment 2 failed to show any significant prior entry effect for either fearful faces (replicating the results of Experiment 1) or angry faces when compared to neutral faces, despite a clear and predictable effect of SOA on TOJs (i.e., inverted S-shaped psychometric function). Unlike previous studies using mainly schematic angry faces (Fecica & Stolz, 2008; West, et al., 2009), here we did not find evidence for the preferential (exogenous) orienting towards photographs of realistic fearful or angry faces when they compete with neutral faces for attention selection. Because our experimental setup was similar to West and colleagues (2009), these results are unlikely to be explained by suboptimal stimulus parameters or task demands. Moreover, since participants of Experiment 2 unambiguously identified the emotion conveyed by fearful and angry faces during a post-experiment rating phase (see also Experiment 1), these results cannot be accounted for by the use of face stimuli providing weak or undifferentiated emotional expressions relative to neutral faces. Finally, the results of the offset task also showed that no facilitated disengagement effect was evidenced for either angry or fearful faces, a phenomenon that could potentially mask any preferential early orienting of attention towards these negative stimuli in the onset task.

An intriguing possibility to account for these null findings (Experiments 1-2) may be related to the specific task set adopted by the participants throughout the experimental session. Given that participants had to focus on spatial and temporal properties to carry out the two-alternative forced choice task (i.e., is it the left or right stimulus appearing/disappearing first?), the actual emotion content of the faces could somehow be filtered out completely in these two experiments. Moreover, previous research showed that early and automatic affective stimulus processing could substantially be reduced when concurrent non-affective (spatial) stimulus dimensions became task-relevant (Everaert, Spruyt, & De Houwer, 2011; Pessoa, Kastner, & Ungerleider, 2002; Spruyt, De Houwer, & Hermans, 2009; Spruyt, De Houwer, Hermans, & Eelen, 2007), consistent with the idea that the (exogenous) capture of attention by emotion is not “magic”, but subject to (state) fluctuations depending on the availability of attentional resources, as well as the specific task set (see Pourtois, et al., 2012). In light of this evidence, we surmised that participants of Experiments 1-2 may have adopted an efficient strategy and eventually allocated attentional resources primarily to the processing of the spatial and temporal properties of the two face stimuli, while actively “ignoring” their emotional content because poorly informative to resolve the task. We have to acknowledge, however, that this account already posits that negative emotional facial expressions do not “automatically” capture attention, because this effect (at least in the case of a TOJ task) may actually depend upon the specific task demands (see Moors & De Houwer, 2006). Accordingly, no prior entry for angry or fearful faces was evidenced in these two earlier experiments, probably because participants could easily ignore the emotional content of the two competing faces and focus on a specific non-affective stimulus feature whose processing was sufficient to perform the task. To address this issue, in Experiment 3 we modified the task instructions and asked participants to judge whether the emotional or the neutral face appeared first (*emotion* TOJ), making the differential emotional content of the two faces in the pair directly task-relevant. Hence, in Experiment 3 a two-alternative forced-choice task was still required, but it concerned the content rather than the spatial position of the stimuli.

4. Experiment 3

4.1. Participants

Thirty-seven psychology students, who did not participate in Experiment 1 or 2, took part in Experiment 3. Using the same exclusion criteria as above (see Experiments 1 and 2), the data of 16 participants had to be removed from the subsequent statistical analyses. The data of 21 participants (19 women, mean age 18 years, range 18-21) were thus included in the final sample.

4.2. Stimuli and procedure

The stimuli were identical to Experiment 2. However, unlike Experiments 1-2, participants were asked to perform a two-alternative forced choice task based on the emotional content of the face stimuli in the pair. More precisely, they were instructed to judge whether the stimulus that appeared first had either a neutral or an emotional expression, thereby making the emotional content of the face stimuli task-relevant. Another notable difference between Experiment 3 and Experiments 1-2 was the use of a block design. In order to facilitate participants' discrimination between emotional and neutral faces, AngerNeut and FearNeut trials were no longer presented in random order throughout the experimental session, but in two separate blocks (counterbalanced across participants).

Finally, ratings for the perceived anger, fear and brightness of the individual face stimuli were collected post-experiment, similarly to Experiments 1-2.

4.3. Results

A total of 0.55% (*SD* = 0.40) of trials were discarded because their RTs were slower than three standard deviations from the mean.

The proportion of "emotion first" responses is plotted in Figure 4A, separately for each condition (AngerNeut vs. FearNeut) and task (onset vs. offset). As expected, the proportion of "emotion first" responses was close to chance level at short compared to long SOAs (both in the onset and offset tasks), as evidenced by a clear inverted S-shaped psychometric function. Moreover, RTs were slower at short relative to long SOAs, as confirmed by a significant main effect of SOA [onset: $F(5, 100) = 9.19$, $p < .001$, $\eta_p^2 = .315$; offset: $F(5, 100) = 10.40$, $p < .001$, $\eta_p^2 = .342$] and a significant quadratic trend [onset: $F(1, 20) = 17.82$, $p < .001$, $\eta_p^2 = .471$; offset: $F(1, 20) = 19.58$, $p < .001$, $\eta_p^2 = .495$] (see Figure 4B). However, no significant SOA x condition interaction was evidenced [onset: $F(5, 100) = 1.16$, $p = .332$, $\eta_p^2 = .055$; offset: $F(5, 100) = 0.73$, $p = .603$, $\eta_p^2 = .035$].

PSS values for each condition (separately for onset and offset tasks) are shown in Table 4. No significant prior entry effect was found, either in the onset [AngerNeut: $t(20) = 0.18$, $p = .858$; FearNeut: $t(20) = -1.27$, $p = .218$] or in the offset task [AngerNeut: $t(20) = 0.24$, $p = .815$; FearNeut: $t(20) = -0.81$, $p = .428$] (Figure 4C).

Post-experiment ratings confirmed that fearful faces were perceived as more fearful compared to neutral faces [$t(20) = 15.84$, $p < .001$] and angry faces [$t(20) = 16.36$, $p < .001$].

In addition, angry faces were perceived as carrying more anger intensity than neutral [$t(20) = 17.00$, $p < .001$] and fearful faces [$t(20) = 16.72$, $p < .001$] (Figure 4D). Finally, participants rated emotional and neutral stimuli as equally bright (ps > .05).

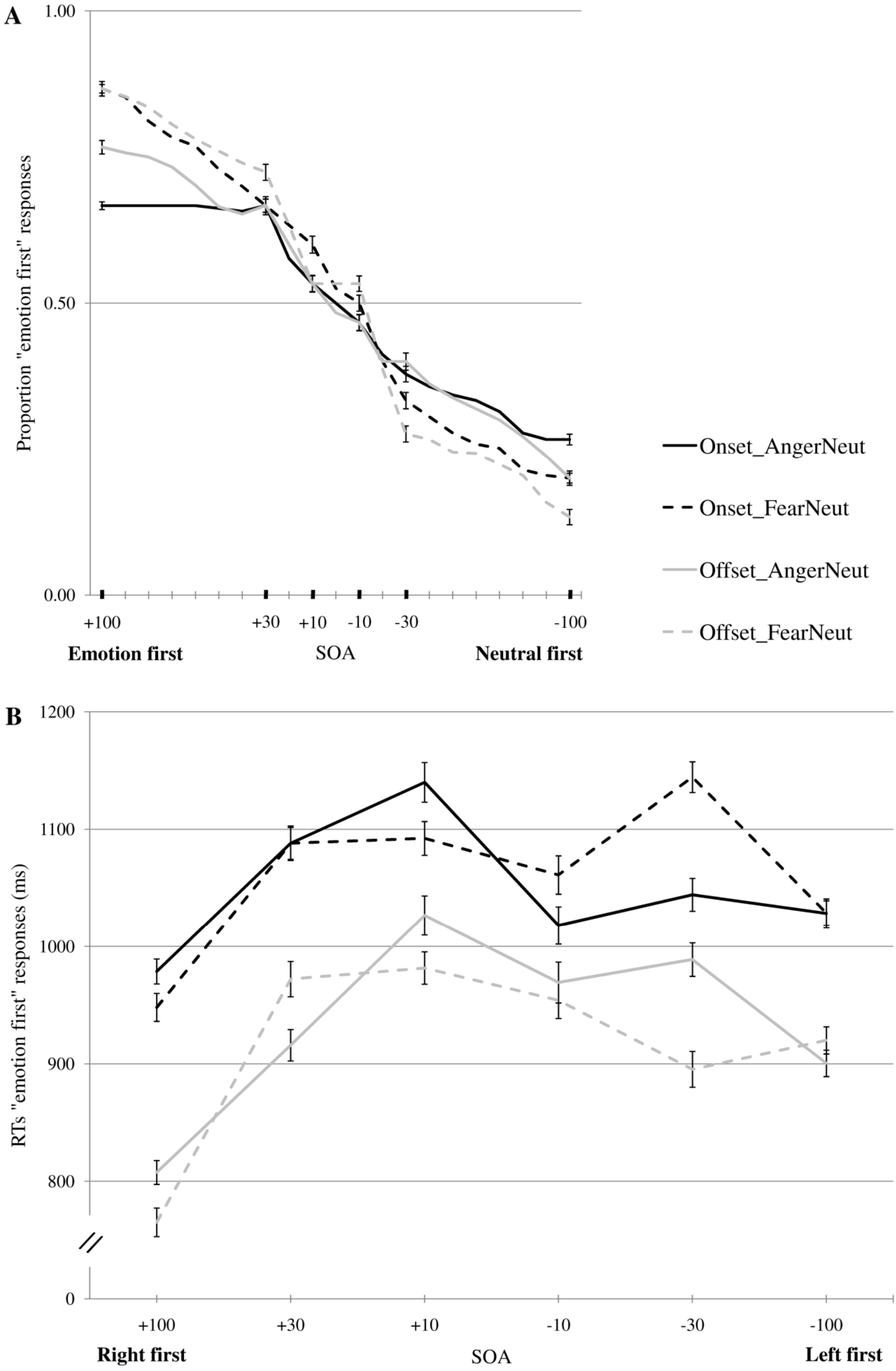
Table 4. Median values and standard deviations (in parenthesis) of the PSS scores obtained in Experiment 3 for each condition, separately for onset and offset tasks.

Task	Condition	PSS
Onset	AngerNeut	1.61 *(34.60)*
	FearNeut	-3.63 *(25.43)*
Offset	AngerNeut	-0.83 *(24.25)*
	FearNeut	-2.55 *(23.79)*

Note. Positive values reflect processing prioritization (i.e., prior entry) for the neutral stimulus in the pair, whereas negative values refer to prior entry for the emotional stimulus in the pair.

4.4. Discussion

Despite the use of an emotion TOJ task (as opposed to a TOJ task based on the location of the face appearing first; see Experiments 1-2), we did not find evidence for a differential prior entry effect for either fearful or angry faces relative to neutral faces. Noteworthy, these non-significant results were obtained despite a clear emotion differentiation of the three emotion categories (as confirmed by post-experiment ratings), as well as the presence of clear inverted S-shaped psychometric functions in 21 participants (unambiguously revealing a clear sensitivity to the main SOA manipulation). Moreover, as was already the case for Experiments 1-2, we could rule out that these null-findings were somehow accounted for by the presence of a facilitated disengagement effect towards either fearful or angry faces in these participants, because no differential effect of the emotional content of the faces was evidenced during the offset task.


A
1.00
0.50
0.00
Proportion "emotion first" responses
+100
+30
+10
-10
-30
-100
Emotion first
SOA
Neutral first
Onset_AngerNeut
Onset_FearNeut
Offset_AngerNeut
Offset_FearNeut
B
1200
1100
1000
900
800
0
RTs "emotion first" responses (ms)
+100
+30
+10
-10
-30
-100
Right first
SOA
Left first

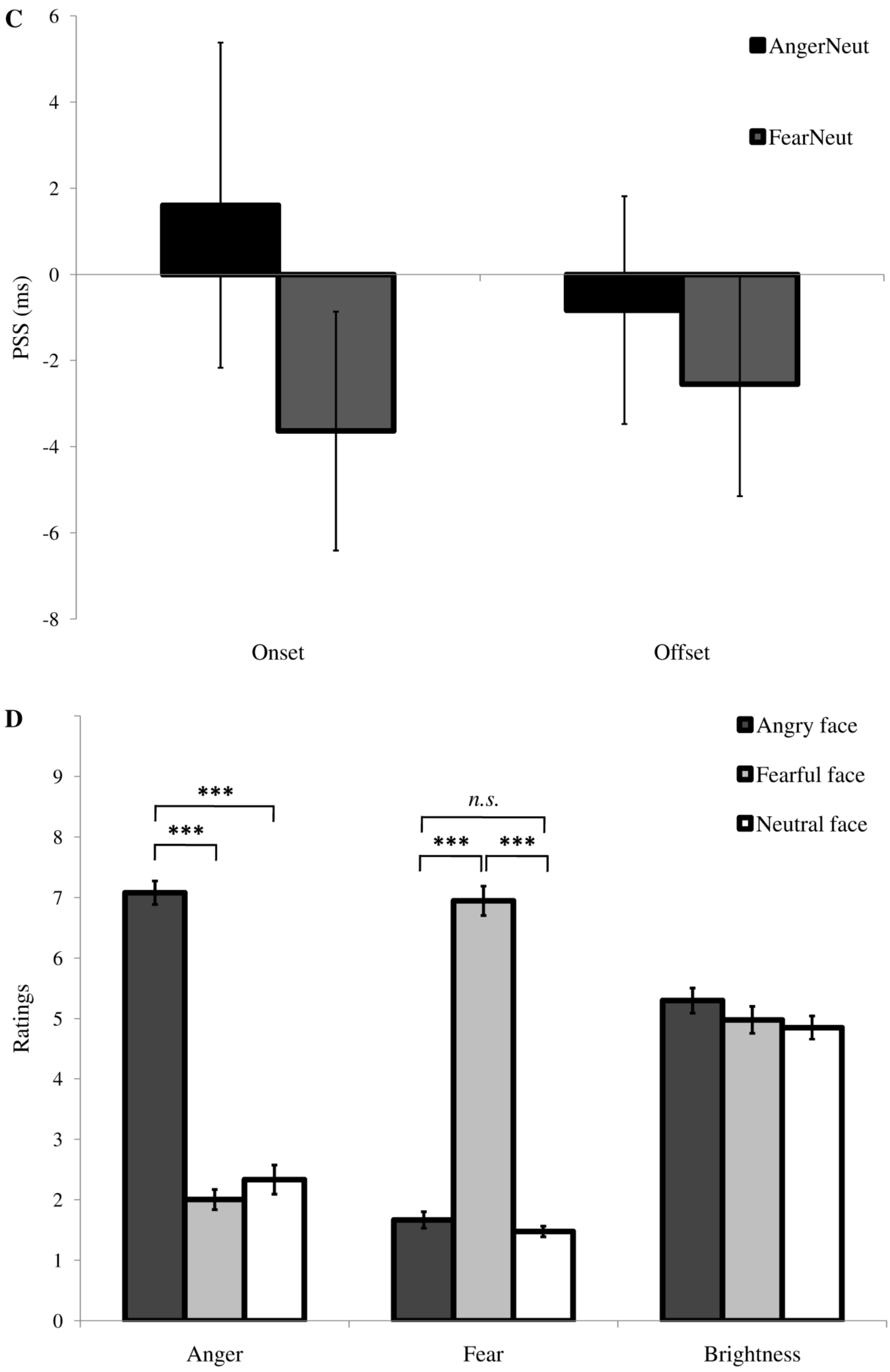


Figure 4. Results of Experiment 3. (A) Proportion of “emotion first” responses, separately for onset and offset tasks, for each condition as a function of SOA (positive SOAs: emotional stimulus appeared first; negative

SOAs: neutral stimulus appeared first). (B) Average RTs of “emotion first” responses, separately for onset and offset tasks, for AngerNeut and FearNeut conditions. Participants were more uncertain at short compared to long SOAs in both tasks, although the distribution of these RTs across the SOAs is noisier. (C) PSS values in onset and offset tasks for AngerNeut and FearNeut conditions. Positive values indicate prior entry for the emotional face, while negative values indicate attentional capture for the neutral face in the pair. However, no prior entry was observed. (D) Mean anger, fear and brightness ratings confirmed that fearful faces were perceived as more fearful, whereas angry faces were rated as more angry. No difference in perceived brightness was reported. *** $p < .001$. Vertical bars correspond to standard errors of the means.

The lack of prior entry effect for angry faces is puzzling to some extent, since participants were asked to process the emotional content of the faces in the pair in order to perform the task. Previous research showed that in these conditions (i.e., when emotion is directly task-relevant), rapid and automatic effects of (negative) emotion on feature-specific attention allocation could be observed in healthy adult participants (Everaert, et al., 2011; Spruyt, De Houwer, Everaert, & Hermans, 2012; Spruyt, et al., 2007). Furthermore, these findings are also at odds with earlier results showing a reliable prior entry effect for angry faces (West, et al., 2009), because similar stimulus parameters were used in these two studies. However, in West, et al. (2009), a rather limited number of face stimuli was used (i.e., four angry and four neutral faces, two male and two female for each category), an element that somewhat casts doubt on the validity of these earlier results and raises the possibility of low-level perceptual confounds. By contrast, in Experiments 1-3, the two faces in the pair were always perceptually dissimilar across the whole experimental session, precluding the use of a systematic strategy (presumably based on the processing of a non-emotional feature) to detect which of the two face stimuli appeared/disappeared first.

Using a stringent and standard exclusion criterion (Perez, et al., 2009; Spence, et al., 2001; Weiß & Scharlau, 2012), we found out that the data of sixteen participants had to be removed from the analysis because they did not show a normal change in TOJ (at least in one experimental condition) as a function of the SOA. This exclusion rate was substantially larger than what we found in Experiments 1-2 (where participants were instructed to focus on *spatial* and *temporal* properties of the two face stimuli in the pair), suggesting that the discrimination of the emotional content of the faces was more demanding than processing the temporal and spatial features of the first face appearing/disappearing on screen. Noteworthy, none of the two previous studies looking at prior entry for angry faces actually used a similar exclusion criterion (Fecica & Stolz, 2008; West, et al., 2009), suggesting that the results reported in these earlier studies might include the data of “poor-performers” who may encounter difficulties to process the (fine-grained) changes in the respective onsets of the two faces. At any rate, in Experiment 4, we aimed at addressing this question and, accordingly, we devised a new modification of the TOJ task enabling to briefly “train” temporal perceptual abilities with low-level geometric stimuli, before the putative effect of the emotional content of the face was systematically explored. We hypothesized that this initial training or familiarization with geometrical figures might later reduce the drop rate for the emotion TOJ. Hence, in Experiment 4, we included at the beginning of the experiment two training blocks during which participants had to perform the TOJ task based on the orientation of line gratings (being either horizontal or vertical). Then, participants performed the emotion TOJ, as described in Experiment 3. Of note, we only used the onset task in Experiment 4, given that no evidence was obtained in Experiments 1-3 for a possible differential disengagement effect towards fearful or angry faces.

5. Experiment 4

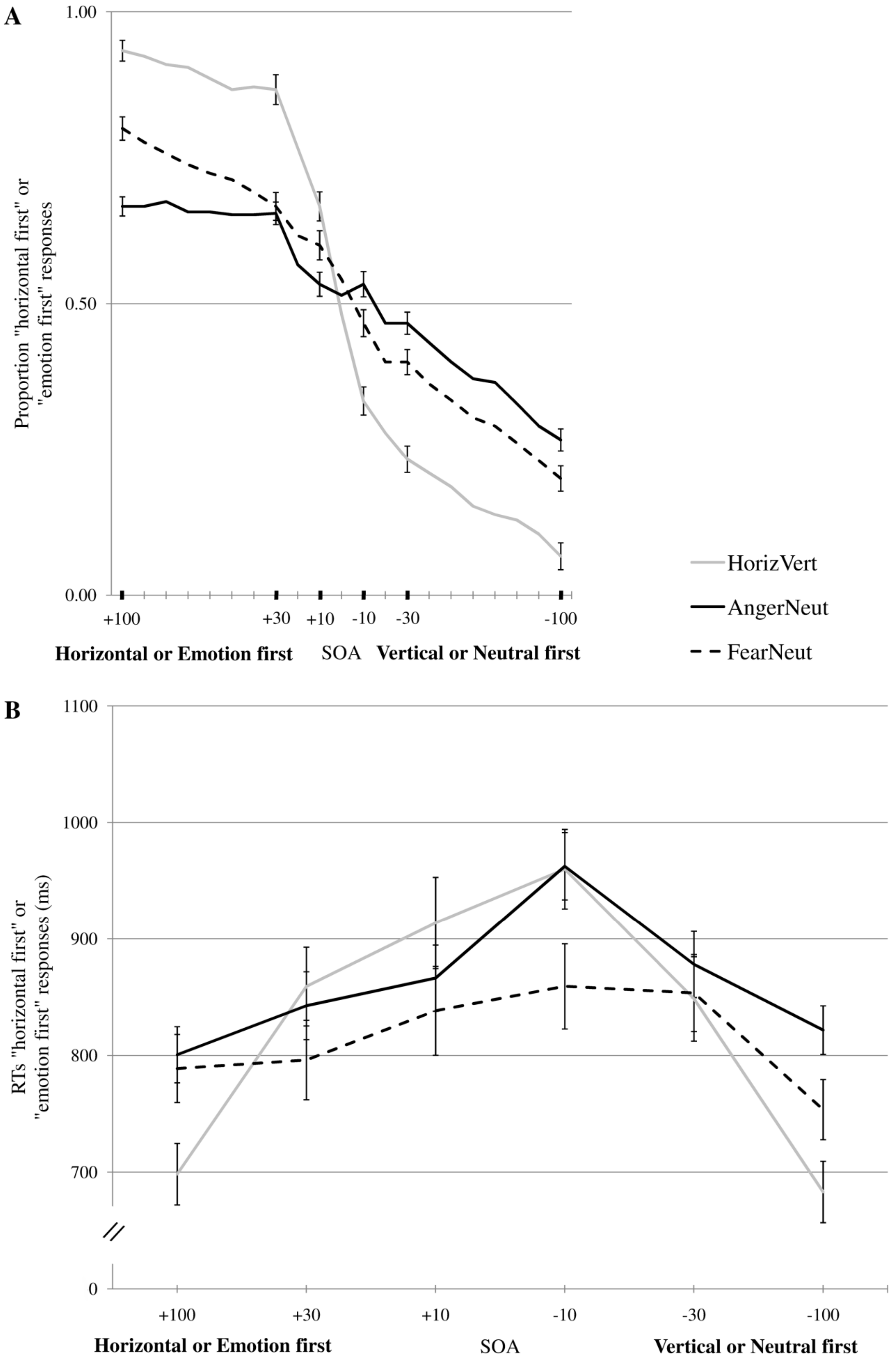
5.1. Participants

Forty psychology students, who did not participate in any of the previous experiments, took part in Experiment 4 for course credits. Using the same exclusion criteria as above (see Experiments 1-3), the data of 23 participants had to be excluded from the subsequent statistical analyses. Hence, the final sample consisted of 17 participants (13 women, mean age 20 years, range 18-30).

5.2. Stimuli and procedure

Face stimuli and procedure were identical to Experiment 3. In addition, before the emotion TOJ task (which was identical to Experiment 3), participants carried out a non-emotion TOJ task aimed at familiarizing them to detect different onsets with the different SOAs. Two blocks were included (each containing 90 trials), in which gratings consisting of either dominant horizontal or vertical black lines on a white background (matched in size with the face stimuli; see Figure 1) were presented equally often on the left and right hemifield for the duration of the SOAs. Participants had to judge whether the horizontal or vertical grating appeared first. Moreover, the offset task was discarded in Experiment 4 in order to shorten the total duration of the experimental session.

Ratings of the individual faces regarding the intensity of anger, fear and brightness were collected at the end of the experiment.


A
1.00
0.50
0.00
Proportion "horizontal first" or "emotion first" responses
+100
+30
+10
-10
-30
-100
Horizontal or Emotion first
SOA
Vertical or Neutral first
HorizVert
AngerNeut
FearNeut
B
1100
1000
900
800
700
0
RTs "horizontal first" or "emotion first" responses (ms)
+100
+30
+10
-10
-30
-100
Horizontal or Emotion first
SOA
Vertical or Neutral first

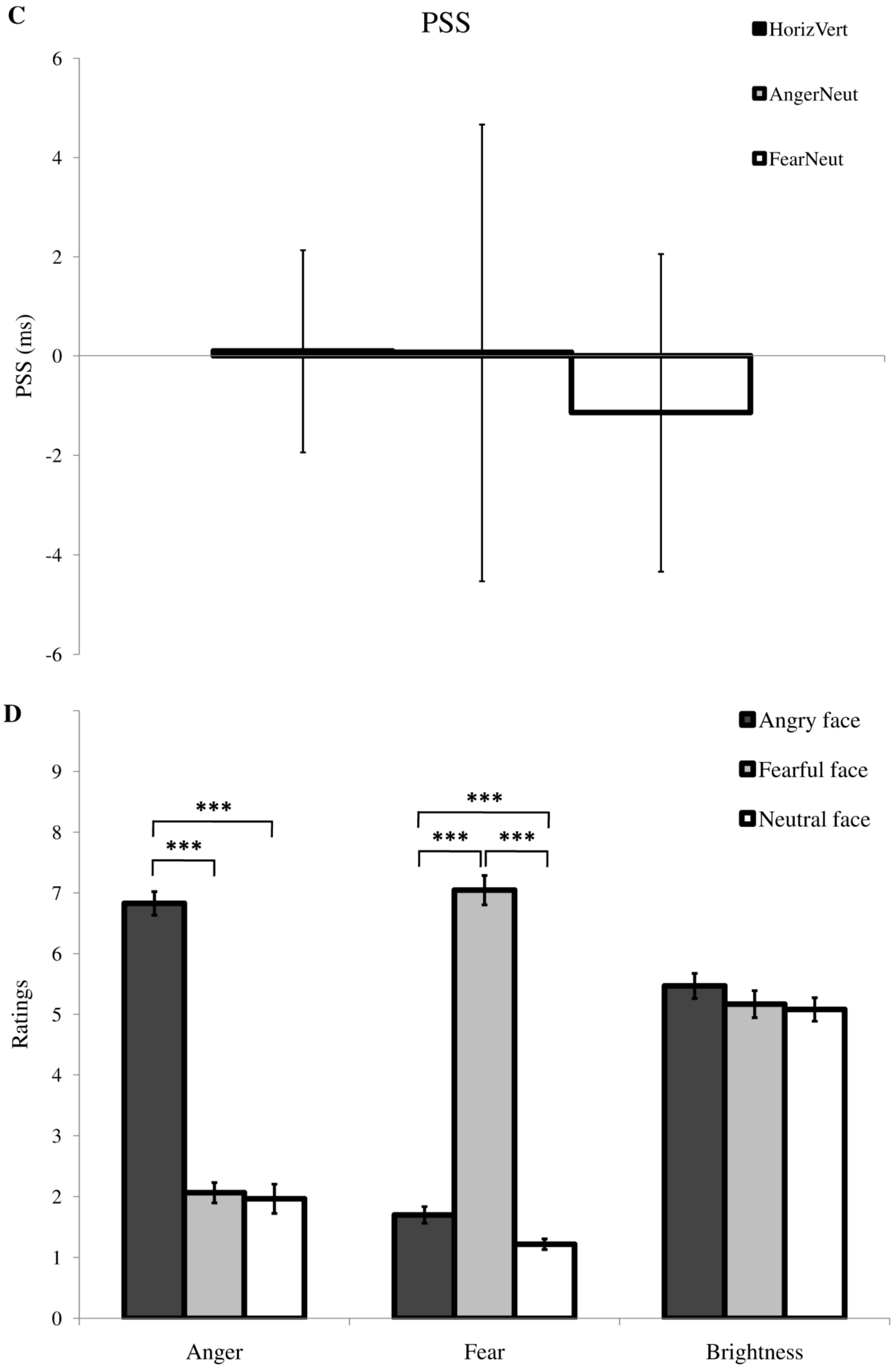


Figure 5. Results of Experiment 4. (A) Proportion of "horizontal first" responses (in the initial orientation tasks) and "emotion first" responses (in the emotion TOJ task), separately for each condition (HorizVert: horizontal-

vertical, solid grey line; AngerNeut: anger-neutral, solid black line; FearNeut: fearful-neutral, dashed black line). Results for the orientation and emotion TOJ tasks are shown together for illustration purposes. Positive SOAs refer to horizontal lines or emotional faces appearing first, whereas negative SOAs indicate that vertical lines or neutral faces appeared first. (B) Average RTs of "horizontal first" and "emotion first" responses for HorizVert, AngerNeut and FearNeut conditions. Participants were more uncertain at short compared to long SOAs, although this effect was more pronounced in the orientation task (presumably because it was easier; see main text). (C) PSS values for HorizVert, AngerNeut and FearNeut conditions. Positive values indicate prior entry for either the horizontal lines or the emotional face in the pair, whereas negative values would indicate prior entry for either the vertical lines or the neutral face. However, no prior entry was observed. (D) Mean anger, fear and brightness ratings collected at the end of the experiment. As expected, fearful faces were rated as more fearful, while angry faces were rated as more angry, with no difference in perceived brightness. *** $p < .001$. Vertical bars correspond to standard errors of the means.

5.3. Results

Trials whose RTs were slower than three standard deviations from the mean were discarded ($M = 0.73\%$, $SD = 0.51$).

Figure 5A shows the proportion of "horizontal first" responses for the non-emotion TOJ task during the two familiarization blocks, as well as the proportion of "emotion first" responses for the subsequent emotion TOJ task. Performance for the non-emotion TOJ task was remarkably accurate, as evidenced by a clear inverted S-shaped psychometric function (HorizVert condition in Figure 5A). By contrast, accuracy was substantially reduced for the emotion TOJ task, as shown by flatter inverted S-shaped psychometric functions for the AngerNeut and FearNeut conditions. Please note that the results reported here are for good performers only (i.e., participants whose PSS was between -100 and +100 ms for all conditions). A 6 (SOA) x 3 (condition: HorizVert, AngerNeut, FearNeut) repeated measures ANOVA on RTs revealed a significant main effect of SOA [$F(5, 80) = 19.42$, $p < .001$, $\eta_p^2 = .548$], as well as significant quadratic trend [$F(1, 16) = 33.12$, $p < .001$, $\eta_p^2 = .674$] (Figure 5B). Moreover, a significant SOA x condition interaction was found [$F(10, 160) = 3.75$, $p = .005$, $\eta_p^2 = .190$]. However, post-hoc t-tests did not show reliable differences across conditions when correcting for multiple comparisons (*p*s > .05).

Table 5 shows the PSS values for each condition separately. No significant prior entry effect was found in the HorizVert condition [$t(16) = 0.58$, $p = .568$], serving therefore as control condition or low-level baseline. However, PSS values were also not significant in the AngerNeut [$t(16) = -0.51$, $p = .616$] and FearNeut [$t(16) = -1.24$, $p = .232$] conditions (Figure 5C).

Post-experiment ratings confirmed that the face stimuli were perceived in line with the *a priori* emotion categories: fearful faces were perceived as more fearful compared to neutral [*t*(16) = 23.66, *p* < .001] and angry faces [*t*(16) = -19.27, *p* < .001]. Likewise, angry faces were perceived as more angry relative to neutral [*t*(16) = 16.99, *p* < .001] and fearful faces [*t*(16) = 14.54, *p* < .001], with no significant difference in perceived brightness across these three conditions (*p*s > .05) (Figure 5D).

Table 5. Median values and standard deviations (in parenthesis) of the PSS scores obtained in Experiment 4 for each condition.

Condition	PSS
Orientation	0.10 *(16.79)*
AngerNeut	0.07 *(37.92)*
FearNeut	-1.14 *(26.36)*

Note. Positive values reflect processing prioritization (i.e., prior entry) for either the vertical lines in the orientation task or the neutral face in the emotional TOJ task. Conversely, negative values refer to prior entry for either the horizontal lines or the emotional face.

5.4. Discussion

Results of Experiment 4 failed to reveal any prior entry effect for either fearful or angry faces, when these stimuli compete with neutral faces for attention selection. As was already the case for Experiments 1-3, this result could not be imputed to a lack of perceived differences between the three emotion categories, since post-experiment ratings showed clear and predictable differences between these three categories. We reasoned that the use of familiarization blocks with horizontal and vertical line gratings (i.e., non-emotional features) might have eased performance during the emotion TOJ task. However, this turned out to be a wrong prediction. Despite the introduction of these two familiarization blocks, in fact, the drop rate was still substantial (23 out of 40 participants, 58%). Hence, 23 participants had PSS values (at least in one condition) exceeding the maximum SOA range (± 100 ms). Unexpectedly, this drop rate was even higher compared to the one found in Experiment 3 (43%), where no familiarization with the vertical and horizontal gratings was introduced. Note that if we only used the data of the TOJ task performed on the gratings, this drop rate would be remarkably lower (10%), suggesting that participants encountered specific difficulties only when asked to decide whether the emotional face in the pair was shown first or not, but not when asked to decide whether horizontal or vertical line gratings appeared first. This conclusion was also reinforced by the direct comparison of the two tasks for the 17 participants included in the analyses (see Figure 5A).

We reasoned that task difficulty during the emotion TOJ might perhaps decrease if we would give more specific instructions to participants. More specifically, while in Experiments 3-4 instructions emphasized the discrimination between “emotional” and neutral faces, the use of specific response labels (angry vs. neutral or fearful vs. neutral) could possibly improve performance. A refined task set suited to bias feature-specific attention allocation towards specific emotional features (Everaert, et al., 2011; Kiefer & Martens, 2010; Spruyt, et al., 2012; Spruyt, et al., 2009), in fact, could facilitate TOJs based on these emotional features. Accordingly, in Experiment 5, we used the same stimuli and setup as in Experiment 4, but asked participants to indicate whether the first stimulus was an angry/fearful (depending on the block) or a neutral face.

6. Experiment 5

6.1. Participants

Thirty-six psychology students, who participated in none of the previous experiments, took part in Experiment 5 in exchange of course credits. Using the same exclusion criterion as above, the data of twenty volunteers were removed from the subsequent statistical analyses, leaving a final sample of 16 participants (9 women, mean age 22 years, range 18-30).

6.2. Stimuli and procedure

Stimuli were identical to Experiment 4. Similarly, two familiarization blocks with horizontal and vertical line gratings were used at the beginning of the experiment in order to familiarize participants with the TOJ task and the different SOAs. Unlike Experiment 4, however, for the subsequent emotion TOJ task participants were specifically asked to decide whether the face that appeared first in the pair was angry vs. neutral or fearful vs. neutral (2 blocks each, counterbalanced across participants). Ten practice trials with either angry/neutral or fearful/neutral stimulus pairs preceded the two experimental blocks.

Ratings for the individual faces regarding the amount of anger, fear or brightness were collected at the end of the experiment.

6.3. Results

Trials whose RTs were slower than three standard deviations from the mean were discarded (M = 0.66%, SD = 0.50).

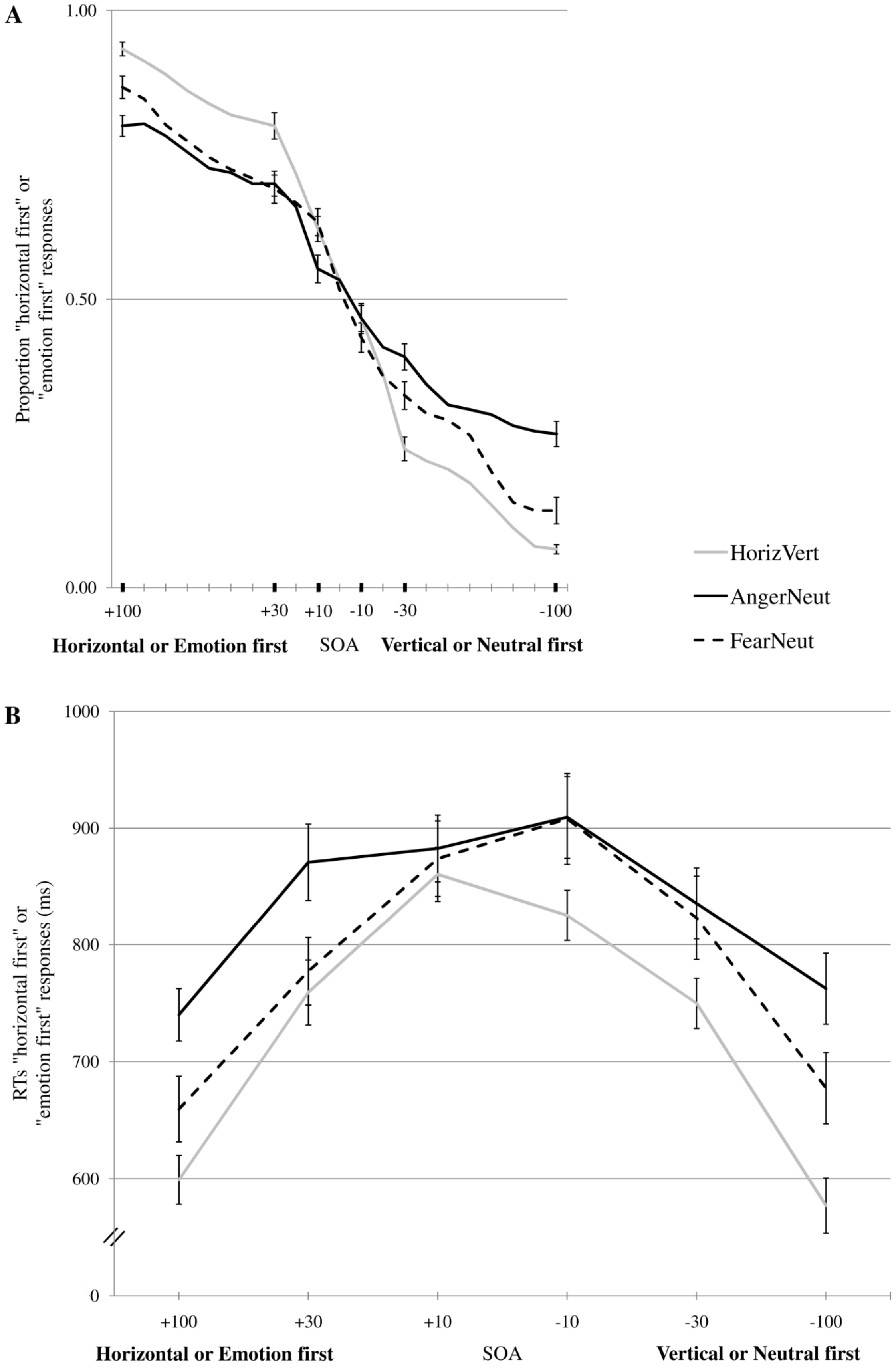
Overall, participants performed better in the familiarization task compared to the emotion TOJ task, as evidenced by flatter inverted S-Shaped psychometric functions for the AngerNeut and FearNeut conditions relative to the HorizVert condition (Figure 6A). A 6 (SOA) x 3 (condition) repeated measures ANOVA on RTs disclosed a significant main effect of SOA [$F(5, 70) = 10.60$, $p < .001$, $\eta_p^2 = .431$], with a significant quadratic trend [$F(1, 14) = 14.05$, $p = .002$, $\eta_p^2 = .501$] (Figure 6B). In addition, a significant SOA x condition interaction was found [$F(10, 140) = 2.08$, $p = .030$, $\eta_p^2 = .129$], although Bonferroni-corrected post-hoc t-tests did not reveal significant differences across conditions.

None of the PSS values (reported in Table 6) was significantly different from zero [HorizVert; $t(15) = -0.65$, $p = .524$; AngerNeut: $t(15) = -1.39$ $p = .184$; FearNeut; $t(15) = -0.68$, $p = .508$] (Figure 6C).

Post-experiment ratings confirmed that emotional faces were perceived as such by participants. Fearful faces were perceived as more fearful compared to neutral [$t(15) = 19.08$, $p < .001$] and angry faces [$t(15) = -13.45$, $p < .001$]. Similarly, angry faces were perceived as more angry than neutral [$t(15) = 15.21$, $p < .001$] and fearful faces [$t(15) = 9.77$, $p < .001$] (Figure 6D). Higher perceived brightness for emotional relative to neutral faces was also reported [anger vs. neutral: $t(15) = 5.54$, $p < .001$; fear vs. neutral: $t(15) = 3.56$, $p = .003$]. However, these subjective differences in brightness did not lead to prior entry effect for either angry or fearful faces relative to neutral faces.

Table 6. Median values and standard deviations (in parenthesis) of the PSS scores obtained in Experiment 5 for each condition.

Condition	PSS
Orientation	-2.56 *(12.78)*
AngerNeut	-9.34 *(35.97)*
FearNeut	0.76 *(22.90)*


A
1.00
0.50
0.00
Proportion "horizontal first" or "emotion first" responses
+100
+30
+10
-10
-30
-100
Horizontal or Emotion first
SOA
Vertical or Neutral first
HorizVert
AngerNeut
FearNeut
B
1000
900
800
700
600
0
RTs "horizontal first" or "emotion first" responses (ms)
+100
+30
+10
-10
-30
-100
Horizontal or Emotion first
SOA
Vertical or Neutral first

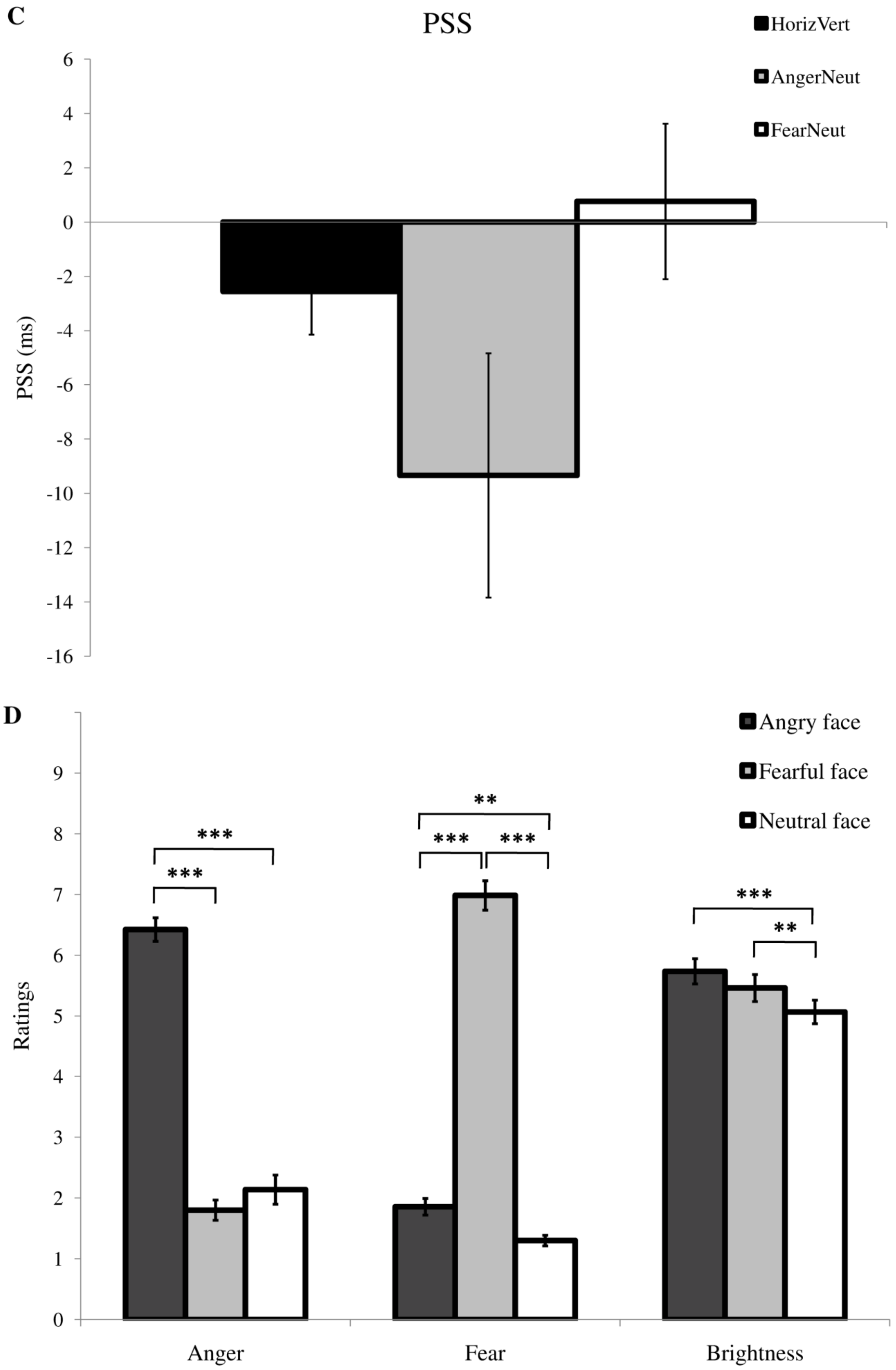


Figure 6. Results of Experiment 5. (A) Average proportion of "horizontal first" and "emotion first" responses, separately for HorizVert, AngerNeut and FearNeut conditions. (B) Average RTs of "horizontal first" and

"emotion first" responses for each condition. Participants were more uncertain at short compared to long SOAs, corroborating the results obtained for Experiments 1-4. (C) PSS values for HorizVert, AngerNeut and FearNeut conditions. No prior entry was observed. (D) Mean anger, fear and brightness ratings confirmed that emotional faces were perceived as such by participants. The subjective difference in perceived brightness did not lead to any differential behavioral effect during the main emotion TOJ task. ** $p < .01$; *** $p < .001$. Vertical bars correspond to standard errors of the means.

6.4. Discussion

Using more specific task instructions than in Experiment 4 (i.e., by explicitly mentioning either anger or fear as target emotion), we still failed to observe a reliable prior entry effect for either fearful or angry faces in this new experiment. Moreover, as was already the case for Experiment 4, the data of a high number of participants had to be discarded (drop rate of 56%) due to PSS values in the AngerNeut and FearNeut conditions that were falling outside the ±100 ms SOA range. Therefore, the use of specific emotion labels during the emotion TOJ did not lead to any gain in accuracy (Experiment 5), compared to more general task instructions only emphasizing the discrimination of emotional vs. neutral faces (Experiments 3-4). Again, these results could not be explained by difficulties to identify or recognize the different emotional facial expressions (see results of the post-experiment ratings), or the use of suboptimal SOAs and/or stimulus parameters (see results for the two familiarization blocks with the line gratings).

7. Additional analyses

When the emotion content became task-relevant (Experiments 3-5), as opposed to the mere appearance of the two faces in the pair (Experiments 1-2), we found out that many participants showed PSS values outside the SOA range (±100 ms). These "poor-performers", therefore, had to be excluded from subsequent statistical analyses (see Figure 7). This suggests that: (i) poor performers could not accurately carry out the emotion TOJ task, even though the majority of them could correctly discriminate which line gratings appeared first (Experiments 4-5), ruling out the possibility of a general perceptual deficit during TOJs; (ii) a temporal discrimination based on emotion was apparently more demanding than a temporal discrimination based on mere visual appearance. However, despite this increase in the drop rate across the five experiments, no prior entry effect for fearful or angry faces was evidenced, compared to neutral faces. We further analyzed the data of Experiments 3-5 to assess whether this increase in the drop rate (compared to Experiments 1-2) might perhaps be explained by specific personality traits and/or differences in perceiving fear or anger intensity in the negative emotional facial expressions selected in our study.

Independent paired t-tests comparing trait anxiety levels (Van der Ploeg, et al., 1979) and Need for Affect scores (Maio & Esses, 2001) between good and poor performers (as identified based on the PSS values) did not show any significant group difference (*p*s > .05) in any of the three experiments. These results could suggest that these specific predispositions did not influence performance during the emotion TOJ task, although a low inter-individual variability with these two trait-related measures might also explain this non-significant finding (see Table 1).

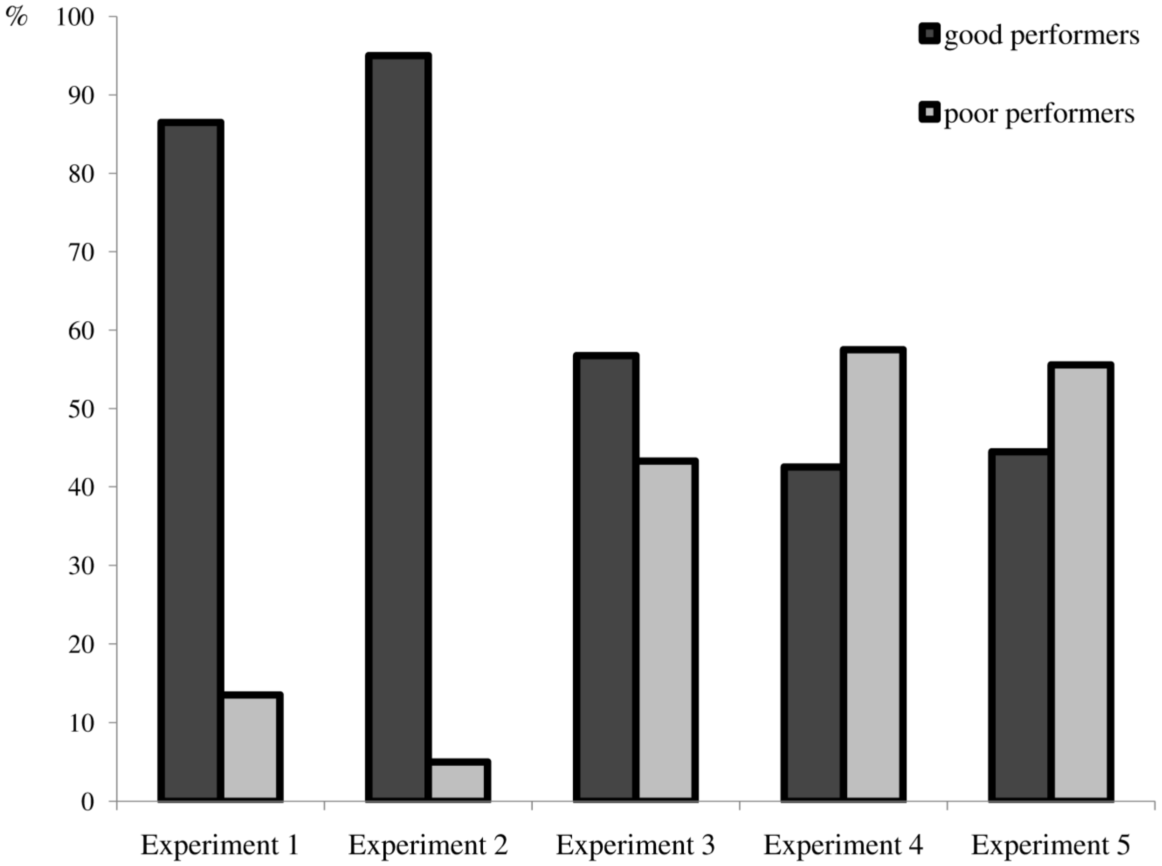


Figure 7. Percentage of good (dark grey bar) vs. poor (light grey bar) performers across the five experiments. The number of poor performers reliably increased in Experiments 3-5 compared to Experiments 1-2, suggesting that the TOJ task based on the emotional content (Experiments 3-5) was more difficult to carry out than TOJ task based on mere spatial location (Experiments 1-2).

By contrast, when comparing good vs. poor performers with regards to the ratings of the emotional faces, we found that -- only in Experiment 4 -- poor performers actually judged *neutral* faces as carrying significantly more anger and fear intensity compared to good performers [anger ratings: $t(38) = -2.48$, $p = .019$; fear ratings: $t(38) = -2.08$, $p = .046$] (Figure

8). Thus, poor performers in Experiment 4 may have perceived neutral faces as less neutral than good performers. Presumably, perceiving neutral faces as slightly more angry or fearful might turn out to be detrimental for performance during the emotion TOJ task, since the relative difference between emotional and neutral faces would be reduced for poor relative to good performers. Given that the perceived emotion intensity in the faces might modulate performance during the emotion TOJ task, we carried out an auxiliary control analysis. More specifically, for the data of Experiment 4, we included the emotional ratings of each neutral, angry and fearful face as covariates in our mixed probit regression model. Two separate analyses were conducted. First, we calculated the *difference* between the emotional ratings of the angry/fearful vs. neutral face on a trial-by-trial basis, to test the hypothesis that a higher difference in the perceived emotional intensity of the stimulus pair would result in facilitated attentional allocation towards the emotional face (i.e., its onset being perceived as first). Nonetheless, this covariate analysis did not reveal any significant PSS, either for the AngerNeut [$t(16) = -2.08$, $p = .285$] or the FearNeut [$t(16) = -0.95$, $p = .357$] condition. Next, we computed the *sum* of the emotional ratings for the two faces in the pair, in order to test whether, at the single trial level, an increased “emotional magnitude” (or overall emotionality) would somehow bias attention allocation towards the emotional faces, and in turn lead to a prior entry for either fearful or angry faces. However, this complementary covariate analysis did not show any significant (i.e., different from zero) PSS values, in any of the experimental conditions [AngerNeut: $t(16) = -0.21$, $p = .983$; FearNeut: $t(16) = -1.25$, $p = .230$]. Based on these results, we can conclude with high confidence that the absence of a reliable prior entry effect for angry or fearful faces compared to neutral faces in Experiment 4 could not be ascribed to uncontrolled trial-by-trial fluctuations in the perceived (negative) emotionality of the two faces in the pair.

8. General discussion

In this study, we used a standard TOJ task to evaluate whether negative emotion (here with a focus on fear and anger) could “automatically” draw attention, and in turn lead to a prior entry effect when competing with neutral stimuli. The added value of this task is that it enables to titrate a more direct effect of the emotional stimulus on (early) attention allocation/orienting mechanisms (Jaskowski, 1993; Stelmach & Herdman, 1991; Stelmach, et al., 1994; Titchener, 1908).

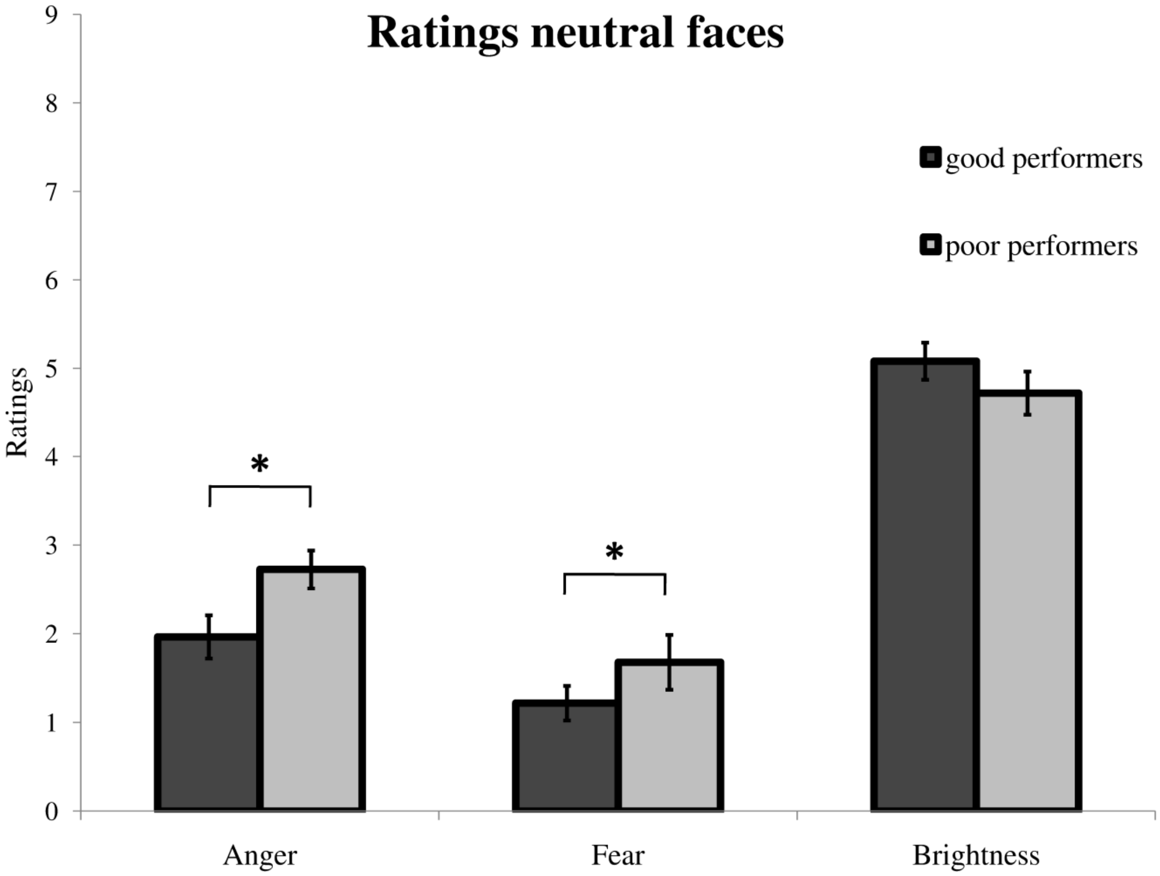


Figure 8. Ratings of perceived anger and fear conveyed by neutral faces in Experiment 4, separately for good and poor performers. This auxiliary analysis showed that poor performers rated neutral faces as significantly more angry and fearful compared to good performers, raising the possibility that poor performers might actually perceive less difference between the two faces of the pair (regarding their emotional content) during the TOJ task, compared to good performers. This might explain why they had abnormal psychometric functions for at least one condition. However, control analyses including the perceived difference in emotional content between the two faces as a covariate in the mixed probit regression model failed to find any differential prior entry effect for emotional relative to neutral faces (see main text).

Previous research using simple non-emotional stimuli already showed that attended stimuli are processed faster than unattended stimuli, an effect that can be captured in this task by a systematic bias in the perception of the onset of the attended stimulus (being perceived as first compared to the unattended one, usually occupying a non-overlapping spatial position in the visual field; see Shore & Spence, 2005; Shore, et al., 2001; Spence & Parise, 2010; Spence, et al., 2001). We sought to assess whether a similar prior entry effect could be obtained when a negative emotional facial expression directly competes for attention with a neutral one. However, results of five experiments clearly failed to corroborate this prediction, despite several incremental changes in task demands and stimulus parameters. Neither fearful nor angry faces were found to exert a systematic and differential influence on TOJs relative to neutral faces, casting doubt on the idea that these negative stimuli would “automatically” or “irrepressibly” draw (exogenous) attention, at least when TOJ tasks are used. Furthermore, this outcome is at variance with two recent studies that did report prior entry for angry faces (Fecica & Stolz, 2008; West, et al., 2009). Before we discuss in depth the possible theoretical reasons for this discrepancy and null findings, we first consider a few methodological elements that might potentially account for these results. A critical analysis of these elements allow us to conclude that there is actually more to find in these results than a simple failure to replicate earlier findings.

8.1. Lack of statistical power?

In each of the five experiments we had a decent sample size -- ranging from N = 36 in Experiment 5 to N = 40 in Experiments 2 and 4. This number is comparable with Fecica & Stolz (2008) (N = 48), and appreciably larger than West, et al. (2009) (who used a sample size ranging from 12 to 16 participants, depending on the experiment). However, the drop rate was substantial in Experiments 3-5. This was based on a stringent (Perez, et al., 2009; Spence, et al., 2001; Weiß & Scharlau, 2012) exclusion criterion (i.e., PSS value exceeding the SOA range), compared to these two previous studies where no such criterion was apparently used (Fecica & Stolz, 2008; West, et al., 2009). Note that, even after excluding these “poor performers” in Experiments 3-5 from the subsequent statistical analyses, the remaining sample size was still similar to West, et al. (2009), who previously reported a prior entry effect for (schematic and realistic) angry faces. Accordingly, an underpowered experiment appears unlikely to account for the present null findings. On the contrary, the use of this stringent criterion enabled us to ascertain that only participants showing a normal sensitivity to the main SOA manipulation (as evidenced by an inverted S-shaped psychometric function) were eventually included in the final sample. It is therefore puzzling to observe that, for these good performers -- who did not differ in terms of personality traits from poor performers --, no differential influence of the emotional content of the faces was evidenced during TOJs with these faces, despite the use of increasingly more specific task demands between Experiment 4 and 5.

8.2. Perceptual confounds?

Fecica & Stolz (2008) and West, et al. (2009) made primarily use of schematic neutral and emotional faces in order to explore whether emotional factors might modulate early attention allocation, as indicated by prior entry effects for these emotion stimuli during the TOJ task.

The use of schematic faces is consistent with earlier studies (e.g., Fox, et al., 2000; Fox, et al., 2001; Öhman, Lundqvist, et al., 2001) that have already investigated (mainly using visual search tasks) the interplay between attention and emotion control systems. While these schematic faces provide the added value to potentially control for perceptual differences between emotional and neutral expressions, these stimuli clearly lack ecological validity (e.g., Horstmann & Bauland, 2006). More importantly, recent research has shown that specific low-level features of schematic face stimuli (e.g., the curvature of the mouth) can actually bias perception and attention in a bottom-up fashion, without the need to postulate any mediation by specific emotion brain mechanisms (Becker, Horstmann, & Remington, 2011; Mak-Fan, et al., 2011; VanRullen, Reddy, & Koch, 2004; VanRullen, Reddy, & Li, 2005; Wolfe & Horowitz, 2004). Moreover, schematic faces are thought to exaggerate facial features, and the representation of the intended emotion may be equivocal (Horstmann & Bauland, 2006). In addition, schematic face stimuli have been shown to produce artificially greater behavioral effects (Hietanen & Leppanen, 2003). Accordingly, the existing evidence of a prior entry effect for angry faces obtained with schematic faces (i.e., Fecica & Stolz, 2008, and Experiments 1-4 in West, et al., 2009) requires some careful evaluation and interpretation regarding the true emotional nature of this early attention orienting effect. Of note, to circumvent these limitations, in Experiments 5-6 West, et al. (2009) confirmed a significant prior entry effect for angry faces using realistic photographs of angry and neutral faces (selected from the same standardized database as used in this study; Ekman & Friesen, 1976). In these conditions, an even larger and significant PSS value was found -- indicating a systematic early attentional capture towards emotional stimuli -- compared to the one obtained with schematic angry faces (Experiments 1-4). However, a careful evaluation of the methods section shows that West, et al. (2009) only used four different face identities (two men and two women) and thus a limited number of face pairs (between 12 and 16 according to our estimation). Presumably, these experimental conditions may have favored the use of a strategy based on perceptual, non-emotional cues to carry out the emotion TOJ task. Importantly, the degree of similarity between stimuli has been shown to reliably influence performance during visual search (Duncan & Humphreys, 1989). In our study, therefore, we opted for the use of a larger number of different face identities (four women and six men), as well as a large number of face pairs (90 per condition) in order to avoid the (implicit) use of a strategy based on specific perceptual cues (e.g., curvature or contrast). The added value of this alternative procedure is that the degree of perceptual dissimilarity between the two faces of the pair was always uninformative for each and every trial, thus preventing participants to use this information to perform the TOJ task. Note that, in Experiment 4, we also ran additional control analyses taking the perceived emotion dissimilarity into account (as a covariate in the probit regression model) but still failed to find evidence for a direct contribution of the perceived (subjective) perceptual difference between the two faces in the pair as potentially accounting for a prior entry for fearful or angry faces. Furthermore, this discrepancy between our results and the findings reported by West, et al. (2009) cannot easily be explained by different stimulus parameters or task demands because, from Experiment 2 onwards, we took special care in matching as much as possible the face stimulus size and (horizontal) eccentricity with the values reported in West, et al. (2009). We also collected additional ratings from the participants in each experiment to make sure that they could reliably perceive fearful, angry and neutral faces as such, and the results for these ratings (across the five experiments) unambiguously confirmed this conclusion. Accordingly, the lack of prior entry for either fearful or angry faces compared to neutral faces cannot easily be ascribed to the use of ambiguous or mildly emotional face stimuli.

8.3. Different experimental procedures?

Given that we explicitly set up our TOJ task based on previous studies (Fecica & Stolz, 2008; West, et al., 2009), it appears unlikely that other uncontrolled factors related to the procedure or the stimulus set could immediately account for the present null findings. Specifically, we included the critical face stimuli in dedicated placeholders located on both sides relative to central fixation, which were subsequently masked by a uniform noise pattern until response (similarly to West, et al., 2009). This procedure ensured that bottom-up effects related to other visual features than the face did not contaminate the performance during the TOJ task. Moreover, the use of placeholders provided spatial cues to participants regarding the two opposite positions in the visual field where the faces would appear each time, limiting drifts of spatial and temporal attention towards non-informative portions of the visual field. Furthermore, we used SOAs of 10, 30, and 100 ms, comparable with 17, 34 and 100 ms in Fecica & Stolz (2008). In addition, by using two response buttons aligned along a vertical axis, we prevented the occurrence of (spatial) stimulus-response compatibility effects (Schneider & Bavelier, 2003), particularly in Experiments 1-2 where a left-right temporal order judgment was required.

It is important to note that the failure to observe reliable prior entry effects for emotional relative to neutral stimuli was not limited to a specific (negative) emotion category. In fact, we observed no attentional capture, neither for fearful (Experiments 1-5) nor for angry faces (Experiments 2-5), despite the fact that several studies, using a variety of experimental paradigms, have reported early orientation of attention towards these stimuli (Anderson, et al., 2003; Bocanegra & Zeelenberg, 2011a; Frischen, Eastwood, & Smilek, 2008; Phelps, et al., 2006; Pourtois, et al., 2004; Pourtois, et al., 2012; Sander, et al., 2007; Whalen, et al., 2004). Accordingly, it is unlikely that the perceived relevance of the threat displayed in the face -- either indirect in the case of fear or more direct in the case of anger -- may have contributed to the differential allocation of attention towards these facial stimuli, and thus this factor cannot immediately account for the non-significant findings reported here. In addition, the results of the offset task (see Experiments 1-3) showed no evidence for a prolonged attentional disengagement for either fearful or angry faces compared to neutral faces, which could potentially have masked a preferential early orienting of attention towards emotional relative to neutral faces in the onset task.

Finally, changes in task instructions did not have any impact on the size or expression of the putative prior entry effect for emotional compared to neutral stimuli. In Experiments 1-2 participants were required to indicate whether the first face in the pair appeared on the left or right side relative to fixation, thereby focusing on spatiotemporal properties of the stimuli exclusively. Thus, the emotional content of the faces was not immediately informative and, as a consequence, it might be strategically useful to actually filter it out in order to resolve the task (Everaert, et al., 2011; Kiefer & Martens, 2010; Pessoa, Kastner, et al., 2002; Pessoa, Kastner, & Ungerleider, 2003; Spruyt, et al., 2012; Spruyt, et al., 2009; Spruyt, et al., 2007). However, no prior entry for emotional faces was observed when participants were explicitly requested to judge whether the emotional or the neutral face appeared first (Experiments 3-4), nor when specific emotion labels (i.e., angry or fearful) had to be used to perform the TOJ task (Experiment 5). Therefore, the use of task sets in which the processing of specific features of the stimuli (i.e., emotional valence) was promoted did not lead to enhanced attentional capture for emotional compared to neutral stimuli.

8.4. PSS as an estimate of prior entry?

In our study, visual prior entry was assessed, following standard practice, by computing the PSS according to the dominant procedure in literature, that is calculating the intercept and slope of a linear regression applied on the inverse normalized proportion of responses (Keetels & Vroomen, 2005; Nava, Bottari, Zampini, & Pavani, 2008; Perez, et al., 2009;

Shore, et al., 2001; Sinnett, et al., 2007; Spence & Parise, 2010; Spence, et al., 2001; Vatakis & Spence, 2006; Zampini, et al., 2005). Importantly, we calculated each participant's intercept and slope with estimated random effects, in order to be able to control for the correlation of measurements within the same subject. By comparison, Fecica & Stolz (2008) did not report the PSS values, making any systematic comparison between their findings and our results (for Experiments 1-2) particularly difficult. Likewise, West, et al. (2009) reported that their PSS was calculated by "determining the intercept at the 50% point on the regression line of each participant's TOJ function" (see p. 1035). However, based on this definition, it is unclear whether these authors initially applied the inverse normalization step described above or not. If we assume that they did not, this could potentially account for the difference between their earlier findings and our news results.

8.5. First (apparent) motion, then emotion?

The results of these five experiments do not support the notion of an automatic capture or early orienting of attention towards the location occupied by threat-related stimuli, as opposed to neutral ones. This outcome is intriguing, especially for Experiments 3-5 where participants were explicitly asked to process the emotional content of the two faces in the pair. However, it turned out that they could still perform the task (at least good performers) without being influenced (in a bottom-up fashion) by this emotion dimension. We hypothesize that these participants perhaps did not show any prior entry effect for negative emotional faces because they first relied on a non-emotional feature to perform the TOJ task. Presumably, the systematic difference between the two face onsets may have produced the compelling impression of apparent motion on the screen, a phenomenon previously described in the literature as "illusory line motion" (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998; Hikosaka, Miyauchi, & Shimojo, 1993; Schmidt, 2000). It appears plausible to consider, given the null findings reported in Experiments 3-5, that participants primarily used this motion cue in order to decide, during a second stage (maybe based on post-perceptual processes, including short term memory), whether the face stimulus triggering this illusory motion (either towards the left or right side) was emotional or not. In other words, as soon as the perceptual system "registered" which stimulus appeared first based on this motion cue, a second, post-perceptual process was perhaps taking place, enabling participants to decide whether this face stimulus had an emotional or neutral content. According to this framework, the processing of the emotional content of the face stimuli would not be early and automatic, but it would likely take place at post-perceptual stages of stimulus processing, once (spatial and temporal) attention has already been allocated either to the left or right side (as determined primarily by the SOA manipulation which created the illusion of apparent motion between the two locations). Thus, in this model, early attention orienting is not biased by negative emotional (i.e., threat-related) stimuli because a concurrent, more advantageous strategy enables participants to resolve the TOJ task using primarily non-emotional, motion-related perceptual cues. One way to test this prediction at the empirical level would be to extensively train participants to actively use the emotion information conveyed by the face stimuli (as well as actively ignore any motion-related cue), similarly to previous psychophysics studies documenting improvements in TOJ performance with experience (Hirsh & Sherrick, 1961) and training-induced improvement (Barsz, 1996; Bernasconi, Grivel, Murray, & Spierer, 2010; Merzenich, et al., 1996; Mossbridge, Fitzgerald, O'Connor, & Wright, 2006; Nickerson & Freeman, 1974; Warren, 1974). Alternatively, a direct disruption of the apparent motion could prevent participants from using (even implicitly) this specific cue in order to resolve the TOJ task. For example, this could be achieved by including a moving dot pattern in the background (as opposed to a uniform static background in the present case).

More generally, the results of this study challenge the notion that threat-related stimuli “automatically” capture attention, and hence lead to a prior entry effect during TOJs when competing with neutral stimuli (West, et al., 2009). Instead, our findings suggest that even though the emotional content of the faces may be directly task-relevant, as long as other perceptual cues (probably motion) can be used by participants to perform the TOJ task, emotion does not bias early stages of attention allocation. Further studies are needed to establish whether, when controlling for these non-emotional perceptual factors, emotion can reliably prioritize the allocation of attention in a reflexive way.

CHAPTER 3

Brain Dynamics of Upstream Perceptual Processes Leading to Visual Object Recognition: A High Density ERP Topographic Mapping Study[1]

Recent studies suggest that visual object recognition is a proactive process through which perceptual evidence accumulates over time before a decision can be made about the object. However, the exact electrophysiological correlates and time-course of this complex process remain unclear. In addition, the potential influence of emotion on this process has not been investigated yet. We recorded high density EEG in healthy adult participants performing a novel perceptual recognition task. For each trial, an initial blurred visual scene was first shown, before the actual content of the stimulus was gradually revealed by progressively adding diagnostic high spatial frequency information. Participants were asked to stop this stimulus sequence as soon as they could correctly perform an animacy judgment task. Behavioral results showed that participants reliably gathered perceptual evidence before recognition. Furthermore, prolonged exploration times were observed for pleasant, relative to either neutral or unpleasant scenes. ERP results showed distinct effects starting 280 ms post-stimulus onset in distant brain regions during stimulus processing, mainly characterized by: (i) a monotonic accumulation of evidence, involving regions of the posterior cingulate cortex/parahippocampal gyrus, and (ii) categorical recognition effects in a region of the

---

[1] Schettino, A., Loeys, T., Delplanque, S., & Pourtois, G. (2011). Brain dynamics of upstream perceptual processes leading to visual object recognition: A high density ERP topographic mapping study. *Neuroimage, 55*(3), 1227-1241.

dorsal anterior cingulate cortex. These findings provide evidence for the early involvement, following stimulus onset, of non-overlapping brain networks during proactive processes eventually leading to visual object recognition.

1. Introduction

Visual object recognition is a fast, accurate and effortless process. Despite swift or dramatic variations in the retinal input (i.e. due to modifications in orientation, size, appearance, viewpoint, or context), human's ability to readily detect and recognize a multitude of visual objects in the environment is hardly challenged (Bar, 2003; Biederman & Bar, 1999). Many studies and models have emphasized the sequential property of visual computations leading to perceptual decision making, from the analysis of sensory information to the selection of the behavioral outcome that best maximizes the expected utility (Biederman, 1987; Marr, 1982; Opris & Bruce, 2005). According to the dominant framework, the visual system evaluates in a probabilistic fashion the available information about various features of the input image, thereby making inferences about its content and preparing possible courses of action (Hegdé, 2008). Classical neurophysiological models postulate that fundamental visual features of the input image are initially processed in lower-level cortical areas of the occipital and inferotemporal cortex, after which they are used to generate an abstract visual representation of the object. Recognition (and subsequent motor execution of the congruent response) is achieved when this representation is successfully matched with templates stored in memory (Biederman, 1987; Palmer, 1999; Ranganath & Rainer, 2003; Riesenhuber & Poggio, 1999; Ungerleider & Mishkin, 1982).

The predominant role of *bottom-up* perceptual processes leading to recognition is clearly stressed in these hierarchical models. However, *top-down* modulatory effects may also influence visual object recognition mechanisms. They include selective attention (Blair, et al., 2009; Treisman & Kanwisher, 1998), task relevance (Egner & Hirsch, 2005), prior probability of encounter (Summerfield & Egner, 2009), working memory (Ranganath, et al., 2004), contextual information (Bar, 2004; Oliva & Torralba, 2007), as well as the monitoring of the decision's outcome (Ridderinkhof, et al., 2004; Ullsperger, et al., 2004). Hence, the rapid bottom-up processes leading to visual object recognition could be assisted by the online activation of abstract information -- stored in higher-level brain regions -- primarily recruited to speed up the concurrent ongoing processing in lower-level visual areas, with the aim to limit the number of computations necessary to eventually identify an object (Bar, et al., 2006; Ganis, et al., 2007).

Recent theoretical accounts emphasized the role of expectations in visual recognition (Bar, 2009b; Kersten, et al., 2004; Summerfield & Egner, 2009; Yuille & Kersten, 2006). In these models, it is hypothesized that the active use of prior information about the most probable visual percept in the forthcoming sensory environment is at work in order to guide the rapid acquisition of diagnostic visual information (i.e., invariant and expected aspects in the environment do not need to be processed thoroughly), as well as to facilitate the interpretation of ambiguous stimuli. Predictive coding models of visual recognition (Friston, 2005; Friston & Kiebel, 2009; Grossberg, 2009; Rao & Ballard, 1999; Serences, 2008) argue that feed-forward information coming from early visual areas is compared, at each stage of the visual processing hierarchy, to top-down expectations whose aim is to predict the responses at the next lower level, primarily through recurrent or feedback connections (see also Di Lollo, Enns, & Rensink, 2000; Enns & Lleras, 2008). If this comparison results in an erroneous output, an error signal is sent back to the higher level via feedforward connections. This error signal is then used to correct the estimate of the input signal at each level (Rao & Ballard, 1999; Serences, 2008). Expected and observed information are iteratively adjusted until the visual system is able to settle on a single perceptual interpretation of the sensory input (Summerfield & Egner, 2009). A plausible neural mechanism underlying the triggering of this top-down facilitation in object recognition has recently been proposed by Moshe Bar (Bar, 2003, 2004, 2007, 2009b). According to this model, a partially analyzed version of the input image, mainly composed of low spatial frequency (LSF) information, is projected rapidly from early visual cortex directly to orbitofrontal (OFC) and parahippocampal (PHC) areas, possibly via a fast dorsal magnocellular route. In the PHC, this blurred image activates the most probable (experience-driven) guesses about the context frame that needs to be triggered. This contextual information is projected to the infero-temporal cortex, where a set of visual associations corresponding to the relevant context is activated. In parallel, the visual information conveyed by the same scene -- here the allocation of attention is on the target object -- would be sufficient to rapidly activate themost likely interpretations of the input image in the OFC. The integration of the representations of the specific context and the candidate interpretations of the target object would in turn result in the reliable selection of a single identity, which can be further refined with specific detail gradually conveyed by high spatial frequency (HSF) information (Bar, 2004). Consistent with this framework, Peyrin, et al. (2010) combined fMRI and ERPs to explore the prioritization of LSF in the processing of visual input, and found that higher order areas in frontal and temporo-parietal regions responded more to LSF stimuli when presented first, whereas occipital visual cortex responded more to HSF stimuli when presented after LSF stimuli (Bar, 2003; Bullier, 2001; Hegdé, 2008). Hence, the use of different imaging methods confirmed a “coarse-to-fine” processing of visual input (Hegdé, 2008; Lomber, 2002; Navon, 1977; Sanocki, 1993; Schyns & Oliva, 1994): the quick processing of LSF in higher order frontal or temporal brain regions could directly influence scene recognition by biasing perceptual processes in object-selective visual areas.

Expectations may lower the threshold that needs to be overcome to make a decision in favor of one option, similar to what accumulator models of decision making propose (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). Such models have proven highly effective in describing performance in recognition memory (Ratcliff, 1978), economic decisions (Sanfey, et al., 2006), semantic (Ratcliff & McKoon, 1982) and lexical (Ratcliff, et al., 2004) tasks, as well as sensory discrimination (Gold & Shadlen, 2007). According to these models, incoming information is accumulated over time, evaluated and assigned to a response option: the actual decision is made when evidence in favor of one of the two responses exceeds a threshold. This framework has been successfully applied in studies on nonhuman primates (Hanes & Schall, 1996; Kim & Shadlen, 1999; Shadlen & Newsome, 2001; Tanaka, 1996), eventually providing a useful mean to link neurophysiology with behavior. Important hints on the brain regions underlying proactive perceptual processes leading to visual object recognition in humans have also been obtained more recently, using mainly fMRI and gradual or progressive stimulus revelation tasks (Carlson, et al., 2006; Heekeren, et al., 2004; Heekeren, et al., 2008; James, et al., 2000; Ploran, et al., 2007; Wheeler, et al., 2008). Among these studies, Ploran, et al. (2007) used an elegant procedure enabling a gradual revelation of masked stimuli over eight discrete and consecutive steps. Participants had to press a button as soon as they could identify the pictures' content with a reasonable degree of confidence. Interestingly, Ploran, et al. (2007) found that different brain areas showed distinct effects during this progressive revelation task. Three different patterns of brain activation were identified. For several posterior occipital regions, activity increased monotonically as a function of the amount of visual information entering the visual system (sensory processors). By contrast, a gradual increase in activity with a distinctive peak at the time of recognition was found in inferior temporal, frontal and parietal regions (including the bilateral fusiform gyrus and the DLPFC), consistent with an accumulation of evidence process which may be necessary to recognize the identity of the target object (accumulators). Finally, in many regions of medial frontal cortex (including the dorsal region of the anterior cingulate cortex, dACC, and the anterior insula), activity remained close to baseline until the moment of actual recognition, suggesting their involvement in decision-related processes that accompany overt visual object recognition. These latter fMRI results are important, as they inform about upstream brain mechanisms leading to visual object recognition and their different functions, with non-overlapping areas involved in accumulation of evidence vs. moment of recognition (Ploran, et al., 2007). However, since these results were obtained using fMRI techniques, the electrophysiological correlates and actual time-course of these upstream perceptual or decision-related effects during visual object recognition remain unclear.

The goal of this study was to use a novel stimulus revelation task and explore, using high density EEG in healthy adult participants, the precise electrophysiological correlates of upstream processes leading to visual object recognition. More specifically, we sought to investigate whether different ERP response profiles could be evidenced, consistent with the assumption of accumulation of evidence prior to visual object recognition subserved by posterior occipital or temporal brain regions, in comparison with medial frontal areas which might show categorical moment-of-recognition effects (see Ploran, et al., 2007). For this purpose, we designed a new task enabling a progressive revelation of the stimulus content, while neural events prior to actual recognition were investigated on a trial-by-trial basis using the millisecond time resolution provided by EEG. Based on the psychological models and neuroscience evidence reviewed above, we hypothesized that upstream perceptual effects leading to visual object recognition could be twofold (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008). First, we surmised ERP effects reflecting the accumulation of perceptual evidence over time, with the main neural generators being localized in posterior brain regions, including the occipital and temporal cortices (Ploran, et al., 2007). Consistent with this view, a linear relationship between the electrophysiological signal and the amount of visual input was expected. In contrast, we predicted that medial frontal brain regions (e.g. dorsal ACC, see Ploran, et al., 2007) might also provide an important mechanism at stake during visual object recognition, showing moment-of-recognition effects characterized by a non-linear and abrupt change in the amplitude of the ERP signal occurring close to actual recognition. As a secondary question, we also investigated whether the emotional content of the stimulus/scene might influence these upstream perceptual processes leading to recognition, consistent with previous imaging studies showing reciprocal interaction effects between regions involved in the processing of emotions (including the amygdala) and more posterior regions implicated in visual object recognition (Sabatinelli, et al., 2009; Vuilleumier, 2005). Hence, we presented participants with standard neutral and emotional scenes (whose content was progressively revealed) and tested for a differential neural effect likely triggered by the emotional content prior to actual recognition. Whereas previous studies have found reliable modulation of early and late ERP components following stimulus onset as a function of the emotional content of the stimulus (either for valence or arousal dimensions; see Carretié, Hinojosa, Martin-Loeches, Mercado, & Tapia, 2004; D'Hondt, et al., 2010; Delplanque, Lavoie, Hot, Silvert, & Sequeira, 2004; Olofsson, Nordin, Sequeira, & Polich, 2008; Schupp, et al., 2000; Schupp, Junghöfer, Weike, & Hamm, 2003a, 2003b), to our knowledge, no ERP study has systematically explored whether upstream or anticipatory perceptual effects during visual object recognition may be influenced by the rapid decoding of the emotional content of the input stimulus, likely based on the selective processing of low-spatial frequency/impoverished visual cues. Accordingly, the goal of our study was also to verify whether the rapid processing of the emotional content of the scene (presumably based on LSF cues; see Bar, 2003, 2004; Pourtois, Dan, et al., 2005; Vuilleumier, Armony, Driver, & Dolan, 2003) could alter upstream brain processes underlying proactive mechanisms of visual object recognition.

2. Methods

2.1. Participants

Nineteen psychology students (14 women, mean age 21 years, range 17-33) were recruited to freely participate in the study, which was approved by the local university ethical committee. All participants were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders. All volunteers gave informed written consent prior to their participation, and were paid 20€.

2.2. Stimuli

Two hundred and thirty-four pictures were selected from the International Affective Picture System (IAPS; Lang, Bradley, & Cuthbert, 2005), a standardized database containing emotionally-evocative pictures that depict objects and scenes across a wide range of categories and situations. The IAPS manual provides normative values for the basic dimensions of emotion -- including arousal and valence -- as rated by the Self-Assessment Manikin (SAM) on a scale from1 to 9 (Bradley & Lang, 1994). Since there are gender differences in both valence and arousal ratings, we selected two sets of pictures in order to balance the arousal levels of the emotional pictures across male vs. female participants. Each set consisted of 138 pictures. Among these, 42 were shared between male and female participants. The pictures were divided into three emotion categories, according to their pre-defined valence scores: *neutral*, *unpleasant* and *pleasant* (see Table 1). For each category (N=46), half of the pictures contained living objects (i.e., human beings or animals) while the other half did not (i.e., landscapes or artifacts)[2].

Table 1. Mean values and standard deviations (in parenthesis) of valence and arousal scores of the selected IAPS pictures.

Picture set	Valence	Arousal
Men		
Neutral	5.12 *(1.29)*	3.90 *(1.88)*
Unpleasant	3.12 *(1.62)*	4.93 *(2.14)*
Pleasant	7.31 *(1.57)*	5.27 *(2.30)*
Women		
Neutral	4.96 *(1.30)*	3.34 *(1.96)*
Unpleasant	2.75 *(1.62)*	5.27 *(2.19)*
Pleasant	7.31 *(1.60)*	4.92 *(2.37)*

Note. Scores range from 1 to 9. Independent samples t-test confirmed no significant difference between women's and men's picture sets, both for valence [$t(274) = -0.31$, $p = .760$] and arousal [$t(274) = -1.18$, $p = .238$] scores. In addition, no difference was found between pictures containing living [valence: $t(136) = 0.10$, $p = .920$; arousal: $t(136) = -0.71$, $p = .478$] and non-living [valence: $t(136) = -0.54$, $p = .590$; arousal: $t(136) = -1.01$, $p = .315$] objects.

We explicitly selected pictures that were neither highly pleasant (i.e., erotic scenes) nor highly unpleasant (i.e., mutilations), because such pictures could lead to specific emotion

---

[2] IAPS image codes. WOMEN PICTURE SET: practice (living) 2220, 2635, 4631, 4651, 4669; practice (non-living) 7002, 7009, 7161, 7590, 7820; neutral (living) 1616, 2190, 2191, 2381, 2383, 2385, 2393, 2480, 2495, 2570, 2595, 2749, 2840, 2890, 4250, 4255, 4310, 9070, 3550.2, 6570.2, 2516, 2702, 8192; neutral (non-living) 6150, 7006, 7010, 7020, 7031, 7035, 7036, 7041, 7050, 7130, 7160, 7175, 7179, 7185, 7187, 7217, 7233, 7235, 7830, 7950, 7025, 7500, 7705; negative (living) 1052, 1201, 1525, 1932, 2276, 2490, 2694, 2715, 2753, 3181, 3300, 4621, 6311, 9041, 9046, 9331, 9404, 9417, 2055.1, 2900.1, 1274, 9160, 9592; negative (non-living) 2692, 5971, 6020, 6230, 6241, 6610, 6800, 9000, 9280, 9320, 9340, 9373, 9470, 9471, 9495, 9611, 9620, 9622, 9630, 9830, 9001, 9290, 9621; positive (living) 1604, 1610, 1721, 2209, 2345, 4510, 4538, 4572, 4626, 4640, 4660, 5621, 8034, 8041, 8080, 8200, 8370, 8470, 8490, 8496, 1740, 4531, 8330; positive (non-living) 2791, 5220, 5450, 5480, 5551, 5594, 5600, 5779, 5780, 5891, 5982, 5994, 7200, 7280, 7350, 7470, 7480, 7545, 7580, 8510, 5300, 7390, 8502. MEN PICTURE SET: practice (living) 1112, 2210, 2214, 2393, 3210; practice (non-living) 7030, 7035, 7050, 7185, 7235; neutral (living) 1101, 1230, 1390, 1935, 1945, 2005, 2220, 2441, 2487, 2514, 2516, 2690, 2749, 2830, 2870, 4503, 4520, 4532, 9700, 2745.1, 1310, 1321, 2635; neutral (non-living) 5390, 5731, 6150, 6800, 7009, 7010, 7020, 7036, 7038, 7041, 7090, 7160, 7161, 7179, 7184, 7186, 7207, 7211, 7233, 7283, 7002, 7100, 7285; negative (living) 2053, 2095, 2141, 2710, 2750, 3181, 6243, 6311, 6312, 6315, 6510, 6530, 6821, 6838, 7380, 9250, 9500, 2352.2, 3550.1, 6570.1, 2683, 2900, 9180; negative (non-living) 6260, 6300, 9000, 9008, 9010, 9090, 9280, 9290, 9301, 9320, 9340, 9470, 9471, 9600, 9611, 9620, 9630, 9830, 9911, 9912, 9001, 9360, 9621; positive (living) 1463, 1811, 1999, 2057, 2208, 2209, 2216, 4001, 4599, 4641, 4653, 4656, 4676, 4681, 4689, 4810, 8116, 8120, 8200, 8496, 1460, 2340, 8470; positive (non-living) 5260, 5270, 5480, 5594, 5600, 5660, 5700, 5760, 5780, 5910, 5982, 7200, 7230, 7270, 7330, 7480, 7580, 8170, 8502, 8510, 7260, 7350, 7470.

reactions which may be different between male and female participants (Lithari, et al., 2010; Proverbio, Adorni, Zani, & Trestianu, 2009; Sabatinelli, Flaisch, Bradley, Fitzsimmons, & Lang, 2004). We also selected ten additional neutral pictures that were used during the practice session (see footnote) and were not included in the subsequent statistical analyses. Furthermore, 18 pictures were scrambled and their content made meaningless. These scrambled pictures were eventually used as “catch” trials to ensure that participants correctly attended to the content of the pictures before taking a decision.

The pre-selected original IAPS pictures (1024×768 pixels, corresponding to 20°×15° of visual angle at a distance of 75 cm) were first converted to grayscale. Each picture was then bandpass-filtered (using ImageJ v1.44 software; http://rsb.info.nih.gov/ij/) according to the spatial frequency bands put forward by Delplanque, N'diaye, Scherer, & Grandjean (2007) (see Table 2). This procedure resulted in six different levels of filtering for each and every picture. Each of these levels reliably differed from one another depending on the actual content of low and high spatial frequency information (see Fig. 1A).

Table 2. The six frequency bands used to filter the IAPS images selected in our study.

Image Level	Frequency Band (pixels/cycle)	
1	256-512	LSF
2	128-512	
3	64-512	↓
4	32-512	
5	16-512	
6	0-512	HSF+LSF

Note. HSF: high spatial frequencies; LSF: low spatial frequencies.

2.3. Procedure

Participants were tested individually in a small, dimly lit room, and seated at 75 cm in front of a 19″ CRT computer screen (refresh rate 100 Hz). They were first asked to complete a practice/familiarization block containing 10 trials. Then, they started the experimental session, which was divided into four blocks (separated by a one-minute pause), each containing 67-68 trials. Each trial had the same underlying structure (see Fig. 1A). It began with a fixation cross displayed for 250 ms in the center of the screen. Then the first (blurred) image level of a given picture was presented for 500 ms, followed by a 250 ms blank screen. Next, the second image level of the same picture (containing more HSF information) was immediately presented for 500 ms, plus 250 ms blank screen, and the same procedure was repeated until the presentation of the sixth image level (i.e., intact/unfiltered picture). Hence, this procedure resulted in a progressive and predictive revelation of the image content by adding, in a stepwise fashion, high spatial frequency information to an initial blurred and meaningless picture. The inter-trial interval (ITI)was set at 1000 ms. Participants were instructed to provide two consecutive responses. First, they were asked to press the spacebar key (on a standard AZERTY keyboard) with their dominant right hand as soon as they felt they could decide, with sufficient confidence, whether the scene contained a living object or not (Response1). Pressing the spacebar key immediately interrupted the presentation of the stimuli. 500 ms after pressing the spacebar, participants were required to validate their choice and to perform a two-alternative forced choice task. They were asked to press the "L" key of the keyboard if the scene contained a living object, or alternatively the "N" key if it did not contain any living object (Response2). This dual registration procedure enabled us to timely separate early recognition effects (Response1) from the overt discrimination of the scene (Response2). Importantly, the actual discrimination (Response2) was required to distinguish correct from incorrect early key presses (when looking retrospectively at ERP data recorded around the onset of Response1; see below). Hence, ERP analyses were primarily focused on neural events taking place prior to Response1, when the actual visual discriminations (based on Response2) turned out to be accurate and errors were removed from the analyses. Because Response1 always required a simple key press, this procedure minimized the potential contamination of ERP data by the activation of competing responses (living vs. non-living). Participants were instructed to find a good balance between high accuracy and the execution of speeded responses. Note that they were not encouraged to respond before the presentation of the last/sixth (unfiltered) picture. For catch trials (i.e., scrambled pictures), a non-living response (Response2) was expected.

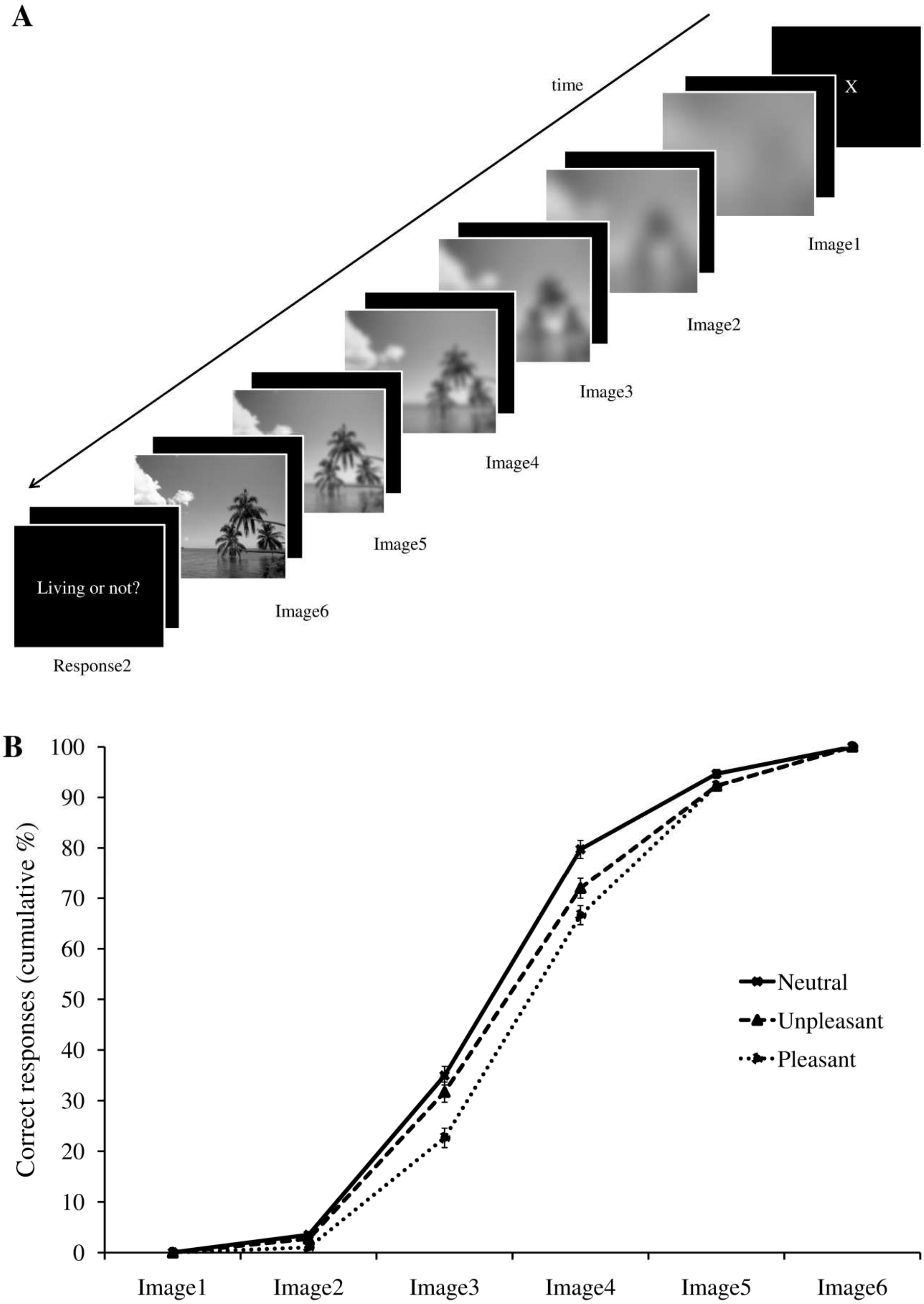


Fig. 1. (A) Example of stimuli, and progressive unfolding task. Neutral, unpleasant and pleasant IAPS scenes (not shown here for copyright reasons) were presented to participants in random order. For each trial, the content of the scene was progressively revealed in six successive steps by adding high spatial frequency information to the initial blurred (meaningless) picture. Each image level was presented for 500 ms, followed by a 250 ms blank screen. Participants were required to press the spacebar key as soon as they could decide whether the scene contained a living object or not, thereby interrupting the sequence (Response1). Next, they

validated their response choice by pressing one out of two predefined response keys (Response2). (B) Cumulative percentages of correct Response1 as a function of the six image levels, separately for each emotion category. Vertical bars correspond to standard errors of the means. The results showed that participants were able to make the animacy judgment task reliably earlier (i.e., less perceptual evidence needed) for neutral (solid line), relative to unpleasant (dotted line) or pleasant (dashed line) scenes. The shift of the psychometric function for unpleasant compared to pleasant scenes was also significant (see behavioral results).

Since we were primarily interested in ERP effects which might foreshadow the actual detection and recognition of a scene, we had, as a prerequisite, to include enough trials per condition (neutral, unpleasant and pleasant) to be able to eventually compute reliable ERP waveforms per condition and for each image level separately. However, we only had a limited set of pictures which were balanced with regard to the living vs. non-living attribute. For this reason, unbeknown to participants, each picture was presented twice during the experimental session. There was a random and unpredictable time lag (i.e., 4-15 intervening images, $M = 10$, $SD = 3$) between first presentations and repetitions. As expected, behavioral results showed significant earlier scene recognition for repetitions, relative to first presentations, and this priming effect was the same for neutral, pleasant and unpleasant scenes (all ps < .05). However, separate analyses of the ERP data for these two conditions (first presentations vs. repetitions) did not reveal any significant difference. Accordingly, they were collapsed in the statistical analyses to increase the signal-to-noise ratio.

Stimulus presentation and behavioral response recordings were controlled using E-Prime software (V2.0.; http://www.pstnet.com/products/e-prime/).

2.4. Questionnaires

At the end of the experimental session, participants were asked to fill out three different questionnaires, in order to assess whether specific affective or personality traits might be related to task performance and/or our ERP effects. Trait anxiety levels were measured by means of the Dutch version of the State-Trait Anxiety Inventory, trait characteristics (Van der Ploeg, et al., 1979). Participants also completed a recent Dutch version of the Need For Closure Scale (Roets & Van Hiel, 2007), which gave an estimate of the participant's tendency or need to obtain any answer, as opposed to tolerating ambiguity. Finally, we also administered the Need For Affect Scale (Maio & Esses, 2001) in order to obtain an independent measure of the general motivation of participants to either approach or avoid situations that are emotion-inducing. However, we found no significant correlation between the scores obtained for each of the three questionnaires and either the behavioral or ERP results obtained during this progressive stimulus revelation task, presumably because of the low standard deviation of the scores obtained for each questionnaire (see Table 3).

Table 3. Mean values and standard deviations (in parenthesis) of the scores obtained for each questionnaire (and relative subscales) administered at the end of the experiment.

Questionnaire		Score
STAI-T		35.58 *(8.73)*
NFCS		3.57 *(0.60)*
	Order	3.83 *(0.83)*
	Predictability	3.51 *(1.07)*
	Decisiveness	3.56 *(0.60)*
	Ambiguity	3.75 *(0.80)*
	Closedmindedness	3.12 *(0.50)*
NFAS		3.71 *(0.55)*
	Approach	4.27 *(0.77)*
	Avoidance	3.15 *(1.28)*

Note. STAI-T: State-Trait Anxiety Inventory, trait version; NFCS: Need for Closure Scale; NFAS: Need for Affect Scale. STAI-T scores range from 20 to 80; NFCS and NFAS use a Likert scale of 6 and 7 points, respectively.

2.5. EEG data acquisition and pre-processing

Electroencephalographic (EEG) activity was continuously recorded using a BIOSEMI Active-Two system (BioSemi, Inc., Netherlands; http://www.biosemi.com) by means of 128 active electrodes fitted into a stretching cap and following the BioSemi ABCD positioning system (i.e., electrode positions are radially equidistant from CZ; http://www.biosemi.com/headcap.htm). Two electrodes, the common mode sense (CMS) active electrode and the driven right leg (DRL) passive electrode, were used as reference and ground electrodes, respectively (http://www.biosemi/faq/cms_and_drl.htm). Vertical electro-oculograms (EOG) were monitored using two additional electrodes placed in the inferior and superior areas of the left orbit. EEG and EOG recordings were sampled at 512 Hz.

ERPs of interest were computed offline using Brain Vision Analyzer 2.0 (Brain Products™ GmbH, Munich, Germany; http://www.brainproducts.com/analyzer2_release.php). First, a common average reference was applied. Next,

−100/+750 ms epochs were created around the onset of the visual stimulus. Afterwards, all the segments were baseline corrected using the pre-stimulus interval (100 ms), before ocular correction was performed (Gratton, Coles, & Donchin, 1983). Artifact rejection was then carried out (mean amplitude of ±76.32 μV scale across participants) to eliminate segments contaminated by artifacts such as residual blinks or muscle activity. Approximately one-third of the data (27.9%) were discarded using these strict criteria. Finally, averaging of the segments was performed, separately for each condition, and a 1-30 Hz bandpass filter was applied to the individual averaged data. Since we were primarily interested in neural processes occurring before the actual recognition of the scene, we used the time of detection (Response1) as the initial reference point in our analyses, and looked at stimulus-locked ERP effects backwards. Only recognitions (Response1) which turned out to be accurate (based on Response2) were included in our ERP analyses (see behavioral results below). Following this procedure, four different epochs were computed for each individual scene: (1) −100/+750 ms around the onset of the stimulus that was recognized during this specific time window (“Recognition”); (2) −100/+750 ms around the onset of the stimulus which immediately preceded recognition (“One image before” recognition); (3) −100/+750 ms around the onset of the stimulus appearing two images before recognition (“Two images before” recognition); (4) −100/+750 ms around the onset of the stimulus appearing three images before recognition (“Three images before” recognition). Epochs of 750 ms following stimulus were used because they encompassed the duration of the stimulus itself (500 ms), as well as the subsequent 250 ms interval. Using this procedure, we could thus look at stimulus-locked ERP effects for image levels that preceded actual recognition, with a gradual distance relative to this event (up to three images before recognition). Note that this specific data analysis allowed us to look at brain processes consistently foreshadowing visual object recognition, bearing in mind that, across trials, the actual decision could be based on different accumulations of perceptual evidence (i.e., the proportion of correct Responses1 turned out to be the largest at image levels three, four and five; see behavioral results below). Furthermore, we computed different ERP waveforms as a function of the emotional valence of the scenes. As a result, 12 individual averages were computed for each participant : (1) Neutral, Recognition (number of segments after pre-processing: M = 50, SD = 9); (2) Neutral, One image before (M = 58, SD = 9); (3) Neutral, Two images before (M = 58, SD = 8); (4) Neutral, Three images before (M = 43, SD = 9); (5) Unpleasant, Recognition (M = 44, SD = 8); (6) Unpleasant, One image before (M = 52, SD = 7); (7) Unpleasant, Two images before (M = 49, SD = 7); (8) Unpleasant, Three images before (M = 37, SD = 10); (9) Pleasant, Recognition (*M* = 47, *SD* = 10); (10) Pleasant, One image before (*M* = 56, *SD* = 10); (11) Pleasant, Two images before (*M* = 52, *SD* = 9); (12) Pleasant, Three images before (*M* = 47, *SD* = 10).

2.6. Analysis of behavioral data

Accuracy was expressed as percentage of correct responses. Since these image levels were not independent of each other (each image relies on the visual information conveyed by previous levels), cumulative percentages were calculated. Using this procedure, we obtained a psychometric curve showing the evolution of the recognition accuracy across the six image levels for each condition separately. Repeated measures analyses of variance (ANOVAs) with two factors (*emotion*: three levels; *image level*: six levels) and post-hoc t-tests were first employed to assess differences in recognition accuracy between conditions. The level of significance for all the analyses was set at $p < .05$. In addition, to verify whether the psychometric curve was shifted as a function of the emotional content of the scene (pleasant, neutral or unpleasant), we also used a proportional odds model (Agresti, 2007), a regression model for ordinal dependent variables (recognition from image level 1, …, recognition from image level 6). This data analysis allows to model the cumulative probability up to and including recognition from image level k (k = 1, ?, 5). The derived odds ratio expresses how much the odds of recognition from image level k or earlier is increased (if larger than 1) or decreased (if smaller than 1) across different emotional contents, and thus provides a single number capturing the shift in psychometric curve. To account for dependencies of trials within the same subject, a multi-level version of the proportional odds model was used here.

2.7. Analysis of ERP data

In this study, reference-free topographic analyses were carried out to objectively characterize differences between conditions and image levels. The basic principles of this method have been described extensively elsewhere (see Lehmann & Skrandies, 1980; Michel, et al., 1999; Michel, et al., 2001; Murray, et al., 2008; Pourtois, et al., 2008). The added value of this method, relative to a more traditional ERP peak analysis (see Picton, et al., 2000), is that it enables to reveal global differences between experimental conditions without *a priori* selecting a few channels or time frames. All channels and time frames are used concurrently in the analysis. Topographic analyses were performed using CARTOOL software (version 3.43; http://brainmapping.unige.ch/Cartool.htm). For each participant separately, we first interpolated noisy channels using a spherical splines transformation (Perrin, Pernier, Bertrand, & Echallier, 1989). Then, the dominant topographic scalp maps were identified in the grand average ERP data (Recognition, One image before, Two images before and Three images before) over a wide time window spanning from 0 to 600 ms after stimulus presentation. To this end, a specific spatiotemporal clustering algorithm, the "Atomize and Agglomerate Hierarchical Clustering" (AAHC; for a detailed description, see Murray, et al., 2008), was used. This clustering method was developed on purpose to reduce complex EEG/ERP data sets. Following standard practice, the optimal number of dominant maps "explaining" the dataset was based on a cross-validation criterion (Pascual-Marqui, Michel, & Lehmann, 1995). The dominant scalp topographies (identified in the group-averaged data) were then fitted to the ERPs of each individual subject using spatial fitting procedures, to determine their representation across subjects and conditions. This procedure provided fine-grained quantitative values, such as the duration of a specific topographic map or its global explained variance (GEV, or goodness of fit), which are critical indices of the significance of a given topography, not available otherwise in a classical component analysis (Picton, et al., 2000). GEV represents the sum of the explained variance weighted by the GFP (Global Field Power) at each moment in time. GEV was entered in repeated measures analyses of variance (ANOVAs) with *image level* and *emotional content* as within-subject factors. Paired t-tests (and a conservative Bonferroni correction) were used as post-hoc comparisons between conditions. When appropriate, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. As for the statistical analysis of the behavioral data, the level of significance for all these analyses was set at $p < .05$.

2.8. *Source localization analysis*

To estimate the likely neural sources underlying the electrical field configurations identified by the previous analysis, we used a specific distributed linear inverse solution, namely standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002). Mathematical validation of this distributed source localization technique has been recently demonstrated (Sekihara, Sahani, & Nagarajan, 2005). The head model for the inverse solution uses the electric potential lead field computed with a boundary element method applied to the MNI152 template (Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002). Scalp electrode coordinates on the MNI brain are derived from the international 5% system (Jurcak, Tsuzuki, & Dan, 2007). The source locations were therefore given as (x, y, z)

coordinates (*x* from left to right; *y* from posterior to anterior; *z* from inferior to superior). The calculation of all reconstruction parameters was based on the computed common average reference. sLORETA units were scaled to amperes per square meter (A/m2). Direct statistical comparisons between conditions were performed in this inverse solution space using paired t-test. The level of significance for all the analyses was set at $p < .05$.

3. Results

3.1. Accuracy

Participants were accurate in this task; mean error rate was 12.84% (*SD* = 5.02). A 3 (emotion)×6 (image level) repeated measures ANOVA revealed a highly significant main effect of image level [$F(5, 108) = 36.09$, $p < .001$, $\eta_p^2 = .626$]. However, there was no evidence of either a significant main effect of emotion [$F(2, 36) = 0.24$, $p = .683$, $\eta_p^2 = .013$] or an emotion x image level interaction [$F(10, 180) = 1.33$, $p = .271$, $\eta_p^2 = .069$]. Post-hoc comparisons showed a significantly higher percentage of errors during the presentation of Image4 compared to the other image levels (all *p*s < .001), because the majority of responses actually occurred during the presentation of this image level (see below). No significant difference was found between errors during Image1 and Image2 ($p > .05$). Furthermore, the percentage of errors was similar for Image1, Image2 and Image6 (all *p*s > .05). In addition, very few errors were committed with catch trials (Image1: *M* = .00, *SD* = .00; Image2: *M* = .00, *SD* = .00; Image3: *M* = .44, *SD* = 1.91; Image4: *M* = 2.90, *SD* = 12.62; Image5: *M* = 5.70, *SD* = 15.73; Image6: *M* = 1.05, *SD* = 4.59), confirming that participants reliably processed the content of the scene before making a response (living vs. non-living object).

Cumulative percentages of correct responses (i.e., Response1 only when Response2 was correct; see Methods) are presented in Table 4. A 3 (emotion) x 6 (image level) repeated measures ANOVA performed on these values showed a significant main effect of emotion [$F(2, 36) = 39.62$, $p < .001$, $\eta_p^2 = .688$], a significant main effect of image level [$F(5, 90) = 637.37$, $p < .001$, $\eta_p^2 = .973$] and a significant emotion x image level interaction [$F(10, 180) = 21.80$, $p < .001$, $\eta_p^2 = .548$]. Post-hoc comparisons confirmed a progressive (although non-linear) gain in recognition from Image1 to Image6 (all *p*s > .05, with the exception of Image1 vs. Image2 for pleasant stimuli, $p < .05$), as shown by an S-shaped psychometric function (see Fig. 1B). Hence, image levels three, four and five presumably provided sufficient diagnostic low and high spatial frequency information to perform the animacy judgment task with high accuracy. More generally, this S-shaped psychometric function confirmed that our progressive stimulus revelation task was successful, since participants did not respond randomly to the different filtered stimuli across trials but, instead, they consistently waited at least until Image3 before interrupting the stimulus sequence. Importantly, the analysis of the cumulative percentage of accuracy of catch trials revealed a similar outcome. A univariate ANOVA disclosed a significant main effect of image level [$F(5, 108) = 152.88$, $p < .001$, $\eta_p^2 = .876$].

Table 4. Mean values and standard deviations (in parenthesis) of cumulative percentages of correct responses, separately for each image level and emotional valence.

Image Level	Neutral	Unpleasant	Pleasant
1	0.07 *(0.29)*	0.00 *(0.00)*	0.00 *(0.00)*
2	3.53 *(4.61)*	2.74 *(4.58)*	1.06 *(2.79)*
3	34.96 *(15.32)*	31.70 *(16.85)*	22.67 *(16.41)*
4	79.71 *(15.23)*	71.98 *(16.74)*	66.66 *(15.94)*
5	94.65 *(7.02)*	92.95 *(7.23)*	92.28 *(4.95)*
6	100.00 *(0.00)*	100.00 *(0.00)*	100.00 *(0.00)*

Post-hoc comparisons confirmed a progressive gain in recognition at each image level (all *p*s < .001, except Image1 vs. Image2, $p > .05$), lending additional support to the assumption that image levels three, four and five contained sufficient diagnostic spatial frequency information to perform the task with high accuracy. Altogether, these results are consistent with the assumption of non-linear accumulation of evidence prior to recognition (Biederman, 1987; Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004).

Next, we assessed whether the emotional content of the scene may have had an influence on the actual recognition of these scenes. To this end, a mixed proportional odds model with emotion as fixed factor (neutral, unpleasant or pleasant), and subject as random effect was carried out (see Table 5). Interestingly, this analysis suggested a highly significant main effect of emotion ($p < .001$). Pairwise comparisons revealed a shift of the distribution as a function of the emotional content of the scene, indicated by reliably earlier animacy judgments when the picture contained a neutral, as opposed to either an unpleasant ($p < .001$) or pleasant ($p < .001$) content (Fig. 1B). The psychometric function was also shifted leftwards (i.e., earlier recognition) for unpleasant relative to pleasant scenes ($p < .001$). Hence, these results showed that pleasant scenes were recognized significantly later than either unpleasant or neutral scenes.

Table 5. Results of the mixed proportional odds model (behavioral results).

Comparison	Odds Ratio (95% CI)	p-value
Neutral vs. Unpleasant	1.41 (1.21, 1.63)	< .001
Neutral vs. Pleasant	2.01 (1.74, 2.33)	< .001
Unpleasant vs. Pleasant	1.43 (1.23, 1.66)	< .001

Note. An odds ratio larger than 1 (smaller than 1, respectively) implies that the probability of recognition at earlier level is higher (smaller, respectively) for the first versus the second condition in the comparison.

3.2. Reaction times

Finally, we calculated the mean reaction times (RTs) for the correct responses at each image level separately (Image1: *M* = .00 ms, *SD* = .00; Image2: *M* = 133.21 ms, *SD* = 217.89; Image3: *M* = 476.43 ms, *SD* = 104.22; Image4: *M* = 362.00 ms, *SD* = 54.59; Image5: *M* = 288.32 ms, *SD* = 40.29; Image6: *M* = 194.18 ms, *SD* = 105.94) and compared RTs at Image3, Image4 and Image5, for which the majority of behavioral responses were recorded. Pairwise comparisons revealed significant differences between Image3 and Image4 [$t(18) = 4.72$, $p < .001$], Image3 and Image5 [$t(18) = 6.98$, $p < .001$] and Image4 and Image5 [$t(18) = 5.58$, $p < .001$], each time indicated by faster decisions for images levels containing more high spatial frequency information than levels providing more degraded visual information.

3.3. ERP results

Following standard practice, a spatiotemporal cluster analysis was applied on the four main ERP conditions (Recognition, One image before, Two images before and Three images before) during a broad temporal window following stimulus onset (0-600 ms post-stimulus). This analysis revealed that five distinct dominant field topographies explained 91.37% of the total variance (see Figs. 2A and B). The two first dominant maps were common to all four conditions, with reliable topographic changes between conditions starting at 280 ms post-stimulus onset. The first dominant map found in the spatial cluster analysis had a prolonged duration (0-216 ms post-stimulus onset) and shared several electrophysiological properties with the visual N1 component (Fig. 2C) (Vogel & Luck, 2000). The fact that this ERP component, which is usually phasic, showed here a sustained effect may be explained by our specific task parameters, in which degraded stimuli are presented and a progressive accumulation of visual information is needed over a prolonged period of time. This N1 scalp map was next replaced by a visual P2 component (Carretié, et al., 2004; Crowley & Colrain, 2004; Freunberger, Klimesch, Doppelmayr, & Holler, 2007; Luck & Hillyard, 1994) which, like the preceding N1 scalp map, was shared across the four image levels (Fig. 2C). The spatiotemporal cluster analysis disclosed that the P2 scalp map had the highest variance during the 216-280 ms time interval post-stimulus onset. Most likely, these two early neural activities reflected the encoding and low-level visual discrimination of the stimulus. Following the offset of the P2 scalp map (280 ms post-stimulus onset), the cluster analysis unambiguously revealed reliable topographic changes across the four image levels, which are necessarily indicative of changes in the configuration of the underlying intracranial generators (Lehmann & Skrandies, 1980; Michel, et al., 1999; Murray, et al., 2008). Whereas another broad occipito-parietal activity was generated following the P2 for image levels one, two and three before recognition, a distinctive occipital negative component (with a concurrent positive, frontal counterpart) was generated during the same latency for the image level corresponding to actual recognition (Fig. 2C). Interestingly, this differential scalp topography was actually generated during the time period (280-360 ms post-stimulus, Recognition level) in which the first decision (Response1) was most likely made by the participant. However, since RTs were variable or jittered across trials and participants, as well as differed as a function of image level (see behavioral results above), it is unlikely that this distinctive scalp map was somehow related only to the preparation or execution of a motor response.

The next step was to verify whether the topographic changes observed across conditions 280 ms after stimulus onset were reliable. Accordingly, these dominant maps were fitted back to the individual ERP data (by using a spatial fitting procedure, as described above) to estimate their representation across time and conditions. For this purpose, three different time intervals were defined based on the outcome of the spatiotemporal cluster analysis, each lasting 80ms: a first interval around the peak (as defined using the GFP) of the visual N1 map, a second one corresponding to the occipital P2 map, and a third one immediately following the P2 map, where topographic differences were found by the preceding cluster analysis.

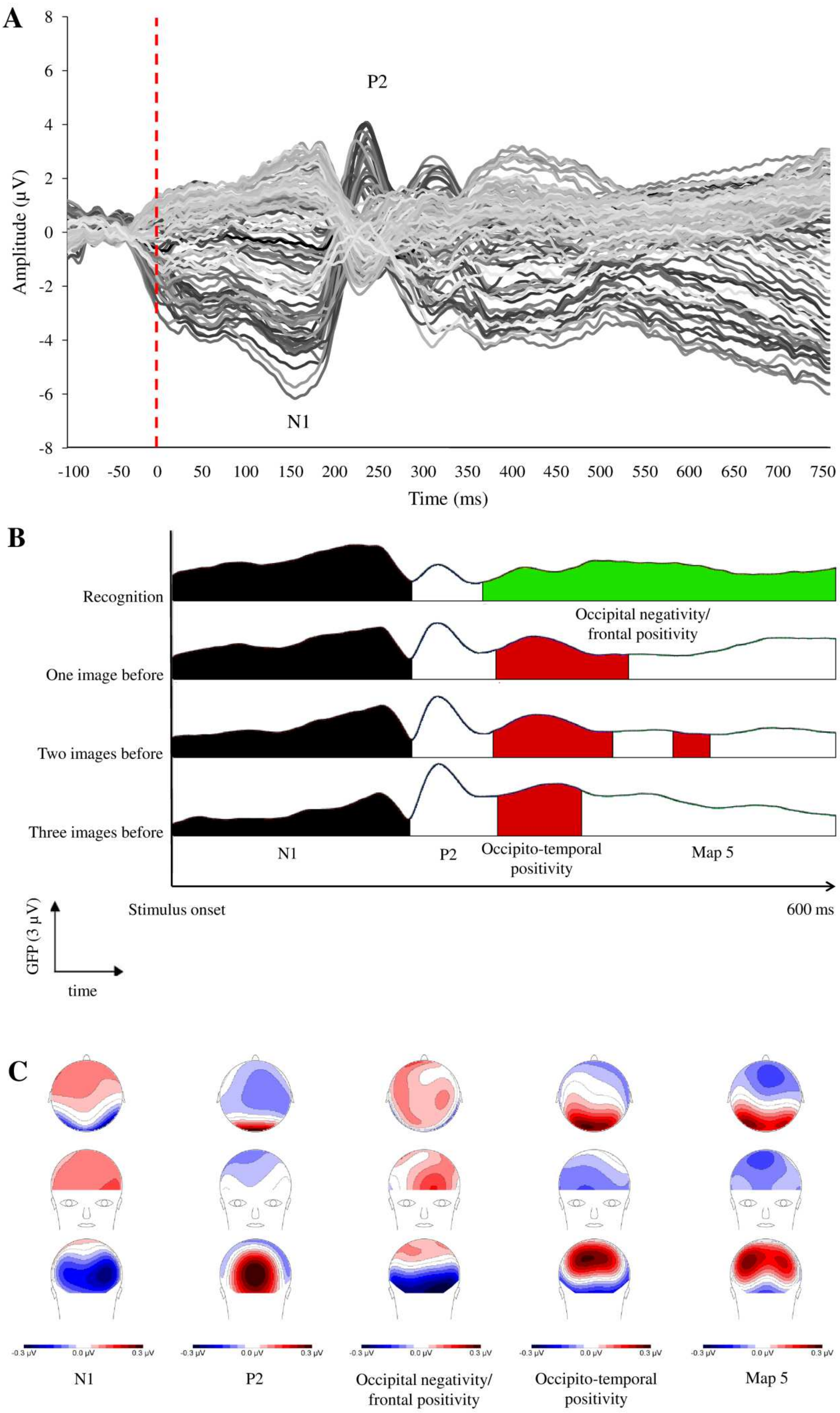

A
8
6
4
2
0
-2
-4
-6
-8
Amplitude (µV)
P2
N1
-100
-50
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
Time (ms)
B
Recognition
One image before
Two images before
Three images before
Occipital negativity/
frontal positivity
N1
P2
Occipito-temporal
positivity
Map 5
Stimulus onset
600 ms
GFP (3 µV)
time
C
-0.3 µV
0.0 µV
0.3 µV
N1
P2
Occipital negativity/
frontal positivity
Occipito-temporal
positivity
Map 5

Fig. 2. (A) Grand-average (N = 19) ERP waveforms (obtained for the image level corresponding to actual recognition) for the 128 electrodes (butterfly). The red dashed vertical line indicates the onset of the visual stimulus. Clear exogenous N1 and P2 ERP components were recorded after stimulus onset. (B) Results of the spatiotemporal cluster analysis (from stimulus onset until 600 ms after stimulus onset) for the four main conditions (Recognition, One image before, Two images before, and Three images before recognition). A main solution with five dominant topographic maps was found to explain > 90% of the variance. Scalp topographies of the N1 and P2 were shared across the four conditions, suggesting a similar low-level encoding and processing of the visual stimulus. By contrast, starting at 280 ms after stimulus onset, reliable topographic alterations were found between the four image levels. Whereas a broad occipito-parietal activity was generated following the P2 for all the image levels preceding recognition (in red), a distinctive occipital negative component (with a concurrent positive, frontal counterpart) was generated at the same latency for the image level corresponding to actual recognition (in green). (C) Horizontal, frontal and occipital views of the five dominant maps (including the N1 and P2) found in the spatial cluster analysis. Amplitude differences were normalized (i.e., the amplitude value at each electrode was divided by the GFP).

For the visual N1 map (136-216ms post-stimulus onset), a univariate ANOVA performed on the GEV values failed to reveal any significant difference between the four image levels [$F(3, 72) = 0.78$, $p = .507$, $\eta_p^2 = .032$], lending support to the assumption that this first occipital map might correspond to the early visual encoding or discrimination of the incoming (blurred) stimulus. For the P2 map (216-296 ms post-stimulus onset), the univariate ANOVA showed a significant main effect of image level [$F(3, 72) = 5.04$, $p = .003$, $\eta_p^2 = .173$]. Post-hoc comparisons revealed a selective decrease of the global explained variance of this P2 map for the image level corresponding to actual recognition, relative to the three other levels (One image before, $p = .044$; Two images before, $p = .010$; Three image before, $p = .008$). The other pairwise comparisons remained non-significant (all *p*s > .05). This latter result suggested a substantial decrease of the P2 at the time of recognition, which could reflect either the processing of low-level visual properties of the stimulus (Luck & Hillyard, 1994) or short-term memory load (Wolach & Pratt, 2001), including priming (Gruber & Muller, 2005; Rugg, Soardi, & Doyle, 1995; Wiggs & Martin, 1998). To ascertain the presence of a reliable topographic change during the third time interval (280-360 ms post-stimulus onset), we then submitted the GEV values obtained after fitting during this time interval to a 2 (*map configuration*: occipito-parietal positive activity vs. occipital negative/frontal positive activity) x 4 (*image level*: Recognition, One image before, Two images before, Three images before) repeated measures ANOVA. This analysis revealed a main effect of map configuration [$F(1, 18) = 4.89$, $p = .040$, $\eta_p^2 = .214$], and a highly significant map configuration x image level interaction [$F(3, 54) = 14.34$, $p < .001$, $\eta_p^2 = .443$], corroborating the assumption of a reliable topographic change across the four image levels during this specific time interval, and in sharp contrast to the results obtained for the preceding visual N1 and P2 scalp maps. Interestingly, post-hoc comparisons showed that the

GEV of the occipital negative/frontal positive activity progressively increased from three images before recognition to actual recognition (see Fig. 3A). More specifically, the GEV of this component was found to reliably increase when comparing actual recognition to one image before [$t(18) = 3.12$, $p = .006$], two images before [$t(18)= 2.61$, $p = .018$] and three images before recognition [$t(18) = 5.02$, $p < .001$]. Similarly, the substantial increase in GEV as a function of the progressive unfolding of the scene content was evident when comparing one image to three images before recognition [$t(18) = 2.56$, $p = .020$], as well as two images to three images before recognition [$t(18) = 2.98$, $p = .008$]. By comparison, the evolution of the GEV values of the concurrent occipito-parietal positive activity over the four image levels showed a different statistical outcome (Fig. 3B). T-tests performed on the GEV values extracted for this topographic activity during the same time interval (280-360ms post-stimulus onset) revealed a significant decrease of the GEV for the image level corresponding to actual recognition, relative to one image [$t(18) = -4.59$, $p < .001$], two images [$t(18) = -3.65$, $p = .002$] and three images before recognition [$t(18) = -4.50$, $p < .001$]. The other pairwise comparisons were not significant. In other words, this significant map configuration x image level interaction indicated a reliable change of microstates during this time interval, when progressively moving from three images before recognition to actual recognition. This topographic change was best explained by a progressive increase of the negative occipital/frontal positive activity when moving closer to recognition, whereas the concurrent occipito-parietal positive activity was found to be stable for the three first levels but showed a sharp decrease at the time corresponding to actual recognition.

We failed to find any significant modulation of these map configurations and their expressions as a function of the emotional content of the scene. We carried out a 2 (map configuration) x 3 (emotional content) x 4 (image level) repeated measures ANOVA on the GEV values extracted during this time interval (280-360 ms post-stimulus onset) that revealed a significant main effect of map configuration [$F(1, 18) = 4.89$, $p = .040$, $\eta_p^2 = .213$] and a significant map configuration x image level interaction [$F(3, 54) = 14.35$, $p < .001$, $\eta_p^2 = .444$], but no significant modulation by the emotional content of the scene. This result suggested that these differential neural processes related to visual object recognition were not influenced by the emotional content of the scenes.

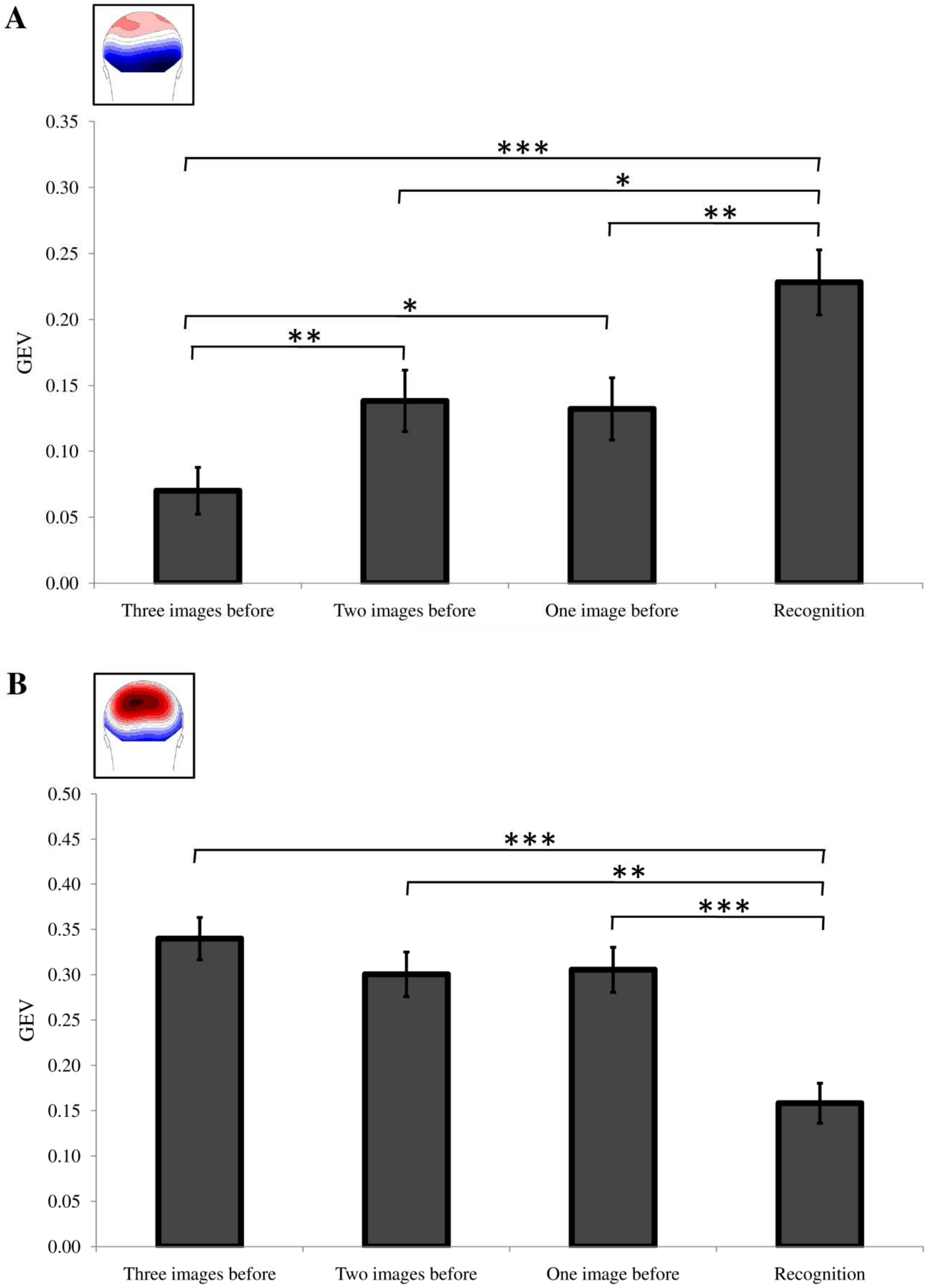


Fig. 3. (A) Statistical results (*$p < .05$; **$p < .01$; ***$p < .001$; vertical bars correspond to standard errors of the mean), obtained after the fitting procedure, for the dominant topography characterized by an occipital negativity and a frontal positivity showed a linear increase when moving closer to recognition. The Global Explained Variance (GEV, arbitrary units) was computed during an 80 ms time interval (280-360 ms post-stimulus onset) and is presented separately for the four conditions. This analysis showed a linear increase of the GEV when moving from three images before recognition to actual recognition. (B) Results obtained for the concurrent

topography of the occipito-parietal positivity identified during the same time interval. Unlike the occipital negative/frontal positive activity, this scalp configuration showed an abrupt decrease for the image level corresponding to recognition, relative to the three preceding images levels where the variance remained stable.

3.4. Source localization results

Next, we used sLORETA (Pascual-Marqui, 2002) to gain insight into the likely neural generators accounting for the different topographies identified by the previous topographic analyses. More specifically, sLORETA was used to explore the brain regions underlying the substantial topographic change found in the previous analysis, which concerned the 280-360 ms post-stimulus onset interval. Since the change in the electric field configuration was most obvious when directly comparing three images before recognition to actual recognition, we first used this contrast in the inverse solution space and performed statistical non-parametric mapping (SnPM) analyses. For this purpose, amplitude data were first normalized (i.e., total average power equal to unity) prior to performing statistical analyses. This procedure revealed a stronger activation for recognition compared to three images before recognition in the left dACC (-15*x*, -10*y*, +45*z*) [$t(18) = 4.16$, $p < .001$] and right dACC (+15*x*, -10*y*, +45*z*) [$t(18) = 4.50$, $p < .001$], extending bilaterally in the supplementary motor area (SMA; left: -25*x*, -10*y*, +45*z*; right: +25*x*, -10*y*, +45*z*; $t(18) = 4.31$, $p < .001$ and $t(18) = 3.01$, $p = .007$, respectively). The reverse contrast (Three images before > Recognition) showed larger activations in the bilateral posterior cingulate cortex (PCC), extending ventrally in the parahippocampal gyrus (PHG) [left: -15*x*, -65*y*, +10*z*; right: -15*x*, -65*y*, +10*z*; $t(18) = -4.56$, $p < .001$ and $t(18) = -2.80$, $p = .012$, respectively] (Fig. 4A). Next, for each of these regions of interest (ROIs) and each subject separately, we extracted the amplitude value (mean amplitude during the 280-360 ms post-stimulus onset interval) to establish how the activity in these regions actually evolved as a function of time of recognition. A 2 (side: left vs. right hemisphere) x 2 (ROI: dACC vs. PCC/PHG) x 4 (image level: Recognition, One image before, Two images before, Three images before) repeated measures ANOVA disclosed a significant main effect of ROI [$F(1, 18) = 22.61$, $p < .001$, $\eta_p^2 = .557$] and a marginally significant effect of image level [$F(3, 54) = 2.71$, $p = .074$, $\eta_p^2 = .131$] but, more importantly, a significant ROI x image level interaction [$F(3, 54) = 14.79$, $p < .001$, $\eta_p^2 = .451$]. For the left dACC (Fig. 4B), post-hoc comparisons revealed a significantly higher activity during actual recognition, relative to one image [$t(18) = 5.29$, $p < .001$], two images [$t(18)=5.22$, $p < .001$] or three images before recognition [$t(18) = 4.90$, $p = .007$]. None of the other pairwise comparisons was significant (all *p*s > .05).

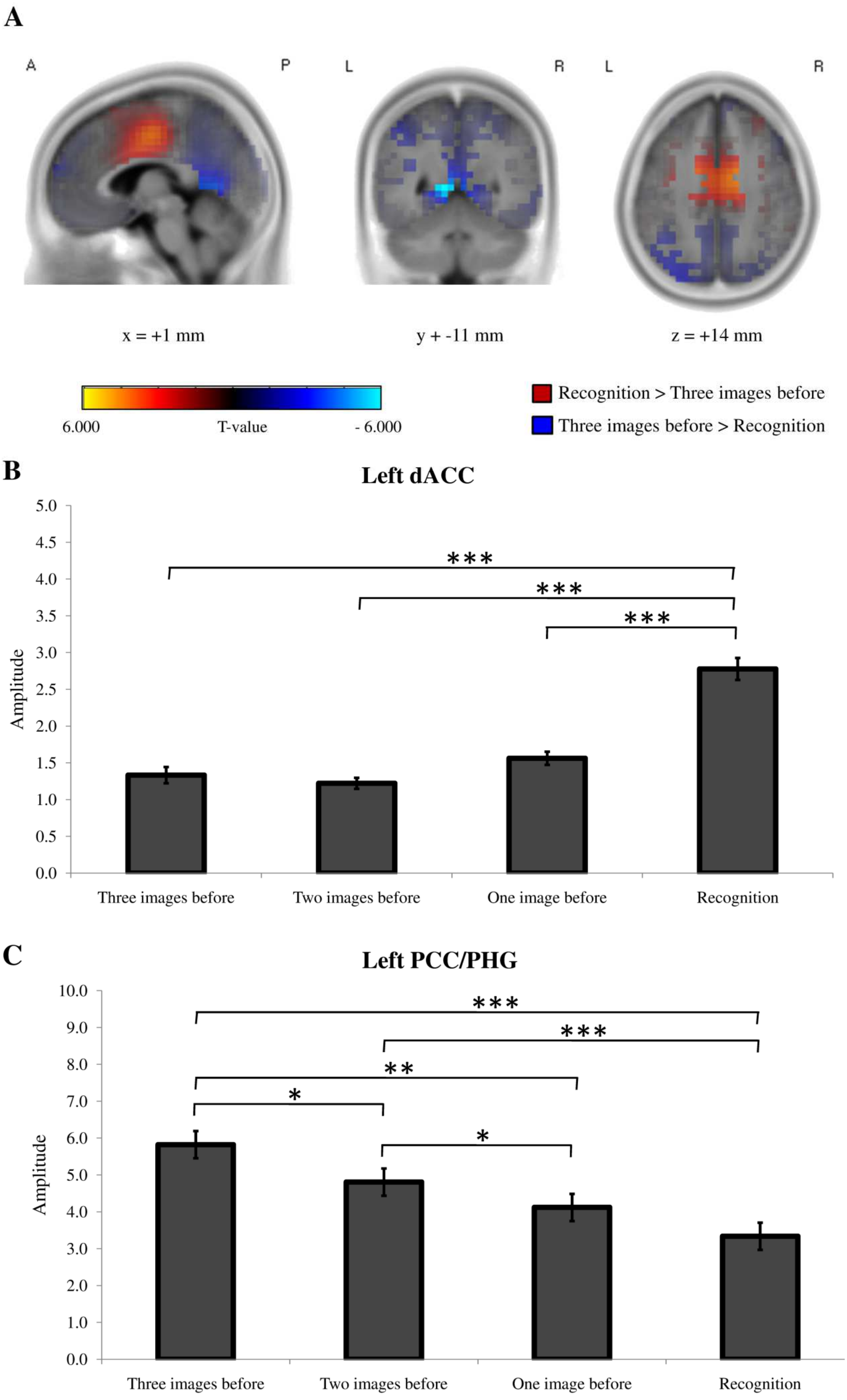

A
A
P
L
R
L
R
x = +1 mm
y + -11 mm
z = +14 mm
6.000
T-value
- 6.000
🟥 Recognition > Three images before
🟦 Three images before > Recognition
B
Left dACC
Amplitude
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
***
***
***
Three images before
Two images before
One image before
Recognition
C
Left PCC/PHG
Amplitude
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
***
***
**
*
*
Three images before
Two images before
One image before
Recognition

Fig. 4. Source localization results. (A) Comparing actual recognition to three images before recognition during the 280-360 ms post-stimulus onset interval disclosed a highly significant effect ($p < .001$ corrected) in the dorsal ACC, bilaterally (±15*x*, +10*y*, +45*z*). The reverse contrast revealed a highly significant effect ($p < .001$ corrected) in the posterior cingulate cortex, extending ventrally towards the parahippocampal gyrus (±15*x*, -65*y*, +10*z*). A: anterior; P: posterior; L: left; R: right. (B) Mean activity extracted from the left dorsal ACC as a function of image level (*$p < .05$; **$p < .01$; ***$p < .001$). Vertical bars correspond to standard errors of the mean. In this dorsal ACC region, a sharp increase was found for recognition, relative to the three preceding levels. (C) By contrast, in the left PPC, a monotonic linear decrease of activity was evidenced when moving towards recognition.

Inverse solution results obtained for the right dACC showed a very similar outcome: stronger activity was evidenced when comparing actual recognition to one image [$t(18) = 5.53$, $p < .001$], two images [$t(18) = 4.09$, $p = .001$] or three images before recognition [$t(18) = 4.21$, $p = .001$]. By contrast, statistical analyses performed on the amplitude values extracted from the PCC/PHG showed a different result, mainly characterized by a linear decrease in activity when moving from three images before recognition to actual recognition. For the left PCC/PHG, paired t-tests showed a significantly lower activity in this region during actual recognition, relative to two images [$t(18) = -4.50$, $p < .001$] and three images before recognition [$t(18) = -5.04$, $p < .001$]. Significantly lower neural activity was also evidenced when comparing one image to two images before recognition [$t(18) = -2.12$, $p = .048$], one image to three images before recognition [$t(18) = -3.39$, $p = .003$], and two images to three images before recognition [$t(18) = -2.51$, $p = .022$], suggesting a linear monotonic decrease of activity in this region as a function of accumulation of perceptual evidence (see Fig. 4C). The activity extracted in the right PCC/PHG showed a similar trend, although attenuated. Statistical analyses revealed significantly higher amplitude during actual recognition compared to three images before recognition [$t(18) = -2.31$, $p = .033$]. The same effect was evidenced when comparing one image to three images before recognition [$t(18) = -2.19$, $p = .042$].

Finally, we verified whether the activity in these ROIs varied with the emotional content of the scenes or not. In none of the four ROIs did the ANOVA reveal any significant effect of the emotional content of the scene, suggesting that neural processing in these four regions was not influenced by the emotional content of the scene during this specific time interval following stimulus onset, consistent with the topographic analyses reported above.

4. Discussion

In this study, we used high density EEG to shed light on the neural events preceding and leading to perceptual decision making during a simple/binary visual categorization task (animacy judgment). For each trial, participants were presented with series of filtered images that were progressively unfolding the content of a complex visual scene, and the participant was asked to discriminate whether this scene contained a living object or not. The sequence started with the presentation of a blurred image whose content was increasingly revealed by adding up, in a non-linear fashion, HSF information, hence providing a temporal decomposition of a “coarse-to-fine” analysis of the incoming visual stimulus. Since previous models have emphasized such a “coarse-to-fine” analysis subtended by different contributions of low vs. high spatial frequency information (Bar, 2004; Bullier, 2001; Hegdé, 2008), this sequential procedure was then developed to investigate upstream evidence accumulation processes leading to recognition. In addition, the visual scenes were neutral, pleasant or unpleasant, to assess whether their emotional content might affect the expression of neural events foreshadowing overt visual object recognition.

Behavioral results confirmed that this new progressive stimulus revelation task was suited to study the temporal dynamic preceding visual object recognition. Participants consistently waited for sufficient perceptual evidence before categorizing the incoming visual stimulus as either living or non-living with high accuracy. Accurate perceptual decisions (mean % response correct > 87) mainly occurred after the presentation of three (spatial filtering: 64-512 pixels/cycle), four (32-512 pixels/cycle) or five (16-512 pixels/cycle) images, suggesting systematic accumulation of evidence before recognition (see Fig. 1B). Consistent with previous studies (Bar, 2004; Delplanque, et al., 2007; Schyns & Oliva, 1994), these three image levels presumably contained adequate diagnostic spatial frequency content to perform the animacy judgment task with high accuracy and confidence. Behavioral results obtained for catch trials corroborated the assumption that participants did not simply guess about the content of the scene but, instead, they reliably accumulated perceptual evidence before making a decision about its content, as shown by progressively higher perceptual decisions made during the presentation of the third, fourth and fifth image levels. Altogether, these behavioral findings are compatible with predictions arising from accumulator models of decision making, which directly emphasize the accumulation of perceptual evidence in favor of each alternative before a response boundary is surpassed (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008).

Interestingly, our behavioral results also showed reliable differences between the three emotion conditions, indicated by earlier recognitions for unpleasant compared to pleasant scenes. Although speculative, this effect obtained for negatively-valenced scenes might be consistent with a negativity bias effect, as described in the Evaluative Space Model (*ESM*: Cacioppo & Gardner, 1999; Cacioppo, et al., 1997, 1999; Norris, et al., 2010). This dominant model makes the assumption of two separable and partially distinct components of the system underlying the evaluation of emotion and affect: (1) positivity, sensitive to appetitive stimuli and promoting approaching behavior; (2) negativity, oriented towards threat or danger, and fostering avoidance (Cacioppo & Gardner, 1999). These components are characterized by distinctive activation functions, a *negativity bias* (strongly aversive stimuli elicit stronger responses than appetitive ones) and a *positivity offset* (when input to the affect system is minimal, positivity outweighs negativity). This negativity bias would lead to slower or diminished responses to non-negative (appetitive/pleasant), relative to negative stimuli (Norris, et al., 2010). However, because neutral scenes were recognized on average earlier than negative scenes, a general negativity bias only does not seem sufficient to account for our behavioral results. Alternatively, we cannot rule out the possibility that picture complexity (or stimulus ambiguity) might actually be different between pleasant, unpleasant and neutral scenes, a factor that could potentially account for the differences in speed of recognition found across these three emotion conditions in our task (see Fig. 1B). However, we first aimed at selecting mildly emotional pictures from the IAPS, which were balanced regarding the animacy dimension but for which clear contrast effects (regarding the valence and arousal dimensions) could be obtained when comparing these neutral, pleasant and unpleasant scenes (see Table 1). This selection procedure resulted in a limited number of individual scenes for each emotion condition. In this context, we could not control for possible low-level visual differences across the three emotion conditions, which may eventually influence performance during the animacy judgment task, even though it remains unclear if a specific low-level visual property may systematically bias the animacy judgments in one direction or the other. Moreover, because the early ERP components (0-200 ms post-stimulus onset, including the N1 and P2) were found to be identical across the three emotion conditions, a differential recognition effect across the three emotion categories triggered by specific low-level visual properties appears unlikely.

At the electrophysiological level, we found evidence for the involvement of different brain regions that differently contributed to mechanisms of perceptual decision making. More specifically, we looked at electrophysiological effects occurring up to three images before recognition, and eventually found reliable topographic modulations of the ERP signal in a window spanning from 280 to 360 ms after stimulus onset across these four image levels. Notably, two different topographic activities showing different evolutions over time (see Figs. 3A and 3B) were evidenced. Whereas the explained variance of an occipital negative/frontal positive activity linearly increased from three images before recognition until time of recognition (Fig. 3A), the explained variance of a concurrent broad occipito-parietal positive activity remained stable for the different image levels before recognition, but substantially dropped at the time of recognition (Fig. 3B). On the other hand, we did not find any significant modulation of these effects by the emotional content of the scene, suggesting a common visual object recognition mechanism for neutral, pleasant and unpleasant pictures during this specific time interval. A possible reason may be that our pre-selected IAPS stimuli were only mildly arousing, in contrast with many previous studies in literature showing reliable visual ERP effects with highly arousing pictures (including the EPN and LPP components, recorded following stimulus onset), relative to neutral, or low arousing scenes (Delplanque, et al., 2004; Junghöfer, Bradley, Elbert, & Lang, 2001; Peyk, Schupp, Keil, Elbert, & Junghöfer, 2009; Schupp, Junghöfer, et al., 2003a, 2003b; Schupp, et al., 2006; Wiens, Peira, Golkar, & Öhman, 2008). In addition, given the specific data analysis used in our study, we cannot rule out the possibility that these neural processes might be triggered with some delay for emotional compared to neutral pictures, an issue that requires further research.

Since these topographic changes observed at the scalp level necessarily denote alterations in the configuration of the intracranial generators (see Lehmann and Skrandies, 1980; Michel, et al., 1999; Murray, et al., 2008; Pourtois, et al., 2008), we performed complementary source localization analyses to gain insight into the putative configuration of the intracranial generators underlying these two specific topographic activities related to visual object recognition. Inverse solutions based on sLORETA (Pascual-Marqui, 2002) confirmed a substantial shift in the localization of the neural generators underlying these two topographic maps. Comparing three images before recognition to time of recognition revealed a highly significant effect ($p < .001$ corrected) in the PCC/PHG complex (Brodmann's areas, BA 29/30). The reverse contrast revealed a highly significant ($p < .001$ corrected) bilateral effect in the dACC (BA 24) (Fig. 4A), with additional effects localized more dorsally, including the

SMA bilaterally. This latter effect in the SMA might reflect a motor preparation component that preceded or was associated with the overt recognition of the scene. While the former regions showed a quasi-linear decrease in activation when moving from three images before recognition to actual recognition (see Fig. 4C), the extracted activity in the latter region showed instead a stable pattern up to the image level corresponding to actual recognition, where an abrupt and substantial increase was evidenced (see Fig. 4B). Hence, distinct effects were found in these two distant regions as a function of accumulation evidence processes preceding actual recognition.

Previous neuroscience studies have linked the PCC, and more specifically its retrosplenial portion (BA 29), to the acquisition of visual memory traces (Dickerson & Eichenbaum, 2010; Shallice, et al., 1994; Valenstein, et al., 1987). Moreover, recent studies have advocated this brain region in the formation of canonical representations for typical contexts encountered during visual object recognition (Aminoff, Gronau, & Bar, 2007; Bar, 2009b; Bar, Aminoff, Mason, & Fenske, 2007), whereas the adjacent PHG would primarily be involved in the generation of associations related to these contexts (Aminoff, et al., 2007; Bar, 2004, 2007; Bar, et al., 2006). Therefore, these different brain areas might work in concert and form a functional network involved in the analysis and generation of contextual information presented together with the actual visual object (Bar, 2009b). Interestingly, here we found a similar network (involving the PCC and PHG) whose activity was found to monotonically decrease when increasing the spatial frequency content of the input image, and hence when progressively revealing important contextual cues about the object's identity (see Fig. 4C). This effect may therefore reflect the processing of diagnostic contextual information (primarily based on LSF information; see Bar, 2009b) needed to optimally categorize the content of the image, and eventually perform the animacy judgment task.

By contrast, source localization results showed the involvement of the bilateral dACC (BA 24) in our task, although the response profile in this region (see Fig. 4B) was reliably different compared to the PCC/PHG (Fig. 4C). The dACC was primarily found to be active during the image level corresponding to the actual recognition of the scene. Activity of the dACC was low and stable during the three image levels before recognition, but substantially increased at the time of recognition, consistent with its involvement in higher-order decision making processes. Various functions have been ascribed to the dACC (Bush, Luu, & Posner, 2000; Ridderinkhof, et al., 2004), including reward evaluation (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003), response conflict/competition (Botvinick, 2007; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Bush, et al., 1998), value judgments (Seitz, Franz, &

Azari, 2009), error detection (Dehaene, Posner, & Tucker, 1994; Holroyd & Coles, 2002), and reward-based decision making (Bush, et al., 2002; Hampton & O'Doherty, 2007; Rushworth & Behrens, 2008; Seo & Lee, 2007, 2009; Watanabe, 2007). A critical function of this region may be the monitoring of choice outcomes, with the aim to further adjust sensory acuity in order to improve the organism's response to the environment (Kable & Glimcher, 2009). Recently, the dACC has also been thought to be responsible for both preparation and online adjustments in response to conflicts, considered either as environmental feedbacks or internally generated signals (Anderson, Anderson, Ferris, Fincham, & Jung, 2009; Fincham & Anderson, 2006; Sohn, Albert, Jung, Carter, & Anderson, 2007). Given the rather categorical/all-or-nothing response profile found for the dACC in our study, this effect may therefore reflect the involvement of this region in monitoring the outcome of a decision, which is based in the present case on the rapid and likely incomplete accumulation of perceptual evidence. Importantly, here perceptual evidence remains partial or incomplete at the time of overt decision (Response1; see Methods). This situation could potentially generate a mild conflict, expressed by the urge to make a rapid decision while the actual accumulation of evidence is not complete yet (Hayden, Pearson, & Platt, 2009; Sarinopoulos, et al., 2010; Scheffers & Coles, 2000). Thus, whereas the PCC/PHG complex may be involved in updating contextual information about the scene (based on the rapid extraction and accumulation of LSF information; see Bar, 2009b), the dACC may contribute to visual object recognition mechanisms in this task by monitoring the ongoing decision's outcome, and serving therefore as an important interface between the accumulation of perceptual evidence on the one hand, and the implementation and execution of a motor plan on the other hand. The rapid acquisition and accumulation of perceptual/contextual evidence taking place in the PCC/PHG would lead in turn to the monitoring of the actual decision's outcome within the dACC, as well as SMA to a smaller degree. This rapid transition might occur by means of reciprocal anatomical connections between these non-overlapping brain regions (Vogt, Finch, & Olson, 1992).

Thus, the experimental design and data analyses used here enabled us to track, using a millisecond time scale, fast evolving upstream perceptual brain processes eventually leading to visual object recognition, whose precise temporal dynamics may be difficult to disclose using the BOLD-fMRI technique only, given its sluggish temporal resolution. However, our ERP are also complementary to previous fMRI results (see Ploran, et al., 2007), as they confirm that non-overlapping brain regions (i.e., dorsal ACC vs. PCC/PHG) were reliably active prior to visual object recognition at different latencies (relative to recognition) and with different response profiles, but they also add new critical information about the precise electrophysiological time-course of these upstream perceptual effects, which primarily concerned an interval spanning from 280 to 360 ms post-stimulus onset.

There are a few restrictions to the present study. First, each visual scene was repeated once after a variable and unpredictable time lag, which may have introduced a slight bias in top-down recognition brain mechanisms for repetitions, relative to first presentations of the scenes. However, our ERP analyses failed to reveal any substantial topographical and source localization difference between first presentations and repetitions, suggesting similar accumulation of evidence processes in these two conditions (though occurring earlier for repetitions relative to first presentations). Although this behavioral advantage in speed of recognition for repetitions relative to first presentations did not yield ERP topographic differences in our study, one may assume that additional brain processes may underlie this differential effect, even if they were not visible in the present case. The use of a more explicit (as opposed to implicit) visual encoding strategy combined with the activation of visual recognition processes based on cues stored in memory might help to reveal different top-down recognition processes for first encounters of the scenes, compared to repetitions. Another limitation concerns the stimuli selected from the IAPS and used in this ERP study. Low-level but uncontrolled differences (e.g., picture complexity, stimulus ambiguity) may exist between neutral, pleasant and unpleasant scenes, which could possibly influence behavioral performance during the animacy judgment task. We cannot exclude the possibility that some of the observed behavioral results (see Fig. 1B) were explained by some uncontrolled “low-level” differences between the three emotion categories (neutral, pleasant and unpleasant scenes).

In sum, our new ERP results provide the first direct electrophysiological evidence of upstream neural events leading to visual object recognition, highlighting distinct effects in the PCC/PHG and dACC during this process rapidly following stimulus onset. Our findings are also in line with recent fMRI results showing that these two regions may differentially contribute to mechanisms of perceptual decision making (Ploran, et al., 2007; Wheeler, et al., 2008). In addition, the use of high density scalp EEG helped us shed light on the distinctive temporal contributions of these two regions during visual object recognition. Whereas the PCC/PHG complex was involved in the accumulation of perceptual evidence prior to actual recognition, the dACC was likely implicated at a later stage in the monitoring of the decision's outcome (see also Ploran, et al., 2007). Future studies are needed to establish whether these two distant regions may exhibit some coupling or reciprocal interactions during upstream perceptual processes leading to visual object recognition.

CHAPTER 4

Multiple synergistic effects of emotion and memory on proactive processes leading to scene recognition[1]

Visual scene recognition is a proactive process through which contextual cues and top-down expectations facilitate the extraction of invariant features. Whether the emotional content of the scenes exerts a reliable influence on these processes or not, however, remains an open question. Here, topographic ERP mapping analysis and a distributed source localization method were used to characterize the electrophysiological correlates of proactive processes leading to scene recognition, as well as the potential modulation of these processes by memory and emotion. On each trial, the content of a complex neutral or emotional scene was progressively revealed, and participants were asked to decide whether this scene had previously been encountered or not (delayed match-to-sample task). Behavioral results showed earlier recognition for old compared to new scenes. Moreover, delayed recognition for emotional relative to neutral scenes was observed. Electrophysiological results revealed that, ~400 ms following stimulus onset, activity in ventral object-selective regions increased linearly as a function of accumulation of perceptual evidence prior to recognition of old scenes. The emotional content of the scenes had an early influence in these areas. By comparison, at the same latency, the processing of new scenes was mostly achieved by dorsal and medial frontal brain areas, including the anterior cingulate cortex and the insula. In the latter region, emotion biased recognition at later stages, likely corresponding to decision-

---

[1] Schettino, A., Loeys, T., & Pourtois, G. (2012). Multiple synergistic effects of emotion and memory on proactive processes leading to scene recognition. Manuscript submitted for publication.

making processes. These findings suggest that emotion can operate at distinct and multiple levels during proactive processes leading to scene recognition, depending on the extent of prior encounter with these scenes.

1. Introduction

Visual scene recognition results from dynamic and reciprocal interactions between bottom-up sensory processing and top-down modulatory influences, including selective attention (Treisman & Kanwisher, 1998), contextual information (Oliva & Torralba, 2007), and prior expectations (Summerfield & Egner, 2009). Predictive coding models of visual perception (Friston, 2005; Grossberg, 2009; Rao & Ballard, 1999) emphasize the predominant role of predictions, i.e., experience-based information about what is possible or probable in the current sensory environment. Such predictions actively guide visual recognition processes, presumably by favoring the rapid selection of perceptual features that eventually help disambiguate the meaning of the retinal input (Clark, 2012). Whenever these top-down predictions happen to be discrepant with bottom-up sensory processing, an error signal (prediction error) is generated and propagated back to higher-level brain regions, with the aim to update or refine the content of the predictions based on the available sensory evidence. If recognition fails, these recursive loops continue to operate until sufficient sensory information is gathered, eventually enabling the activation of the corresponding visual representation of the object or scene stored in memory (Friston, 2005; Rao & Ballard, 1999; Ullman, 1995).

Of note, the differential sensory processing of low (LSF) and high (HSF) spatial frequency information could arguably underlie these reciprocal interaction effects between (bottom-up) sensory processing and (top-down) predictions (Bullier, 2001). In this framework, the rapid extraction of LSF information via dedicated magnocellular pathways may serve to quickly generate coarse predictions regarding the most probable content of the visual scene. This process would later be assisted and refined by the extraction of HSF information (Bar, 2003, 2004).

Whether these dynamic perceptual processes are differentially engaged depending on the (perceived) emotional content of the scene or not, however, has received little attention so far. Interestingly, mounting evidence shows that emotion exerts strong biases on visual perception, including at early stages of recognition (Bocanegra & Zeelenberg, 2009; Öhman, Flykt, et al., 2001; Phelps, et al., 2006; Vuilleumier & Pourtois, 2007; Whalen, et al., 2004). Thus, emotion is no longer seen as a by-product of perception but, instead, as a core determinant of it (Pourtois, et al., 2012). Accordingly, proactive processes during scene recognition are unlikely to be immune to emotion, defined here, following standard practice, as the acquired valence and arousal values of the stimulus (Barrett & Bar, 2009; Bradley & Lang, 1994). More specifically, the speed and extent of iterative processes between bottom-up processing and top-down predictions during scene recognition are probably subject to variations depending on the rapidly extracted emotional meaning of the stimulus. Consistent with this assumption, we recently reported evidence for the modulation of scene recognition by emotion (Schettino, Loeys, Delplanque, & Pourtois, 2011). In this study, participants were presented with series of pictures whose content was progressively revealed by increasing, in up to six sequential steps, the amount of LSF and HSF information. This experimental manipulation was used to mimic a “coarse-to-fine” decomposition of the retinal input, presumably fostering the online generation of guesses or predictions regarding the most likely identity of the stimulus (Bar, 2003; Bullier, 2001; Hegdé, 2008). Participants were asked to discriminate the content of the stimulus progressively revealed using this procedure by performing an orthogonal animacy judgment task. Crucially, the scenes used in this experiment were neutral, pleasant or unpleasant. Behavioral results showed a delayed recognition for emotional compared to neutral scenes, especially for pleasant compared to neutral scenes (see also Schettino, Loeys, Bossi, & Pourtois, 2012). At the electrophysiological level, we found reliable differences starting at approximately 280 ms after stimulus onset, depending on the amount of perceptual evidence accumulated by participants. More specifically, the posterior cingulate cortex (PCC) and the parahippocampal gyrus (PHG) showed a distinctive response profile, characterized by a monotonic accumulation of evidence. Conversely, categorical recognition effects were evidenced in medial frontal regions, including the dorsal anterior cingulate cortex (dACC) (Schettino, et al., 2011). However, none of these non-overlapping brain effects was found to depend upon the emotional content of the scenes. Hence, whereas behavioral results suggested that the emotional content of the scene could reliably slow down the generation of online predictions regarding the actual identity of the stimulus, we did not find a corresponding electrophysiological correlate for this effect. Based on this evidence, we therefore concluded that these proactive brain mechanisms during scene recognition were probably generic, albeit delayed for emotional relative to neutral scenes.

In the present study, we used a modified version of this novel experimental paradigm (Schettino, et al., 2012; Schettino, et al., 2011), with the aim to resolve this apparent discrepancy and assess whether emotion may exert an early modulatory effect on the extraction of diagnostic visual information during proactive processes leading to scene recognition. To further promote the online generation and active use of predictions regarding the likely identity of the scenes progressively revealed, we introduced a standard memory manipulation (Courtney, Ungerleider, Keil, & Haxby, 1997; Goldman-Rakic, 1990; Henson, Hornberger, & Rugg, 2005; Rugg, et al., 1998). For each trial, participants were first asked to encode a new complex colorful scene having either a neutral or emotional meaning. After a constant time interval, the content of either the same or a novel scene was progressively unfolded, similarly to Schettino, et al. (2011). Participants were asked to perform a delayed match-to-sample task, as opposed to a binary animacy judgment task in our previous study (Schettino, et al., 2011). This standard procedure was meant to stimulate the generation of a restricted number of predictions regarding the identity of the scenes progressively revealed, bearing in mind that these predictions were primarily shaped by the initial encoding phase. Moreover, here the emotional content of the scene was task-relevant, as opposed to fully irrelevant in Schettino, et al. (2011). Notably, several lines of evidence suggest that (early and automatic) affective stimulus processing is substantially reduced when concurrent non-affective semantic stimulus dimensions become task-relevant (Everaert, et al., 2011; Pessoa, Kastner, et al., 2002; Spruyt, et al., 2007). We surmised that the lack of clear emotional effect at the electrophysiological level in our previous study was partly related to the use of an orthogonal task, which made the processing of the emotional content of the scenes superficial. To overcome this problem, in the present study, we occasionally asked participants to rate the emotional content of the scene, besides the delayed match-to-sample task. This procedure ensured that participants directly attended to the emotional content of the scenes on each and every trial. To provide additional evidence for the overt processing of the emotional dimension of the stimuli, we also used a standard ERP marker of emotional processing. More specifically, during encoding, we measured and compared the amplitude of the late positive potential (LPP) as a function of the emotional content of the scenes (Foti, Hajcak, & Dien, 2009; Schupp, Cuthbert, et al., 2004; Schupp, Stockburger, et al., 2003).

This experimental design is suited to assess whether the perceived emotional content of the scene could alter brain mechanisms responsible for the online generation of predictions regarding the actual identity of this scene. Furthermore we could compare, using advanced EEG methods, the processing of “old” (previously encountered) vs. “new” scenes, and hence evaluate whether any influence of emotion on brain mechanisms underlying scene recognition was depending on the (short-term) memory status of these complex visual scenes. We anticipated that, if participants successfully encoded the initial picture and later made active predictions regarding the content of the progressively unfolded stimulus, recognition of old scenes should occur earlier than new scenes. In addition, in line with our previous results

(Schettino, et al., 2011), we predicted that participants would recognize emotional scenes later than neutral ones, suggesting that proactive processes during scene recognition are reliably influenced by the perceived emotional content of the scenes. At the electrophysiological level, we expected to observe dissociable response profiles in dorsal/medial frontal regions, as opposed to more ventral brain areas, depending on the extent of prior encounter with these scenes (Schettino, et al., 2011).

2. Methods

2.1. Participants

Twenty-two undergraduate psychology students (all women, mean age 21 years, range 18-26) participated in the study, which was approved by the local Ethics Committee. All participants were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders. The data of one participant could not be saved properly because of technical problems, whereas the EEG data of three other participants were discarded due to excessive artifacts and a low signal-to-noise ratio. Thus, the final sample consisted of 18 participants. All volunteers gave informed written consent prior to their participation, and were compensated 30€.

2.2. Stimuli

The visual stimuli were selected from the International Affective Picture System (IAPS; Lang, et al., 2008). This database provides normative ratings for the basic dimensions of emotion -- including arousal and valence -- obtained with the Self-Assessment Manikin (SAM; Bradley & Lang, 1994). Our stimuli list consisted of 360 pictures, equally divided into three emotion categories according to their pre-defined valence scores: *neutral*, *unpleasant*, and *pleasant* (Table 1) (Schettino, et al., 2012). These pictures were selected on the basis of mean valence and arousal ratings reported by female responders (Lang, et al., 2008), because only women eventually participated in the main experiment (see above). Similar to our previous study (Schettino, et al., 2011), we selected scenes that were neither highly pleasant (i.e., erotic situations) nor highly unpleasant (i.e., mutilations), because these categories were associated with specific emotion reactions in previous research (Schupp, Flaisch, Stockburger, & Junghöfer, 2006; Schupp, et al., 2007). We also included 16 additional neutral pictures that were only used during the practice session (not included in the subsequent statistical analyses). Furthermore, 36 supplementary neutral scenes were scrambled (i.e., each scene was divided into arbitrary grids of 255 x 255 pixels, whose locations were randomly shuffled 10 times) and, as a result, their content was made meaningless. Thus, in total, participants were presented with 412 IAPS scenes[2].

Table 1. Mean values and standard deviations (in parenthesis) of normative valence and arousal scores of the selected IAPS pictures

Emotion category	Valence	Arousal
Neutral	5.14 *(1.38)*	3.68 *(2.05)*
Unpleasant	3.17 *(1.61)*	4.94 *(2.15)*
Pleasant	6.95 *(1.70)*	4.97 *(2.30)*

Note. Scores range from 1 to 9. Independent samples t-test confirmed a significant difference in valence ratings between neutral and unpleasant pictures [$t(119) = 29.34$, $p < .001$], as well as between neutral and pleasant [$t(119) = -26.82$, $p < .001$] and unpleasant and pleasant [$t(119) = -52.58$, $p < .001$] scenes. Significant differences in levels of arousal were also observed between neutral and unpleasant [$t(119) = -29.34$, $p < .001$] and neutral and pleasant [$t(119) = -30.98$, $p < .001$] pictures. However, no difference was evidenced between unpleasant and pleasant scenes [$t(119) = -0.77$, $p = .441$], confirming a balanced level of activation between these two emotion conditions.

Each neutral, unpleasant and pleasant picture was arbitrarily paired with another one from the same emotion category based on low-level visual similarities, as assessed by systematic

---

[2] IAPS scenes selected for the experiment. *Practice*: 2107, 2600, 2980, 5533, 5731, 6837, 7017, 7030, 7036, 7055, 7057, 7140, 7224, 7365, 8121, 8312. *Neutral*: 1350, 1616, 1675, 1903, 1935, 1947, 2025, 2026, 2034, 2191, 2272, 2273, 2279, 2308, 2357, 2377, 2382, 2383, 2390, 2396, 2445, 2446, 2489, 2495, 2514, 2575, 2579, 2593, 2595, 2597, 2606, 2702, 2720, 2749, 2850, 2880, 4090, 4150, 4220, 4250, 4255, 4274, 4275, 4320, 4325, 4605, 4750, 5040, 5395, 5500, 5531, 5532, 5534, 5535, 5900, 6570.2, 7001, 7002, 7003, 7009, 7011, 7014, 7016, 7018, 7019, 7021, 7032, 7033, 7037, 7038, 7042, 7043, 7044, 7045, 7058, 7061, 7062, 7081, 7096, 7130, 7160, 7161, 7170, 7180, 7184, 7186, 7188, 7190, 7207, 7236, 7242, 7247, 7248, 7249, 7255, 7287, 7300, 7354, 7484, 7487, 7493, 7500, 7503, 7506, 7512, 7513, 7546, 7547, 7550, 7590, 7595, 7710, 7820, 7830, 8241, 8311, 8325, 9210, 9260, 9700. *Unpleasant*: 1230, 1240, 1270, 1275, 1280, 1390, 1505, 1617, 1945, 2115, 2130, 2141, 2205, 2276, 2278, 2400, 2455, 2456, 2525, 2681, 2682, 2694, 2695, 2700, 2715, 2716, 2718, 2745.2, 2750, 2752, 2770, 2795, 2799, 2810, 2900.1, 3061, 3160, 3181, 3190, 3210, 3216, 3280, 3300, 3301, 4621, 4635, 4770, 5970, 5973, 6000, 6240, 6241, 6311, 6314, 6561, 6562, 6610, 6800, 6832, 7013, 7023, 7079, 7092, 7136, 7137, 7520, 7521, 8231, 9002, 9005, 9008, 9031, 9041, 9045, 9046, 9080, 9090, 9102, 9145, 9171, 9180, 9182, 9186, 9265, 9270, 9290, 9291, 9295, 9320, 9330, 9331, 9341, 9342, 9390, 9395, 9402, 9404, 9411, 9415, 9417, 9419, 9421, 9435, 9440, 9445, 9469, 9471, 9561, 9584, 9592, 9596, 9635.2, 9830, 9831, 9832, 9912, 9913, 9922, 9926, 9927. *Pleasant*: 1340, 1463, 1540, 1590, 1595, 1640, 1659, 1660, 1720, 1721, 1811, 1999, 2055.2, 2056, 2092, 2151, 2156, 2158, 2224, 2274, 2300, 2331, 2344, 2346, 2352, 2398, 2605, 2616, 2655, 3005.2, 4500, 4530, 4534, 4536, 4559, 4571, 4600, 4601, 4603, 4606, 4610, 4612, 4614, 4616, 4617, 4619, 4623, 4624, 4641, 5199, 5215, 5260, 5301, 5480, 5600, 5622, 5628, 5660, 5700, 5814, 5829, 5831, 5849, 5990, 5994, 6250.2, 7200, 7230, 7250, 7260, 7279, 7281, 7282, 7286, 7289, 7291, 7350, 7352, 7390, 7400, 7410, 7430, 7440, 7460, 7461, 7470, 7477, 7481, 7482, 7488, 7489, 7492, 7496, 7501, 7505, 7508, 7515, 7570, 8032, 8050, 8118, 8120, 8162, 8208, 8220, 8280, 8340, 8350, 8371, 8420, 8460, 8461, 8465, 8467, 8497, 8503, 8510, 8531, 8540, 8620; *Scrambled*: 1112, 1303, 1310, 1645, 1726, 1908, 2002, 2018, 2032, 2038, 2101, 2102, 2104, 2122, 2190, 2220, 2221, 2393, 2440, 2441, 2458, 2480, 2484, 2493, 2506, 2512, 2516, 2518, 2570, 2580, 2635, 2704, 2780, 2830, 2840, 9070.

visual inspection. More specifically, pictures with a clear distinction between a central figure and a homogeneous background were paired together (e.g., a coffee mug on a table vs. a pocket watch on a dark background). The same strategy was applied for more complex scenes (e.g., a traffic jam vs. a woman in the crowd). All the pairs created with this procedure are reported in Table 2.

The selected IAPS pictures were resized to 922 x 691 pixels (90% of the original size) and pre-processed similarly to Schettino, et al. (2011). After grayscale conversion, six bandpass spatial frequency filters were applied on each and every picture (using ImageJ v1.44 software; http://rsb.info.nih.gov/ij/) (for a similar procedure, see Delplanque, et al., 2007). As a result, six distinct levels of filtering were obtained for every IAPS scene, each containing a different amount of low and high spatial frequency information (see also Figure 1A). All these modified pictures were finally resized to 768 x 576 pixels (75% of the original IAPS pictures).

2.1. Procedure

Participants were individually tested in a small, dimly lit room, and seated at a viewing distance of 75 cm in front of a 19" CRT computer screen (refresh rate: 100 Hz). After filling out the informed consent, they were presented with task instructions, followed by a practice block containing 16 trials (with neutral pictures). Then, they moved on to the main experimental session, which was divided into twelve blocks (separated by short breaks), each containing 33 trials (Schettino, et al., 2012; see also Figure 1A). Each trial began with a 1500ms presentation of a colorful, fully detailed picture subtending 18.5° x 13.9° of visual angle, followed by a 2000ms grayscale mask. At the offset of the mask, the actual unfolding sequence started. A fixation cross appeared in the center of the screen for 250 ms. The first grayscale, blurred image level of a given picture (subtending 15.4° x 11.6° of visual angle) was then presented for 500 ms, followed by a 250ms blank screen. Next, the second image level of the same picture (containing slightly more HSF and LSF information) was displayed for 500 ms, plus the 250ms blank screen, and the same procedure was repeated until the presentation of the sixth, non-filtered image level. The inter-trial interval was set at 1000 ms.

Table 2. Stimulus pairs created for the Progressive Unfolding task.

Pair number	Stimulus pairs					
	Neutral		Unpleasant		Pleasant	
	First element	Second element	First element	Second element	First element	Second element
1	2191	7513	2455	9180	1640	7286
2	2272	7500	2525	9635.2	1660	4641
3	2308	4250	3300	2752	2158	2156
4	2357	8311	5970	2694	2274	8208
5	2382	7242	5973	9912	2605	7291
6	2390	5535	6000	2115	2616	2300
7	2514	7061	6241	6832	4530	4500
8	2575	2273	6610	6800	4600	2398
9	2579	2595	7013	9926	4616	4610
10	2606	7037	7079	9041	4619	7260
11	2880	7493	7136	9186	4624	7410
12	4090	7003	7137	7092	5260	7440
13	5040	7161	8231	9440	5622	8620
14	5900	6570.2	9080	2715	5831	2056
15	7009	7190	9102	6314	5849	5628
16	7011	4320	9171	2718	5990	7496
17	7014	2377	9182	2456	5994	8120
18	7021	7248	9265	9031	6250.2	8032
19	7038	5532	9290	9320	7200	8510
20	7042	2034	9291	9342	7279	7489
21	7044	7130	9330	9832	7430	7352
22	7045	2396	9395	3181	7460	5480
23	7062	7186	9415	9471	7477	8465
24	7207	7032	9417	6561	7482	8540
25	7287	2026	9421	2900.1	7501	7505
26	7484	7096	9435	7520	7508	5199
27	7503	1350	9584	9469	7570	5814
28	7590	2850	9592	9270	8460	8497
29	7830	7546	9596	2205	8461	2352
30	9260	4275	9831	9402	8503	7470
31	1616	2445	1270	1275	1340	8420
32	1675	2593	1230	2799	1463	8280
33	1903	7255	1240	1617	1540	1595
34	1947	5531	1280	9830	1590	1720

Pair number	Stimulus pairs					
	Neutral		Unpleasant		Pleasant	
	First element	Second element	First element	Second element	First element	Second element
35	2025	7506	1390	2745	1721	8340
36	2446	2383	1505	9002	2224	4606
37	2489	1935	1945	9419	2331	8350
38	2495	2702	2130	9045	2344	1811
39	2720	7033	2141	9090	3005.2	4571
40	2749	4325	2276	2681	4536	2346
41	4150	2597	2682	2795	4559	2055.2
42	4274	7160	2695	9404	4601	7282
43	4750	4255	2716	2700	4603	8162
44	5534	7547	2810	9913	4612	2151
45	7018	9210	3061	6311	4614	7488
46	7019	7300	3160	9005	4617	2092
47	7043	7016	3190	7521	4623	7481
48	7081	7001	3210	6240	5301	8531
49	7170	7002	3216	4770	5600	7350
50	7180	4605	3280	6562	5660	5215
51	7184	7236	4635	9008	7230	1999
52	7188	7820	9445	9927	7250	7461
53	7247	7249	2278	9295	7281	2655
54	7354	7058	2400	9145	7390	5700
55	7487	8325	2770	9341	7492	5829
56	7512	2279	9390	7023	7515	8467
57	7550	4220	9922	9561	8050	4534
58	7595	5395	9046	4621	8118	1659
59	7710	5500	2750	9411	8220	7289
60	8241	9700	3301	9331	8371	7400

Note. These numbers refer to picture codes, as available in the original database (Lang, et al., 2008).

The aim of this experimental manipulation was to promote a gradual and predictive accumulation of visual information by progressively adding, in a stepwise fashion, diagnostic LSF and HSF information to an initial blurred and undistinguishable picture. Importantly, the grayscale and resize picture conversions (unfolding) relative to the colorful picture (encoding) were performed to discourage participants to use a purely perceptual, pixel-to-pixel matching strategy to retrieve the content of the initial picture during the memory matching task. Participants were required to provide two separate and consecutive manual responses (see also Schettino, et al., 2011, for a similar dual response procedure). First, they were asked to press with their right index finger a pre-defined button on a response box (Cedrus RB-730; http://www.cedrus.com/responsepads/rb730.htm) as soon as they gathered enough perceptual evidence to decide whether the content of the unfolded scene was either the same as the one displayed during the encoding phase, a new one, or a scrambled picture (Response1). These scrambled pictures, for which a separate response was required (see below), were used as "catch" trials to ensure that participants reliably attended to the content of the scenes before responding. Pressing the button immediately interrupted the presentation of the stimulus sequence. After 500 ms, participants were required to validate their first response (Response1). They were asked to press, on a standard AZERTY keyboard, the "O" key if the unfolded scene was the same as the colorful one previously presented ("old" condition), the "N" key if these two scenes were different ("new" condition), or the "S" key if the unfolded scene was displaying a meaningless content ("scrambled" condition). All these responses, for which no time limit was imposed, were coded as Response2. This dual response procedure was used for three main reasons: (i) to dissociate early visual recognition effects (Response1) from the overt discrimination and comparison in short-term memory of the scenes (Response2); (ii) to remove recognition errors (based on accuracy of Response2) from the behavioral and ERP analyses; (iii) to minimize the potential contamination of ERP data by the activation of competing motor responses across the three different conditions, because Response1 always required to press a single button shared across these three conditions. Although instructions emphasized accuracy, participants were encouraged to stop the sequence as soon as they felt they could recognize the content of the scene, which happened to occur before the end of the sequence for a vast majority of trials (see results below). Instructions encouraged participants to perform the memory matching task by using abstract visual representations stored in short-term memory to compare the visual input progressively revealed during unfolding against the colorful scene shown during encoding. The inclusion of "catch" trials (i.e., scrambled scenes) likely prevented the use of a strategy based on low-level details, as confirmed by our behavioral results (see below). In addition, the content of half of the ”old” scenes was unpredictably flipped along the horizontal axis between encoding and retrieval. Participants were informed that an "old" response was to be given for these "flipped" trials, since the memory matching task had to be performed primarily based on the *content* of the scenes. For the statistical analyses of the behavioral and EEG data reported hereafter, "old flipped" and "old unflipped" trials were combined into a single "old" condition, and these "old" trials were eventually compared to "new" trials (see also Schettino, et al., 2012). Hence, using this procedure, for each emotion category (neutral, pleasant, unpleasant), two trial types were compared to each other: “old”, in which the identity of the colorful picture was identical to the scene progressively unfolded; “new”, meaning that the identities of the colorful and unfolded scene were different (although matched as far as possible in terms of low-level visual properties). Importantly, for “new” scenes, no change in terms of emotional content ever occurred between the colorful picture and the gradually unfolded scene. In other words, a neutral colorful picture was always followed by the unfolding of a neutral scene, and the same occurred for emotion-laden stimuli (pleasant-pleasant; unpleasant-unpleasant; see Table 2).

Finally, we sought to verify whether the emotional content of the IAPS pictures selected in our study was actually perceived by our participants in accordance with the normative ratings (Lang, et al., 2008). Moreover, we wanted to make the emotional content of the scenes somehow task-relevant throughout the experiment, in order to increase the likelihood to find reliable differences at the electrophysiological level between emotional and neutral scenes. Therefore, additional ratings of the emotional valence of the colorful scene presented at the beginning of each trial were occasionally asked after the registration of Response2. A standard 9-point SAM (Bradley & Lang, 1994) was used for this purpose, with anchor 1 corresponding to “very unpleasant” and anchor 9 to “very pleasant”. This additional emotion classification task was included in 10% of the total number of trials.

Stimulus presentation and behavioral response recordings were controlled using E-Prime 2.0. (http://www.pstnet.com/products/e-prime/).

2.1. Questionnaires

At the end of the experimental session, participants filled out two questionnaires measuring specific affective dispositions. Levels of trait anxiety were measured by means of the validated Dutch version of the State-Trait Anxiety Inventory (Van der Ploeg, et al., 1979). Participants also completed the Need For Affect Scale (Maio & Esses, 2001), which provides an estimate of the general motivation of participants to either approach or avoid emotion-inducing situations. The results confirmed that our participants had sub-clinical levels of trait anxiety, as well as normal Need For Affect values (Table 3). However, none of these personality measures was found to correlate significantly with either the behavioral or ERP results collected in our study (see also Schettino, et al., 2011, for similar results).

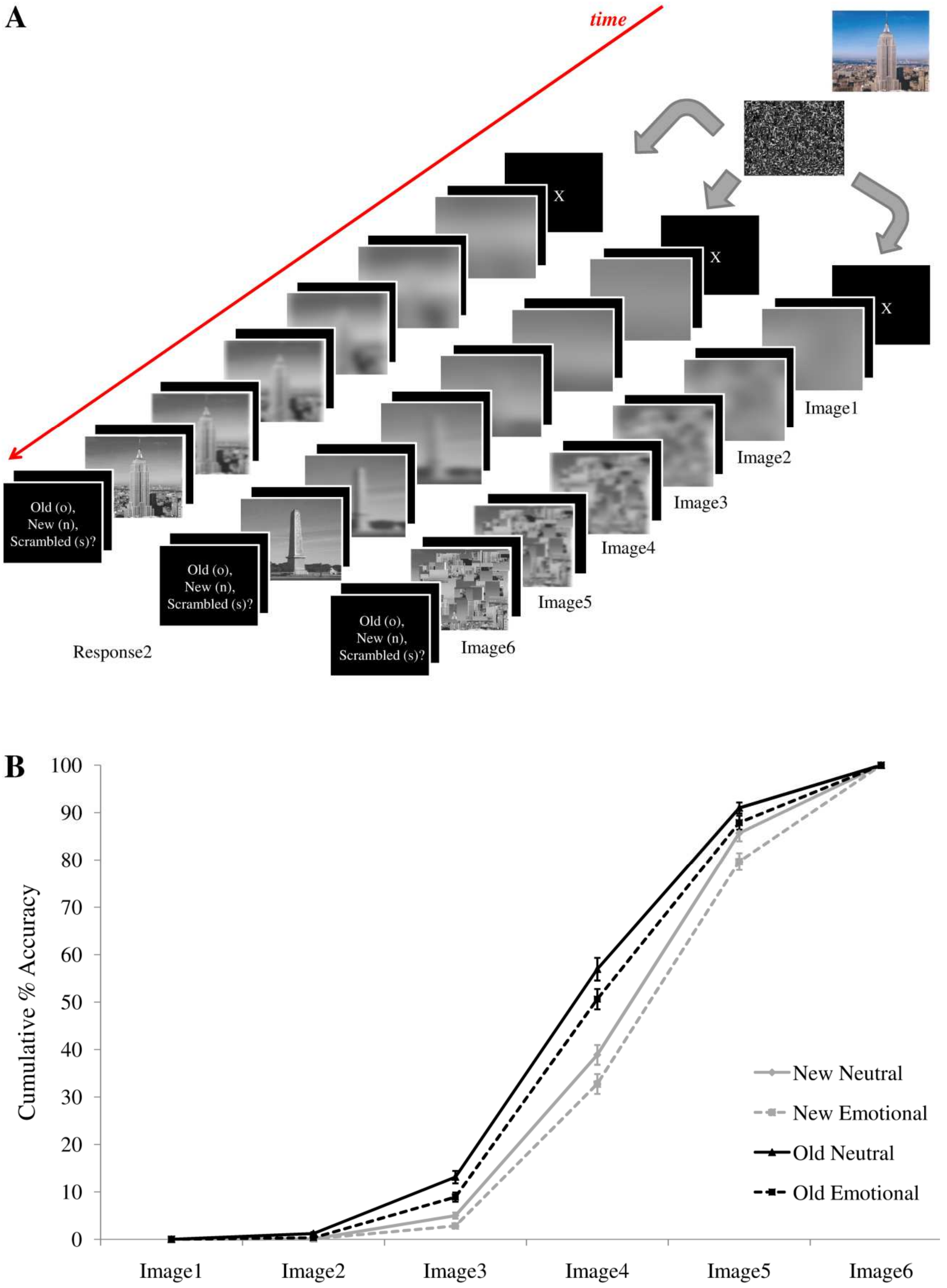


Figure 1. (A) Procedure and task. Colorful neutral, unpleasant and pleasant IAPS scenes (not shown here for copyright reasons) were randomly presented to participants for 1500 ms. After a 2000ms mask, the same scene, a new one, or a scrambled picture was progressively revealed in grayscale, using six successive steps. Each step added new low and high spatial frequency information to the previous one, in a parametric fashion. Each image level was presented for 500 ms, followed by a 250ms blank screen. Participants were required to press a button on the response box (Response1) as soon as they could decide whether the content of the gradually unfolded scene was the one seen at the beginning of the trial (i.e., colorful picture), a new one, or a scrambled picture.

Five hundred milliseconds after Response1, participants validated their choice and confirmed whether the scene was “old”, “new” or “scrambled” (Response2). (B) Behavioral results. Cumulative percentage of correct Responses1 as a function of the six image levels, separately for new (grey lines), old (black lines), neutral (solid lines) and emotional (dashed lines) scenes. Vertical bars correspond to standard errors of the means. Results show that participants were able to perform the memory task reliably earlier (i.e., less perceptual evidence needed) for old relative to new scenes. Moreover, neutral scenes were recognized reliably earlier compared to emotional scenes. No significant interaction effect between memory and emotion was found (see main text).

Table 3. Mean values and standard deviations (in parenthesis) of the scores obtained for each questionnaire (and relative subscales) administered at the end of the experiment.

Questionnaire		Score
STAI-T		40.61 *(10.17)*
NFAS		3.98 *(0.51)*
	Approach	4.79 *(0.61)*
	Avoidance	3.16 *(1.16)*

Note. STAI-T: State-Trait Anxiety Inventory, trait version; NFAS: Need for Affect Scale. STAI-T scores range from 20 to 80. NFAS scores were obtained using a 7-points Likert scale.

2.2. *Recording and pre-processing of EEG data*

Electroencephalographic (EEG) activity was continuously recorded using a BIOSEMI Active-Two system (BioSemi, Inc., Netherlands; http://www.biosemi.com) by means of 128 Ag/AgCl electrodes fitted into a stretching cap and following the BioSemi ABCD position system (i.e., electrode positions are radially equidistant from CZ; http://www.biosemi.com/headcap.htm). Two additional electrodes, the common mode sense (CMS) active electrode and the driven right leg (DRL) passive electrode, were used as reference and ground electrodes, respectively (http://www.biosemi/faq/cms_and_drl.htm). Vertical electro-oculograms (vEOG) were monitored using two additional electrodes placed on the inferior and superior areas of the left orbit, whereas horizontal EOG (hEOG) were recorded by means of two electrodes situated symmetrically on the outer canthus of each eye. EEG and EOG recordings were sampled at 512 Hz.

ERPs of interest were computed offline using Brain Vision Analyser 2.0 (Brain Products™ GmbH, Munich, Germany; http://www.brainproducts.com/analyser2_release.php). First, a topographic interpolation (interpolation by spherical splines; Perrin, et al., 1989) was applied on noisy channels (interpolated channels across 18 subjects: *M* = 3.78, *SD* = 4.28, range 0-11), and a common average reference was then applied. Afterwards, the continuous EEG signal was segmented into individual epochs, excluding trials corresponding to errors (based on accuracy on Response2). Similar to our previous study (Schettino, et al., 2011), four main epochs were computed around stimulus onset (using an interval of 200 ms prior to and 750 ms after this event), enabling us to look backwards at visual ERPs progressively elicited prior to actual recognition (i.e., Response1): (1) segments in which Response1 occurred from 0 to 750 ms after stimulus presentation ("recognition"); (2) segments in which Response1 occurred from 750 to 1500 ms after stimulus presentation ("one image before" recognition); (3) segments in which Response1 occurred from 1500 to 2250 ms after stimulus presentation ("two images before" recognition); (4) segments in which Response1 occurred from 2250 to 3000 ms after stimulus presentation ("three images before" recognition). A segment length of 750 ms was used in order to include stimulus presentation duration (500 ms) as well as the following 250ms blank screen. All these individual segments were then baseline corrected using the entire pre-stimulus interval of 200ms, before a standard ocular correction was performed (Gratton, et al., 1983). Artifact rejection was then carried out (−80/+80µV amplitude scale across participants) to remove segments contaminated by artifacts, including residual eye blinks and muscle activity. Using this procedure, 21.82% of the EEG data were discarded. Finally, stimulus-locked ERP averages were computed, separately for each condition (memory: 2 levels; emotion: 3 levels; recognition times: 4 levels), resulting in 24 individual ERP averages. However, because no difference in recognition between unpleasant and pleasant pictures was observed at the behavioral level, these two emotion conditions were collapsed in order to increase the signal-to-noise ratio, leaving a total number of 16 individual ERP averages (percentage of segments kept after pre-processing: > 70%). These averages were eventually bandpass filtered, using 1-30 Hz cutoffs.

In an auxiliary analysis, we also computed visual ERPs time-locked to the onset of the colorful picture (epoch length: -200/+1500 ms) that was presented at the beginning of the trial and had to be encoded in short-term memory. We sought to establish whether emotional scenes were perceived and processed differently compared to neutral scenes, as indicated by a larger LPP amplitude for emotional relative to neutral scenes (Schupp, et al., 2006; Schupp, et al., 2000; Schupp, Junghöfer, Weike, & Hamm, 2004).

2.3. *Analysis of behavioral data*

Accuracy was expressed as percentage of correct responses. Since the six image levels were not independent of each other (the new visual information provided at each image level relied on perceptual evidence accumulated during previous levels), cumulative percentages were calculated. Using this procedure, we computed psychometric curves showing the evolution of recognition performance across the six image levels for each memory and emotion condition, separately. To statistically verify whether these psychometric curves reliably varied as a function of memory and emotion, we used a proportional odds model (Agresti, 2007). This data analysis allows to model the cumulative probability up to and including recognition from each image level k (k_= 1, ?, 5). The derived odds ratio expresses how much the odds of recognition from image level k or earlier is increased (if larger than 1) or decreased (if smaller than 1) across new, old, neutral and emotional contents, and thus provides a single number capturing the shift in psychometric curve. To account for dependencies of trials within the same subject, a multi-level version of the proportional odds model was used, similarly to our previous studies (Schettino, et al., 2012; Schettino, et al., 2011). The level of significance for these statistical analyses was set at $p < 0.05$.

2.4. *Analysis of ERP data*

First, we analyzed the average amplitude of the LPP component time-locked to the onset of the colorful scene -- extracted from a 500-1000ms time window following stimulus onset -- from an array of centro-parietal electrodes, where this component reached its maximum amplitude (Figure 2). Mean amplitudes of the LPP were analyzed by means of repeated measures ANOVAs, and t-tests were employed as post-hoc comparisons.

To examine the ERP data recorded during the progressive unfolding sequence, reference-free topographic analyses were used, similarly to our previous study (Schettino, et al., 2011). The basic principles of this method have extensively been described elsewhere (Lehmann & Skrandies, 1980; Michel & Murray, 2012; Michel, et al., 1999; Michel, et al., 2001; Murray, et al., 2008; Pourtois, et al., 2008). In short, this method allows to summarize a complex ERP data set into a smaller number of dominant scalp topographies (i.e., global configuration of the electric field across all 128 channels at each time frame). Subsequent analyses further enable to assess how the distribution and expression of these dominant topographies vary in time across experimental conditions, irrespective of changes in the strength of the ERP signal. Topographic analyses were performed using CARTOOL software (version 3.43; http://brainmapping.unige.ch/Cartool.htm; see also Brunet, Murray, & Michel, 2011). The dominant scalp maps were identified in the grand-average ERP data for each main condition (recognition, one image before, two images before and three images before) over a wide time window spanning from 0 to 600 ms after stimulus onset. To this end, a specific spatiotemporal clustering algorithm, the "Atomize and Agglomerate Hierarchical Clustering" (AAHC; for a detailed description, see Murray, et al., 2008; Tibshirani & Walther, 2005), was used. The optimal number of dominant maps best "explaining" the ERP dataset was determined, following standard practice, using a cross-validation criterion (Pascual-Marqui, et al., 1995). These dominant scalp topographies were then fitted back to the ERPs of each individual subject to obtain a quantitative estimate of each map's relative expression across subjects and conditions, as defined by the Global Explained Variance (GEV, or goodness of fit). GEV represents the sum of the explained variance weighted by the Global Field Power (GFP) at each moment in time (Michel & Murray, 2012; Pourtois, et al., 2008). GEV values were entered in repeated measures ANOVAs, with recognition level, memory and emotion as within-subject factors. Paired t-tests were used as post-hoc comparisons between conditions, when required to back up significant interaction effects. Whenever Mauchly's test indicated that the assumption of sphericity had been violated, the degrees of freedom were corrected using Greenhouse-Geisser estimates. The alpha level was set at $p < 0.05$.

2.5. Source localization analysis

To estimate the putative neural generators underlying the dominant voltage topographies identified at the scalp level by the previous analyses, we used sLORETA (Pascual-Marqui, 2002). sLORETA is a distributed linear inverse solution based on the neurophysiological assumption of coherent co-activation of neighboring cortical areas, known to have highly synchronized activity (Silva, et al., 1991). Accordingly, it estimates multiple simultaneously active sources without any *a priori* assumption on the number and position of the underlying dipoles (for a mathematical validation of this localization technique, see Sekihara, et al., 2005). sLORETA solutions are computed within a three-shell spherical head model co-registered to the MNI152 template (Mazziotta, et al., 2001), restricted to the grey matter and the hippocampus. The source locations were therefore given as (x, y, z) coordinates (x from left to right; y from posterior to anterior; z from inferior to superior). The estimation of the three-dimensional intracerebral current density distribution is performed in 6239 voxels (5 mm resolution), each containing an equivalent current dipole. The head model for the inverse solution uses the electric potential lead field computed with a boundary element method applied to the MNI152 template (Fuchs, et al., 2002). Scalp electrode coordinates on the MNI brain are derived from the international 5% system (Jurcak, et al., 2007). The calculation of all reconstruction parameters was based on the computed common average reference. sLORETA units were scaled to amperes per square meter (A/m^2), normalized across subjects, and non-parametric statistics were then applied on log transformed data. Direct statistical comparisons between conditions were performed in this inverse solution space using repeated measures ANOVAs and post-hoc paired t-tests. The level of significance for all theses analyses was set at $p < 0.05$.

3. Results

3.1. *Emotion classification*

Results of the emotion classification task showed high (i.e., more pleasant) ratings for pleasant ($M = 6.14$, $SD = 0.81$), intermediate for neutral ($M = 4.89$, $SD = 0.58$), and low for unpleasant ($M = 3.45$, $SD = 1.06$) scenes. A one-way ANOVA performed on these mean ratings disclosed a highly significant effect of emotion [$F(2, 34) = 39.94$, $p < .001$, $\eta_p^2 = .701$]. Paired samples t-tests carried out on these mean ratings confirmed highly significant differences between neutral and unpleasant pictures [$t(17) = 4.83$, $p < .001$], as well as between neutral and pleasant [$t(17) = -7.47$, $p < .001$] and unpleasant and pleasant [$t(17) = -6.81$, $p < .001$] scenes. Thus, participants rated the emotional valence of the pre-selected stimuli in accordance with the published normative scores (Lang, et al., 2008). These results also confirmed that participants correctly attended to the actual emotional content of the scenes throughout the experiment.

3.2. *Errors and late responses for the progressive unfolding task*

Participants accurately performed the delayed matching task during gradual stimulus revelation. The percentage of errors was negligible ($M = 3.66\%$, $SD = 1.85$). Likewise, an extremely low error rate was obtained for “catch” trials ($M = 1.75\%$, $SD = 1.90$). The percentage of late responses (Responses1 occurring after the last/sixth image level) was also negligible ($M = 1.71\%$, $SD = 1.18$) (see also Table 4).

3.3. Accuracy for the progressive unfolding task

Cumulative percentages of correct responses (i.e., Responses1 only when Responses2 were correct) are presented in Table 4. A mixed proportional odds model (Agresti, 2007; Schettino, et al., 2011) with memory (old, new) and emotion (neutral, emotional) as fixed factors, and participant as random effect was carried out on these values, to verify whether the psychometric curve shifted as a function of memory and/or emotion (see Figure 1B). This analysis revealed an overall earlier recognition for old compared to new scenes, in both neutral and emotional conditions (all *p*s < .001). Interestingly, pairwise comparisons revealed a shift of the psychometric curve as a function of the emotional content of the scene, indicated by an earlier recognition when the scenes contained a neutral as opposed to an emotional content (all *p*s < .01)[3]. No significant interaction was found between memory and emotion (*p* = .632).

3.1. ERP results

3.1.1. LPP during stimulus encoding

Figure 2 shows the grand-average ERPs recorded from an array of medial centro-parietal electrode sites (A8, A20, B5, where A20 roughly corresponds to Pz in the 10/20 international EEG system). These electrodes were selected for illustration purposes after an initial 2 (emotion) x 9 (electrode) repeated measure ANOVA revealed no significant interaction effect between these two factors (*p* > .05). The analysis performed on the mean amplitude of the LPP showed a larger sustained positive component for emotional (*M* = 5.86 µV, *SD* = 4.32) relative to neutral (*M* = 4.91 µV, *SD* = 4.47) scenes. Paired t-tests confirmed a significant amplitude difference between neutral and emotional scenes [*t*(17) = -3.75, *p* = .002]. Thus, these results for the LPP component provided additional evidence for the differential processing of the emotional content of the scenes encoded in short-term visual memory prior to unfolding.

---

[3] The same mixed proportional odds model did not reveal significant differences between pleasant and unpleasant scenes (all *p*s > .05).

Table 4. Mean values and standard deviations (in parenthesis) of cumulative percentages of correct responses, errors and late responses, separately for each image level, emotion, and memory condition.

Image Level	New		Old	
	Neutral	Emotional	Neutral	Emotional
Image1	0.00 *(0.00)*	0.00 *(0.00)*	0.00 *(0.00)*	0.00 *(0.00)*
Image2	0.10 *(0.41)*	0.09 *(0.39)*	1.22 *(2.42)*	0.32 *(0.58)*
Image3	5.01 *(5.36)*	2.81 *(3.16)*	13.12 *(10.40)*	8.90 *(6.97)*
Image4	38.88 *(17.72)*	32.74 *(17.19)*	56.93 *(19.28)*	50.62 *(16.98)*
Image5	85.66 *(14.51)*	79.66 *(14.55)*	90.98 *(8.55)*	87.90 *(11.73)*
Image6	100.00 *(0.00)*	100.00 *(0.00)*	100.00 *(0.00)*	100.00 *(0.00)*
Errors	2.96 *(2.18)*	2.64 *(1.67)*	4.26 *(3.04)*	4.77 *(3.15)*
Late Responses	1.95 *(2.08)*	2.08 *(1.62)*	1.39 *(2.23)*	1.43 *(1.45)*

3.1.1. Topographic analysis of ERPs recorded during the unfolding task

Similarly to our previous study (Schettino, et al., 2011), we found a reliable topographical change across recognition levels immediately following the exogenous N1 and P2 ERP components (Figure 3A). Because this topographical alteration necessarily indicates changes in the configuration of the underlying intracranial generators (Lehmann & Skrandies, 1980; Murray, et al., 2008), a detailed characterization of this topographical change starting ~400 ms following stimulus onset is provided.

At the offset of the P2 scalp map, for three, two and one image before recognition, an occipital positivity with a concurrent, broad frontal negativity was evidenced. Conversely, an occipital positivity accompanied by a more focal frontal positivity was observed one image before recognition, being then predominant at time of recognition. These two distinctive dominant topographical components were then fitted back to the individual ERP data. Based on the outcome of the spatiotemporal analysis, we selected a large time interval lasting 120 ms, namely from 402 to 522 ms after stimulus onset, during which these topographic differences were most obvious (Figure 3A). A 2 (map configuration: occipital positivity/frontal negativity, occipital positivity/frontal positivity) x 4 (recognition level: recognition, one image before, two images before, three images before) repeated measures ANOVA on the GEV values obtained for these two specific topographies revealed significant main effects of map configuration [*F*(1, 17) = 7.40, *p* = .015, η_p^2 = .303] and recognition level [*F*(3, 51) = 10.69, *p* < .001, η_p^2 = .386], as well as a highly significant map configuration x image level interaction [*F*(3, 51) = 9.53, *p* < .001, η_p^2 = .359]. Interestingly, pairwise comparisons showed that the GEV of the occipital positivity/frontal positivity map progressively increased from three images before recognition to actual recognition (all *p*s < .05, except two images vs. three images before recognition, *p* = .079) (see Figure 3C). By comparison, the GEV of the occipital positivity/frontal negativity showed a different pattern (Figure 3B). Paired t-tests revealed a sharp increase in GEV from three to two images before recognition [*t*(17) = -3.88, *p* = .001], followed by a gradual decrease between two images before and time of recognition [*t*(17) = 3.21, *p* = .005], as well as between one image before and actual recognition [*t*(17) = 2.67, *p* = .016]. Thus, this latter topographical scalp map explained most of the variance already two images before recognition, but then gradually decreased and was replaced by the previously described occipital positivity/frontal positivity.

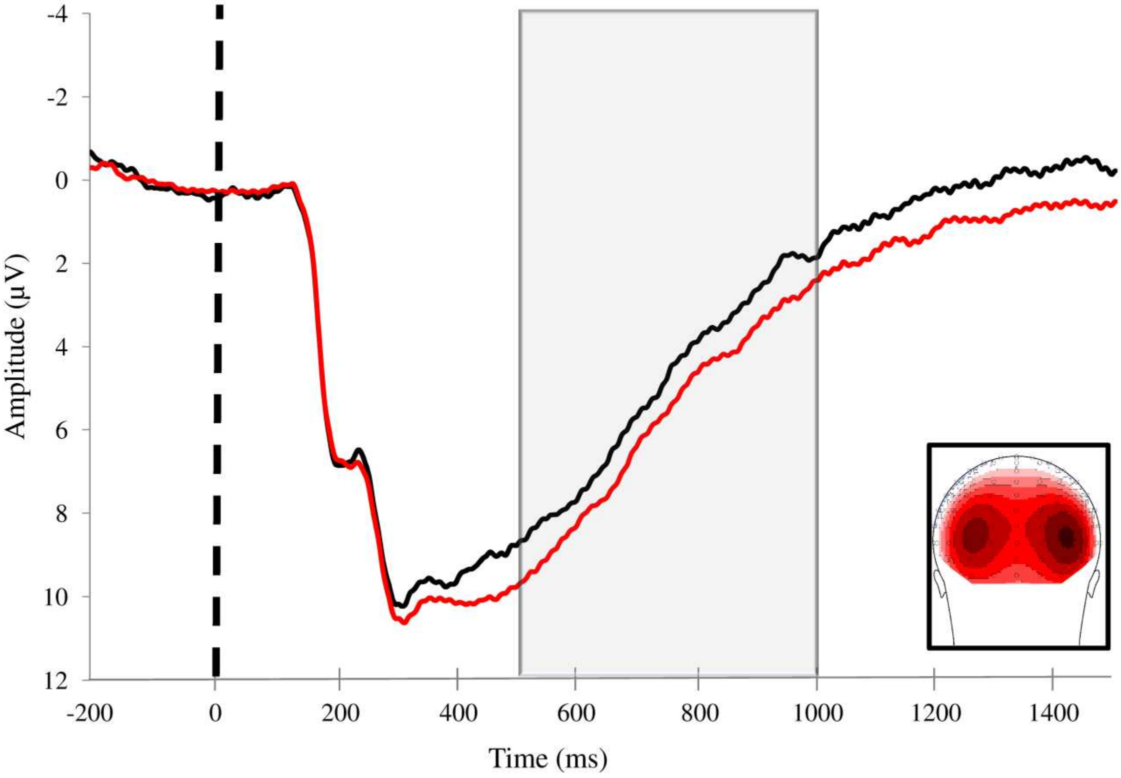
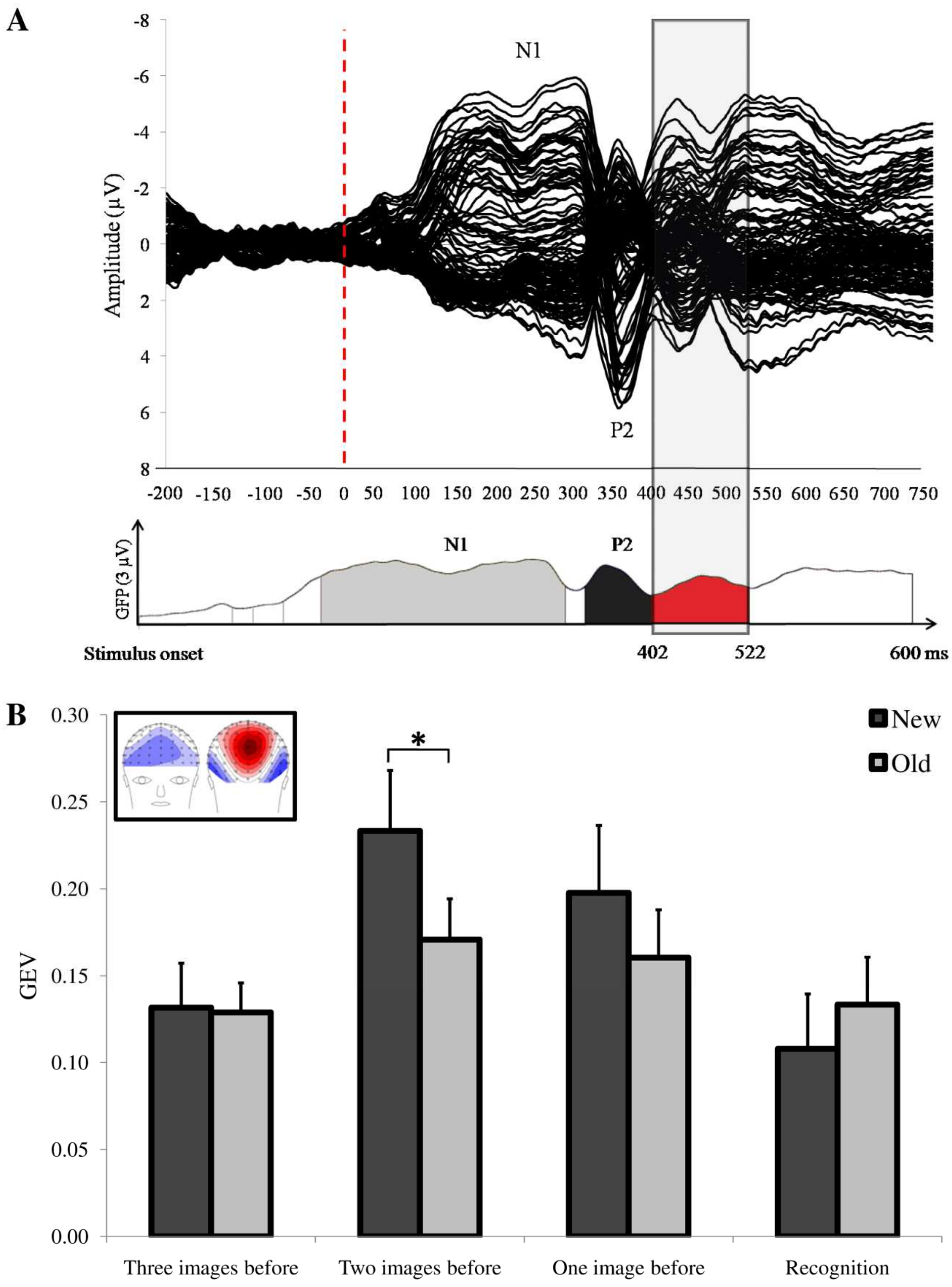


Figure 2. Grand-average LPP recorded from a posterior parietal cluster of electrodes (average of activity recorded from electrodes A8, A20, B5), separately for neutral (black line) and emotional (pleasant and unpleasant merged; red line) scenes. The black vertical dashed line indicates the onset of the colorful picture. The grey area indicates the time window during which the mean amplitude of the LPP was measured (500-1000 ms post-stimulus onset). LPP was significantly larger (*p* = .002) for emotional compared to neutral scenes.

Next, we investigated whether these two dominant topographical maps were differentially influenced by memory, emotional content, or both factors concurrently. Regarding the occipital positivity/frontal negativity topography, a 2 (emotion) x 2 (memory) x 4 (recognition level) repeated measures ANOVA showed a significant memory x recognition level interaction [$F(3, 51) = 3.28$, $p = .028$, $\eta_p^2 = .162$], but no significant effect of emotion[$F(1, 17) = 1.01$, $p = .330$, $\eta_p^2 = .056$]. A similar 2 x 2 x 4 ANOVA on the GEV values of the occipital positivity/frontal positivity map disclosed a significant main effects of memory [$F(1, 17) = 20.30$, $p < .001$, $\eta_p^2 = .544$], and a memory x recognition level interaction [$F(3, 51) = 3.36$, $p = .026$, $\eta_p^2 = .165$] but, again, no significant effect of emotion [$F(1, 17) = 0.29$, $p = .596$, $\eta_p^2 = .017$].

Given the lack of statistical significance for the emotion factor, we collapsed GEV values for neutral and emotional scenes and further investigated the modulatory role of memory. A 2 (memory) x 4 (recognition level) repeated measures ANOVA on the GEV values of the occipital positivity/frontal negativity map disclosed a significant memory x recognition level interaction [$F(3, 51) = 3.28$, $p = .028$, $\eta_p^2 = .162$]. Post-hoc comparison showed higher GEV for new compared to old condition two images before recognition [$t(17) = 2.52$, $p = .022$] (Figure 3B). Thus, this analysis revealed a similar pattern of activity in new and old conditions for this occipital positivity/frontal negativity topographical map, with higher GEV values for new relative to old condition two images before recognition.

A comparable 2 x 4 ANOVA was used for the GEV values obtained for the concurrent occipital positivity/frontal positivity topography. This analysis showed a significant main effect of memory [$F(1, 17) = 20.30$, $p < .001$, $\eta_p^2 = .544$], as well as a significant memory x recognition level interaction [$F(3, 51) = 3.36$, $p = .026$, $\eta_p^2 = .165$]. Post-hoc paired t-test revealed higher GEV values for new relative to old condition, both one image before [$t(17) = 3.41$, $p = .003$] and at recognition level [$t(17) = 4.12$, $p = .001$] (Figure 3C). Thus, a comparable, gradual increase in GEV was observed for new and old scenes, with higher values for new relative to old condition one image before, as well as at time of recognition.


A
-8
-6
-4
-2
0
2
4
6
8
Amplitude (µV)
N1
P2
-200 -150 -100 -50 0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750
GFP (3 µV)
N1
P2
Stimulus onset
402
522
600 ms
B
0.30
0.25
0.20
0.15
0.10
0.05
0.00
GEV
*
New
Old
Three images before
Two images before
One image before
Recognition

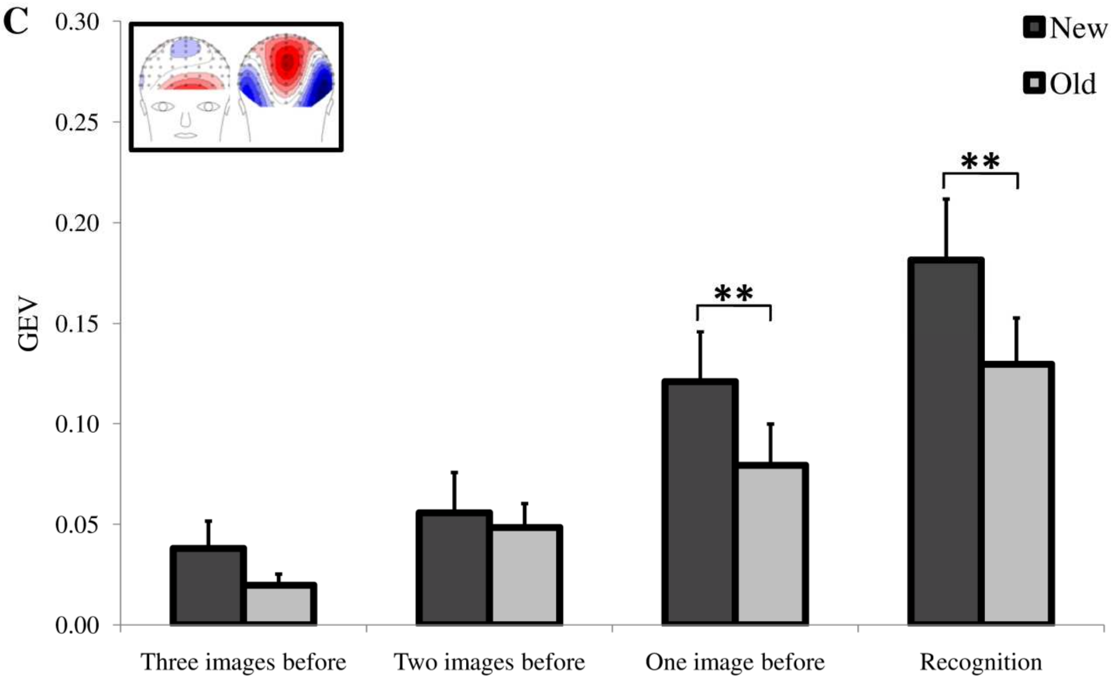


Figure 3. Topographic mapping results. (A, upper panel) Grand-average (N = 18) ERP waveforms, obtained for the level corresponding to actual recognition, for all 128 electrodes concurrently (butterfly). The red dashed vertical line indicates the onset of the visual stimulus. Exogenous N1 and P2 ERP components were recorded after stimulus onset, and were shared across recognition levels. The grey area indicates the time window during which a reliable topographical change occurred. (A, lower panel) The spatiotemporal cluster analysis (0-600 ms after stimulus onset) revealed reliable topographic differences in a time window spanning 402-522 ms following stimulus onset (in red). A dominant occipital positivity/frontal negativity scalp map was recorded for recognition level preceding recognition, whereas a concurrent occipital positivity/frontal positivity map was evidenced mostly one image before recognition, being then predominant at time of recognition. (B) Statistical results, obtained after the fitting procedure, for the GEV of the occipital positivity/frontal negativity scalp map identified in the 402-522ms time window following stimulus onset (separately for new vs. old scenes). Results showed a sharp increase in GEV from three images to two images before recognition, followed by a progressive decrease until recognition. Furthermore, a higher GEV for new relative to old scenes was evidenced two images before recognition. (C) On the other hand, a linear increase in GEV as a function of recognition level was evident for the concurrent occipital positivity/frontal positivity topography. A higher GEV was observed for new compared to old scenes one image before recognition, as well as at recognition. * $p < .05$; ** $p < .01$; *** $p < .001$. Vertical bars correspond to standard errors of the mean.

3.1.2. Source localization results

Using sLORETA (Pascual-Marqui, 2002) we performed statistical non-parametric mapping (SnPM) analyses on the average activity estimated during the 402-522ms post-stimulus onset interval, during which the main topographic change took place across the four recognition levels (from three images before up to recognition). We first compared ERP activity for three images before recognition to actual recognition (see Schettino, et al., 2011, for a similar approach), separately for new and old scenes. This analysis revealed, for new scenes, a stronger activation for recognition relative to three images before in the dACC

(±10*x*, +15*y*, +35*z*) [$t(17) = 2.75$, $p = .014$], whereas symmetrically stronger activity three images before compared to recognition was found in the insula (±40*x*, +15*y*, 0*z*) [$t(17) = -3.29$, $p = .004$] (Figure 4, upper panel). On the other hand, the SnPM analysis for old scenes showed stronger activity during recognition relative to three images before recognition in a non-overlapping network of brain regions, including the PHG (±29*x*, -49*y*, -6*z*) [$t(17) = 4.46$, $p < .001$] and the fusiform gyrus (FG; ±41*x*, -45*y*, -19*z*) [$t(17) = 4.02$, $p = .001$] (Figure 4, lower panel). Therefore, these four brain areas (dACC, insula, PHG and FG) were defined as regions of interest (ROIs) for further analyses[4]. For each ROI, we then extracted the mean amplitude (current density) value during the same time interval (402-522ms following stimulus onset) and statistically assessed potential effects of recognition level and emotion.

3.1.1. dACC and insula (new scenes)

Regarding the dACC (Figure 5A), the 2 (emotion) x 4 (recognition level) ANOVA performed on the mean current density extracted during the 402-522ms interval post-stimulus onset only revealed a significant main effect of recognition level [$F(3, 51) = 5.84$, $p = .002$, $\eta_p^2 = .256$]. Post-hoc t-tests revealed higher activity for late vs. early recognition levels, as evidenced by a significant difference between two images and one image before recognition [$t(17) = -2.74$, $p = .014$]. Activity remained stable between three images and two images before recognition, as well as between one image before and time of recognition (*p*s > .05) (Figure 5A).

Inverse solution results obtained for the insula in the new condition revealed a similar response profile (Figure 5B). However, the emotional content of the scenes visibly influenced reconstructed activity in this ROI. A 2 (emotion) x 4 (recognition level) ANOVA disclosed a significant emotion x recognition level interaction [$F(3, 51) = 3.75$, $p = .016$, $\eta_p^2 = .181$]. Post-hoc analyses showed, for neutral pictures, only a significant difference between three images before and time of recognition [$t(17) = -2.34$, $p = .032$]. Therefore, activity in the insula for neutral pictures was relatively stable until recognition. By contrast, emotional scenes elicited higher activity during late stages of accumulation of evidence, as evidenced by significantly higher activity one image relative to two images before recognition [$t(17) = -3.17$, $p = .006$]. Activity remained stable between three images and two images before

---

[4] None of these ROIs showed differential lateralization effects, as confirmed by 2 (side: left, right) x 4 (recognition level: recognition, one image before, two images before, three images before) repeated measures ANOVAs. Accordingly, for each ROI, activity from both hemispheres was combined.

recognition, as well as between one image before and time of recognition (*p*s > .05) (Figure 5B).

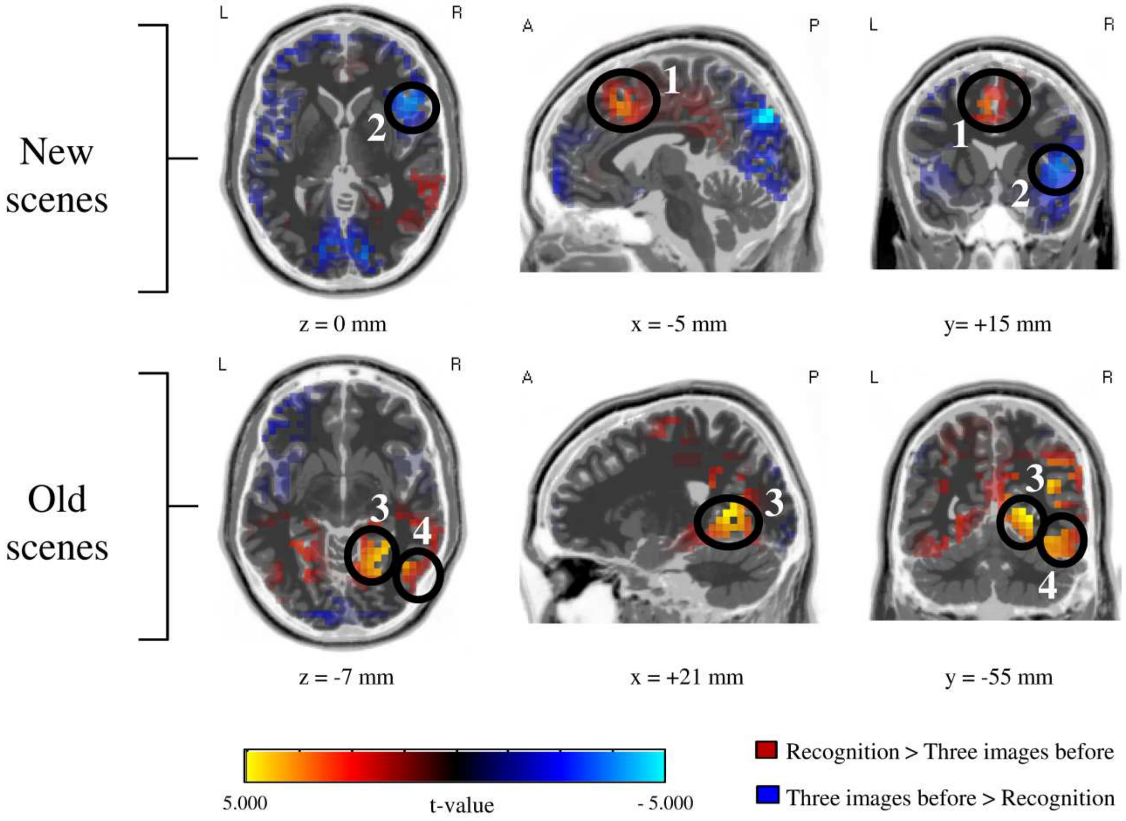


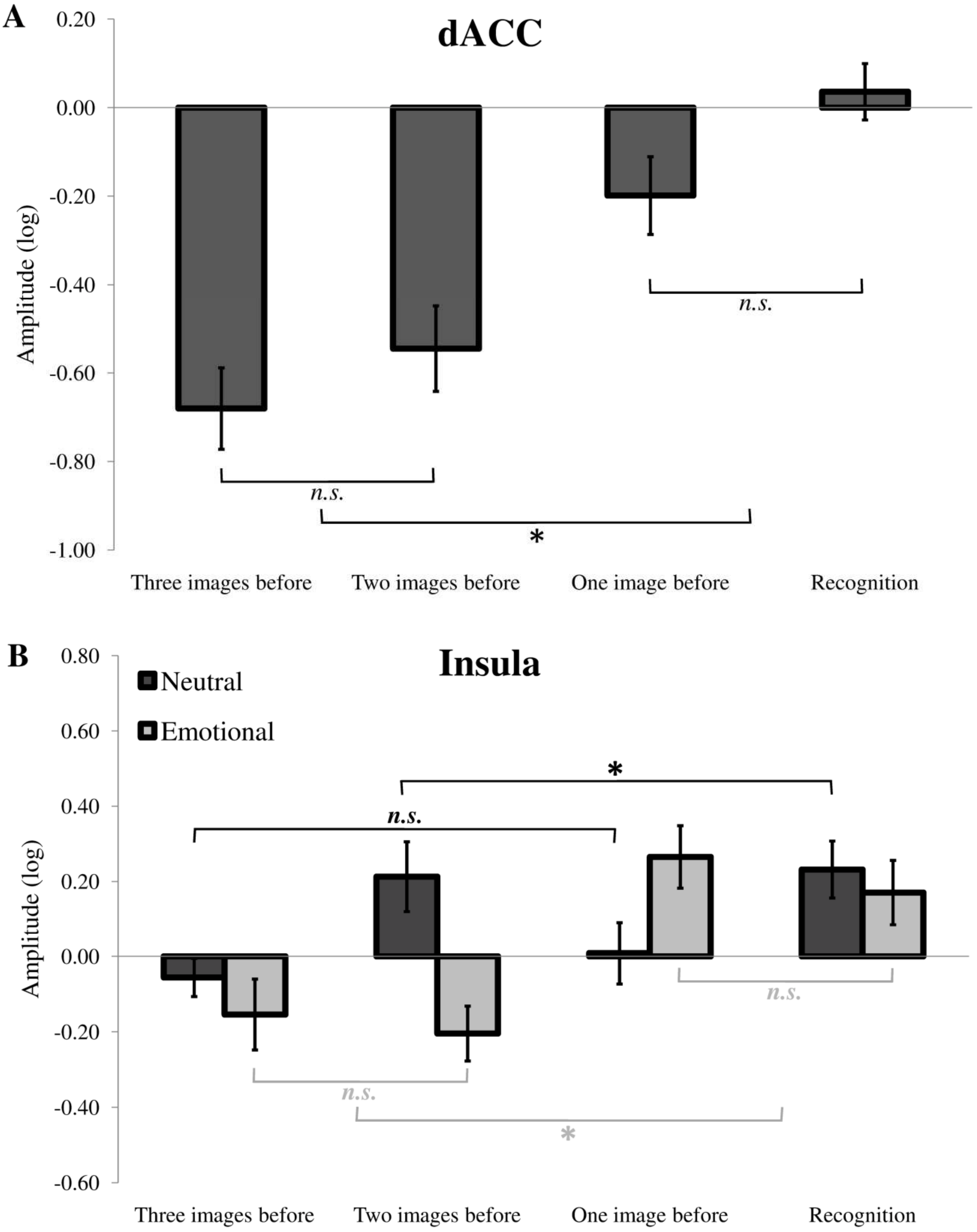
Figure 4. Source localization results. (Upper panel) For new scenes, comparing actual recognition to three images before recognition during the 402-522 ms post-stimulus onset interval revealed a highly significant effect in the dACC (ROI #1: ±10*x*, +15*y*, +35*z*). The reverse contrast (three images before vs. recognition) revealed a highly significant effect in the insula (ROI #2: ±40*x*, +15*y*, 0*z*). (Lower panel) For old scenes, a nonparametric analysis contrasting actual recognition to three images before recognition revealed increased activity in the PHG (ROI #3: ±29*x*, -49*y*, -6*z*), as well as the FG (ROI #4: ±41*x*, -45*y*, -19*z*). L: left; R: right; A: anterior; P: posterior.

3.1.2. PHG and FG (old scenes)

The response profile of the PHG and FG during recognition of old scenes was reliably different compared to the two previous ROIs (i.e., dACC and insula) identified primarily for the processing of new scenes. Unlike a somehow categorical difference between early vs. late stages of accumulation of perceptual evidence, a linear increase in current source density was observed when moving from three images before to actual recognition in the PHG and FG. A 2 (emotion) x 4 (recognition level) ANOVA on the extracted amplitude values for the PHG revealed significant main effects of emotion [$F(1, 17) = 6.33$, $p = .022$, $\eta_p^2 = .271$] and recognition level [$F(3, 51) = 17.06$, $p < .001$, $\eta_p^2 = .501$]. Post-hoc comparisons showed higher activity three images before recognition for neutral compared to emotional scenes [$t(17) = 2.96$, $p = .009$] (Figure 5C). The same analyses performed on the extracted amplitude values for the FG led to comparable results. The 2 (emotion) x 4 (recognition level) ANOVA showed significant main effects of emotion [$F(1, 17) = 8.71$, $p = .009$, $\eta_p^2 = .339$] and recognition level [$F(3, 51) = 11.75$, $p < .001$, $\eta_p^2 = .409$], as well as a significant emotion x recognition level interaction [$F(3, 51) = 3.62$, $p = .019$, $\eta_p^2 = .176$]. As was the case for the PHG, the FG showed higher activity for neutral vs. emotional scenes three images before recognition [$t(17) = 5.62$, $p < .001$] (Figure 5D).

4. Discussion

Using advanced EEG methods, we characterized the electrophysiological correlates of proactive processes active during the recognition of complex visual scenes and conveying either a neutral or an emotional (pleasant or unpleasant) content. We designed a task suited to explore the temporal dynamic of these accumulation of evidence processes, and eventually assessed whether the emotional content of the scenes could reliably modulate its expression or not. Each trial started with the presentation of a colorful neutral or emotional scene to be encoded in short-term memory, followed by the gradual unfolding of the same scene content, a new one, or a scrambled picture. Participants had to decide whether this latter scene progressively unfolded had actually the same identity than the one shown at encoding or not. Moreover, we occasionally asked participants to explicitly rate the emotional valence of the pictures, a manipulation that was meant to promote the overt processing of the emotional content of the scenes throughout the experiment. The acquired ratings confirmed that this manipulation was successful. Likewise, results obtained for the LPP component at encoding further corroborated this conclusion (Figure 2). Interestingly, a number of important new results emerge from this study regarding possible modulatory effects of emotion on the online generation of predictions during overt scene recognition, as discussed here below.


A
dACC
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
Amplitude (log)
n.s.
n.s.
*
Three images before
Two images before
One image before
Recognition
B
Insula
■Neutral
□Emotional
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
Amplitude (log)
*
n.s.
n.s.
n.s.
*
Three images before
Two images before
One image before
Recognition

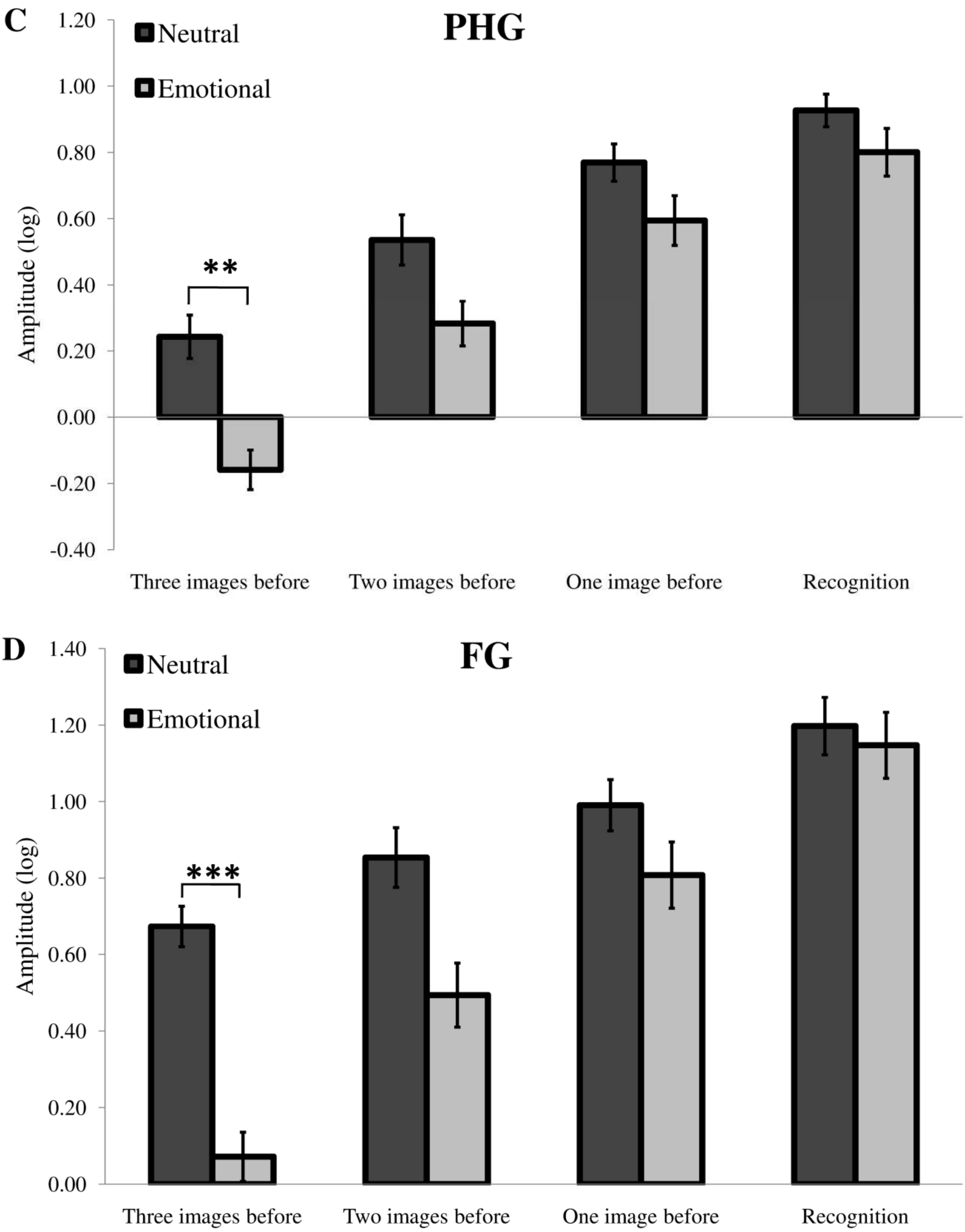


Figure 5. (A) Extracted current density values for the dACC (new scenes) as a function of recognition level. A categorical increase in activity was observed between two images and one image before recognition. (B) Activity in the insula for new scenes, separately for neutral and emotional pictures, as a function of time of recognition. A stable pattern of activity was evidenced in the insula in response to neutral pictures, confirmed by a lack of statistically significant differences among all four recognition levels, except between one image before and recognition. By contrast, insula activity sharply increased between two images and one image before recognition for emotional scenes. (C) and (D) Extracted current density values for the PHG (C) and the FG (D) (old scenes). A monotonic increase of activity in these regions as a function of accumulation of perceptual evidence was evidenced. Moreover, higher activity for neutral (dark grey bars) vs. emotional (light grey bars)

scenes was found three images before recognition. * $p < .05$; ** $p < .01$; *** $p < .001$. Vertical bars correspond to standard errors of the means.

4.1. Prolonged exploration for emotional compared to neutral scenes

Behavioral results confirmed that recognition performance was not random in this task, but based on the accurate gathering of diagnostic perceptual evidence, as provided by the progressive unfolding of the stimulus content and directly constrained by knowledge stored in visual short-term memory. As hypothesized, old scenes were recognized systematically earlier than new scenes (Figure 1B), providing direct evidence for the use of abstract visual representations during unfolding, which were directly shaped by information provided to participants at encoding. Moreover, we replicated our previous results (Schettino, et al., 2011) and found that, regardless of the memory status of the scenes (old vs. new), participants dwelt longer on emotional compared to neutral scenes (Figure 1B). Delayed recognition for emotional relative to neutral scenes could be explained by differential motivational drives between these two categories, including *positivity offset* (Cacioppo & Gardner, 1999; Cacioppo, et al., 1997, 1999; Norris, et al., 2010), as we argued in a recent study (Schettino, et al., 2012). According to this general motivational account, participants are usually inclined to engage in additional exploratory behavior when encountering emotion-laden stimuli, especially stimuli carrying an intrinsic hedonic (or rewarding) value. Interestingly, the use of mildly arousing scenes in our study, as opposed to mutilations or erotica, may have boosted this asymmetry in motivational exploratory strategies between emotional and neutral scenes. We found that this motivational effect found in a sample of healthy (female) individuals was similar for old and new scenes, which suggests that effects of emotion on proactive guesses during scene recognition probably occurred either before or after the input stimulus was compared against existing visual representations stored in (short-term) memory.

4.2. New and old scenes recruit non-overlapping brain networks during proactive visual processes

The use of high-density EEG allowed us to gain insight into the putative brain networks giving rise to proactive processes during scene recognition. Four hundred ms after stimulus onset, we found that the topography of the evoked electric field reliably changed depending on the amount of perceptual evidence accumulated, unambiguously revealing changes in the configuration of the underlying brain networks (Lehmann & Skrandies, 1980; Michel & Murray, 2012; Michel, et al., 1999; Pourtois, et al., 2008). Moreover, these topographical effects were different for old vs. new scenes. Two images before recognition, we found that the dominant occipital positivity/frontal negativity topographical map was decreased for old compared to new scenes, this memory effect later influencing the concurrent occipital positivity/frontal positivity scalp map. These results suggest that memory could rapidly alter the expression of accumulation of evidence processes taking place prior to overt recognition, probably via the activation of non-overlapping brain structures during the processing of old vs. new scenes (Bar, 2004; Ranganath & Rainer, 2003). By contrast, the temporal dynamic of these two dominant topographical maps was not reliably influenced by the emotional content of the scenes, replicating our previous results (Schettino, et al., 2011).

4.3. Modulatory effect of emotion depends on memory

In agreement with the assumption of non-overlapping brain networks between new and old scenes, our complementary source localization analysis (based on sLORETA) confirmed that the processing of old scenes was mostly achieved by a distributed network of ventral brains regions, including the PHG and FG, whereas the processing of new scenes involved mainly neural changes in more dorsal brain regions at the same latency, including the dACC and the insula. Interestingly, we found that activity extracted in these ROIs was influenced by the emotional content of the scenes, though at different stages during the accumulation of evidence depending on the actual memory status of these scenes (old vs. new). These results suggest that emotion may actually exert pervasive effects on these proactive processes during scene recognition. Accordingly, the reconstruction and estimation of the neural sources of the ERP signal using a linear distributed inverse space (containing 6239 nodes) turned out to be more sensitive to capture subtle changes related to the processing of the emotional content of the scenes, compared to the topographical mapping analyses carried out at the scalp level exclusively (using 128 channels in the present case) (Esslen, Pascual-Marqui, Hell, Kochi, & Lehmann, 2004; Pascual-Marqui, Esslen, Kochi, & Lehmann, 2002). This dissociation is not surprising though, given the enhanced spatial sensitivity of distributed source localization methods (including sLORETA) that enables to reveal subtle differences between experimental conditions, which may be otherwise difficult to capture based on local amplitude measurements performed at the sensor level (Brodbeck, et al., 2011; Lantz, Menendez, Andino, & Michel, 2001; Michel & Murray, 2012).

4.4. Enhanced uncertainty during the processing of new scenes

Activity in the dACC for neutral scenes was low and close to baseline until one image before recognition, when a sharp increase took place (see Figure 5A). This specific response profile for the dACC, which is consistent with the involvement of this medial frontal region in higher-order decision making processes (Bush, et al., 2002; Ridderinkhof, et al., 2004; Seo & Lee, 2007), was already found in previous imaging studies looking at accumulation of evidence processes taking place during object or scene recognition (Ploran, et al., 2007; Ploran, et al., 2011; Schettino, et al., 2011). This sharp increase in the dACC close to recognition could reflect either uncertainty or conflict, given the urge to stop the stimulus sequence and take a decision before the sequence comes to an end, while the accumulated sensory evidence may not be completed yet (Anderson, et al., 2009; Philiastides & Sajda, 2007; Sohn, et al., 2007). Interestingly, a similar effect in the dACC was already reported in our previous ERP topographic mapping study, though based on a different task (Schettino, et al., 2011). Hence, this region may be involved in the coding of conflict or uncertainty during perceptual-decision making tasks at a more general level (Ridderinkhof, et al., 2004; Ullsperger, et al., 2004), particularly when recognition targets are embedded in an emotional context (Kanske & Kotz, 2011a, 2011b).

Not only the dACC, but also the insula showed increased activity in response to new scenes, although it additionally showed a modulation according to the emotional content of the scenes. In this region, emotion influenced the response profile during the accumulation of evidence prior to recognition. Activation of the insula has previously been reported for a variety of tasks and stimuli (Craig, 2009), including when exogenous sensory stimuli acquire salience (Seeley, et al., 2007), as well as in response to challenging and uncertain perceptual tasks (Grinband, Hirsch, & Ferrera, 2006; Philiastides & Sajda, 2007). Likewise, previous ERP work based on a similar approach as used here already source-localized the insula during the early processing of response errors during a speeded go/nogo task (Dhar, Wiersema, & Pourtois, 2011). Presumably, similar effects could explain the involvement of the insula in our task, during the accumulation of perceptual evidence prior to recognition. Whereas uncertainty likely decreased as a function of accumulation of evidence, at time of recognition this process is presumably not completed yet. It is interesting to note that emotional stimuli could somehow exacerbate this process in the insula, thereby shifting the neural response earlier in time -- probably because of their intrinsic motivational salience -- compared to neutral stimuli. Presumably, the emotional content of the stimulus might actively interfere with task demands (memory matching task), resulting in an earlier experience of uncertainty or conflict for emotional compared to neutral scenes.

More generally, the parallel involvement of the insula and dACC in our study during the processing of new scenes is not surprising, but consistent with many previous findings showing that these two regions likely operate together and eventually form a functional network activated across a variety of stimuli and tasks (Ploran, et al., 2007; Seeley, et al., 2007; Sterzer & Kleinschmidt, 2010), including the implementation of task sets (Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Nelson, et al., 2010), awareness processes (Craig, 2009; Dhar, et al., 2011; Mayr, 2004), as well as performance monitoring (Ito, Stuphorn, Brown, & Schall, 2003; Ridderinkhof, et al., 2004; Wheeler, et al., 2008).

4.5. Emotion-dependent monotonic accumulation of perceptual evidence in ventral object-selective regions during the processing of old scenes

While the processing of new scenes was selectively associated with specific neural effects in the dACC and insula starting 400 ms post-stimulus onset, a different picture emerged for the processing of old scenes at the same latency, where ventral object-sensitive brain regions were found to be significantly more active. Among them, activity in the PHG increased linearly as a function of accumulation of evidence, with the lowest amplitude values three images before recognition and the highest at time of recognition (see Figure 5C). Given the ubiquitous involvement of this specific brain area in the processing of contextual information needed for successful scene recognition (Bar & Aminoff, 2003; Bar, Aminoff, & Schacter, 2008; Kveraga, et al., 2011), we interpreted our results as reflecting similar mechanisms. The linear trend found in this region during accumulation of evidence prior to recognition (see also Schettino, et al., 2011) is consistent with the assumption that this region is rapidly involved in the analysis of diagnostic contextual information, primarily conveyed by magnocellular input (Bar, 2004, 2007, 2009b). Interestingly, we found that this response profile was not identical for emotional compared to neutral scenes. More specifically, during early stages of accumulation of evidence (i.e., three images before recognition), the amplitude of the PHG was significantly lower for emotional compared to neutral scenes. An early interference effect probably took place for emotional scenes, eventually preventing the initiation of the generative accumulation of evidence process (presumably based on contextual information carried by LSF cues) taking place in this region. However, because at time of recognition this amplitude difference between neutral and emotional scenes was no longer significant, it is likely that a more rapid accumulation of evidence took place in this region for emotional compared to neutral scenes. Hence, emotional factors would initially interfere with the processing of diagnostic contextual information (based on the extraction of LSF cues), but later boost this same process when more evidence is timely gathered and accumulated (Figure 5C), presumably based on the concurrent processing of HSF information. Such an interpretation is compatible with mounting evidence showing the importance of contextual visual information during emotional scene recognition (Barrett & Kensinger, 2010; Barrett, Mesquita, & Gendron, 2011; de Gelder, et al., 2006; Righart & de Gelder, 2006, 2008a, 2008b). Thus, due to their enhanced biological relevance, emotional scenes led to a steeper accumulation of evidence in the PHG relative to neutral scenes.

A similar outcome was found for the FG. Previous imaging studies already reported the involvement of the FG in the gradual accumulation of perceptual evidence (James, et al., 2000; Malach, et al., 1995). Presumably, a similar accumulation of evidence mechanism took place in the FG in our study, although one may argue that the diagnostic visual information used by this region may be different compared to the PHG. More specifically, the FG could be involved in the accumulation of perceptual evidence based on the rapid extraction of invariant object features or viewpoint information, as opposed to contextual information in the PHG (Epstein, Harris, Stanley, & Kanwisher, 1999; Grill-Spector & Malach, 2004; Malach, et al., 1995; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005; Pourtois, et al., 2010a).

4.6. Conclusion

The results of this study shed light on the electrophysiological correlates (and the putative underlying brain sources) of accumulation of perceptual evidence prior to scene recognition. Four hundred ms following stimulus onset, regions in the dACC and insula were selectively active prior to overt recognition of new scenes, likely translating either enhanced uncertainty or conflict monitoring during perceptual decision making for this specific condition. By contrast, at the same latency following stimulus onset, old scenes recruited primarily ventral object-selective regions during accumulation of perceptual evidence, including the PHG and FG. However, an early modulation of their response profile as a function of the emotional content of the scenes was observed. Accordingly, even though neural mechanisms underlying accumulation of evidence during scene recognition are multiple and can be dissociated based on memory (i.e., old vs. new scenes), it is striking to observe that emotion exerts pervasive interference effects on these proactive processes during fairly early stages of accumulation of perceptual evidence (when the retinal input is still minimal or impoverished), in addition to these strong modulatory effects imposed by memory on perception.

CHAPTER 5

Valence-specific modulation in the accumulation of perceptual evidence prior to visual scene recognition[1]

Visual scene recognition is a dynamic process through which incoming sensory information is iteratively compared with predictions regarding the most likely identity of the input stimulus. In this study, we used a novel progressive unfolding task to characterize the accumulation of perceptual evidence prior to scene recognition, and its potential modulation by the emotional valence of these scenes. Our results show that emotional (pleasant and unpleasant) scenes led to slower accumulation of evidence compared to neutral scenes. In addition, when controlling for the potential contribution of non-emotional factors (i.e., familiarity and complexity of the pictures), our results confirm a reliable shift in the accumulation of evidence for pleasant relative to neutral and unpleasant scenes, suggesting a valence-specific effect. These findings indicate that proactive iterations between sensory processing and top-down predictions during scene recognition are reliably influenced by the rapidly extracted (positive) emotional valence of the visual stimuli. We interpret these findings in accordance with the notion of a genuine positivity offset during emotional scene recognition.

---

[1] Schettino, A., Loeys, T., Bossi, M., & Pourtois, G. (2012). Valence-Specific Modulation in the Accumulation of Perceptual Evidence Prior to Visual Scene Recognition. *PloS One, 7*(5), e38064.

1. Introduction

Visual object recognition has classically been conceived as resulting from a set of serial computations performed by dedicated ventral object-selective brain regions located in the infero-temporal cortex, eventually enabling to progressively extract the precise meaning of the retinal input (Palmer, 1999; Riesenhuber & Poggio, 1999). Whereas bottom-up processes are typically emphasized in these hierarchical models, the visual computations performed by these object-selective areas are nonetheless susceptible to top-down modulatory effects, including selective attention (Blair, et al., 2009; Kim & Rehder, 2011; Treisman & Kanwisher, 1998), prior expectations (Rahnev, Lau, & de Lange, 2011; Summerfield & Egner, 2009), contextual information (Bar, 2004; Oliva & Torralba, 2007), or decision-making (Heekeren, et al., 2008; Ridderinkhof, et al., 2004). Therefore, visual object recognition processes are not limited to the analysis of sensory information, but they are further shaped by higher order (i.e., not strictly perceptual) processes.

Interestingly, an alternative view has been put forward to account for these complex interaction effects between bottom-up sensory processing and top-down modulatory influences during recognition. Namely, predictive coding models (Enns & Lleras, 2008; Friston, 2005; Friston & Kiebel, 2009; Grossberg, 2009; Lochmann & Deneve, 2011; Mumford, 1992; Rao & Ballard, 1999; Spratling, 2008; Ullman, 1995) advocate that visual object recognition processes taking place within the infero-temporal cortex result from the dynamic interplay between (top-down) predictions and (bottom-up) errors (Alink, et al., 2010). Predictions reflect prior knowledge related to probable events in the sensory environment, and they are employed to reduce the computational burden of visual perception by guiding attention towards salient aspects of the environment, as well as facilitating the interpretation of ambiguous visual input (Summerfield & Egner, 2009). Whenever a discrepancy is detected between these top-down predictions and bottom-up sensory processing, an error signal (prediction error) is generated and propagated back to higher-level brain regions, with the aim to update or refine the content of the predictions, and in turn accommodate online sensory processing with the current specificities carried by the visual input (Di Lollo, et al., 2000; Friston, 2005; Rao & Ballard, 1999; Ullman, 1995). In this framework, the expected and actual sensory input are dynamically compared at each stage of processing by means of recursive loops, until the system is able to generate the most likely interpretation of the target object (Friston, 2009; Summerfield & Egner, 2009).

Of note, asymmetries in speed of processing and visual pathways between low spatial frequency (LSF) and high spatial frequency (HSF) information could potentially provide a mechanistic account to explain predictive coding effects during early stages of recognition of single objects and complex visual scenes (Bar, et al., 2006). Several studies have already established the differential contribution of LSF vs. HSF input in face recognition (Goffaux, et al., 2011; Pourtois, Dan, et al., 2005; Winston, et al., 2003), as well as in the processing of complex visual scenes (Carretié, Hinojosa, Lopez-Martin, & Tapia, 2007; Oliva & Schyns, 1997; Schyns & Oliva, 1994; Torralba & Oliva, 2003). More precisely, LSF information seems more useful in identifying the gist of the scene in conditions of fast stimulus presentation (i.e., 30 ms), whereas for longer durations (150 ms) observers rely more on HSF information (Schyns & Oliva, 1994). Interestingly, because LSF information travels rapidly from early sensory visual areas to prefrontal and anterior temporal regions (via magnocellular projections), this early coarse analysis of the visual input might actually serve to generate predictions about its content (Bar, 2003, 2004, 2007, 2009b; Kveraga, Ghuman, & Bar, 2007).

Nonetheless, all the models reviewed so far have dealt with the processing of neutral visual stimuli, exclusively. Hence, the question remains whether, when encountering emotion-laden objects or scenes, perceptual processes underlying these proactive guesses are comparable to those involved during the processing of neutral stimuli or not. In fact, given the accumulating empirical evidence showing that emotion does not simply add a specific flavor to perception but can have profound influences in stimulus processing, both at the behavioral and neural levels (Barrett & Bar, 2009; Bocanegra & Zeelenberg, 2009; Öhman, Flykt, et al., 2001; Phelps, et al., 2006; Vuilleumier, 2005; Vuilleumier & Pourtois, 2007), one can argue that predictive coding during visual scene recognition may reliably be influenced by emotional factors. However, this question has received little empirical support, and it is therefore still unclear whether mechanisms of predictive coding may change during visual scene recognition or not, depending on the extracted emotional content or value of the incoming stimulus. To address this issue, we recently developed and validated a new experimental paradigm. It enables us to study effects of emotion (i.e., valence and/or arousal) on the speed of proactive guesses during scene recognition, both at the behavioral and electrophysiological (event-related brain potentials, ERPs) levels (Schettino, et al., 2011). For each individual trial, participants were presented with series of filtered images that were gradually unfolding the content of a complex visual scene while they had to perform an (orthogonal) animacy judgment task. Each trial began with the presentation of a blurred image, whose content was progressively revealed by increasing, in up to six sequential, parametric and predictive steps, the amount of diagnostic LSF and HSF information. Therefore, this procedure mimicked a “coarse-to-fine” accumulation of perceptual evidence (Bar, 2003; Bullier, 2001; Hegdé, 2008; Navon, 1977; Schyns & Oliva, 1994). Importantly, the visual scenes used in this study (extracted from a standard database) could be neutral, pleasant or unpleasant, based on independent arousal and valence ratings obtained for these visual stimuli. Behavioral results confirmed that this task was suited to study predictive coding effects during scene recognition because participants did not respond randomly, but they accumulated sufficient perceptual evidence before deciding, with high accuracy, whether the content of the scene was living or not (Gold & Shadlen, 2007; Heekeren, et al., 2008; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). Importantly, this effect was not identical for the three emotion categories. Participants probably accumulated perceptual evidence less rapidly (reflected in prolonged recognition) for emotional compared to neutral scenes, this effect being most obvious for pictures having a pleasant content. These results could be interpreted as reflecting a *negativity bias* during scene recognition (Cacioppo & Ito, 1999; Ito, et al., 1998). Negativity bias refers to the fact that aversive stimuli usually elicit stronger responses compared to appetitive ones, leading in turn to a faster recognition for unpleasant relative to pleasant pictures. However, the prolonged exploration for pleasant scenes was also consistent with the concurrent activation of *positivity offset* during scene recognition (Ito, et al., 1998; Norris, et al., 2010). In this view, when input to the affect system is minimal, positivity may outweigh negativity. Due to their intrinsic hedonic value, pleasant scenes could therefore be associated with prolonged exploration. Hence, behavioral results of this study (Schettino, et al., 2011) were equivocal with regard to the underlying emotional or motivational drive accounting for these findings.

Whereas these results shed light on mechanisms underlying the generation of proactive guesses during scene recognition and how emotion may influence these complex processes, a main question also arose regarding the specificity of these effects. For instance, considering the fact that we used an orthogonal task (i.e., animacy judgment), it is conceivable that the emotional content of the scene had little or no direct impact on the expression of processes involved in accumulation of perceptual evidence (Pessoa, 2008; Pessoa, Kastner, et al., 2002; Spruyt, et al., 2009; Spruyt, et al., 2007). Moreover, it was unclear from these results alone whether emotion as such, or other non-controlled factors, may actually have produced the change in the rate of accumulation of evidence between emotional and neutral scenes. Presumably, the selected visual scenes did not differ only regarding the actual emotional content, but also their intrinsic picture complexity and/or familiarity, even though we took special care to minimize obvious perceptual and structural differences across the three emotion categories (Schettino, et al., 2011). However, if the neutral vs. emotional scenes selected in our study were not properly balanced along these two specific non-emotional dimensions (i.e., picture complexity and familiarity), we cannot exclude the possibility that the reported behavioral effects may be imputed to these factors, rather than the differential processing of the emotional content during scene recognition. Presumably, more complex or less familiar visual scenes might be associated with delayed accumulation of evidence in our task. Hence the question remains whether the prolonged accumulation of evidence found in our study for emotional relative to neutral scenes may (at least partly) be explained by changes in picture complexity and/or familiarity across the three emotion categories, rather than the emotional content *per se* (Schettino, et al., 2011). Therefore, the goal of the present study was to assess whether trial-by-trial variations along these two dimensions may overshadow or confound genuine effects of emotion during the accumulation of perceptual evidence prior to scene recognition or not.

We referred to *picture complexity* as the extent to which a target object in the foreground can be easily segregated from its background (Bradley, et al., 2007). Figure-ground segregation is a fundamental process in visual scene recognition (Craft, et al., 2007; Driver & Baylis, 1996; Roelfsema, Lamme, Spekreijse, & Bosch, 2002). Following initial sensory registration of contours, the visual system automatically groups regions adjacent to each contour with either the main figure in the foreground or the background, thereby prioritizing, in the subsequent analysis, all regions grouped with the figure (Albright & Stoner, 2002). However, despite the ubiquitous importance of this gestalt mechanism in vision, previous research has found only weak correlations between picture complexity (e.g., figure-ground segregation) and visual emotion processing (Bradley, et al., 2007; Carretié, et al., 2007; Carretié, et al., 2004; Wiens, Sand, & Olofsson, 2011). As a matter of fact, motivationally relevant stimuli, particularly emotional scenes, usually influence late perceptual or even post-perceptual stages of processing, presumably after earlier mechanisms contributing to figure-ground segregation come into play (Schupp, et al., 2006; Schupp, et al., 2007). However, all these studies used (relatively) brief and static presentations of fully detailed neutral vs. emotional stimuli, therefore strongly limiting the online generation of predictions about the actual identity of the incoming visual input. Therefore, these earlier studies did not allow to titrate the potential influence of picture complexity on the accumulation of evidence leading to (emotion) scene recognition. We predicted that, in our experiment, picture complexity might actually influence accumulation of evidence, indicated by slower accumulation rates for pictures characterized by a more complex, as opposed to less complex content (i.e., a less vs. a more obvious figure-ground segregation).

Familiarity, on the other hand, was defined as the frequency of encounter associated with a given stimulus (picture content), following standard practice (Libkuman, et al., 2007). Familiarity is a relevant construct to take into account in the present case, given its potential overlap with emotion processes. In fact, novelty has been found to elicit threat-like cardiovascular responses in social situations involving the violation of stereotypical expectations (Mendes, et al., 2007). Moreover, a comparable startle reflex was observed for novel and emotional pictures (Bradley, et al., 1993). These negative evaluations of novel/unfamiliar stimuli could be due to the difficulty with which individuals extract diagnostic information necessary for a quick and efficient recognition (Winkielman, Schwarz, Fazendeiro, & Reber, 2003). Specifically, high *fluency* (i.e., enhanced processing facilitation) is accompanied by an increase of positive affective reactions, as evidenced by more positive judgments of neutral pictures presented for a prolonged period of time (Winkielman & Cacioppo, 2001). This effect could potentially explain well-known psychological phenomena such as “mere-exposure”, that is people’s general tendency to prefer stimuli they are repeatedly exposed to (Monahan, Murphy, & Zajonc, 2000; Zajonc, 1968, 2001). Accordingly, it is important to establish whether familiarity, rather than emotion (i.e., valence and/or arousal), may account for changes in accumulation of evidence prior to recognition. Given the evidence reviewed here above, we predicted more familiar scenes to be recognized earlier than less familiar scenes in our experiment.

To address these questions, we designed a new experiment based on the previously validated progressive unfolding task (Schettino, et al., 2011) and collected data in a sample of healthy adult participants. Noteworthy, in addition to the main memory matching task (old-new judgment; see below), we instructed participants to directly attend to the emotional content of the stimuli by occasionally asking them to rate the emotional valence of the scenes. These instructions are at variance with the animacy judgment task used in our previous study (Schettino, et al., 2011). We reasoned that this manipulation should augment the relevance of emotional features during the task (Spruyt, et al., 2009), and hence the likelihood to observe reliable differences between the three emotion categories (neutral, pleasant and unpleasant) during accumulation of evidence prior to scene recognition. Furthermore, each and every scene used during the main experiment was subsequently rated in terms of familiarity and picture complexity by two independent samples of participants, using standard 9-point Likert scales. Afterwards, we used these independent ratings in a single-trial analysis to assess whether systematic changes in accumulation of evidence prior to recognition (as measured in the main progressive unfolding experiment) might be confounded by variations along picture complexity and/or familiarity. More specifically, we assessed whether the prolonged exploration for emotional compared to neutral scenes (see results) might be due to systematic changes in picture complexity and/or familiarity across these categories.

2. Methods

2.1. Ethics statement

The study was approved by the ethics committee of the Faculty of Psychological and Educational Sciences, Ghent University. All participants were required to give written informed consent prior to their participation.

2.2. Participants

Eighteen psychology students (all women, mean age 21 years, range 18-26) participated in the main experiment (progressive unfolding task) in exchange of 30€. In addition, 20 volunteers (15 women, mean age 23 years, range 18-34) participated in the picture complexity rating experiment, whereas another sample of 21 participants (17 women, mean age 23 years, range 19-37) completed the familiarity rating experiment. Each participant of the two rating experiments received 8€. All individuals were native Dutch speaking, right-handed, had normal or corrected-to-normal vision, with no history of neurological or psychiatric disorders.

2.3. Stimuli

The visual stimuli were selected from the International Affective Picture System (IAPS) (Lang, et al., 2008), a standard database containing neutral and emotionally-evocative pictures depicting objects and scenes across various ecological situations. This database provides normative ratings for the basic dimensions of emotion -- including arousal and valence -- using the Self-Assessment Manikin (SAM) (Bradley & Lang, 1994). The stimulus list consisted of 360 pictures, equally divided into three emotion categories according to their standardized valence scores: *neutral*, *unpleasant* and *pleasant* (Table 1). Notably, these pictures were selected on the basis of mean valence and arousal ratings reported by female responders (Lang, et al., 2008), because only women eventually participated in the main experiment (see above). Since the main purpose was to assess valence-specific effects during scene recognition, the selected pleasant and unpleasant scenes were properly balanced with regard to levels of arousal (see Table 1). Similarly to our previous study (Schettino, et al., 2011), highly pleasant (i.e., erotic situations) or highly unpleasant (i.e., mutilations) scenes were not included in the stimulus set, given the specific emotion responses often associated with these two categories (Schupp, et al., 2006, 2007). Moreover, we included 16 additional neutral pictures that were only used during the practice session (therefore not considered in the subsequent statistical analyses). Finally, 36 supplementary neutral scenes were scrambled (i.e., each picture was divided into grids of 255 x 255 pixels, which were randomly shuffled 10 times), thereby disrupting the content of the scene. Thus, a total of 412 IAPS pictures (including practice and scrambled trials) were shown to participants of the main experiment, while participants of the two rating experiments were presented with the 360 main pictures (excluding practice and scrambled scenes)[2].

---

[2] Number codes of pictures selected from the database (Lang, et al., 2008) are provided, for each category separately. Practice: 2107, 2600, 2980, 5533, 5731, 6837, 7017, 7030, 7036, 7055, 7057, 7140, 7224, 7365, 8121, 8312. Neutral: 1350, 1616, 1675, 1903, 1935, 1947, 2025, 2026, 2034, 2191, 2272, 2273, 2279, 2308, 2357, 2377, 2382, 2383, 2390, 2396, 2445, 2446, 2489, 2495, 2514, 2575, 2579, 2593, 2595, 2597, 2606, 2702, 2720, 2749, 2850, 2880, 4090, 4150, 4220, 4250, 4255, 4274, 4275, 4320, 4325, 4605, 4750, 5040, 5395, 5500, 5531, 5532, 5534, 5535, 5900, 6570.2, 7001, 7002, 7003, 7009, 7011, 7014, 7016, 7018, 7019, 7021, 7032, 7033, 7037, 7038, 7042, 7043, 7044, 7045, 7058, 7061, 7062, 7081, 7096, 7130, 7160, 7161, 7170, 7180, 7184, 7186, 7188, 7190, 7207, 7236, 7242, 7247, 7248, 7249, 7255, 7287, 7300, 7354, 7484, 7487, 7493, 7500, 7503, 7506, 7512, 7513, 7546, 7547, 7550, 7590, 7595, 7710, 7820, 7830, 8241, 8311, 8325, 9210, 9260, 9700. Unpleasant: 1230, 1240, 1270, 1275, 1280, 1390, 1505, 1617, 1945, 2115, 2130, 2141, 2205, 2276, 2278, 2400, 2455, 2456, 2525, 2681, 2682, 2694, 2695, 2700, 2715, 2716, 2718, 2745.2, 2750, 2752, 2770, 2795, 2799, 2810, 2900.1, 3061, 3160, 3181, 3190, 3210, 3216, 3280, 3300, 3301, 4621, 4635, 4770, 5970, 5973, 6000, 6240, 6241, 6311, 6314, 6561, 6562, 6610, 6800, 6832, 7013, 7023, 7079, 7092, 7136, 7137, 7520, 7521, 8231, 9002, 9005, 9008, 9031, 9041, 9045, 9046, 9080, 9090, 9102, 9145, 9171, 9180, 9182, 9186, 9265, 9270, 9290, 9291, 9295, 9320, 9330, 9331, 9341, 9342, 9390, 9395, 9402, 9404, 9411, 9415, 9417, 9419, 9421, 9435, 9440, 9445, 9469, 9471, 9561, 9584, 9592, 9596, 9635.2, 9830, 9831, 9832, 9912, 9913, 9922, 9926, 9927. Pleasant: 1340, 1463, 1540, 1590, 1595, 1640, 1659, 1660, 1720, 1721, 1811, 1999, 2055.2, 2056, 2092, 2151, 2156, 2158, 2224, 2274, 2300, 2331, 2344, 2346, 2352, 2398, 2605, 2616, 2655, 3005.2, 4500, 4530, 4534, 4536, 4559, 4571, 4600, 4601, 4603, 4606, 4610, 4612, 4614, 4616, 4617, 4619, 4623, 4624, 4641, 5199, 5215, 5260, 5301, 5480, 5600, 5622, 5628, 5660, 5700, 5814, 5829, 5831, 5849, 5990, 5994, 6250.2, 7200, 7230, 7250, 7260, 7279, 7281, 7282, 7286, 7289, 7291, 7350, 7352, 7390, 7400, 7410, 7430, 7440, 7460, 7461, 7470, 7477, 7481, 7482, 7488, 7489, 7492, 7496, 7501, 7505, 7508, 7515, 7570, 8032, 8050, 8118, 8120, 8162, 8208, 8220, 8280, 8340, 8350, 8371, 8420, 8460, 8461, 8465, 8467, 8497, 8503, 8510, 8531, 8540, 8620; Scrambled: 1112, 1303, 1310, 1645, 1726, 1908, 2002, 2018, 2032, 2038, 2101, 2102, 2104, 2122, 2190, 2220, 2221, 2393, 2440, 2441, 2458, 2480, 2484, 2493, 2506, 2512, 2516, 2518, 2570, 2580, 2635, 2704, 2780, 2830, 2840, 9070.

Table 1. Mean values and standard deviations (in parenthesis) of normative valence and arousal scores for the selected IAPS pictures.

Emotion category	Valence	Arousal
Neutral	5.14 *(1.38)*	3.68 *(2.05)*
Unpleasant	3.17 *(1.61)*	4.94 *(2.15)*
Pleasant	6.95 *(1.70)*	4.97 *(2.30)*

Note. Scores range from 1 to 9. Independent samples t-test confirmed a highly significant difference in valence between neutral and unpleasant [$t(119) = 29.34$, $p < .001$], neutral and pleasant [$t(119) = -26.82$, $p < .001$] and unpleasant and pleasant [$t(119) = -52.58$, $p < .001$] scenes. Significant differences were also observed in levels of arousal, specifically between neutral and unpleasant [$t(119) = -29.34$, $p < .001$] and neutral and pleasant [$t(119) = -30.98$, $p < .001$] pictures. However, no significant arousal difference was evidenced between unpleasant and pleasant scenes [$t(119) = -0.77$, $p = .441$], confirming a balanced level of activation between these two emotion conditions.

Each neutral, unpleasant and pleasant scene was arbitrarily paired with another one from the same emotion category based on low-level visual similarities, assessed by systematic visual inspection. More specifically, for each emotion category separately, pictures with a clear distinction between a central figure and a homogeneous background were paired together (e.g., a coffee mug on a table vs. a pocket watch on a dark background), and the same strategy was applied for more complex scenes (e.g., a traffic jam vs. a woman in the crowd). These pairs were used during the main task to minimize the use of purely perceptual, pixel-to-pixel matching strategies (see here below). All the pairs created with this procedure are reported in Table 2.

The selected IAPS scenes were resized to 922 x 691 pixels (90% of the original size) and pre-processed similarly to our previous study (Schettino, et al., 2011): after grayscale conversion, six bandpass spatial frequency filters were applied on every picture (using ImageJ v1.44 software; http://rsb.info.nih.gov/ij/) (Delplanque, et al., 2007). As a result, six distinct levels of filtering were obtained for every IAPS scene, each containing a different amount of low and high spatial frequency information (Schettino, et al., 2011). All these modified pictures were finally resized to 768 x 576 pixels (75% of the original IAPS pictures).

Table 2. Stimulus pairs created for the progressive unfolding task.

Pair number	Stimulus pairs					
	Neutral		Unpleasant		Pleasant	
	First element	Second element	First element	Second element	First element	Second element
1	2191	7513	2455	9180	1640	7286
2	2272	7500	2525	9635.2	1660	4641
3	2308	4250	3300	2752	2158	2156
4	2357	8311	5970	2694	2274	8208
5	2382	7242	5973	9912	2605	7291
6	2390	5535	6000	2115	2616	2300
7	2514	7061	6241	6832	4530	4500
8	2575	2273	6610	6800	4600	2398
9	2579	2595	7013	9926	4616	4610
10	2606	7037	7079	9041	4619	7260
11	2880	7493	7136	9186	4624	7410
12	4090	7003	7137	7092	5260	7440
13	5040	7161	8231	9440	5622	8620
14	5900	6570.2	9080	2715	5831	2056
15	7009	7190	9102	6314	5849	5628
16	7011	4320	9171	2718	5990	7496
17	7014	2377	9182	2456	5994	8120
18	7021	7248	9265	9031	6250.2	8032
19	7038	5532	9290	9320	7200	8510
20	7042	2034	9291	9342	7279	7489
21	7044	7130	9330	9832	7430	7352
22	7045	2396	9395	3181	7460	5480
23	7062	7186	9415	9471	7477	8465
24	7207	7032	9417	6561	7482	8540
25	7287	2026	9421	2900.1	7501	7505
26	7484	7096	9435	7520	7508	5199
27	7503	1350	9584	9469	7570	5814
28	7590	2850	9592	9270	8460	8497
29	7830	7546	9596	2205	8461	2352
30	9260	4275	9831	9402	8503	7470
31	1616	2445	1270	1275	1340	8420
32	1675	2593	1230	2799	1463	8280
33	1903	7255	1240	1617	1540	1595
34	1947	5531	1280	9830	1590	1720
35	2025	7506	1390	2745	1721	8340

Pair number	Stimulus pairs					
	Neutral		Unpleasant		Pleasant	
	First element	Second element	First element	Second element	First element	Second element
36	2446	2383	1505	9002	2224	4606
37	2489	1935	1945	9419	2331	8350
38	2495	2702	2130	9045	2344	1811
39	2720	7033	2141	9090	3005.2	4571
40	2749	4325	2276	2681	4536	2346
41	4150	2597	2682	2795	4559	2055.2
42	4274	7160	2695	9404	4601	7282
43	4750	4255	2716	2700	4603	8162
44	5534	7547	2810	9913	4612	2151
45	7018	9210	3061	6311	4614	7488
46	7019	7300	3160	9005	4617	2092
47	7043	7016	3190	7521	4623	7481
48	7081	7001	3210	6240	5301	8531
49	7170	7002	3216	4770	5600	7350
50	7180	4605	3280	6562	5660	5215
51	7184	7236	4635	9008	7230	1999
52	7188	7820	9445	9927	7250	7461
53	7247	7249	2278	9295	7281	2655
54	7354	7058	2400	9145	7390	5700
55	7487	8325	2770	9341	7492	5829
56	7512	2279	9390	7023	7515	8467
57	7550	4220	9922	9561	8050	4534
58	7595	5395	9046	4621	8118	1659
59	7710	5500	2750	9411	8220	7289
60	8241	9700	3301	9331	8371	7400

Note. These numbers refer to picture codes, as available in the original database (Lang, et al., 2008).

2.4. *Procedure*

Participants were individually tested in a small, dimly lit room, and seated at a viewing distance of 75 cm in front of a 19" CRT computer screen (refresh rate: 100 Hz). After filling out the informed consent, they were presented with task instructions, followed by a practice block containing 16 neutral pictures. Then, they moved on to the experimental session, divided into twelve blocks, each containing 33 trials. Each trial had the following structure. A colorful, fully detailed picture (922 x 691 pixels, subtending 18.5° x 13.9° of visual angle) was first presented on the screen for 1500 ms, followed by a grayscale mask displayed for 2000 ms. Then, the actual unfolding sequence (Schettino, et al., 2011) began. A fixation cross appeared in the center of the screen for 250 ms. The first grayscale, blurred image level of a given picture (768 x 576 pixels, 15.4° x 11.6°) was then presented for 500 ms, followed by a 250ms blank screen. Next, the second image level of the same picture (identical pixel size, but containing slightly more HSF and LSF information) was displayed for 500 ms, plus the 250ms blank screen, and the same procedure was repeated until the presentation of the sixth, non-filtered image level. The inter-trial interval was constant and set at 1000 ms (Figure 1A). This experimental manipulation was used to promote a gradual and predictive accumulation of perceptual evidence by progressively adding, in a stepwise fashion, high and low spatial frequency information to the first undistinguishable picture (Schettino, et al., 2011). Importantly, the grayscale and resize conversions relative to the original colorful scene (presented at the beginning of each trial) were applied to discourage participants to use a pixel-to-pixel matching strategy to perform the task. Two separate and consecutive responses were required. First, participants were asked to press a button on a response box (Cedrus RB-730; http://www.cedrus.com/responsepads/rb730.htm) with their right index finger as soon as they felt they gathered enough perceptual evidence to decide, with sufficient confidence, whether the content of the unfolded scene was either the same as the one displayed at the beginning of the trial, a new one, or a new scrambled picture (Response1). These scrambled pictures, for which a separate response was required (see below), were used as “catch” trials to ensure that participants reliably attended to the content of the scenes. Pressing the button immediately interrupted the presentation of the stimulus sequence. After 500 ms, participants were required to perform a three-alternative forced choice delayed matching task, in order to validate their first response (Response1). Specifically, they had to press, on a standard AZERTY keyboard, the “O” key if the unfolded scene was the same as the colorful one previously presented (“old” condition), the “N” key if these two scenes were different (“new” condition), or the “S” key if the unfolded scene was displaying a meaningless content (“scrambled” condition). All these responses, for which no time constraint was established, were coded as Response2. The main purpose of this dual response procedure was to dissociate early visual detection (Response1) from the overt discrimination of the scene requiring a specific stimulus-response mapping (Response2) (Schettino, et al., 2011). Participants were asked to focus on accuracy, but at the same time they were encouraged not to wait until the end of the unfolding sequence to decide about the content of the visual scene

(Response1). Responses1 occurring after the presentation of the last/sixth image level were therefore classified as late responses and analyzed separately.

To promote the use of abstract visual representations during overt scene recognition, another experimental manipulation was applied besides the aforementioned inclusion of scrambled pictures as “catch” trials. Specifically, half of the ”old” scenes (i.e., unfolded pictures that were identical to the previously encountered colorful scenes) were unpredictably flipped along the horizontal axis between encoding (colorful picture) and retrieval (unfolding). Participants were informed that an “old” response was expected for these “flipped” pictures, since the main task required them to focus on the *content* of each scene to perform the matching task. In the subsequent behavioral analyses, “old flipped” and “old unflipped” trials were combined into a single “old” condition, to be compared to “new” trials. In sum, for each emotion category (neutral, pleasant, unpleasant), two trial types were contrasted: “old” (N = 180), in which the identity of the initial colorful picture was identical to the scene progressively unfolded, and “new” (N = 180), meaning that the identities of the colorful and unfolded scene were different (although matched as far as possible in terms of low level visual properties using specific stimulus pairs; see the *Stimuli* section). Hence, for “new” scenes, we used the pairs previously created (see Table 2), with one picture of the pair used as colorful image (encoding) and the other used during unfolding (counterbalanced across participants). We created several stimulus lists such that, across participants, each picture appeared equally often in the “new” and “old” conditions. Importantly, for “new” scenes, no change in terms of emotional content ever occurred between the colorful picture and the scene gradually revealed during unfolding. Accordingly, a neutral colorful picture was always followed by the unfolding of a neutral scene, and the same occurred for emotion-laden stimuli (pleasant-pleasant; unpleasant-unpleasant; see also Table 2). The order of “old”, “new” and “scrambled” trials was randomized.

Finally, in order to verify whether the emotional content of the selected IAPS pictures was actually perceived as such and in line with the normative ratings (Lang, et al., 2008), as well as to keep the emotional content task-relevant throughout the experiment, participants were occasionally asked, after the registration of Response2, to also rate the emotional valence of the colorful scene presented at the beginning of each trial by means of a standard 9-point SAM (Bradley & Lang, 1994), with anchor 1 corresponding to “very unpleasant” and anchor 9 to “very pleasant”. This additional emotion classification task concerned 10% of the total number of trials. Such manipulation was also employed to increase the likelihood to detect reliable differences between emotional and neutral scenes during accumulation of evidence prior to scene recognition because, with these specific task demands, participants had to attend to the emotional content of the scene (Spruyt, et al., 2009; Spruyt, et al., 2007).

Stimulus presentation and behavioral response recordings were controlled using E-Prime 2.0. (http://www.pstnet.com/products/e-prime/).

2.5. *Rating experiments*

Participants were tested in pairs in a dimly lit room, seated at a viewing distance of 75 cm in front of individual 19" CRT screens. In each pair, one member was assigned to rate familiarity while the other was asked to focus on picture complexity of the pre-selected IAPS scenes. After completing the informed consent, they were presented with task instructions, including examples. Then, they moved on to the experimental session, divided into six blocks of 60 trials, separated by short breaks. After an initial fixation cross displayed for 500 ms, neutral, pleasant and unpleasant colorful pictures (hence corresponding to the picture presented at the beginning of each trial of the main progressive unfolding experiment) were presented on the screen in randomized order for 2000 ms. Participants were asked to ignore the hedonic valence of the scenes and provide either familiarity or picture complexity ratings (depending on the condition they were assigned to) using 9-point Likert scales. In case of familiarity judgments, the question was: ”How often have you encountered a scene like the one depicted in the picture?”. Scores ranged from 1 (never) to 9 (very often). Raters judging picture complexity, on the other hand, were presented with the question: “Do you consider this picture as having a homogeneous background and an obvious central figure or do you perceive it as more ‘noisy?’”, with “clear figure-ground” anchoring the lower end of the scale and “complex scene” anchoring the upper end. The visual stimuli were never displayed on the screen during the rating phase.

E-Prime 2.0 was used for stimulus presentation and response recordings.

2.6. *Analysis of behavioral data*

One-way ANOVAs and post-hoc t-tests were used to verify that the emotional content of the scenes was perceived by our participants in agreement with the normative ratings, as well as to explore differences between neutral, unpleasant and pleasant pictures in terms of familiarity and picture complexity.

Accuracy on the progressive unfolding task was expressed as percentage of correct responses. Moments of recognition (Responses1) across the six image levels were not independent of each other: in fact, perceptual evidence was gradually accumulating based on visual input provided by previous image levels. Therefore, cumulative percentages were calculated. This procedure resulted in six psychometric curves showing the evolution of recognition performance across the six image levels, separately for each memory (old, new) and emotion (neutral, unpleasant, pleasant) condition. To characterize effects of emotion and memory on recognition performance, we used a proportional odds model with memory and emotion as predictors (Agresti, 2007). This complex model provides a regression analysis for ordinal dependent variables (recognition from image level 1,..., recognition from image level 6). This data analysis, performed at the single-trial level, allows to model the cumulative probability up to and including recognition from each image level k ($k = 1, ..., 5$). The derived odds ratio expresses how much the odds of recognition from image level k or earlier is increased (if larger than 1) or decreased (if smaller than 1) across new, old, neutral and emotional (unpleasant and pleasant) contents, and thus provides a single number capturing the shift in psychometric curve. To account for dependencies of trials within the same subject, a multi-level version of the proportional odds model was used, similarly to our previous study (Schettino, et al., 2011).

Next, we included the mean scores (averaged across raters) of familiarity and picture complexity obtained for each individual picture as additional predictors in the proportional odds model. We verified whether any effect of emotion and/or memory on recognition performance obtained during the main progressive unfolding experiment could be explained by a concurrent effect of familiarity and/or picture complexity.

The level of significance for all these analyses was set at $p < 0.05$. To control for Type I error, a conservative Bonferroni correction was applied to each of the six pairwise comparisons of interest (i.e., emotion, 3 levels; memory, 2 levels) evaluated in each statistical model for the accuracy.

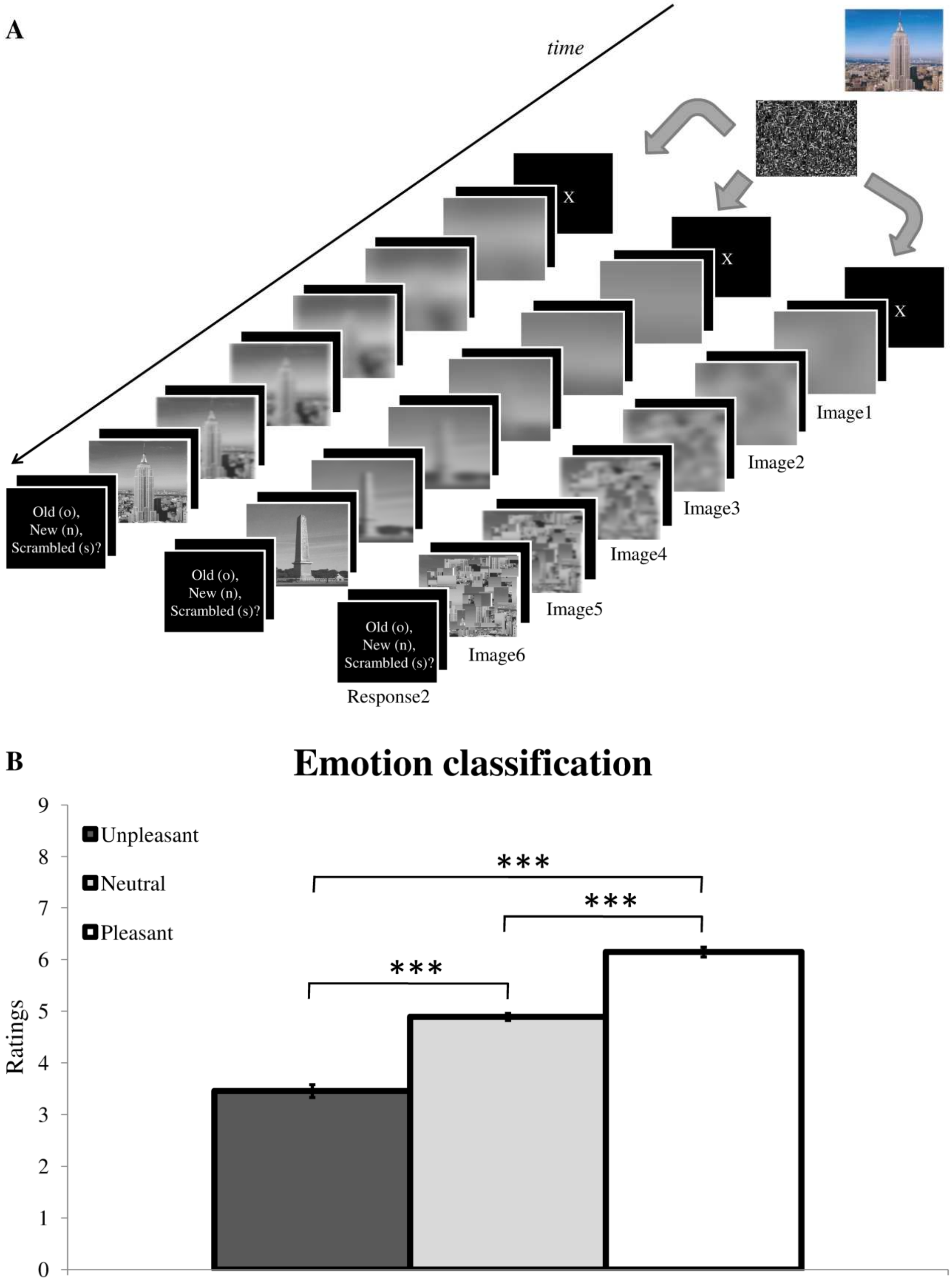


Figure 1. Trial presentation and results of the emotional classification task. (A) Main trial types during the progressive unfolding experiment. A colorful neutral, unpleasant or pleasant IAPS scene (not shown here for copyright reasons) was first presented for 1500 ms, in random order. Following a 2000 ms uniform mask, the same scene (45%), a new one (45%), or a scrambled picture (10%) was progressively revealed in grayscale, using six successive steps varying in a monotonic fashion regarding the content of LSF and HSF information. Each image level was presented for 500 ms, followed by a 250ms blank screen. Participants had to press a pre-defined button as soon as they could decide whether the gradually unfolded scene was the one seen at the beginning of the trial, a new one, or a scrambled picture (Response1). Five hundred milliseconds after

Response1, participants validated their choice and indicated whether the scene was ‘‘old’’, ‘‘new’’ or ‘‘scrambled’’ by pressing one out of three buttons (Response2). (B) Results of the emotion classification task (occurring after Response2 on 10% of the trials) showed higher scores (corresponding to more pleasant pictures) for pleasant scenes (white bar), followed by neutral (light grey bar) and unpleasant (dark grey bar) scenes. ***: $p < .001$. Vertical bars correspond to standard errors of the means.

3. Results

3.1. Emotion classification task during the progressive unfolding experiment

Results showed higher ratings for pleasant scenes (M = 6.14, SD = 0.81), followed by neutral (M = 4.89, SD = 0.58) and unpleasant (M = 3.45, SD = 1.06) pictures. A one-way ANOVA on these ratings disclosed a highly significant effect of emotion [$F(2, 34) = 39.94$, $p < .001$, $\eta_p^2 = .701$]. Post-hoc t-tests confirmed highly significant differences between neutral and unpleasant pictures [$t(17) = 4.83$, $p < .001$], as well as between neutral and pleasant [$t(17) = -7.47$, $p < .001$] and unpleasant and pleasant [$t(17) = -6.81$, $p < .001$] scenes (Figure 1B). These results confirmed that participants perceived and identified the emotional content of the pre-selected stimuli in accordance with the published normative ratings (Lang, et al., 2008).

3.2. Accuracy for the progressive unfolding experiment

The percentage of errors remained low in this task (M = 3.66%, SD = 1.85). Likewise, very few errors were committed with “catch” trials (M = 1.75%, SD = 1.90). In addition, the percentage of late responses (Responses1 occurring after the last/sixth image level) was negligible (M = 1.71%, SD = 1.18), providing additional evidence that participants accurately performed the matching task during the gradual stimulus revelation and did not wait until the presentation of the last, fully detailed image level to stop the stimulus sequence (Response1).

Table 3 shows the cumulative percentages of correct responses (i.e., Responses1 only when Responses2 were correct).

Table 3. Mean values and standard deviations (in parenthesis) of cumulative percentages of correct responses, separately for each image level, emotion and memory condition.

	New			Old		
Image Level	Neutral	Unpleasant	Pleasant	Neutral	Unpleasant	Pleasant
Image1	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*
Image2	0.10	0.19	0.00	1.22	0.46	0.19
	(0.41)	*(0.79)*	*(0.00)*	*(2.42)*	*(0.00)*	*(0.79)*
Image3	5.01	3.28	2.35	13.12	10.26	7.55
	(5.36)	*(3.57)*	*(3.00)*	*(10.40)*	*(9.77)*	*(8.98)*
Image4	38.88	32.07	33.40	56.93	51.62	49.61
	(17.72)	*(17.19)*	*(18.24)*	*(19.28)*	*(16.48)*	*(17.68)*
Image5	85.66	77.94	81.38	90.98	87.95	87.85
	(14.51)	*(15.87)*	*(14.05)*	*(8.55)*	*(8.70)*	*(9.40)*
Image6	100.00	100.00	100.00	100.00	100.00	100.00
	(0.00)	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*	*(0.00)*

A mixed proportional odds model (Agresti, 2007; Schettino, et al., 2011) with memory (old, new) and emotion (neutral, unpleasant, pleasant) as fixed factors, and participant as random effect was carried out on these values, to verify whether the obtained psychometric curves shifted as a function of memory and/or emotion (Figure 2A and 2B). This analysis revealed, as expected, an overall earlier recognition for old compared to new scenes in all emotion conditions (all *p*s < .001). More interestingly, pairwise comparisons revealed a shift of the distribution as a function of the emotional content of the scenes. Specifically, an earlier recognition (i.e., less accumulation of evidence) was observed when the picture contained a neutral as opposed to an emotional content (all *p*s < .01), with no significant difference between pleasant and unpleasant scenes (all *p*s > .05) (see Table 4). The interaction between these two effects (memory and emotion) showed a trend towards significance (*p* = .064), indicating that the observed delay in recognition for emotional compared to neutral scenes was slightly more pronounced for old relative to new scenes.

Table 4. Results of the mixed proportional odds model.

Memory condition	Comparison	Odds ratio (95% CI)	p-value
New	pleasant vs. neutral	0.76 (0.63,0.91)	0.006*
	pleasant vs. unpleasant	0.76 (0.63,0.92)	0.008
	unpleasant vs. neutral	0.99 (0.81,1.20)	0.939
Old	pleasant vs. neutral	0.62 (0.51,0.74)	< 0.001*
	pleasant vs. unpleasant	0.61 (0.50,0.73)	< 0.001*
	unpleasant vs. neutral	1.01 (0.84,1.22)	0.901

Note. An odds ratio larger than 1 (smaller than 1, respectively) implies that the probability of recognition at earlier levels is higher (smaller, respectively) for the first vs. the second condition included in the comparison. CI indicates confidence interval. * indicates significant difference after Bonferroni correction.

3.1. Rating experiments

Familiarity ratings of the pre-selected IAPS pictures revealed lower scores for unpleasant scenes (*M* = 3.48, *SD* = 0.98), followed by neutral (*M* = 4.87, *SD* = 0.89) and pleasant (*M* = 4.93, *SD* = 1.05) scenes. A one-way ANOVA on these values disclosed a highly significant effect of emotion [$F(2, 40) = 58.64$, $p < .001$, $\eta_p^2 = .746$]. Post-hoc t-tests showed significant differences between unpleasant and neutral [$t(20) = -8.51$, $p < .001$], as well as unpleasant and pleasant [$t(20) = -7.70$, $p < .001$] scenes (Figure 3A). Mean familiarity was similar for pleasant and neutral scenes [$t(20) = -0.78$, $p = .445$].

Results of picture complexity ratings, on the other hand, showed higher scores for unpleasant (*M* = 4.70, *SD* = 0.81), relative to neutral (*M* = 4.13, *SD* = 0.64) and pleasant (*M* = 4.16, *SD* = 0.74) pictures. A one-way ANOVA carried out on these ratings revealed a highly significant effect of emotion [$F(2, 38) = 16.12$, $p < .001$, $\eta_p^2 = .459$]. Pairwise comparisons confirmed significant differences between unpleasant and neutral [$t(19) = 4.19$, $p < .001$] and unpleasant and pleasant [$t(19) = 5.87$, $p < .001$] scenes (Figure 3B), whereas no difference was observed between pleasant and neutral scenes [$t(19) = -0.28$, $p = .779$]. Thus, unpleasant pictures were characterized by lower familiarity and higher picture complexity compared to neutral and pleasant scenes.

Familiarity and picture complexity were found to be anti-correlated, as confirmed by a significant negative correlation [Pearson's $r(360) = -0.40$, $p < .001$].

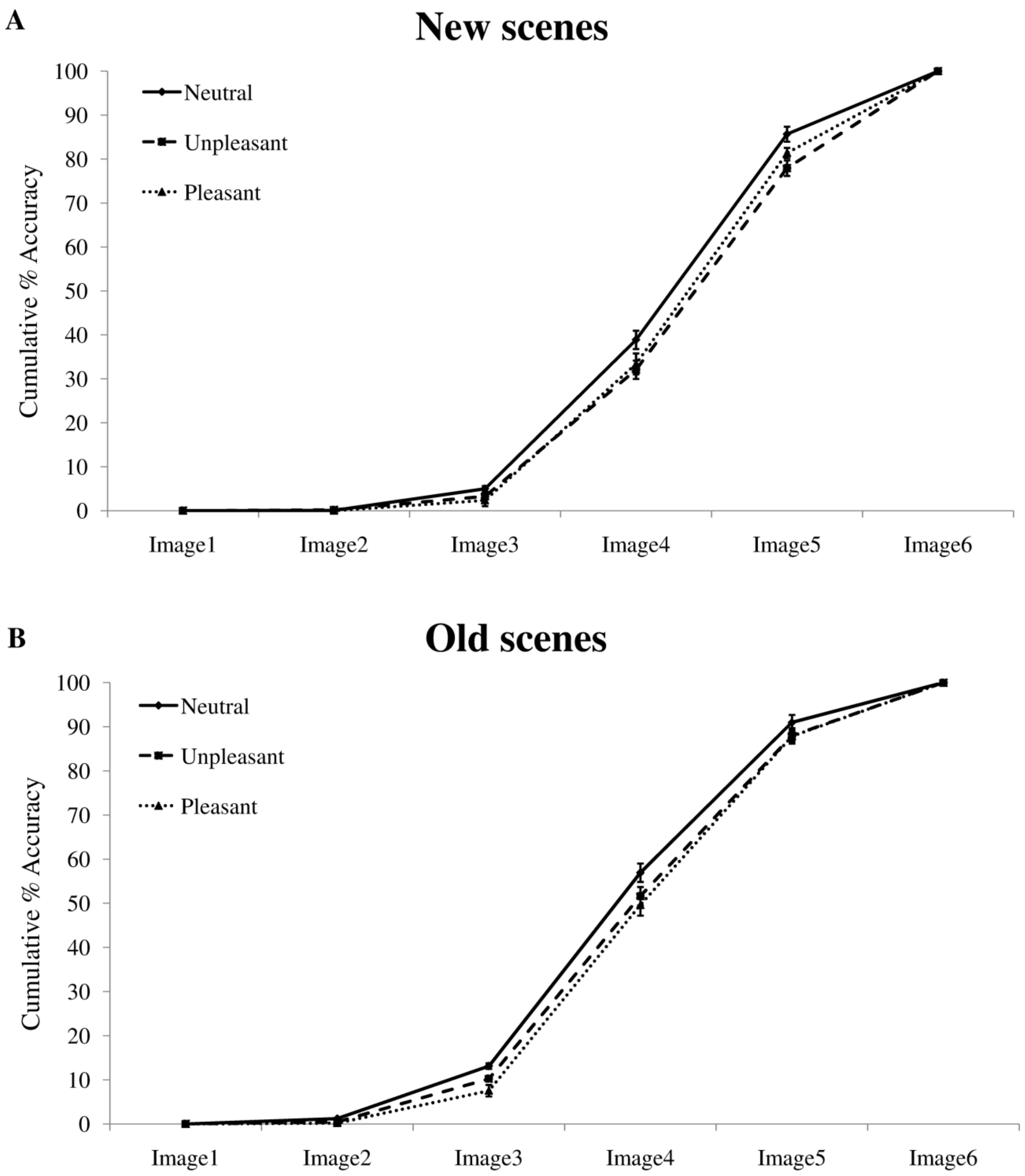


Figure 2. Accuracy in the main progressive unfolding task. Cumulative percentage of correct Responses1 as a function of the six image levels, in the (A) new and (B) old condition, separately for neutral (solid line), unpleasant (dashed line) and pleasant (dotted line) scenes. The shape and variation of the psychometric function according to the main experimental factors (memory and emotion) confirmed that: (i) participants gathered perceptual evidence prior to recognition; (ii) they had a significantly earlier recognition (i.e., less perceptual evidence needed) for old compared to new scenes. Moreover, for each of these two memory levels, emotional scenes led to a delayed recognition relative to neutral scenes. Vertical bars correspond to standard errors of the means.

3.2. Accuracy for progressive unfolding experiment when controlling for familiarity and picture complexity of the visual scenes

Next, we included the average familiarity and picture complexity ratings, obtained for each visual scene separately, as concurrent predictors in the proportional odds model, in order to statistically assess whether the significant effects of memory (i.e., prolonged explorations for new relative to old scenes) and emotion (i.e., prolonged explorations for emotional relative to neutral scenes) might be confounded by trial-to-trial fluctuations along these non-emotional dimensions.

Main effects of familiarity and picture complexity were significant (all *p*s < .001), indicating earlier recognition for more familiar and less complex pictures, in line with our predictions. However, and crucially, the analysis revealed that, after having modeled the potential contribution of these two factors (Table 5), pleasant scenes in the new condition were still associated with a delayed recognition relative to neutral pictures (*p* = .006) (see also Figure 4A). Pleasant scenes were also recognized later compared to unpleasant pictures (*p* = .034). However, this difference was no longer considered significant after correction for multiple comparisons (see Table 5). Interestingly, the difference between neutral and unpleasant scenes was no longer significant in this analysis (*p* = .621), suggesting that familiarity and picture complexity might have accounted for the difference between neutral and emotional scenes in our first analysis (see Table 4). A very similar statistical outcome was observed for old scenes: pleasant pictures led to a prolonged recognition compared to either neutral (*p* < .001) or unpleasant (*p* < .001) scenes (see also Figure 4B), whereas the difference between neutral and unpleasant pictures was no longer significant (*p* = .671). Importantly, the interaction effect between emotion and memory was not significant (*p* = .102), indicating that the delay in recognition for pleasant scenes was comparable in the new and old conditions. The shift found for pleasant relative to neutral scenes before correcting for complexity and familiarity (see Figure 2) did not therefore appear to be related exclusively to these two specific factors (unlike the case of unpleasant scenes), because the refined analysis controlling for variations along these factors still confirmed this shift (Figure 4).

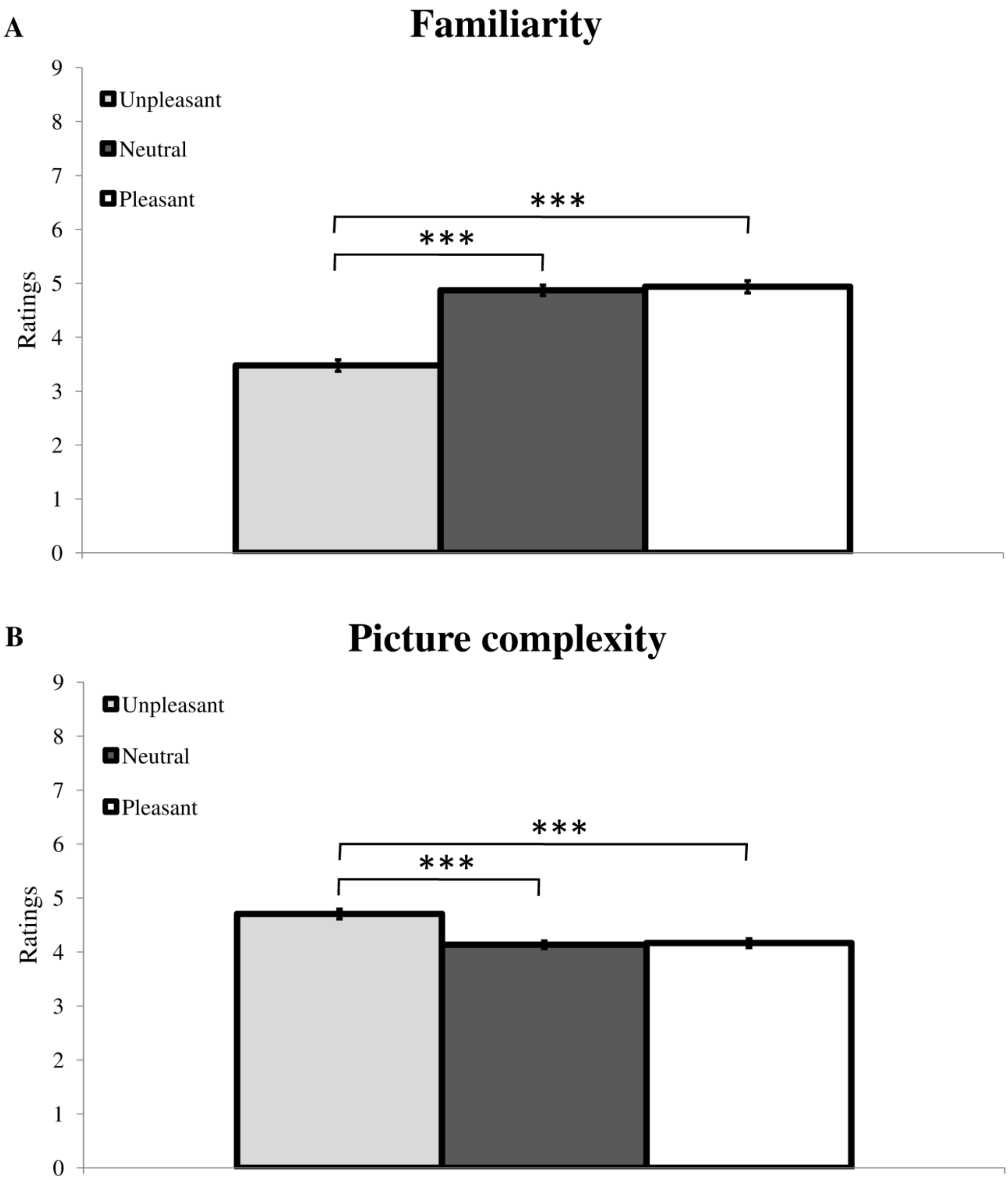


Figure 3. Familiarity and picture complexity ratings. Mean (A) familiarity and (B) picture complexity ratings, separately for neutral (dark grey bar), unpleasant (light grey bar) and pleasant (white bar) scenes. On average, unpleasant scenes were rated as less familiar and perceptually more complex (i.e., less obvious figure-ground segregation) relative to either neutral or pleasant scenes. ***: $p < .001$. Vertical bars correspond to standard errors of the means

In order to assess whether familiarity and picture complexity had different influences on accumulation of evidence processes in our experiment, we next modeled recognition performance separately for familiarity and picture complexity. Including effects of familiarity in the model (Table 5) revealed, in the new condition, a significantly delayed recognition for pleasant relative to neutral scenes (*p* = .002). The difference between pleasant and unpleasant scenes (*p* = .179), and between unpleasant and neutral scenes (*p* = .092) were not significant. In the old condition, pleasant scenes were also recognized reliably later compared to neutral (*p* < .001) and unpleasant (*p* = .001) ones, whereas the difference between unpleasant and neutral scenes was not significant (*p* = .228).

When modeling the specific contribution of picture complexity (Table 5), the analysis revealed, in the new condition, a delayed recognition for pleasant relative to neutral scenes (*p* = .006), whereas the unpleasant vs. neutral comparison was not significant (*p* = .939). The difference between recognition of pleasant vs. unpleasant pictures (*p* = .008) was marginally significant after Bonferroni correction. The analysis of recognition performance in the old condition revealed that pleasant scenes were recognized significantly later relative to neutral (*p* < .001) and unpleasant (*p* < .001) scenes, whereas the difference between unpleasant and neutral scenes was not significant (*p* = .901).

Although these analyses led to the same conclusions, it is interesting to note that -- based on the standard Akaike information criterion (AIC) (Akaike, 1974) -- the model including both familiarity and picture complexity was providing the best statistical fit. More specifically, the AIC was 13488 for the model including only familiarity, 12619 for the model with only picture complexity, and 12615 for the model with both factors. Since a lower AIC value is considered to fit the data better (Akaike, 1974), these results suggest that familiarity explained some of the variability over and beyond picture complexity, the inclusion of this latter factor providing a better fit than the former.

4. Discussion

The aim of our study was twofold: (i) to investigate whether the emotional valence of complex visual scenes could have an impact on the accumulation of perceptual evidence prior to their recognition, in line with previous findings showing a delayed recognition (i.e., prolonged accumulation of evidence) for emotional compared to neutral stimuli (Schettino, et al., 2011); (ii) to verify whether these effects may be explained by trial-to-trial fluctuations along other non-emotional variables, with a focus on familiarity and picture complexity.

We used a progressive unfolding task that proved to be useful to explore accumulation of evidence processes prior to scene recognition (Schettino, et al., 2011). After a standard picture encoding phase, participants were presented with series of filtered images that were progressively unfolding the same picture content, a new one, or a scrambled one relative to encoding, and the task was to decide whether this scene had previously been presented or not (delayed match-to-sample task). Of note, the content of either the same scene or a new one was progressively revealed by adding up, in a non-linear fashion, LSF and HSF information, providing a “coarse-to-fine” temporal decomposition of the visual stimulus (Bullier, 2001; Hegdé, 2008; Marr, 1982). We reasoned that the use of impoverished LSF information (and HSF information to a lesser extent), largely predominating at the beginning of the unfolding sequence, could foster the generation of proactive guesses about the actual identity of the scene progressively revealed (Bar, 2003, 2004, 2007).

Results showed a delayed recognition for new compared to old scenes, as well as for emotional relative to neutral pictures, consistent with our previous results (Schettino, et al., 2011). While the former memory effect confirms that participants used abstract visual representations stored in short-term memory to perform the task (Alvarez & Cavanagh, 2004; Vogel, Woodman, & Luck, 2001), the latter effect indicates that these predictive coding mechanisms during scene recognition were not immune to the rapidly perceived emotional content of the input stimulus. Specifically, pleasant and unpleasant scenes were associated with a delayed recognition relative to neutral pictures, suggesting an emotion-specific modulation of predictive coding effects during scene recognition. Moreover, this systematic time lag for recognizing emotional scenes was similar in the new and old conditions, suggesting a general effect taking place irrespective of the memory status of the perceived scenes.

Table 5. Results of the alternative mixed proportional odds model, once item-specific values along familiarity and picture complexity were included in the model.

Memory condition	Comparison	Familiarity and picture complexity		Only familiarity		Only picture complexity	
		Odds ratio (95% CI)	*p*-value	Odds ratio (95% CI)	*p*-value	Odds ratio (95% CI)	*p*-value
New	pleasant vs. neutral	0.76 (0.63,0.91)	0.006*	0.74 (0.62,0.88)	0.002*	0.76 (0.63,0.91)	0.006*
	pleasant vs. unpleasant	0.80 (0.65,0.98)	0.034	0.88 (0.72,1.07)	0.179	0.76 (0.63,0.92)	0.008
	unpleasant vs. neutral	0.95 (0.78,1.17)	0.621	0.84 (0.69,1.03)	0.092	0.99 (0.81,1.20)	0.939
Old	pleasant vs. neutral	0.62 (0.51,0.74)	< 0.001*	0.63 (0.52,0.76)	< 0.001*	0.62 (0.51,0.74)	< 0.001*
	pleasant vs. unpleasant	0.64 (0.53,0.78)	< 0.001*	0.71 (0.58,0.86)	0.001*	0.61 (0.50,0.73)	< 0.001*
	unpleasant vs. neutral	0.96 (0.79,1.16)	0.671	0.89 (0.74,1.08)	0.228	1.01 (0.84,1.22)	0.901

Note. *: significant difference after Bonferroni correction.

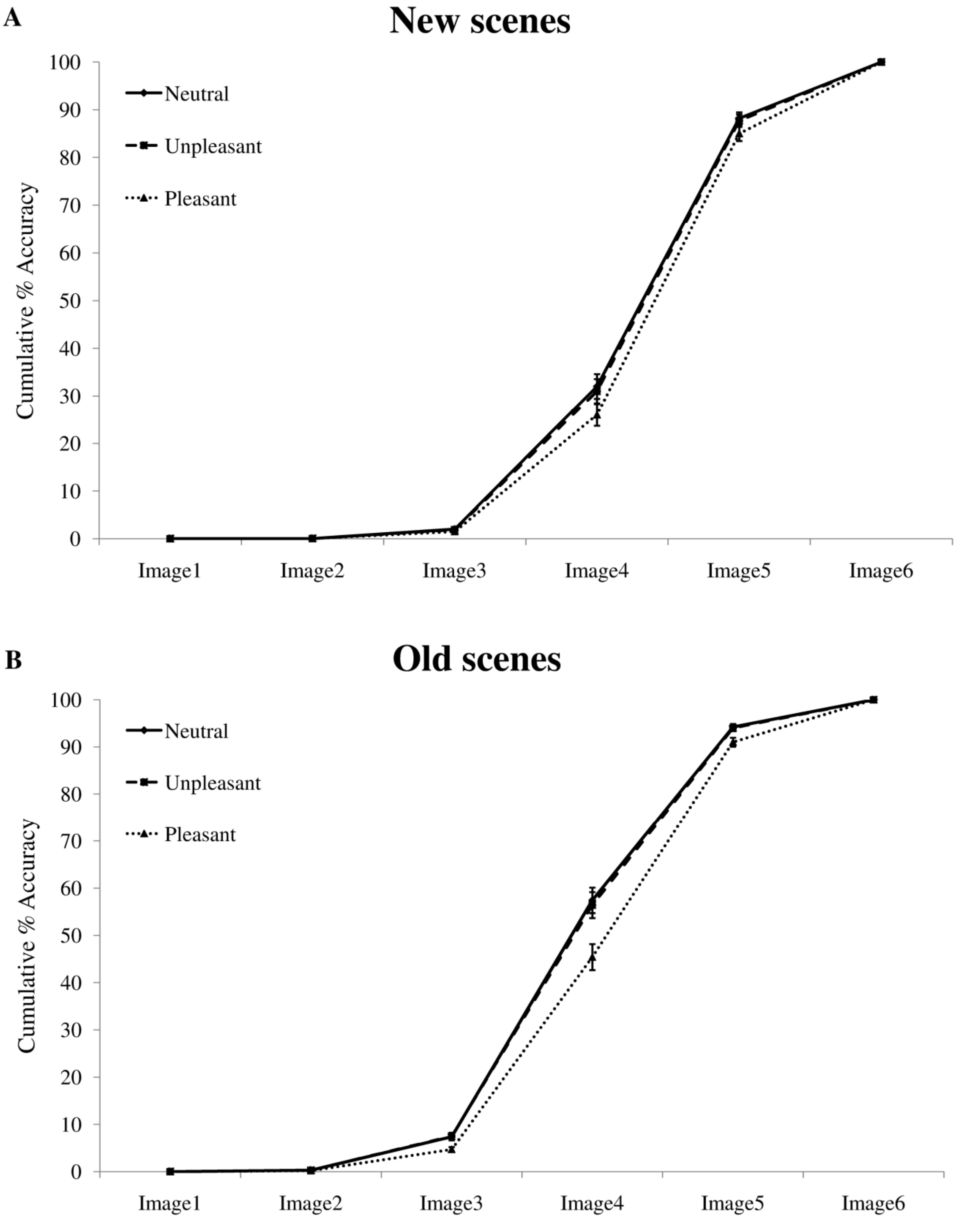


Figure 4. Accuracy in the main progressive unfolding task, adjusted for non-emotional factors. Cumulative percentage of correct Responses1 as a function of the six image levels, in the (A) new and (B) old condition, separately for neutral (solid line), unpleasant (dashed line) and pleasant (dotted line) scenes, once these values were adjusted for familiarity and picture complexity. A significant shift of the psychometric function (corresponding to prolonged accumulation of evidence) was observed for pleasant compared to either neutral or unpleasant scenes, regardless of memory (old vs. new). No significant difference was found between neutral and unpleasant scenes.

However, we also found that familiarity and picture complexity each had a substantial influence on accumulation of evidence processes prior to scene recognition. First, results of the additional rating experiments showed that the selected unpleasant scenes were rated as less familiar than either neutral or pleasant scenes (Figure 3A), consistent with previous work (Bradley, et al., 1993; Winkielman & Cacioppo, 2001). This result is in line with the well-known “mere-repeated-exposure” phenomenon (Monahan, et al., 2000; Zajonc, 1968, 2001), showing that human beings tend to develop a preference towards objects deemed familiar. Therefore, unpleasant objects or events that are typically avoided are considered as less familiar, exactly as found in our rating experiment. Second, our results showed that unpleasant pictures were also considered to be perceptually more complex compared to either neutral or pleasant scenes (Figure 3B). More specifically, unpleasant scenes were systematically associated with a less evident figure-ground segregation in the auxiliary rating experiment, an effect which might lead to a decreased fluency to process these scenes and hence confer them a negative valence (Winkielman & Cacioppo, 2001; Winkielman, et al., 2003). Thus, based on the lower familiarity and higher picture complexity scores obtained for the unpleasant relative to the neutral and pleasant scenes selected in our study, one would predict a change in the speed of accumulation of perceptual evidence for this specific class of emotional stimuli, when compared to the two other conditions. Likewise, given the balanced mean ratings for pleasant and neutral scenes, one could anticipate that accumulation of perceptual evidence would be similar for these two conditions. Instead, our single-trial analysis, in which we included familiarity and complexity ratings -- obtained for each and every scene separately -- as concurrent regressors (in addition to emotion and memory), revealed that *pleasant* scenes were associated with a distinctive delayed accumulation of evidence relative to the two other conditions, regardless of the memory status (old vs. new) and hence presumably ease of recognition of these scenes. Thus, at first sight, familiarity and complexity ratings alone could not account for the shift obtained for pleasant relative to neutral scenes during the main task. These results provide evidence for the contribution of *positivity offset* during emotion scene recognition (Cacioppo & Gardner, 1999; Cacioppo, et al., 1997, 1999; Norris, et al., 2010). This concept refers to the fact that, when inputs to the affect system are minimal, positivity outweighs negativity. As a consequence, organisms may engage in exploratory behavior under conditions in which no immediate threat is detected, with the aim to gain knowledge about novel stimuli in the environment and their potential value, an effect that is usually exacerbated for pleasant/positive compared to neutral or unpleasant stimuli [53]. Accordingly, the results of our study show that participants were prone to gather additional evidence about pictures carrying intrinsic reinforcing hedonistic values (in this case, pleasant pictures), probably because these pictures better matched their actual motivational dispositions. This latter observation also suggests that the influence of positive emotion on perception in our task was probably operating at an abstract level of stimulus representation, before or after specific short-term memory traces came into play. Of note, a prolonged exploration for pleasant relative to neutral or unpleasant scenes in our experiment may alternatively be explained by the differential motivational relevance of this specific emotion stimulus category (Bradley, 2009; Cacioppo & Gardner, 1999; Lang, et al., 1990; Norris, et al., 2010; Schupp, et al., 2000). This general account appears unlikely though, because we did not observe any gain or change during accumulation of evidence for unpleasant compared to neutral scenes, despite the obvious motivational and/or evolutionary relevance of these negative stimuli (Cacioppo & Ito, 1999; Ito, et al., 1998).

The prolonged accumulation of evidence for pleasant relative to neutral and unpleasant scenes may stem from an increase in the number of actual iterations made between updated predictions (initially shaped or constrained by the encoding of the scene in short-term memory) and the progressively accumulated degraded sensory evidence during unfolding, with the aim to minimize prediction errors and favor the most likely interpretation concerning the actual identity of the scene (Friston, 2002, 2005, 2009; Serences, 2008). Alternatively, rather than a quantitative change in the ratio between predictions and errors during accumulation of perceptual evidence, the processing of pleasant scenes may be associated with an overall shift in the decision criterion, relative to neutral or unpleasant scenes. In this view, accumulation of sensory evidence would occur equally fast for neutral and unpleasant scenes, but the delayed decision-making process for pleasant scenes would primarily stem from an enhanced competition between (two or more) choices or alternatives at the decision level *per se*. The use of computational modeling, and more specifically diffusion models, might turn out to be valuable in this context to tease apart these two accounts (Gold & Shadlen, 2007; Heekeren, et al., 2004; Heekeren, et al., 2008; Philiastides, et al., 2011). According to these models, decision making is achieved after having accumulated sufficient sensory evidence, and eventually gathered information in favor of one out of two (or more) alternatives, hence reaching a decision threshold (Gold & Shadlen, 2007; Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). The speed of accumulation of perceptual evidence (also termed *drift rate*) heavily depends on the strength of the sensory signal, as well as the signal-to-noise ratio. Thus, the aforementioned computational models provide useful hints to better explain how specific dispositions to engage in exploratory or approach-related behavior in non-threatening environments (i.e., positivity offset) may ultimately influence proactive processes leading to perceptual decision making. Further studies are needed to assess whether the processing of pleasant scenes is accompanied by a change in the drift rate compared to neutral or unpleasant scenes, or whether genuine post-perceptual processes may account for this emotion effect. Likewise, additional neuroimaging and/or neurophysiological studies might help clarify whether accumulation of evidence processes are actually generic but vary in speed -- depending on the emotional content of the input stimulus -- or, instead, several non-overlapping accumulation of evidence brain process may co-exist and can be activated predominantly depending on the valence of this input stimulus.

We have to acknowledge some limitations related to our experimental design and specific data analysis. Familiarity and visual complexity ratings of the pre-selected scenes were collected from two independent samples of participants, while another sample of participants completed the unfolding experiment. It would probably have been more optimal, from a statistical point of view, to use a full within-subject design. However, we did not want to create any bias or expectation regarding the content of the pictures that were progressively revealed during the main experiment. Therefore, we could not ask the same participants to rate the pre-selected visual scenes along the familiarity and picture complexity dimensions before the unfolding experiment. Conversely, ratings obtained for these stimuli would probably be influenced by prior exposure and unbalanced explorations during the unfolding experiment, as revealed for pleasant relative to neutral and unpleasant scenes in our study. Another limitation lies in the possible specificity of these effects for women, because we included mainly female participants and a differential processing of the emotional content of visual stimuli for men and women has previously been reported (Bradley, Codispoti, Sabatinelli, & Lang, 2001; Hamann & Canli, 2004; Wager, Phan, Liberzon, & Taylor, 2003). However, the pictures were carefully selected according to the normative ratings published in the manual for this specific gender (Lang, et al., 2008). Moreover, we purposefully decided not to include highly arousing pictures (e.g., mutilations or erotica) in our stimulus set, because these extreme pictures were found to elicit the largest differences between male and female participants in previous research (Schupp, et al., 2006, 2007).

To sum up, the results of our study show that accumulation of evidence prior to scene recognition is substantially influenced by the perceived emotional content of the visual stimulus. More specifically, emotional scenes were associated with a prolonged accumulation of evidence relative to neutral scenes. Controlling for non-emotional dimensions (i.e., familiarity and picture complexity) further revealed a delayed recognition for pleasant compared to unpleasant and neutral scenes, suggesting a valence-specific influence on the speed of proactive guesses prior to perceptual decision-making. More generally, these findings are consistent with a positivity offset during complex scene recognition. The propensity to dwell longer on pleasant compared to neutral or unpleasant scenes may be explained by a change in the ratio between predictions and errors during accumulation of evidence, while participants actively make guesses and computed online the most probable interpretation regarding the identity of the incoming and progressively unfolded visual scene. Finally, given the evidence showing a strong positivity offset during emotional scene recognition (that cannot easily be accounted for by systematic trial-to-trial fluctuations along familiarity or picture complexity), we believe that this specific experimental paradigm and stimulus set may eventually turn out to be valuable to shed light on possible qualitative alterations during visual emotion perception typically observed in specific psychopathological conditions. For example, this task appears useful to explore possible changes between the expression of positivity offset vs. negativity bias during scene or object recognition, a modification that might characterize exploration strategies preferentially used by depressed or high anxious individuals (Conklin, Strunk, & Fazio, 2009; Shook, Fazio, & Vasey, 2007).

CHAPTER 6

General Discussion

The main goal of this dissertation was to characterize the influence of emotion (here mainly operationalized as changes in valence and arousal values of pre-selected faces or complex visual scene) on visual perception. In particular, I sought to investigate the extent to which emotional factors influence perceptual processing during (either face or scene) recognition in a bottom-up, “reactive” as opposed to a top-down, “proactive” way. To answer these questions, I first looked at bottom-up effects of emotion on the temporal perception of competing events by means of dedicated emotional attention mechanisms that are thought to prioritize the early detection (and hence onset discrimination) of motivationally salient stimuli (here angry or fearful faces), relative to neutral stimuli. Next, I explored how emotion could bias visual perception via top-down mechanisms, operating here as proactive guesses or expectations generated “online” by the participants about the most probable identity of the visual stimulus under conditions of impoverished retinal input (see Chapters 3-5). Two distinct experimental paradigms have been devised and used to address these complementing questions: a standard Temporal Order Judgment (TOJ) task (Chapter 2) and a progressive unfolding paradigm (Chapters 3-5). Regarding the first question, the results obtained reveal that motivationally significant stimuli do not “magically” capture attention and in turn bias temporal perception, but other factors (e.g., distinctive low-level perceptual cues such as motion or, alternatively, the actual task set) may compete for bottom-up attention selection with (higher-level) emotional cues, and eventually prevent the latter to exert a systematic influence on the “automatic” guidance of attention (Chapter 2). With regard to the second question, results show that the generation and use of specific predictions regarding the most probable content of complex visual scenes -- progressively revealed across several successive steps -- are not immune to emotional factors, lending support to the assumption that emotion, besides other components (e.g., attention or memory), can proactively influence the selection of diagnostic perceptual cues eventually used by the visual system to extract the meaning of the visual input (Chapters 3-5).

1. “Reactive” influences of emotion on perception: No prior entry for threat-related faces

In Chapter 2 I tested, across five experiments, the hypothesis of an early attentional capture by motivationally relevant (i.e., threat-related) face stimuli by means of a TOJ task: participants were presented with a pair of faces, one on the left and one on the right side of fixation, with variable SOAs between their respective occurrences. Critically, one of the faces was either angry or fearful, whereas the other one was neutral. Participants were asked to judge which face stimulus appeared first (Experiments 1-2), or whether the emotional face was presented first (Experiments 3-5). I surmised that, due to their enhanced motivational relevance (Anderson, et al., 2003; Whalen, et al., 1998), emotional faces would reflexively attract attention and in turn bias the perception of the two competing onsets in favor of the location where the negative emotional face was presented. More specifically, I predicted that the angry or fearful face in the pair would systematically be perceived as appearing first relative to the neutral face, particularly in conditions of uncertainty (namely when the SOA was short). I opted for the use of the TOJ task because it has previously been shown to selectively measure the early orienting of (spatial) attention, with little confound or contribution of other attentional components, e.g., disengagement or re-orienting (Jaskowski, 1993; Stelmach & Herdman, 1991).

However, the results of five experiments (Chapter 2) clearly failed to provide evidence for this hypothesis. Importantly, these null findings could not easily be accounted for by participants’ inability to make use of the information about the onset of the two faces, because higher accuracy and faster reaction times for long (e.g., 100 ms) relative to short (e.g., 10 ms) SOAs were observed, meaning that they made active use of this temporal cue in order to resolve the TOJ task. In addition, by asking participants to rate the emotional content of the face stimuli at the end of each experimental session, I could rule out the possibility that fearful or angry expressions were not correctly recognized as such by the participants, because these ratings unambiguously showed strong asymmetries in the amount of perceived fear or anger across the different face conditions in the expected direction. Furthermore, suboptimal task parameters such as inappropriate stimulus size and/or distance from fixation appeared equally unlikely to account for these null findings, given that I used a standard experimental set-up previously used in the literature (e.g., West, et al., 2009). Moreover, the use of different task instructions across different experiments did not alter the pattern of results, namely no significant prior entry effect was observed for threat-related faces despite task demands becoming more specific across the five experiments: no reliable attentional bias for emotional vs. neutral faces was found, either when participants had to perform a simple left-right discrimination task (hence irrespective of the emotional content of the faces), or when they were requested to base their decisions on the perceived emotional content of the faces (thereby making the emotional facial expression directly task-relevant). In addition, I devised a variant of the TOJ paradigm and asked participants to focus on the offsets of the two competing faces rather than their onsets, to assess whether the null findings obtained with the original TOJ task could be explained by possible confounds related to disengagement effects (i.e., attention was held, rather than captured, on the spatial location where the threatening face stimulus was presented; Fox, et al., 2001; Koster, et al., 2004). However, I still failed to find evidence for a differential effect triggered by the negative emotional facial expressions. Hence, obvious uncontrolled confounds (such as the inability to identify the emotional content conveyed by the face stimuli), inappropriate task demands, or other methodological issues (e.g., lack of statistical power) did not provide plausible alternative explanations to account for these null results.

1.1. Abrupt stimulus onset is prioritized over emotional information during attention selection

Even though some caution is needed when trying to interpret null findings, given the (carefully controlled) five different experiments conducted in large samples of participants (ranging from 36 to 40) it appears reasonable to conclude that these results do not support the assumption of an “automatic” or “magical” attentional capture by motivationally significant stimuli (i.e,. conveying either a direct or indirect threat). An intriguing possibility to account for these results is that participants probably (implicitly) used a specific strategy that implied an active filtering of the emotional content of the faces, even when, surprisingly, task instructions required them to selectively attend to this specific feature in order to perform the TOJ task (see Experiments 3-5 in Chapter 2). Even in these conditions, participants presumably relied on non-emotional cues to guide their behavior and eventually resolve the task, particularly in conditions of uncertainty (i.e., short, as opposed to long SOAs). More specifically, one could hypothesize that the systematic variation in the SOA between the two competing face onsets created a compelling impression of apparent motion on the screen (Goebel, et al., 1998), compatible with the “illusory line motion” phenomenon (a line is perceived as dynamically stretched away from a visual focal point, although in reality it is quickly presented at full length; Hikosaka, et al., 1993; Schmidt, 2000). This conjecture, albeit speculative, is also consistent with a large body of studies in the literature showing that abrupt onset/offset effects provide compelling cues that somehow reflexively capture attention (Jonides & Yantis, 1988; Schreij, Owens, & Theeuwes, 2008; Theeuwes, 1995; Yantis & Jonides, 1984). For example, abrupt, unattended, and unexpected onsets of distracter stimuli in the visual array reliably slow down target detection during visual search tasks, lending support to the notion that onset is an important stimulus feature because its occurrence may be associated with important changes in the environment that need to be readily detected by the organism (Jonides & Yantis, 1988; Theeuwes, 1995; Todd & Gelder, 1979), e.g., the appearance of a new object in the environment (Davoli, Suszko, & Abrams, 2007; Yantis & Hillstrom, 1994). Similarly, it could be argued that, in the TOJ task, participants’ attention was first reflexively drawn to the direction (either leftward or rightward) of this apparent motion artificially created by the two asynchronous onsets, rather than the emotional content of the faces. Only during a second (and perhaps post-perceptual) stage, participants “retrospectively” processed the emotional content of the faces based on this apparent motion cue. Of note, the lack of prior entry effect for faces conveying either a direct or indirect threat (i.e., anger or fear, respectively) suggests that changes in the actual motivational significance of the faces (e.g., Adams, Gordon, Baird, Ambady, & Kleck, 2003; Anderson, et al., 2003; Sander, et al., 2007) were not accompanied by systematic variations in early attention selection processes. As my results suggest, attentional capture towards both direct and indirect threat seems to be only concomitant to the prioritization of another low-level perceptual cue, of which (apparent) motion appears to be the most plausible candidate at this stage.

1.2. Capture by perceptual dissimilarity rather than emotion per se?

At first sight, the results of the five experiments reported in Chapter 2 appear at odds with two earlier studies published in the literature that already used the TOJ task and showed prior entry effects for emotional faces . However, a careful consideration of several methodological elements casts doubt on the robustness of these effects (see Chapter 2 for details). Among them, a brief analysis of the stimuli used in these two studies may potentially provide an alternative explanation as to why prior entry effects were observed for emotional relative to neutral faces in these earlier studies, in contrast to my results.

In the study by Fecica & Stolz (2008), as well as in Experiments 1-4 in West, et al. (2009)'s study, schematic neutral and emotional faces were used to explore whether emotion could lead to a genuine prior entry effect during early attention selection or orienting. Many studies in the literature, mainly using visual search tasks, have investigated attentional biases towards emotional (in particular threat-related) stimuli by employing schematic faces, in order to be able to better control for perceptual differences between emotional expressions (e.g., Fox, et al., 2000; Fox, et al., 2001; Öhman, Lundqvist, et al., 2001); for a recent review, see Horstmann, 2007). However, these schematic face stimuli have been extensively criticized for lacking ecological validity (e.g., Horstmann & Bauland, 2006). In addition, an alternative interpretation stipulates that specific low-level features embedded in these schematic face stimuli may be sufficient to promote differences in detection speed, rather than a processing of their emotional content (Becker, et al., 2011; Mak-Fan, et al., 2011; VanRullen, et al., 2004; Wolfe & Horowitz, 2004). More specifically, the orientation of the internal features (e.g., the curvature of the mouth or eyebrows) relative to the external circular edge delimiting the stimulus would be the crucial element that allows the visual system to identify an emotional face target among neutral distracters, without the need to postulate any mediation by specific emotion brain mechanisms (Coelho, Cloete, & Wallis, 2010; Purcell & Stewart, 2010). Therefore, TOJ tasks using schematic faces (i.e., Fecica & Stolz, 2008, and Experiments 1-4 in West, et al., 2009) should be treated carefully with regards to the potential generalizability of their results, as well as interpretations postulating a direct mediation by emotional attention brain mechanisms.

Interestingly, Experiments 5 and 6 in West, et al. (2009) confirmed prior entry effects for motivationally relevant stimuli when using realistic emotional and neutral faces, selected from the same standardized database (Ekman & Friesen, 1976) as used in Chapter 2. Importantly, West, et al. (2009) reported PSS values for these realistic faces which were even larger than the ones obtained with schematic faces, suggesting that the use of more ecologically valid stimuli further enhanced attentional capture for emotion-laden objects. However, in the Methods section of Experiments 5 (which extends to Experiment 6) it is reported that only four face identities (two men and two women) were used as stimulus set and, as a consequence, a limited number of face pairs (between 12 and 16, depending on the inclusion of trials with neutral and emotional faces of the same identity) were used in this earlier study. Although this strategy perhaps eases the burden of having to control for perceptual confounds, on the other hand it likely compromised the normal (ecologically valid) variability across the different encounters of the face stimuli (Pinkham, Griffin, Baron, Sasson, & Gur, 2010). More importantly, however, these experimental conditions may have favored the use of a perceptual strategy based on the detection of the degree of similarity between the faces in the pair (rather than any difference between the two faces along a genuine emotion dimension), this factor generally being known to influence performance during visual search tasks (Duncan & Humphreys, 1989). Specifically, neutral and emotional faces may have remarkably differed not only in terms of emotional expression, but with regards to other factors, such as first order configuration (e.g., the contrast ratio between the sclera and pupil) or second order configuration elements (e.g., the distance of the eyes from the nose). These differences may ultimately have guided attention allocation and, in turn, artificially created a bias towards emotional faces, without corresponding to a true capture of attention by emotion. To avoid these confounds and provide sufficient variability across the different pairs, I opted for the use of ten different identities (four women) taken from the same database (Ekman & Friesen, 1976), which allowed me to create a large number of face pairs (90 per condition, excluding pairs with neutral and emotional faces expressed by the same identity). The added value of this alternative procedure is that, in every trial, the degree of perceptual dissimilarity between the two faces in the pair was always uninformative, thus preventing participants to use this cue to perform the TOJ task. However, in these conditions, the results reported in Chapter 2 revealed no reliable attentional capture for threat-related compared to neutral faces. Thus, I surmise that West, et al. (2009)'s results could be explained (at least partly) by a systematic imbalance in terms of perceptual (dis)similarity between emotional and neutral faces (Duncan & Humphreys, 1989). At any rate, future studies are needed in order to assess whether the degree of visual (dis)similarity, rather than the emotional expression, is eventually the critical feature accounting for a prior entry effect for threat-related faces when they compete with neutral faces for attention selection and access to awareness.

1.3. Task set

In addition, it should be noted that task instructions (despite being adapted across different experiments; see Chapter 2) always emphasized the processing of the respective onsets in the pair in order to resolve the task. More specifically, participants were required either to indicate whether the face on the left or right relative to central fixation appeared first (Experiments 1-2), or whether the emotional or neutral face appeared first (Experiments 3-5). This procedure (which is, notably, identical to the one employed in previous studies; Fecica & Stolz, 2008; West, et al., 2009), likely imposed a strong top-down attentional set towards the detection and processing of the face onsets, perhaps at the cost of the reflexive processing of the emotional content of the faces (Banich, et al., 2000). Hence, even when the negative emotional content of the face became task-relevant (Experiments 3-5) as opposed to the mere spatial location of the two onsets (Experiments 1-2), still in these conditions the perceptual system was probably prioritizing the onset information (corresponding to the most compelling and useful cue immediately available), and only integrating at a later stage of processing (possibly when iconic or short-term memory processes were active) the emotional information in order to eventually resolve the task.

1.4. The possible contribution of inter-individual differences

Another potential reason as to why threat-related faces were not prioritized over neutral faces during the TOJ tasks could be related to “flattened” personality traits, more specifically the fact that I primarily tested non-anxious or non-dysphoric participants (as verified using standard personality questionnaires). Earlier studies based on other experimental paradigms (usually cueing or dot probe tasks) already showed stronger attentional capture for negative emotional (face) stimuli in participants having specific negative affect traits or states (Bar-Haim, et al., 2007; Bishop, 2007, 2008; Cisler & Koster, 2010; Fox, 1993; Mogg & Bradley, 1998). Accordingly, future studies are needed in order to assess whether a prior entry effect for threat-related faces could be found in high anxious or depressed participants, who usually show generalized attentional biases towards this specific category of visual stimuli.

1.5. Implications for models of emotional attention

The results obtained in Chapter 2 suggest that motivationally relevant stimuli (i.e., angry and fearful faces) do not “automatically” capture attention when competing with neutral stimuli for early attention selection. Instead, as proposed above, participants perhaps prioritized non-emotional cues to successfully perform the task (particularly in conditions of uncertainty, i.e., at short SOAs), given the enhanced perceptual salience of these non-emotional cues. Although speculative, apparent motion, as well as the perceptual (dis)similarity of the two face stimuli in the pair, appear to offer plausible (and useful) explanations for the lack of prior entry effect for threat-related faces. In this model, there is therefore no “reflexive” capture of attention by negative emotion during TOJs because compelling perceptual cues may guide initial attention selection processes and in turn override a genuine emotional effect.

Recent models emphasize the pivotal role of emotional factors in guiding attention and in turn influencing visual perception, via brain mechanisms that do not overlap with those involved in the exogenous or endogenous control of attention (Brosch, et al., 2011; Keil, et al., 2005; Pourtois, et al., 2012; Todd, et al., 2012; Vuilleumier, 2005). Given this architecture, effects of emotion on attention are thought to occur in parallel to other attentional effects mediated by dorsal frontoparietal networks involved in the endogenous and exogenous control of attention (Corbetta & Shulman, 2002). The results of Chapter 2 seem, at first glance, to disconfirm this model: the use of low-level perceptual cues (e.g., apparent motion) appears to be sufficient to reflexively capture attention, with the analysis of the emotional content of the face stimuli presumably occurring at a later (post-perceptual) stage of processing, after the rapid initial orienting of attention. However, I believe it is still possible, despite these null findings, to accommodate the results of Chapter 2 to the framework of emotional attention outlined here above, namely by considering the distinctive temporal characteristics of exogenous, endogenous, and emotional effects during the guidance of attention.

In a recent study, Brosch, et al. (2011) recorded EEG during a modified version of the dot-probe task (for a description of the paradigm, see paragraph 2.6 of Chapter 1) devised to orthogonally test the influence of emotional, exogenous, and endogenous factors during attention selection. In a typical dot-probe task, attention is quickly oriented towards the location where the non-informative cue appears, thereby facilitating the subsequent processing of a target stimulus if it appears at the same cued location (MacLeod, et al., 1986; see also Pourtois, et al., 2004). For example, contrast sensitivity is augmented for Gabor patches appearing at the location previously occupied by emotional (i.e., fearful) faces (Phelps, et al., 2006; see also Bocanegra & Zeelenberg, 2009, 2011b). Brosch, et al. (2011) found selective effects of exogenous attention (here changes in luminance) at the level of the *cue-locked* N2pc, a lateralized component sensitive to fast shifts of spatial attention towards stimuli appearing in the left or right visual field (Woodman & Luck, 1999). Conversely, larger amplitudes of the *target-locked* P1 component were found for targets following an emotional compared to a neutral cue, indicating an increase of sensory gain for stimuli presented in the location previously occupied by an emotional object (Luck, et al., 2000). To interpret these results, these authors proposed a "cascade" of attention mechanisms, in which the effects of exogenous attention are thought to quickly influence early stages of stimulus processing, with effects of emotional attention occurring slightly later and presumably "tuning" the visual system to specific features of the target depending on the emotional information conveyed by the preceding cue (Brosch, et al., 2011). Similarly, during the TOJ task (Chapter 2) it could be hypothesized that the abrupt onset of the first face stimulus would swiftly capture attention towards its location, but the "late" effects of emotional attention would be disrupted by the onset of the second face in the pair. Only after the second stimulus has disappeared, the visual system would process the emotional features of the two stimuli, and this analysis would necessarily occur at a post-perceptual stage (presumably in short-term visual memory). It should also be mentioned that endogenous factors such as task instructions, i.e., requesting participants to judge which face in the pair appeared first, would bias participants to specifically focus on the stimulus onset, thereby working in synergy with the exogenous attentional effects described above and likely overshadowing or tempering a guidance of early attention processes by the emotional features of the face stimuli (see Figure 1).

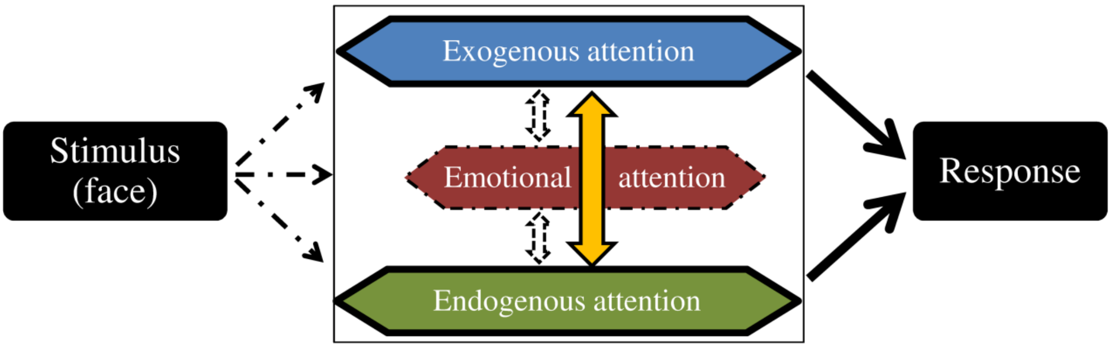


Figure 1. Tentative model accounting for the results reported in Chapter 2. Five experiments (based on the use of a stringent TOJ task) failed to confirm the assumption that threat-related faces reflexively capture attention. Instead, other non-emotional, exogenous cues, among which apparent motion and the degree of visual (dis)similarity seem to be plausible candidates, are likely used by the visual system in order to resolve the task. In addition, task instructions (i.e., requesting participants to judge which stimulus appeared first) might contribute to endogenously bias attention to the preferential processing of the onset of the faces. The synergic

interplay of exogenous and endogenous effects would eventually prevent emotion to exert a strong and specific influence on the early guidance of attention following stimulus onset. (Adapted from Pourtois, et al., 2012).

In this view, specific characteristics of the task set could also play a role and eventually explain why attention selection was not based on the distinctive emotional content of the face stimuli. The most important difference between the dot-probe task used in Brosch, et al. (2011) and the TOJ task used in Chapter 2 lies on the fact that, in the latter, there is no emotional cue preceding the target. The cue thought to guide early attention selection (i.e., a threat-related facial expression) *is* the target. It follows that the visual system probably tunes to specific features that would enhance the processing of a target appearing in the same location (Bocanegra & Zeelenberg, 2009, 2011b; Brosch, et al., 2011; Phelps, et al., 2006), although the following stimulus would always appear on the *opposite* side.

Alternatively, it could be hypothesized that performing TOJs based on the fast presentation of the face stimuli -- ranging from 110 to 200 ms depending on the SOA (10 and 100 ms, respectively; see Chapter 2 for details) -- would leave little resources available for the processing of their emotional expression. Previous studies have already proposed that cognitive load can have a profound impact on the processing of the emotional content of the stimuli (Eimer, Holmes, & McGlone, 2003; Holmes, Vuilleumier, & Eimer, 2003; Pessoa, Kastner, et al., 2002; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002), in that the processing of unattended stimulus features would be possible only when attention is not fully depleted (Lavie, 1995; Treisman, 1969). In this case, the unattended stimulus feature would be the emotional connotation of the faces in the pair, given that the task required participants to primarily focus on the onset of the stimuli. However, this explanation seems unlikely, given that no prior entry was found even when the instructions explicitly requested participants to detect whether the emotional or neutral face appeared first.

Hopefully, these interpretations might provide valuable avenues for future research, and at this stage they could help test specific predictions regarding the specificity or overlap between different components of attention selection (particularly regarding their temporal dynamic), when they concern emotional or non-emotional factors.

2. ”Proactive” influences of emotion on perception: the generation of prediction during emotional scene recognition

In the second part of this thesis (Chapters 3-5), I investigated whether emotion could influence visual scene recognition in a proactive (as opposed to reactive) way. To tackle this issue, I validated and used a new experimental paradigm (progressive unfolding task) in which the content of natural visual scenes was progressively revealed to participants by gradually adding LSF and HSF information to an initial blurred and undistinguishable image. Participants had to report whether the scene contained a living object or not (Chapter 3), or whether the content of the scene matched the one briefly presented at the beginning of the trial or not (Chapters 4-5). The main purpose of this procedure was to simulate a “coarse-to-fine” decomposition of the retinal input eventually fostering the online generation of guesses about its content, keeping in mind that LSF and HSF cues differentially contribute to these proactive processes during object and scene recognition (Bar, 2003; Bullier, 2001; Hegdé, 2008; Navon, 1977; Schyns & Oliva, 1994). Moreover, in Chapters 3-4, high density EEG was simultaneously recorded in order to characterize the exact spatio-temporal dynamics associated with these complex perceptual processes. Overall, results suggest that the motivational relevance of the scenes reliably influenced recognition speed, emotional scenes being recognized later compared to neutral scenes. Distinct neural networks were found to underlie different proactive perceptual processes, with ventral occipito-temporal regions gradually accumulating perceptual evidence while medial frontal areas subtended categorical, decision-related processes (Chapter 3). Furthermore, the reconstructed neural activity in these networks was modulated by the emotional content of the scenes as well as the extent of previous experience with these stimuli, presumably influencing the generation of predictions regarding the most likely content of the scene (Chapter 4). Importantly, I also ascertained that the delayed recognition for emotional compared to neutral scenes observed at the behavioral level could not be solely ascribed to other non-emotional factors, such as familiarity and picture complexity (Chapter 5).

2.1. Chapter 3: Behavioral effects and electrophysiological correlates of proactive processes during scene recognition

The aim of this Chapter was twofold: (i) to identify, using high-density EEG, the electrophysiological correlates of accumulation of evidence processes likely operating prior to overt recognition, keeping in mind that earlier (fMRI) studies (employing neutral objects or scenes) already found out that these processes involved non-overlapping brain networks showing different tuning functions (Ploran, et al., 2007; Ploran, et al., 2011); (ii) to explore whether the emotional content of the scenes might influence these upstream perceptual processes leading to visual scene recognition, given the strong reciprocal interaction effects seen between regions involved in emotion processing (such as the amygdala and OFC) and visual areas in the occipital and infero-temporal cortex implicated in visual object or scene recognition (Pourtois, et al., 2012; Rauss, Schwartz, & Pourtois, 2011; Vuilleumier, 2005).

Behavioral results confirmed the validity of this paradigm to titrate generative effects of predictions or guesses occurring prior to overt scene recognition: participants did not respond randomly on a trial-by-trial basis, but instead they systematically accumulated sufficient perceptual evidence (i.e., they waited, on average, until the third out of six image levels) before accurately deciding whether the content of the scene was living or not. Interestingly, this process was systematically modulated by the emotional content of the scenes. Specifically, visual information was accumulated less rapidly for unpleasant compared to neutral scenes, while the slowest accumulation of evidence was found for pleasant scenes. These results were interpreted under the framework of the Evaluative Space Model (Cacioppo & Berntson, 1994, 1999; Norris, et al., 2010). A faster accumulation of evidence for unpleasant relative to pleasant scenes could reflect a *negativity bias* effect, i.e., stronger responses for aversive compared to appetitive stimuli in the environment (Cacioppo & Ito, 1999; Ito, et al., 1998). On the other hand, a prolonged exploration for pleasant compared to neutral scenes could be explained by a *positivity offset* effect (Ito, et al., 1998; Norris, et al., 2010). Based on the assumption that engaging in exploratory behavior is promoted in non-threatening situations, pleasant scenes could be associated with a delayed recognition compared to neutral or unpleasant scenes, given their intrinsic hedonic (and approach-related) value. Functionally, this effect could also reflect a change in the rate of accumulation of perceptual evidence made by the participants prior to overt recognition. At any rate, these behavioral results shed light on the modulatory role of emotional factors during the accumulation of perceptual evidence prior to scene recognition.

Source localization analyses performed with sLORETA (Pascual-Marqui, 2002) identified several brain structures that differed regarding their actual response profile: (i) a monotonic accumulation of evidence, involving regions of the posterior cingulate cortex (PCC) and parahippocampal gyrus (PHG); (ii) categorical recognition effects in medial frontal regions, including the dorsal anterior cingulate cortex (dACC). I proposed that the PCC/PHG complex could be involved in updating contextual information about the scene in which the main objects were embedded, presumably by rapidly extracting and analyzing LSF information (Bar, 2004, 2009b; Bar & Aminoff, 2003). On the other hand, the dACC could contribute to scene recognition mechanisms by monitoring the ongoing decision’s outcome (Anderson, et al., 2009; Fincham & Anderson, 2006; Sohn, et al., 2007), serving therefore as a key structure integrating the accumulation of perceptual evidence with the implementation and execution of a (motor) response.

The pattern of recognition behavior was congruent with accumulator models of perceptual decision making (Gold & Shadlen, 2007; Heekeren, et al., 2008; Wang, 2008). It can be hypothesized that accumulation of perceptual evidence in favor of one out of two alternative responses (i.e., “living” vs. “non-living”) would mainly take place when the ambiguity of the visual input is (at least partially) resolved, that is when distinctive features belonging to one of the two categories emerged after the processing of diagnostic LSF and HSF information (Bar, 2003; Hegdé, 2008). Complementary source localization analyses showed that the accumulation of visual information was primarily subserved by inferior temporal regions such as the PCC and PHG, presumably reflecting the analysis of the contextual cues surrounding the main object embedded in the scene (Bar, 2004; Bar, et al., 2008; Fenske, Aminoff, Gronau, & Bar, 2006). After the perceptual evidence accumulated in favor of one alternative surpassed a recognition threshold, medial frontal areas such as the dACC monitored the response and provided online adjustments based on feedback loops (Kable & Glimcher, 2009; Sohn, et al., 2007).

The progressive unfolding task used in Chapter 3 contributed to characterize the electrophysiological correlates of perceptual and decision-related processes occurring prior to scene recognition. However, I had only very little grasp on the number and content of expectations and guesses generated online by the participants on a trial-by-trial basis. Because each scene was new and complex, and no cue was given to participants regarding the most probable scene content progressively unfolded, no control on the number and content of the proactive guesses (presumably generated by participants during unfolding) could be achieved. This is an important limitation, because recent research found that the number and selectivity of guesses associated with a particular object reliably influenced visual recognition processes: a stimulus associated with a low number of guesses belonging to the same semantic category was recognized more easily than a stimulus generating many heterogeneous guesses (Torfs, Panis, Bartlema, & Wagemans, submitted; see also Bar, 2009a). To overcome this problem, in Chapter 4 I modified the progressive unfolding task and introduced a standard memory manipulation with the aim to standardize the actual content of the proactive guesses generated by the participants on a trial-by-trial basis during emotional scene recognition.

2.2. Chapter 4: Multiple synergistic effects of emotion and memory during proactive visual scene recognition

In Chapter 4 I attempted to better control the content and specificity of prior expectations generated by participants during the gradual unfolding of complex scenes. This was achieved by introducing a standard memory manipulation: for each trial I primed participants with a specific stimulus content, and I subsequently asked them, during the unfolding phase, to perform a matching task based on the processing of this initial cue (see also Goldman-Rakic, 1990; Rugg, et al., 1998). The purpose of this manipulation was to constrain the predictions proactively made by participants during unfolding, an effect that should augment the specificity of these generative processes and in turn their expression at the electrophysiological level. Each trial began with the presentation of a colorful (pleasant, unpleasant, or neutral) IAPS scene (Lang, et al., 2008), and participants were required to encode and remember it for the duration of the trial. After a constant time interval, the content of either the same or a novel visual scene was gradually revealed, similarly to Chapter 3. Participants were required to judge whether the progressively unfolded scene had the same content than the one presented at the beginning of the trial or, instead, a new scene was gradually unfolded (delayed match-to-sample task). With this experimental manipulation, I promoted the likely generation of a restricted number of guesses regarding the most probable content of the unfolded scenes. Moreover, participants were occasionally asked to rate the emotional valence of the colorful scenes (besides the delayed match-to-sample task), thereby making the emotional content task-relevant. Further, the differential processing of emotional vs. neutral scenes was also verified at the electrophysiological level by recording and analyzing amplitude changes of the late positive potential (LPP; Foti, et al., 2009; Schupp, Cuthbert, et al., 2004; Schupp, Stockburger, et al., 2003) in response to the initial intact scene.

Behavioral results showed faster accumulation of evidence for old relative to new scenes, confirming that participants were actively using the colorful scene as a cue to guide their recognition processes later on during unfolding. In addition, replicating the behavioral results of Chapter 3, I found that emotional scenes were on average associated with a slower accumulation of evidence compared to neutral scenes. However, this effect was similar for old and new scenes; thus, it did not interact with the differential memory status of the stimuli. Both negativity bias (Cacioppo & Ito, 1999) and positivity offset (Ito, et al., 1998) were still assumed to mediate these behavioral effects: motivationally relevant scenes were apparently explored more thoroughly compared to neutral scenes, suggesting that proactive guesses were somehow delayed for the former compared to the latter. Moreover, the fact that this strong motivational effect on visual scene recognition was similar for old and new scenes suggested that memory and emotion likely exerted independent effects on these perceptual mechanisms.

The estimated neural generators of the dominant voltage maps recorded on the scalp (see Chapter 4 for details) supported the assumption of distinct brain networks active during the processing of old vs. new scenes. More specifically, recognition of old scenes was primarily subtended by ventral occipito-temporal regions, including the parahippocampal gyrus (PHG) and the fusiform gyrus (FG). Activity in these areas increased linearly as a function of accumulation of evidence, likely reflecting the processing of contextual cues in the PHG (presumably based on LSF input; Bar, 2004; Bar & Aminoff, 2003; Bar, et al., 2008) as well as invariant object features or maybe viewpoint information in the FG (Epstein, et al., 1999; Malach, et al., 1995; Pourtois, et al., 2010a). Importantly, activity in these two ventral occipito-temporal regions was reliably modulated by the emotional content of the (old) scenes. Emotion primarily interfered with the normal accumulation of contextual information in these brain areas, whereas it speeded up this process when additional perceptual evidence was aggregated. These results are consistent with an increasingly larger literature pointing at the pivotal role of contextual factors (broadly defined) in emotion perception and recognition (Aviezer, et al., 2008; Barrett, et al., 2011; Carroll & Russell, 1996), such as stimulus-based (e.g., body posture, voice), perceiver-based (subjects' intact access to conceptual knowledge about emotion), and culture-based contextual effects (e.g., cultural differences in the encoding or the understanding of facial expressions) (Barrett, et al., 2011).

The processing of new scenes, on the other hand, was mainly subtended by more dorsomedial regions, including the dACC and the insula. With regards to the dACC, activity in this region was at baseline until one image before recognition, when it substantially increased. However, no difference was observed between emotional and neutral scenes. This pattern of activity during accumulation of evidence, consistent with the involvement of this region in higher-order decision making processes (Bush, et al., 2002; Ridderinkhof, et al., 2004; Seo & Lee, 2007), not only corroborated previous (fMRI) results (Carlson, et al., 2006; Ploran, et al., 2007; Ploran, et al., 2011; Wheeler, et al., 2008), but also replicated the findings reported in Chapter 3. Noteworthy, in the dACC, roughly the same pattern of activation during unfolding was observed regardless of the specific task demands, i.e., a delayed match-to-sample task (Chapter 4) vs. an animacy judgment task (Chapter 3). Altogether, these findings provide further evidence for the task-aspecificity of this perceptual decision making mechanism within the dACC (Ridderinkhof, et al., 2004; Ullsperger, et al., 2004). Specifically, the dACC likely monitors the response outcome that allows the selection of the best response at a given moment in time and in light of the current (un)certainty and/or volatility (Behrens, Woolrich, Walton, & Rushworth, 2007), regardless of the specific task at hand (i.e., either perception-based or memory-based). As already pointed out in Chapter 3, this categorical increase in dACC activity close to the actual moment of recognition might reflect perceptual decision making processes based on uncertainty and/or conflict. Because participants were encouraged to perform the task before the end of the six-step unfolding sequence (and they apparently did so, as confirmed by the behavioral results), these conditions might foster the experience of a mild conflict at the time of (overt) decision making, namely the urge to take a decision despite insufficient or incomplete perceptual evidence (Anderson, et al., 2009; Sohn, et al., 2007). The results showed, however, that this effect was not enhanced when the scenes had a clear emotional content (but see Kanske & Kotz, 2011a, 2011b). Complementary source localization analyses also revealed categorical effects in the insula in response to new scenes. Interestingly, the emotional content of the scenes modulated the reconstructed activity in this region: increased amplitude was observed between one image before and actual recognition for neutral scenes, a categorical effect which was shifted earlier in time (i.e., between two and one image before recognition) during the processing of emotional scenes. The insula has consistently been found active in response to salient exogenous sensory stimuli (Craig, 2009; Seeley, et al., 2007) as well as challenging and uncertain perceptual tasks (Grinband, et al., 2006; Philiastides & Sajda, 2007). As proposed above for the dACC, during the progressive unfolding task the level of uncertainty (and the resulting conflict among multiple interpretations for the visual percept) abruptly increased close to recognition, when a response selection (old vs. new scene) had to be made. Alternatively, the emotional scenes, given their enhanced motivational salience, could trigger a specific activation in the insula early on during unfolding, with the aim to monitor whether the impoverished content was potentially threatening and, in turn, may have led to a change in homeostasis (Sterzer & Kleinschmidt, 2010).

More generally, the results of Chapter 4 shed new light on the impact and relative contribution of memory and emotion factors on proactive processes during visual scene recognition. However, the question arose as to whether some of these effects (in particular at the behavioral level) may not be explained by uncontrolled non-emotional factors that would eventually create asymmetries between the different experimental conditions (neutral, pleasant or unpleasant). Among them, picture complexity and familiarity appeared especially important to consider, given earlier studies showing a strong modulation of these variables on electrical brain activity recorded in response to emotional stimuli (see Chapter 1 for details). Therefore, it seemed critical to rule out the possibility that the effects observed in Chapters 3-4 between the three stimulus categories were (at least partly) explained by uncontrolled variations along these specific non-emotional dimensions.

2.3. Chapter 5: Emotion-specific influences on proactive visual scene recognition

In Chapter 5 I asked two independent samples of participants to rate either the familiarity or picture complexity of each and every scene used in Chapter 4. Results of the rating study showed: (i) lower familiarity scores for unpleasant relative to neutral and pleasant scenes, in line with the “mere-repeated-exposure” effect (i.e., participants usually tend to develop a preference for objects deemed familiar and, as a consequence, unpleasant objects are typically considered as less familiar; Monahan, et al., 2000; Zajonc, 1968, 2001); (ii) higher picture complexity scores for unpleasant compared to neutral and pleasant scenes, consistent with the observation that objects more difficult to process (i.e., having a lower “fluency”) are more likely seen as unpleasant and thus are somehow devaluated (Winkielman & Cacioppo, 2001; Winkielman, et al., 2003).

Afterwards, I used the behavioral data collected during the progressive unfolding task described in Chapter 4 and I included these familiarity and picture complexity ratings as predictors (besides the factors memory and emotion) in a proportional odds model (see

Chapter 5 for details), to verify whether the effects of memory and emotion during proactive scene recognition found in Chapter 4 were somehow confounded by trial-to-trial fluctuations along these non-emotional dimensions. Results revealed that the previously seen shift in the accumulation of perceptual evidence for unpleasant compared to neutral scenes was partly confounded by changes in picture complexity and familiarity: when controlling for these two factors, in fact, this shift was no longer significant. Hence, familiarity and picture complexity could in part explain the pattern of behavioral results for unpleasant scenes reported in Chapter 4. Conversely, a delayed accumulation of evidence for pleasant compared to neutral scenes was still evident in this control analysis, with no differences between old and new scenes. These findings corroborated the assumption that positive information (and not familiarity or picture complexity exclusively) truly promotes exploratory behavior in non-threatening situations, leading to a slower or delayed accumulation of evidence process relative to neutral or unpleasant scenes. More generally, these results fit nicely with many earlier studies in the literature showing an independent contribution of emotion (besides other non-emotional factors) on visual recognition processes (Bradley, et al., 2007; Cano, Class, & Polich, 2009; Carretié, et al., 2007; Carretié, et al., 2004; De Cesarei & Codispoti, 2006; Delplanque, et al., 2007; Junghöfer, et al., 2001; Müller, Andersen, & Keil, 2008; Wiens, et al., 2011).

3. Emotion influences proactive processes leading to visual scene recognition at multiple levels: an integrative account

3.1. Positivity offset effects during visual scene perception

The main behavioral outcome of Chapters 3 and 4 is the prolonged accumulation of evidence for emotional compared to neutral scenes, with pleasant scenes recognized even later than unpleasant ones. These results were initially interpreted as reflecting both negativity bias (Cacioppo & Ito, 1999) and positivity offset (Ito, et al., 1998) effects, since it was not possible to disambiguate between these two concurrent phenomena. However, results of Chapter 5 additionally showed that this general shift in the accumulation of perceptual evidence for pleasant scenes could not be explained by uncontrolled variations along the familiarity and/or picture complexity dimensions. Interestingly, lower familiarity and higher picture complexity scores were obtained for unpleasant scenes, with pleasant and neutral scenes showing comparable ratings. Therefore, one would expect changes in the accumulation of perceptual evidence for unpleasant scenes. Instead, when controlling for non-emotional factors, pleasant scenes were still associated with a distinctive delayed accumulation of evidence relative to the two other conditions, clearly demonstrating a valence-specific effect of emotion operating during proactive processes prior to visual scene recognition[1]. Therefore, based on the findings reported in Chapter 5, I argued that a positivity offset effect appeared to be the most plausible explanation to account for the pattern of behavioral results obtained in Chapters 3-5 using the progressive unfolding task. The general tendency of healthy adult individuals to engage in exploratory behavior in non-threatening events or situations can easily explain the specific delayed recognition observed for pleasant scenes: participants in Chapters 3-5, tested in a quiet room and with the perspective of a monetary reward, were undergraduates with no history of neurologic or psychiatric disorders and with anxiety levels in the normal range (as confirmed by the scores of the STAI-Trait scale; Van der Ploeg, et al., 1979). This situation likely promoted a careful analysis of the visual scenes gradually revealed during the task, and this process was arguably even more thorough during the unfolding of scenes with a clear hedonic content (e.g., babies, food, romance) likely matching the goals and needs of these participants.

As a next step, one could for example assess the magnitude of positivity offset and negativity bias effects at a single-subject level prior to testing, with the aim to verify (and even predict), during a second step, whether these parameters can influence recognition speed of neutral, pleasant, and unpleasant scenes gradually revealed during a progressive unfolding task (similar to the task used in Chapter 4). A valuable way to model these concurrent activation functions is provided by regression parameters, with the positivity offset representing a higher intercept value for the positive activation function and the negativity bias representing a steeper slope for the negative activation function (Cacioppo, et al., 1997; Ito & Cacioppo, 2005; see also Figure 2). Using this framework, previous studies (e.g., Ito, et al., 1998) already assessed the strength of participants' reactions to positive and negative IAPS scenes by including normative ratings of arousal, valence, and dominance as predictors in separate regression analyses, and found higher intercept value predicting positivity (positivity offset) and a steeper slope predicting negativity (negativity bias). Translated to the

---

[1] Interestingly, this specific delay in recognition for pleasant scenes seems to be independent from the ease or difficulty to generate likely candidates regarding the scene content. This delay was in fact observed for new and old scenes alike, meaning that the failure to confirm a given set of predictions (i.e., new scene) did not interfere with the propensity to dwell longer on these positive scenes.

progressive unfolding task, it could be hypothesized that higher intercept values would predict slower recognition times for pleasant scenes, whereas higher slope values would predict faster recognition times for unpleasant scenes, hence revealing a latency shift during proactive scene recognition depending on the actual underlying psychological or motivational process.

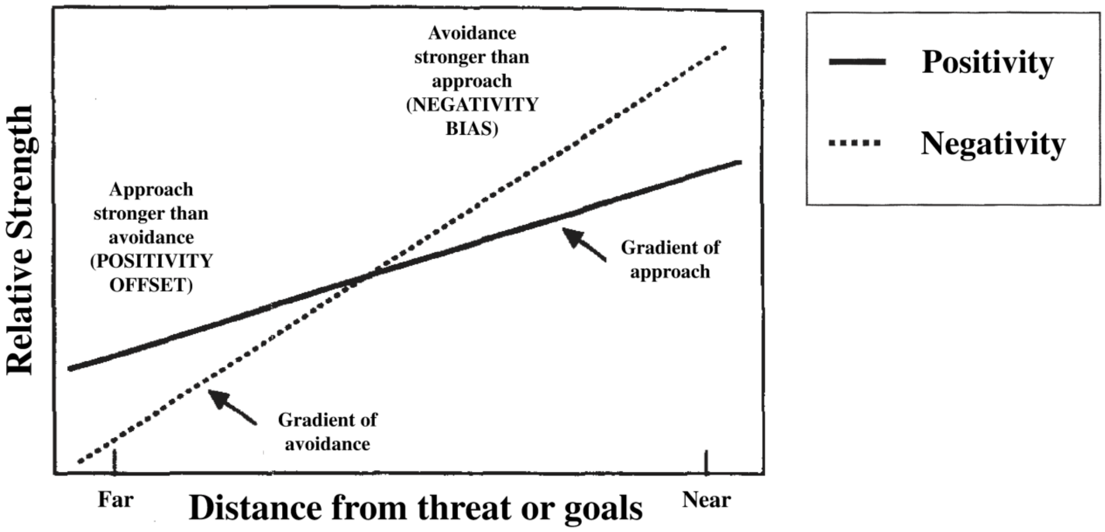


Figure 2. Visual representation of positivity offset and negativity bias activation functions. At far distances from threat (or, more generally, goals), positivity (solid line) is stronger than negativity (dotted line), as reflected by higher intercept values. As a result, under such conditions the tendency to approach novel situations or stimuli is stronger than avoidance. Conversely, when the motivation to pursue a goal is pressing or threat is imminent, negativity is stronger (as represented by a steeper slope) than positivity, leading in turn to avoidance. (Adapted from Cacioppo, et al., 1997).

3.2. Emotion as a top-down predictive coding mechanism during visual scene perception

As described above, the main behavioral outcome of Chapters 3-4 is the prolonged accumulation of evidence for emotional compared to neutral scenes, especially for pleasant ones. Results of Chapter 5 additionally showed that this general shift in the accumulation of perceptual evidence for pleasant scenes could not be explained by uncontrolled variations along the familiarity and/or picture complexity dimensions. This effect is in line with a positivity offset, namely the general predisposition or motivation of healthy adult participants to engage in the exploration of non-threatening events or situations, because the long-term benefits to explore might turn out to be higher for the organism (e.g., mapping environments, learning new contingencies, satisfying current goals and needs) than the gain associated with safety when fleeing (Cacioppo & Gardner, 1999; Norris, et al., 2010). Functionally, I proposed that this prolonged accumulation of evidence for pleasant compared to neutral scenes may result from an increased number of iterations made by the perceptual system between predictions and the bottom-up sensory input (see Chapter 5). These recursive loops are meant to eventually minimize prediction errors and identify the most likely interpretation of the stimulus content, as subsumed by modern models of predictive coding of human cognition (Friston, 2002, 2005; Rao & Ballard, 1999; Serences, 2008; see also Figure 3). But what may be a plausible mechanistic account at the base of this theoretical framework?

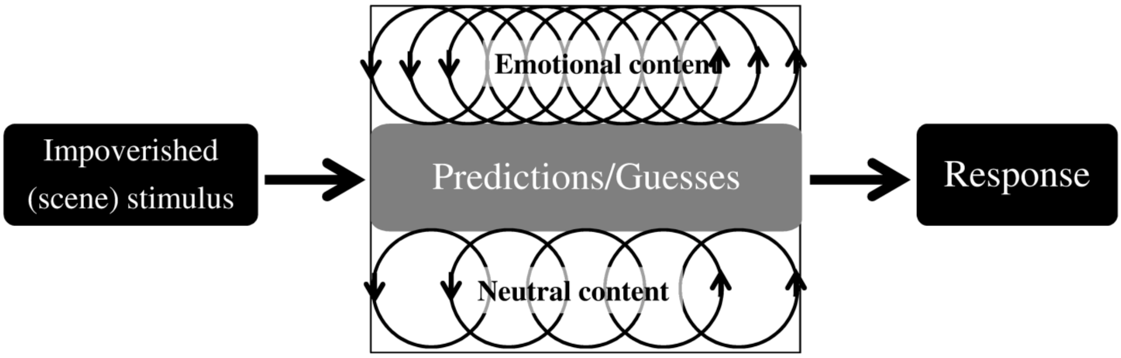


Figure 3. Graphical representation of the critical components involved in proactive processes during visual scene recognition, as well as their sensitivity to emotional factors. The results of Chapter 3-5 revealed that emotion (particularly pleasant) scenes were recognized later compared to neutral scenes, presumably due to prolonged exploration originated by the highly reinforcing hedonic value of this class of stimuli. At the functional level, this delayed recognition for emotional scenes might result from a slower or increased number of iterations between predictions and the parametric increase of visual input as a function of image level during the unfolding procedure, with the final aim to eventually minimize prediction error and to identify the most likely interpretation of the stimulus content.

As argued in Chapter 1, the rapid extraction of LSF information via dedicated magnocellular pathways may serve to quickly generate coarse predictions regarding the most probable object in the visual scene (Bar, 2003), as well as diagnostic contextual cues (Bar, 2004). This process would later be refined by the (slower) extraction of HSF information conveyed by parvocellular pathways (Bar, 2007). A recent update of the original framework has been proposed, in which emotion is included as a relevant modulatory factor (Barrett & Bar, 2009). Here, affective properties are thought to be intrinsically embedded in the prediction originated by the processing of the gist of the scene, given that specific medial frontal brain areas (including the ACC and the insula) active during recognition processes directly project to autonomic and endocrine output centers in the hypothalamus, midbrain, and brainstem (Barrett & Bar, 2009). Therefore, this model implies that internal affective states of the organism might potentially be one feature of the set of predictions used to facilitate recognition (Bar, 2004; Bar & Aminoff, 2003; Barrett & Bar, 2009).

Based on the new results obtained in Chapters 3-5, I propose a revision of this model that enables to better account for modulations of visual scene recognition by emotion and motivation (Bar, 2003, 2004, 2007, 2009; see also Figure 3). The electrophysiological findings reported in Chapters 3-4 show that activity in occipito-temporal regions, such as the PCC/PHG and FG, was modulated by the emotional content of the stimulus. Given the involvement of the PCC/PHG complex in the processing of contextual information (Bar, 2004; Bar, et al., 2008; Fenske, et al., 2006), the results of Chapters 3-4 show that the extraction of diagnostic contextual cues during early stages of recognition is influenced by the motivational significance of these cues, hence the processing of contextual cues during proactive vision is shaped by emotion (Barrett & Kensinger, 2010; Barrett, et al., 2011; Righart & de Gelder, 2006, 2008a, 2008b). Interestingly, such modulation occurred early in the unfolding process relative to the moment of recognition, suggesting that emotional factors influenced the processing of contextual cues in conditions of impoverished visual perception, when presumably guesses or predictions were still not specific yet. However, this hypothesis requires further empirical validation, for example by crossing fMRI and EEG data (Kanske & Kotz, 2011a; Peyrin, et al., 2010). FG activation, on the other hand, presumably reflected the processing of invariant object features or viewpoint information (Epstein, et al., 1999; James, et al., 2000; Malach, et al., 1995; Pourtois, et al., 2010a) or, alternatively, subordinate level judgments (as compared to basic-level decisions) required to shape and refine online the early guesses or predictions regarding the most probable stimulus content progressively revealed to participants (Gauthier, Anderson, Tarr, Skudlarski, & Gore, 1997; Palmeri & Gauthier, 2004; Tarr & Gauthier, 2000). It could be hypothesized that, in Chapter 4, the gradual accumulation of perceptual evidence served to perform a coarse or categorical judgment (i.e., old-new), thereby modulating activity in FG (Kosslyn, Thompson, Gitelman, & Alpert, 1998). These hypotheses alike, however, need further empirical support.

Activity in medial frontal brain areas and the insula was also reliably modulated as a function of the emotional content of the scenes. Importantly, this categorical modulation in the insula occurred close to the time of overt recognition of new scecnes, as opposed to earlier effects in the PHG and FG for old scenes. Therefore, it could be hypothesized that activity in this brain region might reflect verification processes aiming at assessing whether the stimulus content, or more specifically the most plausible candidate (presumably based on a selective and earlier effect occurring in the OFC via magnocellular vision), is potentially threatening (or, more generally, motivationally significant) and, in turn, may lead to a change in homeostasis (Sterzer & Kleinschmidt, 2010). Alternatively, this region might work in concert with the dACC to resolve the conflict arising because a response needed to be produced in conditions of impoverished visual perception (Ploran, et al., 2007; Ploran, et al., 2011; Sterzer & Kleinschmidt, 2010). Future studies are needed to elucidate whether conflict detection *per se* or performance monitoring account for the observed categorical effects found in these regions during proactive vision (Ito, et al., 2003; Ridderinkhof, et al., 2004; Wheeler, et al., 2008).

It should be pointed out that source localization analyses of the electrical brain activity recorded on the scalp during the progressive unfolding task (Chapters 3-4) did not show any differential activation in the OFC, an area thought to be involved (as early as 130 ms after stimulus onset, 50 ms before any activity recorded from object-selective regions; Bar, et al., 2006) in the generation of the most likely interpretations of the main object or scene (Bar, 2007). However, as a matter of fact, this OFC activation was observed in tasks where the visual stimuli depicted scenes providing a clear segregation between the main object in the foreground and the background (see, for example, the stimuli used in Bar, et al., 2006), whereas more complex (neutral and emotional) IAPS scenes were employed in Chapters 3-5. Alternatively, the main time window of interest selected in Chapters 3-4 for the analysis of the ERP data may have overlooked a selective contribution from the OFC, bearing in mind that ventral or lateral prefrontal regions are notoriously hard to identify with EEG source localization methods (Kringelbach, 2005; but see Dhar, et al., 2011), unlike MEG (Bar, et al., 2006; Gamond, et al., 2011). However, the selection of this time interval was based, following standard practice, on the main outcome of the spatiotemporal clustering analysis showing different topographical maps (denoting reliable changes in the underlying neural networks) 280-360 ms (Chapter 3) and 402-522 ms (Chapter 4) post stimulus onset.

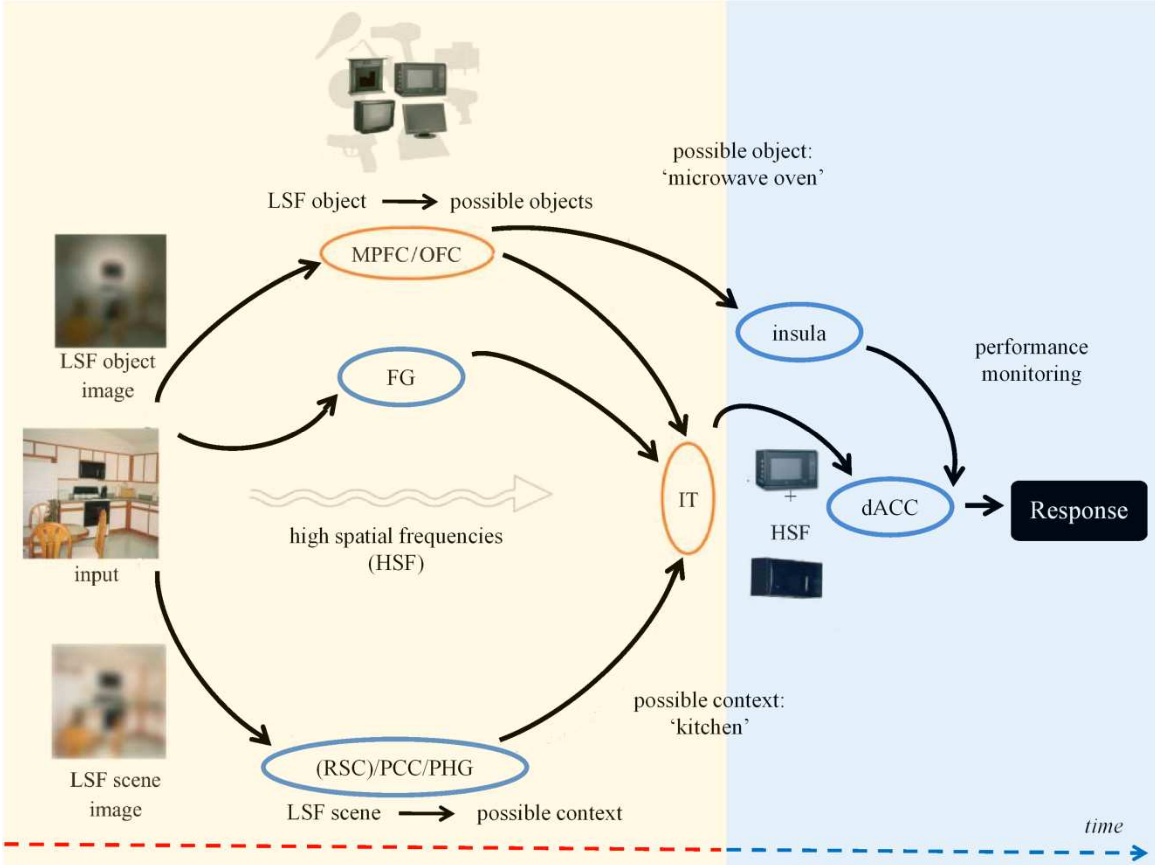


Figure 4. Adaptation of Moshe Bar (2009b)'s neuroanatomical model of visual object and scene recognition, taking into account the new EEG results reported in Chapters 3-4. In parallel to the bottom-up transfer of the image details conveyed by high spatial frequency information (HSF) along the visual pathways (undulated arrow in the centre), low spatial frequency (LSF) information is quickly projected from early visual areas to ventral occipito-temporal and medial prefrontal regions, possibly via magnocellular pathways. This coarse but rapid information is sufficient to generate "early guesses" about the most likely identity of the main objects in the scene and diagnostic contextual cues useful for recognition. These predictions are validated and refined in infero-temporal regions with the gradual arrival of HSF input. The EEG results of Chapters 3-4 further specify that the motivational relevance of complex visual scenes has an impact on recognition at different latencies in different regions (circled in blue). Specifically, early modulation by emotion was found in ventral areas such as the PHG and FG, involved in the analysis of contextual and object (or viewpoint)-dependent features, respectively. On the other hand, categorical changes in neural activity in the insula and dACC were found close to the moment of overt recognition, likely indicating the involvement of these areas in conflict and uncertainty resolution, as well as response monitoring. The red and blue background indicate the differential patterns of activation observed at the electrophysiological level, namely a linear accumulation of evidence prior to recognition and a categorical increase close to overt recognition, respectively. MPFC: medial prefrontal cortex; OFC: orbitofrontal cortex; FG: fusiform gyrus; RSC: retrosplenial complex; PCC: posterior cingulate cortex; PHG: parahippocampal gyrus; IT: infero-temporal cortex; dACC: dorsal anterior cingulate cortex. The arrows are unidirectional in the figure to emphasize the proposed flow of visual information, although all these connections are bidirectional in nature. (Adapted from Bar, 2009b).

3.3. Modeling the contribution of emotional factors during perceptual decision making

Neurobiological accounts and computational models of perceptual decision making postulate that the brain is in essence a proactive machine (i.e., constructivist view), in the sense that it progressively gathers perceptual evidence in favor of one out of two (or more) alternatives until a certain threshold is surpassed, thereby resulting in a decision and producing a (motor) response (Gold & Shadlen, 2007; Heekeren, et al., 2004; Heekeren, et al., 2008). As described in Chapter 1, start points (the initial amount of evidence in favor of each response, presumably influenced by prior knowledge) and drift rates (the different average speed of evidence accumulation for each response) are important parameters to model and explain differences in speed of recognition (Brown & Heathcote, 2008; Usher & McClelland, 2001). These models, however, do not explicitly include emotion as a potential source of variability. It could be hypothesized, in fact, that emotional factors would modulate the rate of accumulation of information, in that smaller drift rates would lead to slower responses (Philiastides, et al., 2011). Alternatively, decision-related processes may be slowed down when diagnostic emotional features are detected and, as a consequence, it would take longer for the system to reach the decision threshold (Philiastides, et al., 2011; Ratcliff, 1978; Ratcliff & McKoon, 2008). At this stage, the new results reported in Chapters 3-4 do not fully disambiguate whether emotion influences accumulation of evidence processes (i.e., the number of iterations between predictions and errors), decision-related processes, or both. However, the electrophysiological results reported in Chapter 4 provide first hints on the response profile of different brain systems performing complex visual computations during proactive emotional scene recognition. Whereas the response profile of some of these brain regions is consistent with a genuine accumulation of evidence (PHG and FG), others (dACC and insula) appear compatible with subsequent decision-related processes (see also Ho, et al., 2012; Ho, Brown, & Serences, 2009). Hence, the new findings reported in Chapter 4 suggest that emotion can exert complex modulatory influences at each of these two levels (see Figure 5).

4. Conclusions

4.1. *Future perspectives*

The results of Chapter 2 suggest that motivationally relevant stimuli do not reflexively capture attention when competing with neutral stimuli for attention selection. Presumably, other non-emotional cues are prioritized by the visual system in order to resolve the TOJ task in the most efficient way. Among these, I proposed that the systematic variation in the SOA between the two competing face onsets created a compelling impression of apparent motion on the screen (Hikosaka, et al., 1993). However, this tentative post hoc explanation needs to be put to the test in future studies, given that I did not assess directly whether apparent motion was indeed the (non-emotional) cue actively used by participants and able to somehow override effects of (negative) emotion during early attention selection. For example, stimulus parameters could be altered in such a way that any impression of apparent motion would be made impossible or alleviated, e.g., by adding a moving random-dot pattern in the background of the display. Given that motion speeds up attentional orienting when clearly informative (Hillstrom & Yantis, 1994; Raymond, 2000), disrupting any apparent motion cues would normally promote the use of other strategies to perform the task, in this case the analysis of the emotional content of the face stimuli.

Furthermore, the TOJ and progressive unfolding tasks devised and employed in this thesis could be used in the context of experimental psychopathology in order to further explore and characterize possible biases in early attention and perceptual processes in specific populations. More specifically, if prior entry effects for threatening vs. neutral faces would be observed when testing high anxious, dysphoric, or even clinically depressed patients with the experimental design used in Chapter 2, this would add to the existing literature and provide further evidence for early attentional biases towards threat-related material in these individuals (Bar-Haim, et al., 2007; Cisler & Koster, 2010; Davidson, et al., 2002; Fox, 1993; MacLeod, et al., 1986; Mogg & Bradley, 1998; Price & Drevets, 2012). Because the TOJ task, unlike other spatial cueing or orienting tasks, provides a “pure” measure of early attentional capture (i.e., not confounded by other attention processes, including shifting or disengagement), such an outcome would unequivocally show that early attention processes are indeed influenced by motivational and situational factors, besides more classical structural components or effects (task set or exogenous effects).

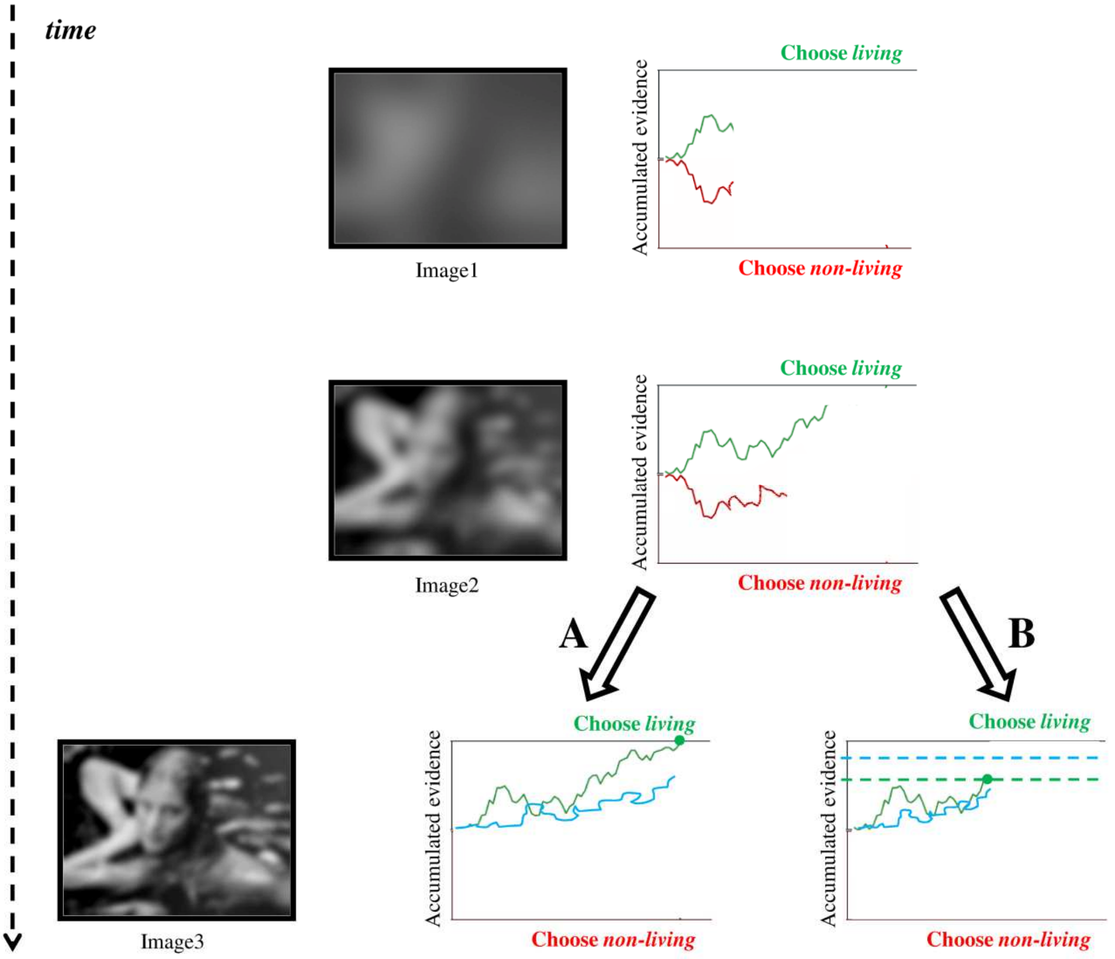


Figure 5. According to accumulator models of perceptual decision making, perceptual evidence in favor of one or more alternative candidate representations of a (visual) stimulus content is accrued over time until a threshold is surpassed, thereby leading to a decision and (potentially) eliciting a response. During the progressive unfolding task, the accumulation of visual information in favor of either “living” (solid green line) or “non-living” (solid red line) responses would occur only when the ambiguity of the visual input is (at least partially) resolved, i.e., when diagnostic features of the scene (in this example, the detection of two crossed arms below a human face) slowly emerged from the picture series as a consequence of the parametric increase of HSF and LSF information. Importantly, when participants are presented with emotional scenes, recognition is delayed compared to neutral scenes at the behavioral level. Two alternative mechanisms are proposed to mediate this effect (blue solid line): (A) slower accumulation of evidence when the scene has an emotional content; (B) a lower response threshold for neutral (dashed blue line) compared to emotional (dashed green line) scenes.

The progressive unfolding task, on the other hand, could prove useful in assessing the alteration of a positivity offset effect in high anxious or depressed individuals. Whereas an increased sensitivity to negativity is shared between these two psychopathological conditions (with an exacerbated negativity bias), they could however be partly dissociated regarding the residual sensitivity to positive emotions that should normally be absent in depressed (but not anxious) individuals (Conklin, et al., 2009; Shook, et al., 2007). More specifically, it has been proposed that, with equal levels of (high) negativity bias, anxiety is the more likely outcome if positivity is intact, whereas depression may arise in people with lower positivity (Norris, et al., 2010; Watson, Clark, & Tellegen, 1988; Watson & Tellegen, 1985). As proposed here above (see paragraph 3.1), parameters of positivity and negativity can be calculated for each participant separately via linear regressions: intercept values indicate positivity offset, whereas slope values refer to negativity bias (Ito, et al., 1998; Ito & Cacioppo, 2005). If this hypothesis holds true, a first prediction would be that, with comparable slope values, lower intercept values would be observed for depressed relative to anxious individuals. In a second phase, intercept and slope values could be included as predictors in separate linear regressions to investigate whether they could account for recognition speed during the progressive unfolding task (see Chapters 3-5): hypothetically, higher intercept values could in fact predict slower recognition times for pleasant scenes. Hence, the use of computational modeling in the context of experimental psychopathology might help better tease apart the actual distinctive properties vs. shared components of different psychopathological conditions, here with a focus on depression and anxiety.

4.2. Summary and conclusion

The new results reported in this dissertation show that emotion can bias visual perception at multiple levels and through different brain mechanisms. In Chapter 2, I did not find evidence for an “automatic” effect of emotion on attention selection during the temporal processing of competing onsets. Based on these results, I concluded that (negative) emotional stimuli do not “magically” capture attention (Pourtois, et al., 2012), but their early preferential processing can somehow be bypassed when the perceptual system can exploit other low-level cues in order to guide attention allocation in the environment, or the current task set fosters the use of these non-emotional cues (see also Todd, et al., 2012). I proposed a new model assuming multiple and separate attentional components (see Figure 1) that seems able to account for these new findings without challenging the notion of specific emotional attention brain processes, although further empirical validation is needed. By contrast, in Chapters 3-5 I found clear evidence for a modulation of visual scene recognition by their emotional content, when proactive processes and predictions were promoted and participants made active use of them in order to guide their behavior. In this condition, I argued that motivational drives (e.g., positivity offset), rather than basic emotional features or dimensions (e.g., arousal or valence), actually influenced the rate at which perceptual information was accumulated by the visual system and compared with existing representations (stored in memory) before a decision was made regarding the most likely identity of the scene. The new findings reported in Chapters 3-5 show that multiple brain systems operate in parallel in order to achieve these computations, with several traceable contributions of emotion. Remarkably, these effects took place before overt recognition, suggesting that emotion does not only "reactively" alter visual perception, but it can proactively exert strong influences on the guidance and selection of (diagnostic) perceptual cues eventually used by the visual system to take a decision about the content of the visual percept. As such, these results lend support to pioneering accounts proposed in the literature (Bruner & Postman, 1949; Niemi & Näätänen, 1981), as well as more recent models (Bar, 2003; Gosselin & Schyns, 2003) that posit that the phenomenology of visual perception is eventually based not only on the efficient processing capabilities of encapsulated modules dealing with a bottom-up analysis of the retinal input, but prior knowledge, expectations and motivation profoundly shape and influence these processes with the aim to construct and constantly update the meaning of our external visual world.

Taken together, these new results provide evidence for the assumption that emotion is not a simple by-product of visual perception, but instead it proactively shapes it. More generally, these finding lend support to the notion that the human visual system is a remarkable proactive machine, which is able to "effortlessly" perform complex computations that eventually enable us to make sense of our environment, within a few hundreds of milliseconds following light changes captured by the retina.

NEDERLANDSE SAMENVATTING

Het belangrijkste doel van dit proefschrift was om de invloed van emotie (wat voornamelijk geoperationaliseerd werd als veranderingen in valentie en opwindingswaarden van op voorhand geselecteerde gezichten of complexe visuele scenes) op visuele perceptie. Ik ging meer bepaald op zoek naar de mate waarin emotionele factoren perceptuele verwerking beïnvloedde tijdens (gezichts of scene) herkenning op een ‘bottom-up’, reactieve vs. een ‘top-down’, proactieve manier. Om deze vragen te beantwoordden, heb ik eerst naar ‘bottom-up’ effecten van emotie gekeken op de temporele perceptie van competitieve gebeurtenissen d.m.v. specifieke emotionele aandachtsmechanismen. Van deze mechanismen wordt gedacht dat ze voorrang geven aan de vroege detectie van motivationeel opvallende stimuli (in dit proefschrift kwade of angstige gezichten), vergeleken met neutrale stimuli. Vervolgens heb ik onderzocht hoe emotie visuele perceptie kon beïnvloeden via ‘top-down’ mechanismen die hier werkzaam zijn als proactieve gokken of verwachtingen die online gegenereerd werden over de meest waarschijnlijke identiteit van een visuele stimulus wanneer de retinale input verminderd is. Om deze vragen te beantwoorden werden twee verschillende experimentele paradigma’s ontwikkeld en gebruikt: Een standaard temporele volgorde beoordelingtaak (Hoofdstuk 2) en een progressief ontwikkelingsparadigma (Hoofdstuk 3-5).

De nieuwe bevindingen die werden gerapporteerd in dit proefschrift tonen dat emotie de visuele perceptie kan beïnvloeden op verschillende niveau’s en via verschillende hersenmechanismen. De resultaten m.b.t. de eerste vraag tonen dat motivationeel significante stimuli de aandacht niet trekken op een “magische” wijze en vervolgend de temporele perceptie beïnvloeden, maar andere factoren (vb. aparte lagere niveau perceptuele ‘cues’ zoals beweging of de actuele taak set) zouden kunnen wedijveren met emotionele (hogere orde) ‘cues’ om geselecteerd te worden via ‘bottom-up’ aandacht, om zo uiteindelijk tegen te gaan dat emotionele cues systematisch en automatische de aandacht trekken (Hoofdstuk 2).

M.b.t. de tweede vraag tonen de resultaten dat de generatie en het gebruik van specifieke verwachtingen m.b.t. de meest waarschijnlijke inhoud van complexe visuele scenes – die progressief worden getoond op verschillende opeenvolgende stappen – niet immuun zijn voor emotionele factoren. Deze resultaten steunen de assumptie dat emotie, naast andere componenten (vb. aandacht of geheugen), proactief de selectie van diagnostische perceptuele ‘cues’ kan beïnvloeden die uiteindelijk gebruikt worden door het visueel systeem om de betekenis van de visuele input af te leiden (Hoofdstuk 3-5). Ik argumenteer dat motivationele krachten (vb. positiviteits ‘offset’), meer dan basis emotionele kenmerken of dimensies (vb. opwinding of valentie), de snelheid eigenlijk beïnvloedde waarmee perceptuele informatie werd opgebouwd door het visuele systeem en werd vergeleken met bestaande voorstellingen (opgeslagen in het geheugen), voordat er een beslissing werd gemaakt m.b.t. de eigenlijke identiteit van de scene. Elektrofysiologische resultaten toonden verder dat verschillende hersensystemen parallel werken om tot deze berekeningen te komen, met verschillende te achterhalen contributies van emotie. Opvallend was ook dat deze effecten plaatsvonden voor de openlijke herkenning. Dit wekt de suggestie dat emotie niet enkel “reactief” de visuele perceptie verandert, maar ook “proactief” een sterke invloed kan uitoefenen op het leiden en selectie van (diagnostische) perceptuele cues die uiteindelijk gebruikt worden door het visueel systeem om een beslissing te nemen over de inhoud van het visuele percept.

We kunnen dus zeggen dat deze nieuwe resultaten evidentie geven voor de aanname dat emotie niet simpelweg een bijproduct is van visuele perceptie, maar dat het vormgeeft aan de visuele perceptie. Meer algemeen kunnen we concluderen dat deze bevindingen de notie steunen dat het menselijke visuele systeem een opmerkelijke proactieve machine is, die “zonder moeite” complexe berekeningen kan uitvoeren die ons uiteindelijk toestaan om onze omgeving zin te geven, en dit binnen enkele honderden milliseconden volgend op lichtveranderingen die opgevangen worden door de retina.

References

Adams, R. B., Jr., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science, 300*(5625), 1536.

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*. New York: Wiley.

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron, 21*(2), 373-383.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723.

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience, 25*, 339-379.

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology-General, 133*(3), 323-338.

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus Predictability Reduces Responses in Primary Visual Cortex. *Journal of Neuroscience, 30*(8), 2960-2966.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science, 15*(2), 106-111.

Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience, 118*(4), 1099-1120.

Aminoff, E., Gronau, N., & Bar, M. (2007). The parahippocampal cortex mediates spatial and nonspatial associations. *Cerebral Cortex, 17*(7), 1493-1503.

Anderson (2005). Affective influences on the attentional dynamics supporting awareness. *J Exp Psychol Gen, 134*(2), 258-281.

Anderson, Anderson, J. F., Ferris, J. L., Fincham, J. M., & Jung, K. J. (2009). Lateral inferior prefrontal cortex and anterior cingulate cortex are engaged at different stages in the solution of insight problems. *Proceedings of the National Academy of Sciences of the United States of America, 106*(26), 10799-10804.

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*(13), 5627-5633.

Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature, 411*(6835), 305-309.

Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia, 40*(7), 817-826.

Arnell, K. M., Killman, K. V., & Fijavz, D. (2007). Blinded by emotion: target misses follow attention capture by arousing distractors in RSVP. *Emotion, 7*(3), 465-477.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724-732.

Bacon, W. F., & Egeth, H. E. (1994). Overriding Stimulus-Driven Attentional Capture. *Perception & Psychophysics, 55*(5), 485-496.

Banich, M. T., Milham, M. P., Atchley, R. A., Cohen, N. J., Webb, A., Wszalek, T., et al. (2000). Prefrontal regions play a predominant role in imposing an attentional 'set': evidence from fMRI. *Cognitive Brain Research, 10*(1-2), 1-9.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1-24.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience, 15*(4), 600-609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*(8), 617-629.

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn Sci, 11*(7), 280-289.

Bar, M. (2009a). A cognitive neuroscience hypothesis of mood and depression. *Trends in Cognitive Sciences, 13*(11), 456-463.

Bar, M. (2009b). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1235-1243.

Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron, 38*(2), 347-358.

Bar, M., Aminoff, E., Mason, M., & Fenske, M. (2007). The units of thought. *Hippocampus, 17*(6), 420-428.

Bar, M., Aminoff, E., & Schacter, D. L. (2008). Scenes unseen: The parahippocampal cortex intrinsically subserves contextual associations, not scenes or places per se. *Journal of Neuroscience, 28*(34), 8539-8544.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America, 103*(2), 449-454.

Barrett, L. F. (2006a). Are Emotions Natural Kinds? *Perspectives on Psychological Science, 1*(1), 28-58.

Barrett, L. F. (2006b). Valence is a basic building block of emotional life. *Journal of Research in Personality, 40*(1), 35-55.

Barrett, L. F., & Bar, M. (2009). See it with feeling: affective predictions during object perception. *Philos Trans R Soc Lond B Biol Sci, 364*(1521), 1325-1334.

Barrett, L. F., & Kensinger, E. A. (2010). Context Is Routinely Encoded During Emotion Perception. *Psychological Science, 21*(4), 595-599.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in Emotion Perception. *Current Directions in Psychological Science*, *20*(5), 286-290.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373-403.

Barsz, K. (1996). Accuracy of same different judgments of sequences of complex tones differing in tonal order under various levels of fundamental frequency range, listener training, and type of standard sequence. *Journal of the Acoustical Society of America, 99*(3), 1660-1669.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.

Becker, S. I., Horstmann, G., & Remington, R. W. (2011). Perceptual Grouping, Not Emotion, Accounts for Search Asymmetries With Schematic Faces. *Journal of Experimental Psychology-Human Perception and Performance, 37*(6), 1739-1757.

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience, 10*(9), 1214-1221.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*(6), 551-565.

Bernasconi, F., Grivel, J., Murray, M. M., & Spierer, L. (2010). Plastic brain mechanisms for attaining auditory temporal order judgment proficiency. *Neuroimage, 50*(3), 1271-1279.

Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review, 94*(2), 115-147.

Biederman, I., & Bar, M. (1999). One-shot viewpoint invariance in matching novel objects. *Vision Research, 39*(17), 2885-2899.

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences, 11*(7), 307-316.

Bishop, S. J. (2008). Neural mechanisms underlying selective attention to threat. *Molecular and Biophysical Mechanisms of Arousal, Alertness, and Attention, 1129*, 141-152.

Blair, M. R., Watson, M. R., Walshe, R. C., & Maj, F. (2009). Extremely Selective Attention: Eye-Tracking Studies of the Dynamic Allocation of Attention to Stimulus Features in Categorization. *Journal of Experimental Psychology-Learning Memory and Cognition, 35*(5), 1196-1206.

Bocanegra, B. R., & Zeelenberg, R. (2009). Emotion Improves and Impairs Early Vision. *Psychological Science, 20*(6), 707-713.

Bocanegra, B. R., & Zeelenberg, R. (2011a). Emotion-Induced Trade-Offs in Spatiotemporal Vision. *Journal of Experimental Psychology-General, 140*(2), 272-282.

Bocanegra, B. R., & Zeelenberg, R. (2011b). Emotional cues enhance the attentional effects on spatial and temporal resolution. *Psychonomic Bulletin & Review, 18*(6), 1071-1076.

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive Affective & Behavioral Neuroscience, 7*(4), 356-366.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652.

Boucher, J., & Osgood, C. E. (1969). The Pollyanna Hypothesis. *Journal of Verbal Learning and Verbal Behavior, 8*(1), 1-&.

Bower, G. H. (1981). Mood and Memory. *American Psychologist, 36*(2), 129-148.

Bradley, M. M. (2009). Natural selective attention: Orienting and emotion. *Psychophysiology, 46*(1), 1-11.

Bradley, M. M., Codispoti, M., Sabatinelli, D., & Lang, P. J. (2001). Emotion and motivation II: sex differences in picture processing. *Emotion, 1*(3), 300-319.

Bradley, M. M., Hamby, S., Low, A., & Lang, P. J. (2007). Brain potentials in perception: Picture complexity and emotional arousal. *Psychophysiology, 44*(3), 364-373.

Bradley, M. M., & Lang, P. J. (1994). Measuring Emotion: The Self-Assessment Mannequin and the Semantic Differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49-59.

Bradley, M. M., Lang, P. J., & Cuthbert, B. N. (1993). Emotion, Novelty, and the Startle Reflex: Habituation in Humans. *Behavioral Neuroscience, 107*(6), 970-980.

Brandeis, D., & Lehmann, D. (1986). Event-Related Potentials of the Brain and Cognitive-Processes: Approaches and Applications. *Neuropsychologia, 24*(1), 151-168.

Broadbent, D. E. (1957). A Mechanical Model for Human Attention and Immediate Memory. *Psychological Review, 64*(3), 205-215.

Broadbent, D. E., & Broadbent, M. H. P. (1987). From Detection to Identification: Response to Multiple Targets in Rapid Serial Visual Presentation. *Perception & Psychophysics, 42*(2), 105-113.

Brodbeck, V., Spinelli, L., Lascano, A. M., Wissmeier, M., Vargas, M. I., Vulliemoz, S., et al. (2011). Electroencephalographic source imaging: a prospective study of 152 operated epileptic patients. *Brain, 134*, 2887-2897.

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition & Emotion*, *24*(3), 377-400.

Brosch, T., Pourtois, G., Sander, D., & Vuilleumier, P. (2011). Additive effects of emotional, endogenous, and exogenous attention: Behavioral and electrophysiological evidence. *Neuropsychologia*, *49*(7), 1779-1787.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: linear ballistic accumulation. *Cogn Psychol, 57*(3), 153-178.

Bruner, J. S., & Postman, L. (1949). On the Perception of Incongruity: A Paradigm. *J Pers, 18*, 206-223.

Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal Analysis of Multichannel EEG: CARTOOL. *Computational Intelligence and Neuroscience, 2011*.

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews, 36*(2-3), 96-107.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4*(6), 215-222.

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D., Jenike, M. A., et al. (2002). Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America, 99*(1), 523-528.

Bush, G., Whalen, P. J., Rosen, B. R., Jenike, M. A., McInerney, S. C., & Rauch, S. L. (1998). The counting stroop: An interference task specialized for functional neuroimaging - Validation study with functional MRI. *Human Brain Mapping, 6*(4), 270-282.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between Attitudes and Evaluative Space: A Critical Review, with Emphasis on the Separability of Positive and Negative Substrates. *Psychological Bulletin, 115*(3), 401-423.

Cacioppo, J. T., & Berntson, G. G. (1999). The affect system: Architecture and operating characteristics. *Current Directions in Psychological Science, 8*(5), 133-137.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In R. Lewis & J. M. Haviland-Jones (Eds.), *The Handbook of Emotion, 2nd edition* (pp. 173-191). New York: Guilford Press.

Cacioppo, J. T., & Gardner, W. L. (1999). Emotion. *Annual Review of Psychology, 50*, 191-214.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond Bipolar Conceptualizations and Measures: The Case of Attitudes and Evaluative Space. *Personality and Social Psychology Review, 1*(1), 3-25.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*(5), 839-855.

Cacioppo, J. T., & Ito, T. A. (1999). Activation functions for positive and negative affective processes: The negativity bias. *Psychophysiology, 36*, S4.

Calvo, M. G., & Nummenmaa, L. (2011). Time course of discrimination between emotional facial expressions: The role of visual saliency. *Vision Research, 51*(15), 1751-1759.

Cannon, W. B. (1932). *The wisdom of the body*. New York: Norton.

Cano, M. E., Class, Q. A., & Polich, J. (2009). Affective valence, stimulus attributes, and P300: Color vs. black/white and normal vs. scrambled images. *International Journal of Psychophysiology, 71*(1), 17-24.

Carlson, T., Grol, M. J., & Verstraten, F. A. J. (2006). Dynamics of visual recognition revealed by fMRI. *Neuroimage, 32*(2), 892-905.

Carretié, L., Hinojosa, J. A., Lopez-Martin, S., & Tapia, M. (2007). An electrophysiological study on the interaction between emotional content and spatial frequency of visual stimuli. *Neuropsychologia, 45*(6), 1187-1195.

Carretié, L., Hinojosa, J. A., Martin-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping, 22*(4), 290-299.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205-218.

Catani, M., Jones, D. K., Donato, R., & ffytche, D. H. (2003). Occipito-temporal connections in the human brain. *Brain, 126*, 2093-2107.

Chun, M. M., & Potter, M. C. (1995). A Two-Stage Model for Multiple-Target Detection in Rapid Serial Visual Presentation. *Journal of Experimental Psychology-Human Perception and Performance, 21*(1), 109-127.

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review, 30*(2), 203-216.

Cisler, J. M., Wolitzky-Taylor, K. B., Adams, T. G., Babson, K. A., Badour, C. L., & Willems, J. L. (2011). The emotional Stroop task and posttraumatic stress disorder: A meta-analysis. *Clinical Psychology Review, 31*(5), 817-828.

Clark, A. (2012). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 3*, 1-86.

Coelho, C. M., Cloete, S., & Wallis, G. (2010). The face-in-the-crowd effect: When angry faces are just cross(es). *Journal of Vision, 10*(1).

Compton, R. J. (2003). The Interface Between Emotion and Attention: A Review of Evidence from Psychology and Neuroscience. *Behavioral and Cognitive Neuroscience Reviews, 2*(2), 115-129.

Conklin, L. R., Strunk, D. R., & Fazio, R. H. (2009). Attitude formation in depression: Evidence for deficits in forming positive attitudes. *Journal of Behavior Therapy and Experimental Psychiatry, 40*(1), 120-126.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58*(3), 306-324.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 201-215.

Courtney, S. M., Ungerleider, B. G., Keil, K., & Haxby, J. V. (1997). Transient and sustained activity in a distributed neural system for human working memory. *Nature, 386*(6625), 608-611.

Craft, E., Schutze, H., Niebur, E., & von der Heydt, R. (2007). A neural model of figure-ground organization. *Journal of Neurophysiology, 97*(6), 4310-4326.

Craig, A. D. (2009). How do you feel - now? The anterior insula and human awareness. *Nature Reviews Neuroscience, 10*(1), 59-70.

Cristinzio, C., N'Diaye, K., Seeck, M., Vuilleumier, P., & Sander, D. (2010). Integration of gaze direction and facial expression in patients with unilateral amygdala damage. *Brain, 133*, 248-261.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology, 115*(4), 732-744.

D'Hondt, F., Lassonde, M., Collignon, O., Dubarry, A. S., Robert, M., Rigoulot, S., et al. (2010). Early brain-body impact of emotional arousal. *Frontiers in Human Neuroscience, 4*.

Damaraju, E., Huang, Y. M., Barrett, L. F., & Pessoa, L. (2009). Affective learning enhances activity and functional connectivity in early visual cortex. *Neuropsychologia, 47*(12), 2480-2487.

Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: G.P. Putnam.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society B-Biological Sciences, 351*(1346), 1413-1420.

Davidson, R. J. (2000). Affective style, psychopathology, and resilience: Brain mechanisms and plasticity. *American Psychologist, 55*(11), 1196-1214.

Davidson, R. J., Lewis, D. A., Alloy, L. B., Amaral, D. G., Bush, G., Cohen, J. D., et al. (2002). Neural and behavioral substrates of mood and mood regulation. *Biol Psychiatry, 52*(6), 478-502.

Davoli, C. C., Suszko, J. W., & Abrams, R. A. (2007). New objects can capture attention without a unique luminance transient. *Psychon Bull Rev, 14*(2), 338-343.

De Boeck, P. (2008). Random Item IRT Models. *Psychometrika, 73*(4), 533-559.

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology, 43*(2), 207-215.

de Gelder, B., Meeren, H. K. M., Righart, R., van den Stock, J., van de Riet, W. A. C., & Tamietto, M. (2006). Beyond the face: exploring rapid influences of context on face processing. *Visual Perception, Pt 2: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception, 155*, 37-48.

De Ruiter, C., & Brosschot, J. F. (1994). The Emotional Stroop Interference Effect in Anxiety: Attentional Bias or Cognitive Avoidance? *Behaviour Research and Therapy, 32*(3), 315-319.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a Neural System for Error-Detection and Compensation. *Psychological Science, 5*(5), 303-305.

Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., & Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neuroscience Letters, 356*(1), 1-4.

Delplanque, S., N'diaye, K., Scherer, K., & Grandjean, D. (2007). Spatial frequencies or emotional effects? A systematic measure of spatial frequencies for IAPS pictures by a discrete wavelet analysis. *Journal of Neuroscience Methods, 165*(1), 144-150.

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience, 18*, 193-222.

Dhar, M., Wiersema, J. R., & Pourtois, G. (2011). Cascade of Neural Events Leading from Error Commission to Subsequent Awareness Revealed Using EEG Source Imaging. *Plos One, 6*(5).

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology-General, 129*(4), 481-507.

Di Lollo, V., Kawahara, J., Ghorashi, S. M. S., & Enns, J. T. (2005). The attentional blink: Resource depletion or temporary loss of control? *Psychological Research-Psychologische Forschung, 69*(3), 191-200.

Dickerson, B. C., & Eichenbaum, H. (2010). The Episodic Memory System: Neurocircuitry and Disorders. *Neuropsychopharmacology, 35*(1), 86-104.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science, 298*(5596), 1191-1194.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., et al. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature, 389*(6651), 596-599.

Dolcos, F., & Denkova, E. (2008). Neural correlates of encoding emotional memories: A review of functional neuroimaging evidence. *Cell Science Reviews, 5*(2), 78-122.

Dosenbach, N. U. F., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends in Cognitive Sciences, 12*(3), 99-105.

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92,* 53-78.

Driver, J., & Baylis, G. C. (1996). Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognitive Psychology, 31*(3), 248-306.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Duncan, J., & Humphreys, G. W. (1989). Visual Search and Stimulus Similarity. *Psychological Review, 96*(3), 433-458.

Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention Perception & Psychophysics, 71*(8), 1683-1700.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics, 63*(6), 1004-1013.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics, 65*(3), 352-358.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48,* 269-297.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8*(12), 1784-1790.

Eimer, M., Holmes, A., & McGlone, F. P. (2003). The role of spatial attention in the processing of facial expression: an ERP study of rapid brain responses to six basic emotions. *Cogn Affect Behav Neurosci, 3*(2), 97-110.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion, 6*(3-4), 169-200.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*. Oxford, England: Pergamon Press.

Enns, J. T., & Lleras, A. (2008). What's next? New evidence for prediction in human vision. *Trends in Cognitive Sciences, 12*(9), 327-333.

Epp, A. M., Dobson, K. S., Dozois, D. J. A., & Frewen, P. A. (2012). A systematic meta-analysis of the Stroop task in depression. *Clinical Psychology Review, 32*(4), 316-328.

Epstein, R., Harris, A., Stanley, D., & Kanwisher, N. (1999). The parahippocampal place area: recognition, navigation, or encoding? *Neuron, 23*(1), 115-125.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392*(6676), 598-601.

Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K., & Lehmann, D. (2004). Brain areas and time course of emotional processing. *Neuroimage, 21*(4), 1189-1203.

Esteves, F., Dimberg, U., & Öhman, A. (1994). Automatically Elicited Fear: Conditioned Skin-Conductance Responses to Masked Facial Expressions. *Cognition & Emotion, 8*(5), 393-413.

Everaert, T., Spruyt, A., & De Houwer, J. (2011). On the (Un)conditionality of Automatic Attitude Activation: The Valence Proportion Effect. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 65*(2), 125-132.

Farivar, R. (2009). Dorsal-ventral integration in object recognition. *Brain Research Reviews, 61*(2), 144-153.

Fecica, A. M., & Stolz, J. A. (2008). Facial affect and temporal order judgments: emotions turn back the clock. *Exp Psychol, 55*(1), 3-8.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex, 1*(1), 1-47.

Fenske, M. J., Aminoff, E., Gronau, N., & Bar, M. (2006). Top-down facilitation of visual object recognition: object-based and context-based contributions. *Visual Perception, Pt 2: Fundamentals of Awareness: Multi-Sensory Integration and High-Order Perception, 155*, 3-21.

Fincham, J. M., & Anderson, J. R. (2006). Distinct roles of the anterior cingulate and prefrontal cortex in the acquisition and performance of a cognitive skill. *Proceedings of the National Academy of Sciences of the United States of America, 103*(34), 12941-12946.

Finney, D. J. (1964). *Probit analysis: Statistical treatment of the sigmoid curve*. London: Cambridge University Press.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary Covert Orienting Is Contingent on Attentional Control Settings. *Journal of Experimental Psychology-Human Perception and Performance, 18*(4), 1030-1044.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science, 18*(12), 1050-1057.

Forgas, J. P., & Fiedler, K. (1996). Us and them: Mood effects on intergroup discrimination. *Journal of Personality and Social Psychology, 70*(1), 28-40.

Foti, D., Hajcak, G., & Dien, J. (2009). Differentiating neural responses to emotional pictures: Evidence from temporal-spatial PCA. *Psychophysiology, 46*(3), 521-530.

Fox, E. (1993). Attentional bias in anxiety: selective or not? *Behav Res Ther, 31*(5), 487-493.

Fox, E., & Damjanovic, L. (2006). The eyes are sufficient to produce a threat superiority effect. *Emotion, 6*(3), 534-539.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14*(1), 61-92.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *J Exp Psychol Gen, 130*(4), 681-700.

Freunberger, R., Klimesch, W., Doppelmayr, M., & Holler, Y. (2007). Visual P2 component is related to theta phase-locking. *Neuroscience Letters, 426*(3), 181-186.

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition & Emotion, 1*(2), 115-143.

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual search for faces with emotional expressions. *Psychological Bulletin, 134*(5), 662-676.

Friston, K. (2002). Functional integration and inference in the brain. *Progress in Neurobiology, 68*(2), 113-143.

Friston, K. (2003). Learning and inference in the brain. *Neural Networks, 16*(9), 1325-1352.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B-Biological Sciences, 360*(1456), 815-836.

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences, 13*(7), 293-301.

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1211-1221.

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., & Ebersole, J. S. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology, 113*(5), 702-712.

Gainotti, G. (1972). Emotional behavior and hemispheric side of the lesion. *Cortex, 8*(1), 41-55.

Gainotti, G. (2012). Unconscious processing of emotions and the right hemisphere. *Neuropsychologia, 50*(2), 205-218.

Gamond, L., George, N., Lemarechal, J. D., Hugueville, L., Adam, C., & Tallon-Baudry, C. (2011). Early influence of prior experience on face perception. *Neuroimage, 54*(2), 1415-1426.

Ganis, G., Schendan, H. E., & Kosslyn, S. M. (2007). Neuroimaging evidence for object model verification theory: Role of prefrontal control in visual object categorization. *Neuroimage, 34*(1), 384-398.

Gauthier, I., Anderson, A. W., Tarr, M. J., Skudlarski, P., & Gore, J. C. (1997). Levels of categorization in visual recognition studied using functional magnetic resonance imaging. *Current Biology, 7*(9), 645-651.

George, N., Dolan, R. J., Fink, G. R., Baylis, G. C., Russell, C., & Driver, J. (1999). Contrast polarity and face recognition in the human fusiform gyrus. *Nat Neurosci, 2*(6), 574-580.

Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z. S., et al. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia, 46*(8), 2129-2139.

Giesbrecht, B., & Di Lollo, V. (1998). Beyond the attentional blink: Visual masking by object substitution. *Journal of Experimental Psychology-Human Perception and Performance, 24*(5), 1454-1466.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience, 10*(5), 1563-1573.

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*(2), 467-476.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*, 535-574.

Goldman-Rakic, P. S. (1990). Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates. *Prog Brain Res, 85*, 325-335; discussion 335-326.

Goodale, M. A., & Milner, A. D. (1992). Separate Visual Pathways for Perception and Action. *Trends in Neurosciences, 15*(1), 20-25.

Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological Science, 14*(5), 505-509.

Gotlib, I. H., & McCann, C. D. (1984). Construct Accessibility and Depression: An Examination of Cognitive and Affective Factors. *Journal of Personality and Social Psychology, 47*(2), 427-439.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A New Method for Off-Line Removal of Ocular Artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468-484.

Grech, R., Cassar, T., Muscat, J., Camilleri, K., Fabri, S., Zervakis, M., et al. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of NeuroEngineering and Rehabilitation, 5*(1), 25.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends Cogn Sci, 6*(12), 517-523.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annu Rev Neurosci, 27*, 649-677.

Grinband, J., Hirsch, J., & Ferrera, V. P. (2006). A neural representation of categorization uncertainty in the human brain. *Neuron, 49*(5), 757-763.

Grossberg, S. (2009). Cortical and subcortical predictive dynamics and learning during perception, cognition, emotion and action. *Philosophical Transactions of the Royal Society B-Biological Sciences, 364*(1521), 1223-1234.

Gruber, T., & Muller, M. M. (2005). Oscillatory brain activity dissociates between associative stimulus content in a repetition priming task in the human EEG. *Cerebral Cortex, 15*(1), 109-116.

Gschwind, M., Pourtois, G., Schwartz, S., de Ville, D. V., & Vuilleumier, P. (2012). White-Matter Connectivity between Face-Responsive Regions in the Human Brain. *Cerebral Cortex, 22*(7), 1564-1576.

Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex, 10*(1), 69-81.

Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology, 14*(2), 233-238.

Hampton, A. N., & O'Doherty, J. P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. *Proceedings of the National Academy of Sciences of the United States of America, 104*(4), 1377-1382.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science, 274*(5286), 427-430.

Hansen, C. H., & Hansen, R. D. (1988). Finding the Face in the Crowd: An Anger Superiority Effect. *Journal of Personality and Social Psychology, 54*(6), 917-924.

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2009). Fictive Reward Signals in the Anterior Cingulate Cortex. *Science, 324*(5929), 948-950.

Heekeren, H. R., Marrett, S., Bandettini, P. A., & Ungerleider, L. G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature, 431*(7010), 859-862.

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience, 9*(6), 467-479.

Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology, 84*(4), 405-439.

Hellige, J. B. (1983). Feature similarity and laterality effects in visual masking. *Neuropsychologia, 21*(6), 633-639.

Hellige, J. B., Cox, P. J., & Litvac, L. (1979). Information processing in the cerebral hemispheres: selective hemispheric activation and capacity limitations. *J Exp Psychol Gen, 108*(2), 251-279.

Henson, R. N. A., Hornberger, M., & Rugg, M. D. (2005). Further dissociating the processes involved in recognition memory: An fMRI study. *Journal of Cognitive Neuroscience, 17*(7), 1058-1073.

Hietanen, J. K., & Leppanen, J. M. (2003). Does facial expression affect attention orienting by gaze direction cues? *Journal of Experimental Psychology-Human Perception and Performance, 29*(6), 1228-1243.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal Visual Attention Produces Illusory Temporal Order and Motion Sensation. *Vision Research, 33*(9), 1219-1240.

Hillstrom, A. P., & Yantis, S. (1994). Visual Motion and Attentional Capture. *Perception & Psychophysics, 55*(4), 399-411.

Hillyard, S. A. (1985). Electrophysiology of Human Selective Attention. *Trends in Neurosciences, 8*(9), 400-405.

Hirsh, I. J., & Sherrick, C. E. (1961). Perceived Order in Different Sense Modalities. *Journal of Experimental Psychology, 62*(5), 423-&.

Ho, T., Brown, S., van Maanen, L., Forstmann, B. U., Wagenmakers, E. J., & Serences, J. T. (2012). The Optimality of Sensory Processing during the Speed-Accuracy Tradeoff. *Journal of Neuroscience, 32*(23), 7992-8003.

Ho, T. C., Brown, S., & Serences, J. T. (2009). Domain General Mechanisms of Perceptual Decision Making in Human Cortex. *Journal of Neuroscience, 29*(27), 8675-8687.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron, 36*(5), 791-804.

Holländer, A., Corballis, M. C., & Hamm, J. P. (2005). Visual-field asymmetry in dual-stream RSVP. *Neuropsychologia, 43*(1), 35-40.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research, 16*(2), 174-184.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679-709.

Hoorens, V., & Buunk, B. P. (1993). Social-Comparison of Health Risks: Locus of Control, the Person-Positivity Bias, and Unrealistic Optimism. *Journal of Applied Social Psychology, 23*(4), 291-302.

Hopfinger, J. B., & West, V. M. (2006). Interactions between endogenous and exogenous attention on cortical visual processing. *Neuroimage, 31*(2), 774-789.

Horstmann, G. (2007). Preattentive face processing: What do visual search experiments with schematic faces tell us? *Visual Cognition, 15*(7), 799-833.

Horstmann, G., & Bauland, A. (2006). Search asymmetries with real faces: Testing the anger-superiority effect. *Emotion, 6*(2), 193-207.

Ihssen, N., & Keil, A. (2009). The costs and benefits of processing emotional stimuli during rapid serial visual presentation. *Cognition & Emotion, 23*(2), 296-326.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science, 302*(5642), 120-122.

Ito, T., Cacioppo, J. T., & Lang, P. J. (1998). Eliciting affect using the international affective picture system: Trajectories through evaluative space. *Personality and Social Psychology Bulletin, 24*(8), 855-879.

Ito, T. A., & Cacioppo, J. T. (2001). Affect and attitudes: A social neuroscience approach. In J. P. Forgas (Ed.), *Handbook of affect and social cognition* (pp. 50-74). Mahwah, NJ: Erlbaum.

Ito, T. A., & Cacioppo, J. T. (2005). Variations on a human universal: Individual differences in positivity offset and negativity bias. *Cognition & Emotion, 19*(1), 1-26.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology, 75*(4), 887-900.

Izard, C. E. (1993). Four systems for emotion activation: cognitive and noncognitive processes. *Psychol Rev, 100*(1), 68-90.

James, T. W., Humphrey, G. K., Gati, J. S., Menon, R. S., & Goodale, M. A. (2000). The effects of visual object priming on brain activation before and after recognition. *Current Biology, 10*(17), 1017-1024.

Jaskowski, P. (1993). Selective Attention and Temporal-Order Judgment. *Perception, 22*(6), 681-689.

Jemel, B., Schuller, A. M., & Goffaux, V. (2010). Characterizing the Spatio-temporal Dynamics of the Neural Events Occurring prior to and up to Overt Recognition of Famous Faces. *Journal of Cognitive Neuroscience, 22*(10), 2289-2305.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Percept Psychophys, 43*(4), 346-354.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. *Nature, 387*(6635), 805-807.

Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology, 38*(2), 175-178.

Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *Neuroimage, 34*(4), 1600-1611.

Kable, J. W., & Glimcher, P. W. (2009). The Neurobiology of Decision: Consensus and Controversy. *Neuron, 63*(6), 733-745.

Kanske, P., & Kotz, S. A. (2011a). Emotion Speeds up Conflict Resolution: A New Role for the Ventral Anterior Cingulate Cortex? *Cerebral Cortex, 21*(4), 911-919.

Kanske, P., & Kotz, S. A. (2011b). Emotion Triggers Executive Attention: Anterior Cingulate Cortex and Amygdala Responses to Emotional Words in a Conflict Task. *Human Brain Mapping, 32*(2), 198-208.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*(11), 4302-4311.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philos Trans R Soc Lond B Biol Sci, 361*(1476), 2109-2128.

Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience, 23,* 315-341.

Kawasaki, H., Adolphs, R., Kaufman, O., Damasio, H., Damasio, A. R., Granner, M., et al. (2001). Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. *Nature Neuroscience, 4*(1), 15-16.

Keetels, M., & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research, 167*(4), 635-640.

Keil, A., & Ihssen, N. (2004). Identification facilitation for emotionally arousing verbs during the attentional blink. *Emotion, 4*(1), 23-35.

Keil, A., Ihssen, N., & Heim, S. (2006). Early cortical facilitation for emotionally arousing targets during the attentional blink. *Bmc Biology, 4*.

Keil, A., Moratti, S., Sabatinelli, D., Bradley, M. M., & Lang, P. J. (2005). Additive effects of emotional content and spatial selective attention on electrocortical facilitation. *Cerebral Cortex, 15*(8), 1187-1197.

Keil, A., Sabatinelli, D., Ding, M., Lang, P. J., Ihssen, N., & Heim, S. (2009). Re-entrant projections modulate visual cortex in affective perception: evidence from Granger causality analysis. *Human Brain Mapping, 30*(2), 532-540.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271-304.

Kiefer, M., & Martens, U. (2010). Attentional Sensitization of Unconscious Cognition: Task Sets Modulate Subsequent Masked Semantic Priming. *Journal of Experimental Psychology-General, 139*(3), 464-489.

Kim, J. N., & Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience, 2*(2), 176-185.

Kim, S., & Rehder, B. (2011). How prior knowledge affects selective attention during category learning: An eyetracking study. *Memory & Cognition, 39*(4), 649-665.

Kim, Y. H., Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., & Mesulam, M. M. (1999). The large-scale neural network for spatial attention displays multifunctional overlap but differential asymmetry. *Neuroimage, 9*(3), 269-277.

Klar, Y., & Giladi, E. E. (1997). My group can be below the group's average: A robust positivity bias in favor of anonymous peers. *Journal of Personality and Social Psychology, 73*(5), 885-901.

Kosslyn, S. M., Thompson, W. L., Gitelman, D. R., & Alpert, N. M. (1998). Neural systems that encode categorical versus coordinate spatial relations: PET investigations. *Psychobiology, 26*(4), 333-347.

Koster, E. H., Crombez, G., Verschuere, B., Van Damme, S., & Wiersema, J. R. (2006). Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance. *Behav Res Ther, 44*(12), 1757-1771.

Koster, E. H. W., Crombez, G., Van Damme, S., Verschuere, B., & De Houwer, J. (2004). Does imminent threat capture and hold attention? *Emotion, 4*(3), 312-317.

Koster, E. H. W., Verschuere, B., Crombez, G., & Van Damme, S. (2005). Time-course of attention for threatening pictures in high and low trait anxiety. *Behaviour Research and Therapy, 43*(8), 1087-1098.

Kourtzi, Z., & Connor, C. E. (2011). Neural Representations for Object Perception: Structure, Category, and Adaptive Coding. *Annual Review of Neuroscience, Vol 34, 34*, 45-67.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: Linking reward to hedonic experience. *Nature Reviews Neuroscience, 6*(9), 691-702.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and Cognition, 65*(2), 145-168.

Kveraga, K., Ghuman, A. S., Kassam, K. S., Aminoff, E. A., Hamalainen, M. S., Chaumon, M., et al. (2011). Early onset of neural synchronization in the contextual associations network. *Proceedings of the National Academy of Sciences of the United States of America, 108*(8), 3389-3394.

Lang, P. J. (1985). The cognitive psychophysiology of emotion: Fear and anxiety. In A. H. Tuma & J. D.Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 131-170). Hillsdale, NJ: Erlbaum.

Lang, P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *American Psychologist, 50*(5), 372-385.

Lang, P. J., & Bradley, M. M. (2010). Emotion and the motivational brain. *Biological Psychology, 84*(3), 437-450.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1990). Emotion, Attention, and the Startle Reflex. *Psychological Review, 97*(3), 377-395.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. *Attention and Orienting: Sensory and Motivational Processes*, 97-135.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Digitized photographs, instruction manual, and affective ratings (Tech. Rep. No. A-6).* Gainesville: University of Florida, Center for Research in Psychophysiology.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8*: University of Florida, Gainesville, FL.

Lantz, G., Menendez, R. G. D., Andino, S. G., & Michel, C. M. (2001). Noninvasive localization of electromagnetic epileptic activity. II. Demonstration of sublobar accuracy in patients with simultaneous surface and depth recordings. *Brain Topography, 14*(2), 139-147.

Lavie, N. (1995). Perceptual Load as a Necessary Condition for Selective Attention. *Journal of Experimental Psychology-Human Perception and Performance, 21*(3), 451-468.

LeDoux, J. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon & Schuster.

Lehmann, D., & Skrandies, W. (1980). Reference-Free Identification of Components of Checkerboard-Evoked Multichannel Potential Fields. *Electroencephalography and Clinical Neurophysiology, 48*(6), 609-621.

Libkuman, T. M., Otam, H., Kern, R., Viger, S. G., & Novak, N. (2007). Multidimensional normative ratings for the international affective picture system. *Behavior Research Methods, 39*(2), 326-334.

Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., et al. (2010). Are Females More Responsive to Emotional Stimuli? A Neurophysiological Study Across Arousal and Valence Dimensions. *Brain Topography, 23*(1), 27-40.

Lochmann, T., & Deneve, S. (2011). Neural processing as causal inference. *Current Opinion in Neurobiology, 21*(5), 774-781.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*(6843), 150-157.

Lomber, S. G. (2002). Learning to see the trees before the forest: Reversible deactivation of the superior colliculus during learning of local and global visual features. *Proceedings of the National Academy of Sciences of the United States of America, 99*(6), 4049-4054.

Lucas, N., & Vuilleumier, P. (2008). Effects of emotional and non-emotional cues on visual search in neglect patients: Evidence for distinct sources of attentional guidance. *Neuropsychologia, 46*(5), 1401-1414.

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

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological Correlates of Feature Analysis during Visual Search. *Psychophysiology, 31*(3), 291-308.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4*(11), 432-440.

Luo, Q., Holroyd, T., Jones, M., Hendler, T., & Blair, J. (2007). Neural dynamics for facial threat processing as revealed by gamma band synchronization using MEG. *Neuroimage, 34*(2), 839-847.

Luo, Q., Holroyd, T., Majestic, C., Cheng, X., Schechter, J., & Blair, R. J. (2010). Emotional Automaticity Is a Matter of Timing. *Journal of Neuroscience, 30*(17), 5825-5829.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional Bias in Emotional Disorders. *Journal of Abnormal Psychology, 95*(1), 15-20.

Macleod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin, 109*(2), 163-203.

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality, 69*(4), 583-615.

Mak-Fan, K. M., Thompson, W. F., & Green, R. E. A. (2011). Visual search for schematic emotional faces risks perceptual confound. *Cognition & Emotion, 25*(4), 573-584.

Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Sciences, 8*(5), 204-210.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-Related Activity Revealed by Functional Magnetic-Resonance-Imaging in Human Occipital Cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*(18), 8135-8139.

Mammucari, A., Caltagirone, C., Ekman, P., Friesen, W., Gainotti, G., Pizzamiglio, L., et al. (1988). Spontaneous facial expression of emotions in brain-damaged patients. *Cortex, 24*(4), 521-533.

Maratos, F. A. (2011). Temporal Processing of Emotional Stimuli: The Capture and Release of Attention by Angry Faces. *Emotion, 11*(5), 1242-1247.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences, 9*(6), 296-305.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* New York: Freeman.

Mathewson, K. J., Arnell, K. M., & Mansfield, C. A. (2008). Capturing and holding attention: the impact of emotional words in rapid serial visual presentation. *Mem Cognit, 36*(1), 182-200.

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

Mayer, A. R., Dorflinger, J. M., Rao, S. M., & Seidenberg, M. (2004). Neural networks underlying endogenous and exogenous visual-spatial orienting. *Neuroimage, 23*(2), 534-541.

Mayr, U. (2004). Conflict, consciousness, and control. *Trends in Cognitive Sciences, 8*(4), 145-148.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society B-Biological Sciences, 356*(1412), 1293-1322.

McKenna, F. P., & Sharma, D. (2004). Reversing the emotional stroop effect reveals that it is not what it seems: The role of fast and slow components. *Journal of Experimental Psychology-Learning Memory and Cognition, 30*(2), 382-392.

McMains, S., & Kastner, S. (2011). Interactions of Top-Down and Bottom-Up Mechanisms in Human Visual Cortex. *Journal of Neuroscience, 31*(2), 587-597.

Melloni, L., Schwiedrzik, C. M., Muller, N., Rodriguez, E., & Singer, W. (2011). Expectations Change the Signatures and Timing of Electrophysiological Correlates of Perceptual Awareness. *Journal of Neuroscience, 31*(4), 1386-1396.

Mendes, W. B., Blascovich, J., Hunter, S. B., Lickel, B., & Jost, J. T. (2007). Threatened by the unexpected: Physiological responses during social interactions with expectancy-violating partners. *Journal of Personality and Social Psychology, 92*(4), 698-716.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*(5245), 77-81.

Mesulam, M. M. (1998). From sensation to cognition. *Brain, 121 (Pt 6)*, 1013-1052.

Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage, 61*(2), 371-385.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology, 115*(10), 2195-2222.

Michel, C. M., Seeck, M., & Landis, T. (1999). Spatiotemporal dynamics of human cognition. *News in Physiological Sciences, 14*, 206-214.

Michel, C. M., Thut, G., Morand, S., Khateb, A., Pegna, A. J., de Peralta, R. G., et al. (2001). Electric source imaging of human brain functions. *Brain Research Reviews, 36*(2-3), 108-118.

Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.

Mogg, K., & Bradley, B. P. (1998). A cognitive-motivational analysis of anxiety. *Behaviour Research and Therapy, 36*(9), 809-848.

Mogg, K., & Bradley, B. P. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition & Emotion, 13*(6), 713-740.

Mogg, K., Mathews, A., & Weinman, J. (1989). Selective processing of threat cues in anxiety states: a replication. *Behav Res Ther, 27*(4), 317-323.

Monahan, J. L., Murphy, S. T., & Zajonc, R. B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science, 11*(6), 462-466.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297-326.

Moritz, S., Jacobsen, D., Kloss, M., Fricke, S., Rufer, M., & Hand, I. (2004). Examination of emotional Stroop interference in obsessive-compulsive disorder. *Behav Res Ther, 42*(6), 671-682.

Moskowitz, G. B. (2002). Preconscious effects of temporary goals on attention. *Journal of Experimental Social Psychology, 38*(4), 397-404.

Mossbridge, J. A., Fitzgerald, M. B., O'Connor, E. S., & Wright, B. A. (2006). Perceptual-learning evidence for separate processing of asynchrony and order tasks. *Journal of Neuroscience, 26*(49), 12708-12716.

Most, S. B., Chun, M. M., Widders, D. M., & Zald, D. H. (2005). Attentional rubbernecking: Cognitive control and personality in emotion-induced blindness. *Psychonomic Bulletin & Review, 12*(4), 654-661.

Most, S. B., & Wang, L. L. (2011). Dissociating Spatial Attention and Awareness in Emotion-Induced Blindness. *Psychological Science, 22*(3), 300-305.

Mulckhuyse, M., & Theeuwes, J. (2010). Unconscious attentional orienting to exogenous cues: A review of the literature. *Acta Psychologica, 134*(3), 299-309.

Müller, M. M., Andersen, S. K., & Keil, A. (2008). Time course of competition for visual processing resources between emotional pictures and foreground task. *Cerebral Cortex, 18*(8), 1892-1899.

Mumford, D. (1992). On the Computational Architecture of the Neocortex. 2. The Role of Corticocortical Loops. *Biological Cybernetics, 66*(3), 241-251.

Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography, 20*(4), 249-264.

Nava, E., Bottari, D., Zampini, M., & Pavani, F. (2008). Visual temporal order judgment in profoundly deaf individuals. *Experimental Brain Research, 190*(2), 179-188.

Navon, D. (1977). Forest Before Trees: The Precedence of Global Features in Visual Perception. *Cognitive Psychology, 9*(3), 353-383.

Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton.

Nelson, S. M., Dosenbach, N. U. F., Cohen, A. L., Wheeler, M. E., Schlaggar, B. L., & Petersen, S. E. (2010). Role of the anterior insula in task-level control and focal attention. *Brain Structure & Function, 214*(5-6), 669-680.

Nickerson, R. S., & Freeman, B. (1974). Discrimination of Order of Components of Repeating Tone Sequences: Effects of Frequency Separation and Extensive Practice. *Perception & Psychophysics, 16*(3), 471-477.

Niemi, P., & Näätänen, R. (1981). Foreperiod and Simple Reaction Time. *Psychological Bulletin, 89*(1), 133-162.

Nieuwenstein, M. R., Potter, M. C., & Theeuwes, J. (2009). Unmasking the Attentional Blink. *Journal of Experimental Psychology-Human Perception and Performance, 35*(1), 159-169.

Norris, C. J., Gollan, J., Berntson, G. G., & Cacioppo, J. T. (2010). The current status of research on the structure of evaluative space. *Biological Psychology, 84*(3), 422-436.

Öhman, A. (2001). Nonconscious emotion: Evolutionary, neural, and psychophysiologic perspectives. *Psychophysiology, 38*, S5-S5.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology-General, 130*(3), 466-478.

Öhman, A., Hamm, A., & Hugdahl, K. (2000). Cognition and the autonomic nervous system: Orienting, anticipation, and conditioning. In J. Cacioppo, L. Tassinary & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 533-575). New York: Cambridge University Press.

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: a threat advantage with schematic stimuli. *J Pers Soc Psychol, 80*(3), 381-396.

Olatunji, B. O., Ciesielski, B. G., Armstrong, T., & Zald, D. H. (2011). Emotional Expressions and Visual Search Efficiency: Specificity and Effects of Anxiety Symptoms. *Emotion, 11*(5), 1073-1079.

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*(1), 72-107.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11*(12), 520-527.

Olivers, C. N. L., & Nieuwenhuis, S. (2006). The beneficial effects of additional task load, positive affect, and instruction on the attentional blink. *Journal of Experimental Psychology-Human Perception and Performance, 32*(2), 364-379.

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology, 77*(3), 247-265.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719.

Opris, I., & Bruce, C. J. (2005). Neural circuitry of judgment and decision mechanisms. *Brain Research Reviews, 48*(3), 509-526.

Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.

Palmeri, T. J., & Gauthier, I. (2004). Visual object understanding. *Nature Reviews Neuroscience, 5*(4), 291-U217.

Pascual-Marqui, R. D. (1999). Review of methods for solving the EEG inverse problem. *International Journal of Bioelectromagnetism, 1,* 75-86.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology, 24,* 5-12.

Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): A review. *Methods and Findings in Experimental and Clinical Pharmacology, 24*, 91-95.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1995). Segmentation of Brain Electrical Activity into Microstates: Model Estimation and Validation. *Ieee Transactions on Biomedical Engineering, 42*(7), 658-665.

Peelen, M. V., Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *Neuroimage, 22*(2), 822-830.

Perez, A., Peers, P. V., Valdes-Sosa, M., Galan, L., Garcia, L., & Martinez-Montes, E. (2009). Hemispheric modulations of alpha-band power reflect the rightward shift in attention induced by enhanced attentional load. *Neuropsychologia, 47*(1), 41-49.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical Splines for Scalp Potential and Current-Density Mapping. *Electroencephalography and Clinical Neurophysiology, 72*(2), 184-187.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*(2), 148-158.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2002). Attentional control of the processing of neutral and emotional stimuli. *Cognitive Brain Research, 15*(1), 31-45.

Pessoa, L., Kastner, S., & Ungerleider, L. G. (2003). Neuroimaging studies of attention: From modulation of sensory processing to top-down control. *Journal of Neuroscience, 23*(10), 3990-3998.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of the United States of America, 99*(17), 11458-11463.

Peyk, P., Schupp, H. T., Keil, A., Elbert, T., & Junghöfer, M. (2009). Parallel processing of affective visual stimuli. *Psychophysiology, 46*(1), 200-208.

Peyrin, C., Michel, C. M., Schwartz, S., Thut, G., Seghier, M., Landis, T., et al. (2010). The Neural Substrates and Timing of Top-Down Processes during Coarse-to-Fine Categorization of Visual Scenes: A Combined fMRI and ERP Study. *Journal of Cognitive Neuroscience, 22*(12), 2768-2780.

Phaf, R. H., & Kan, K. J. (2007). The automaticity of emotional Stroop: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 184-199.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala *Annual Review of Psychology* (Vol. 57, pp. 27-53). Palo Alto: Annual Reviews.

Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science, 17*(4), 292-299.

Philiastides, M. G., Auksztulewicz, R., Heekeren, H. R., & Blankenburg, F. (2011). Causal Role of Dorsolateral Prefrontal Cortex in Human Perceptual Decision Making. *Current Biology, 21*(11), 980-983.

Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex, 16*(4), 509-518.

Philiastides, M. G., & Sajda, P. (2007). EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *Journal of Neuroscience, 27*(48), 13082-13091.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology, 37*(2), 127-152.

Pinkham, A. E., Griffin, M., Baron, R., Sasson, N. J., & Gur, R. C. (2010). The face in the crowd effect: anger superiority when using real faces and multiple identities. *Emotion, 10*(1), 141-146.

Plonsey, R. (1963). Reciprocity Applied to Volume Conductors and Ecg. *Ieee Transactions on Biomedical Engineering, Bm10*(1), 9-&.

Ploran, E. J., Nelson, S. M., Velanova, K., Donaldson, D. I., Petersen, S. E., & Wheeler, M. E. (2007). Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *Journal of Neuroscience, 27*(44), 11912-11924.

Ploran, E. J., Tremel, J. J., Nelson, S. M., & Wheeler, M. E. (2011). High Quality but Limited Quantity Perceptual Evidence Produces Neural Accumulation in Frontal and Parietal Cortex. *Cerebral Cortex, 21*(11), 2650-2662.

Posner, M. I. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology, 32*(Feb), 3-25.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the Detection of Signals. *Journal of Experimental Psychology-General, 109*(2), 160-174.

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*(1), 65-79.

Pourtois, G., Delplanque, S., Michel, C., & Vuilleumier, P. (2008). Beyond conventional event-related brain potential (ERP): Exploring the time-course of visual emotion processing using topographic and principal component analyses. *Brain Topography, 20*(4), 265-277.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex, 14*(6), 619-633.

Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*.

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). Portraits or people? Distinct representations of face identity in the human visual cortex. *Journal of Cognitive Neuroscience, 17*(7), 1043-1057.

Pourtois, G., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010a). Modulation of Face Processing by Emotional Expression and Gaze Direction during Intracranial Recordings in Right Fusiform Cortex. *Journal of Cognitive Neuroscience, 22*(9), 2086-2107.

Pourtois, G., Spinelli, L., Seeck, M., & Vuilleumier, P. (2010b). 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*(1), 83-93.

Pratto, F., & John, O. P. (1991). Automatic vigilance: the attention-grabbing power of negative social information. *J Pers Soc Psychol, 61*(3), 380-391.

Price, J. L., & Drevets, W. C. (2012). Neural circuits underlying the pathophysiology of mood disorders. *Trends in Cognitive Sciences, 16*(1), 61-71.

Proverbio, A. M., Adorni, R., Zani, A., & Trestianu, L. (2009). Sex differences in the brain response to affective scenes with or without humans. *Neuropsychologia, 47*(12), 2374-2388.

Pulford, B. D., & Colman, A. M. (1996). Overconfidence, base rates and outcome positivity/negativity of predicted events. *British Journal of Psychology, 87*, 431-445.

Purcell, D. G., & Stewart, A. L. (2010). Still another confounded face in the crowd. *Attention Perception & Psychophysics, 72*(8), 2115-2127.

Rahnev, D., Lau, H., & de Lange, F. P. (2011). Prior Expectation Modulates the Interaction between Sensory and Prefrontal Regions in the Human Brain. *Journal of Neuroscience, 31*(29), 10741-10748.

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience, 24*(16), 3917-3925.

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews Neuroscience, 4*(3), 193-202.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79-87.

Ratcliff, R. (1978). Theory of Memory Retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R., & McKoon, G. (1982). Speed and accuracy in the processing of false statements about semantic information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 8*, 16-36.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873-922.

Ratcliff, R., Thapar, A., Gomez, P., & McKoon, G. (2004). A diffusion model analysis of the effects of aging in the lexical-decision task. *Psychology and Aging, 19*(2), 278-289.

Rauss, K., Schwartz, S., & Pourtois, G. (2011). Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience and Biobehavioral Reviews, 35*(5), 1237-1253.

Raymond, J. E. (2000). Attentional modulation of visual motion perception. *Trends in Cognitive Sciences, 4*(2), 42-50.

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*(3), 849-860.

Reeves, A., & Sperling, G. (1986). Attention Gating in Short-Term Visual Memory. *Psychological Review, 93*(2), 180-206.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuiss, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306*(5695), 443-447.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(11), 1019-1025.

Righart, R., & de Gelder, B. (2006). Context influences early perceptual analysis of faces - An electrophysiological study. *Cerebral Cortex, 16*(9), 1249-1257.

Righart, R., & de Gelder, B. (2008a). Rapid influence of emotional scenes on encoding of facial expressions: an ERP study. *Social Cognitive and Affective Neuroscience, 3*(3), 270-278.

Righart, R., & De Gelder, B. (2008b). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive Affective & Behavioral Neuroscience, 8*(3), 264-272.

Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H., & Bosch, H. (2002). Figure-ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience, 14*(4), 525-537.

Roets, A., & Van Hiel, A. (2007). Separating ability from need: Clarifying the dimensional structure of the need for closure scale. *Personality and Social Psychology Bulletin, 33*(2), 266-280.

Rosen, A. C., Rao, S. M., Caffarra, P., Scaglioni, A., Bobholz, J. A., Woodley, S. J., et al. (1999). Neural basis of endogenous and exogenous spatial orienting: A functional MRI study. *Journal of Cognitive Neuroscience, 11*(2), 135-152.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage, 20*(3), 1609-1624.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature, 392*(6676), 595-598.

Rugg, M. D., Soardi, M., & Doyle, M. C. (1995). Modulation of event-related potentials by the repetition of drawings of novel objects. *Cognitive Brain Research, 3*(1), 17-24.

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience, 11*(4), 389-397.

Russell, J. A. (1980). A Circumplex Model of Affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145-172.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion, 23*(7), 1259-1283.

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin, 125*(1), 3-30.

Sabatinelli, D., Flaisch, T., Bradley, M. M., Fitzsimmons, J. R., & Lang, P. J. (2004). Affective picture perception: gender differences in visual cortex? *Neuroreport, 15*(7), 1109-1112.

Sabatinelli, D., Lang, P. J., Bradley, M. M., Costa, V. D., & Keil, A. (2009). The timing of emotional discrimination in human amygdala and ventral visual cortex. *J Neurosci, 29*(47), 14864-14868.

Salemink, E., van den Hout, M. A., & Kindt, M. (2007). Selective attention and threat: quick orienting versus slow disengagement and two versions of the dot probe task. *Behav Res Ther, 45*(3), 607-615.

Salters-Pedneault, K., Gentes, E., & Roemer, L. (2007). The role of fear of emotion in distress, arousal, and cognitive interference following an emotional stimulus. *Cogn Behav Ther, 36*(1), 12-22.

Sander, D., Grandjean, D., Kaiser, S., Wehrle, T., & Scherer, K. R. (2007). Interaction effects of perceived gaze direction and dynamic facial expression: Evidence for appraisal theories of emotion. *European Journal of Cognitive Psychology, 19*(3), 470-480.

Sanfey, A. G., Loewenstein, G., McClure, S. M., & Cohen, J. D. (2006). Neuroeconomics: cross-currents in research on decision-making. *Trends in Cognitive Sciences, 10*(3), 108-116.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755-1758.

Sanocki, T. (1993). Time-Course of Object Identification - Evidence for a Global-to-Local Contingency. *Journal of Experimental Psychology-Human Perception and Performance, 19*(4), 878-898.

Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., et al. (2010). Uncertainty during Anticipation Modulates Neural Responses to Aversion in Human Insula and Amygdala. *Cerebral Cortex, 20*(4), 929-940.

Scheffers, M. K., & Coles, M. G. H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology-Human Perception and Performance, 26*(1), 141-151.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion, 23*(7), 1307-1351.

Scherg, M., & Voncramon, D. (1986). Evoked Dipole Source Potentials of the Human Auditory-Cortex. *Electroencephalography and Clinical Neurophysiology, 65*(5), 344-360.

Schettino, A., Loeys, T., Bossi, M., & Pourtois, G. (2012). Valence-Specific Modulation in the Accumulation of Perceptual Evidence Prior to Visual Scene Recognition. *PloS One, 7*(5), e38064.

Schettino, A., Loeys, T., Delplanque, S., & Pourtois, G. (2011). Brain dynamics of upstream perceptual processes leading to visual object recognition: A high density ERP topographic mapping study. *Neuroimage, 55*(3), 1227-1241.

Schmidt, W. C. (2000). Endogenous attention and illusory line motion reexamined. *Journal of Experimental Psychology-Human Perception and Performance, 26*(3), 980-996.

Schneider, K. A., & Bavelier, D. (2003). Components of visual prior entry. *Cognitive Psychology, 47*(4), 333-366.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Percept Psychophys, 70*(2), 208-218.

Schupp, H., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: Event-related brain potential studies. *Journal of Psychophysiology, 20*(2), 115-115.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology, 37*(2), 257-261.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Hillman, C. H., Hamm, A. O., & Lang, P. J. (2004). Brain processes in emotional perception: Motivated attention. *Cognition & Emotion, 18*(5), 593-611.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003a). Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport, 14*(8), 1107-1110.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003b). Emotional facilitation of sensory processing in the visual cortex. *Psychological Science, 14*(1), 7-13.

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2004). The selective processing of briefly presented affective pictures: An ERP analysis. *Psychophysiology, 41*(3), 441-449.

Schupp, H. T., Stockburger, J., Bublatzky, F., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). The selective processing of affective pictures: Effects of emotion and attention. *Journal of Psychophysiology, 17*(3), 178-178.

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2006). Stimulus novelty and emotion perception: the near absence of habituation in the visual cortex. *Neuroreport, 17*(4), 365-369.

Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *Journal of Neuroscience, 27*(5), 1082-1089.

Schwabe, L., Merz, C. J., Walter, B., Vaitl, D., Wolf, O. T., & Stark, R. (2011). Emotional modulation of the attentional blink: The neural structures involved in capturing and holding attention. *Neuropsychologia, 49*(3), 416-425.

Schwartz, C. E., Wright, C. I., Shin, L. M., Kagan, J., & Rauch, S. L. (2003). Inhibited and uninhibited infants "grown up": Adult amygdalar response to novelty. *Science, 300*(5627), 1952-1953.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & Emotion, 14*(4), 433-440.

Schyns, P. G., & Oliva, A. (1994). From Blobs to Boundary Edges: Evidence for Time- and Spatial-Scale-Dependent Scene Recognition. *Psychological Science, 5*(4), 195-200.

Sears, D. O. (1983). The Person Positivity Bias. *Journal of Personality and Social Psychology, 44*(2), 233-250.

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., et al. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience, 27*(9), 2349-2356.

Seitz, R. J., Franz, M., & Azari, N. P. (2009). Value judgments and self-control of action: The role of the medial frontal cortex. *Brain Research Reviews, 60*(2), 368-378.

Sekihara, K., Sahani, M., & Nagarajan, S. S. (2005). Localization bias and spatial resolution of adaptive and non-adaptive spatial filters for MEG source reconstruction. *Neuroimage, 25*(4), 1056-1067.

Seo, H., & Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *Journal of Neuroscience, 27*(31), 8366-8377.

Seo, H., & Lee, D. (2009). Behavioral and Neural Changes after Gains and Losses of Conditioned Reinforcers. *Journal of Neuroscience, 29*(11), 3627-3641.

Serences, J. T. (2008). Value-Based Modulations in Human Visual Cortex. *Neuron, 60*(6), 1169-1181.

Serences, J. T., Shomstein, S., Leber, A. B., Golay, X., Egeth, H. E., & Yantis, S. (2005). Coordination of voluntary and stimulus-driven attentional control in human cortex. *Psychological Science, 16*(2), 114-122.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences, 10*(1), 38-45.

Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology, 86*(4), 1916-1936.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S. J., & Dolan, R. J. (1994). Brain Regions Associated with Acquisition and Retrieval of Verbal Episodic Memory. *Nature, 368*(6472), 633-635.

Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences, 1*(8), 291-296.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1994). Attention to Visual Pattern Information Produces the Attentional Blink in Rapid Serial Visual Presentation. *Journal of Experimental Psychology-Human Perception and Performance, 20*(2), 357-371.

Sharot, T., Riccardi, A. M., Raio, C. M., & Phelps, E. A. (2007). Neural mechanisms mediating optimism bias. *Nature, 450*(7166), 102-+.

Sheldon, K. M., Ryan, R., & Reis, H. T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin, 22*(12), 1270-1279.

Shook, N. J., Fazio, R. H., & Vasey, M. W. (2007). Negativity bias in attitude learning: A possible indicator of vulnerability to emotional disorders? *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 144-155.

Shore, D. I., & Spence, C. (2005). Prior entry. In L. Itti, G. Rees & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 89-95). North Holland: Elsevier.

Shore, D. I., Spence, C., & Klein, R. M. (2001). Visual prior entry. *Psychol Sci, 12*(3), 205-212.

Silva, L. R., Amitai, Y., & Connors, B. W. (1991). Intrinsic oscillations of neocortex generated by layer 5 pyramidal neurons. *Science, 251*(4992), 432-435.

Sinnett, S., Juncadella, M., Rafal, R., Azanon, E., & Soto-Faraco, S. (2007). A dissociation between visual and auditory hemi-inattention: Evidence from temporal order judgements. *Neuropsychologia, 45*(3), 552-560.

Smith, F. W., & Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proceedings of the National Academy of Sciences of the United States of America, 107*(46), 20099-20103.

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*(3), 161-168.

Sohn, M. H., Albert, M. V., Jung, K. J., Carter, C. S., & Anderson, J. R. (2007). Anticipation of conflict monitoring in the anterior cingulate cortex and the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 104*(25), 10330-10334.

Spence, C., & Parise, C. (2010). Prior-entry: A review. *Consciousness and Cognition, 19*(1), 364-379.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology-General, 130*(4), 799-832.

Spratling, M. W. (2008). Predictive coding as a model of biased competition in visual attention. *Vision Research, 48*(12), 1391-1408.

Spruyt, A., De Houwer, J., Everaert, T., & Hermans, D. (2012). Unconscious semantic activation depends on feature-specific attention allocation. *Cognition, 122*(1), 91-95.

Spruyt, A., De Houwer, J., & Hermans, D. (2009). Modulation of automatic semantic priming by feature-specific attention allocation. *Journal of Memory and Language, 61*(1), 37-54.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology, 54*(1), 44-53.

Stelmach, L. B., & Herdman, C. M. (1991). Directed Attention and Perception of Temporal Order. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 539-550.

Stelmach, L. B., Herdman, C. M., & Mcneil, K. R. (1994). Attentional Modulation of Visual Processes in Motion Perception. *Journal of Experimental Psychology-Human Perception and Performance, 20*(1), 108-121.

Sterzer, P., & Kleinschmidt, A. (2010). Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Structure & Function, 214*(5-6), 611-622.

Storbeck, J., & Clore, G. L. (2007). On the interdependence of cognition and emotion. *Cognition & Emotion, 21*(6), 1212-1237.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences, 13*(9), 403-409.

Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron, 59*(2), 336-347.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19*, 109-139.

Tarr, M. J., & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience, 3*(8), 764-769.

Theeuwes, J. (1994). Endogenous and Exogenous Control of Visual Selection. *Perception, 23*(4), 429-440.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Percept Psychophys, 57*(5), 637-644.

Theeuwes, J. (2005). Irrelevant singletons capture attention. In L. Itti, G. Rees & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 418-424). San Diego, CA: Elsevier.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520-522.

Tibboel, H., De Houwer, J., & Crombez, G. (2008). The emotional modulation of the attentional blink: Effects of taboo and non-taboo arousing stimuli in a dual and single task paradigm. *International Journal of Psychology, 43*(3-4), 338-338.

Tibboel, H., Van Bockstaele, B., & De Houwer, J. (2011). Is the emotional modulation of the attentional blink driven by response bias? *Cognition & Emotion, 25*(7), 1176-1183.

Tibshirani, R., & Walther, G. (2005). Cluster validation by prediction strength. *Journal of Computational and Graphical Statistics, 14*(3), 511-528.

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan.

Todd, J. T., & Gelder, P. V. (1979). Implications of a transient-sustained dichotomy for the measurement of human performance. *J Exp Psychol Hum Percept Perform, 5*(4), 625-638.

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-biased attention as emotion regulation. *Trends Cogn Sci, 16*(7), 365-372.

Torfs, K., Panis, S., Bartlema, A., & Wagemans, J. (submitted). Temporal dynamics of within-trial, top-down facilitation of object recognition.

Torfs, K., Panis, S., & Wagemans, J. (2010). Identification of fragmented object outlines: A dynamic interplay between different component processes. *Visual Cognition*, *18*(8), 1133-1164.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network-Computation in Neural Systems, 14*(3), 391-412.

Treisman, A. M. (1969). Strategies and Models of Selective Attention. *Psychological Review, 76*(3), 282-&.

Treisman, A. M., & Gelade, G. (1980). Feature-Integration Theory of Attention. *Cognitive Psychology, 12*(1), 97-136.

Treisman, A. M., & Kanwisher, N. G. (1998). Perceiving visually presented objects: recognition, awareness, and modularity. *Current Opinion in Neurobiology, 8*(2), 218-226.

Ullman, S. (1995). Sequence Seeking and Counter Streams: A Computational Model for Bidirectional Information Flow in the Visual Cortex. *Cerebral Cortex, 5*(1), 1-11.

Ullsperger, M., Volz, K. G., & von Cramon, D. Y. (2004). A common neural system signaling the need for behavioral changes. *Trends in Cognitive Sciences, 8*(10), 445-446.

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*. Cambridge, MA: The MIT Press.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol Rev, 108*(3), 550-592.

Valenstein, E., Bowers, D., Verfaellie, M., Heilman, K. M., Day, A., & Watson, R. T. (1987). Retrosplenial Amnesia. *Brain, 110*, 1631-1646.

Van der Ploeg, H., Defares, P. B., & Spielberger, C. D. (1979). *Self Report Questionnaire STAI, Version DY-1 and DY-2*. Lisse, The Netherlands: Swets & Zeitlinger.

VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research, 46*(18), 3017-3027.

VanRullen, R., Reddy, L., & Koch, C. (2004). Visual search and dual tasks reveal two distinct attentional resources. *Journal of Cognitive Neuroscience, 16*(1), 4-14.

VanRullen, R., Reddy, L., & Li, F. F. (2005). Binding is a local problem for natural objects and scenes. *Vision Research, 45*(25-26), 3133-3144.

Vatakis, A., & Spence, C. (2006). Temporal order judgments for audiovisual targets embedded in unimodal and bimodal distractor streams. *Neuroscience Letters, 408*(1), 5-9.

Verleger, R., Smigasiewicz, K., & Moller, F. (2011). Mechanisms underlying the left visual-field advantage in the dual stream RSVP task: evidence from N2pc, P3, and distractor-evoked VEPs. *Psychophysiology, 48*(8), 1096-1106.

Verleger, R., Sprenger, A., Gebauer, S., Fritzmannova, M., Friedrich, M., Kraft, S., et al. (2009). On why left events are the right ones: neural mechanisms underlying the left-hemifield advantage in rapid serial visual presentation. *J Cogn Neurosci, 21*(3), 474-488.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology, 37*(2), 190-203.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology-Human Perception and Performance, 27*(1), 92-114.

Vogt, B. A., Finch, D. M., & Olson, C. R. (1992). Functional Heterogeneity in Cingulate Cortex: The Anterior Executive and Posterior Evaluative Regions. *Cerebral Cortex, 2*(6), 435-443.

Vogt, J., De Houwer, J., Moors, A., Van Damme, S., & Crombez, G. (2010). The automatic orienting of attention to goal-relevant stimuli. *Acta Psychologica, 134*(1), 61-69.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends Cogn Sci, 9*(12), 585-594.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron, 30*(3), 829-841.

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

Vuilleumier, P., & Huang, Y. M. (2009). Emotional Attention: Uncovering the Mechanisms of Affective Biases in Perception. *Current Directions in Psychological Science, 18*(3), 148-152.

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia, 45*(1), 174-194.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., & Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience, 7*(11), 1271-1278.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, in press*.

Wager, T. D., Phan, K. L., Liberzon, I., & Taylor, S. F. (2003). Valence, gender, and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *Neuroimage, 19*(3), 513-531.

Wang, X. J. (2008). Decision Making in Recurrent Neuronal Circuits. *Neuron, 60*(2), 215-234.

Warren, R. M. (1974). Auditory Temporal Discrimination by Trained Listeners. *Cognitive Psychology, 6*(2), 237-256.

Watanabe, M. (2007). Role of anticipated reward in cognitive behavioral control. *Current Opinion in Neurobiology, 17*(2), 213-219.

Watson, D. (2000). *Mood and temperament. Emotions and social behavior*. New York, NY: Guilford Press.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and Validation of Brief Measures of Positive and Negative Affect: The Panas Scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a Consensual Structure of Mood. *Psychological Bulletin, 98*(2), 219-235.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology, 76*(5), 820-838.

Watts, F. N., McKenna, F. P., Sharrock, R., & Trezise, L. (1986). Color Naming of Phobia-Related Words. *British Journal of Psychology, 77*, 97-108.

Weierich, M. R., Wright, C. I., Negreira, A., Dickerson, B. C., & Barrett, L. F. (2010). Novelty as a dimension in the affective brain. *Neuroimage, 49*(3), 2871-2878.

Weiß, K., & Scharlau, I. (2012). At the mercy of prior entry: Prior entry induced by invisible primes is not susceptible to current intentions. *Acta Psychologica, 139*(1), 54-64.

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78*(6), 1024-1037.

West, G. L., Anderson, A. A., & Pratt, J. (2009). Motivationally significant stimuli show visual prior entry: evidence for attentional capture. *J Exp Psychol Hum Percept Perform, 35*(4), 1032-1042.

Whalen, P. J., Kagan, J., Cook, R. G., Davis, F. C., Kim, H., Polis, S., et al. (2004). Human amygdala responsivity to masked fearful eye whites. *Science, 306*(5704), 2061-2061.

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

Wheeler, M. E., Petersen, S. E., Nelson, S. M., Ploran, E. J., & Velanova, K. (2008). Dissociating Early and Late Error Signals in Perceptual Recognition. *Journal of Cognitive Neuroscience, 20*(12), 2211-2225.

Wiens, S., Peira, N., Golkar, A., & Öhman, A. (2008). Recognizing Masked Threat: Fear Betrays, But Disgust You Can Trust. *Emotion, 8*(6), 810-819.

Wiens, S., Sand, A., & Olofsson, J. K. (2011). Nonemotional features suppress early and enhance late emotional electrocortical responses to negative pictures. *Biological Psychology, 86*(1), 83-89.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8*(2), 227-233.

Wilson, T. D., & Bar-Anan, Y. (2008). The unseen mind. *Science, 321*(5892), 1046-1047.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81*(6), 989-1000.

Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. M. K. C. Klauer (Ed.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189-217). Mahwah, NJ: Lawrence Erlbaum.

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*(20), 1824-1829.

Wittling, W., & Roschmann, R. (1993). Emotion-related hemisphere asymmetry: subjective emotional responses to laterally presented films. *Cortex, 29*(3), 431-448.

Wolach, I., & Pratt, H. (2001). The mode of short-term memory encoding as indicated by event-related potentials in a memory scanning task with distractions. *Clinical Neurophysiology, 112*(1), 186-197.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics, 63*(3), 381-389.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495-501.

Woodman, G. F., & Luck, S. J. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature, 400*(6747), 867-869.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII* (pp. 73-103). Cambridge, MA: MIT Press.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: evidence from equiluminant visual objects. *J Exp Psychol Hum Percept Perform, 20*(1), 95-107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. *J Exp Psychol Hum Percept Perform, 10*(5), 601-621.

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition & Emotion, 24*(1), 3-47.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: analysis by synthesis? *Trends in Cognitive Sciences, 10*(7), 301-308.

Zajonc, R. B. (1968). Attitudinal effects of mere exposures. *Journal of Personality and Social Psychology, 9*(2, Pt. 2), 1-27.

Zajonc, R. B. (1980). Feeling and Thinking: Preferences Need No Inferences. *American Psychologist, 35*(2), 151-175.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science, 10*(6), 224-228.

Zampini, M., Brown, T., Shore, D. I., Maravita, A., Roder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica, 118*(3), 277-291.