

Beyond the face: Exploring rapid influences of context on face processing.

Beatrice de Gelder^{1,2,3,*}, Hanneke K.M. Meeren^{1,2,3}, Ruthger Righart¹, Jan Van den Stock¹, Wim A.C. van de Riet^{1,3} and Marco Tamietto⁴

- 1) Cognitive and Affective Neurosciences Laboratory, Tilburg University, The Netherlands.
- 2) Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA.
- 3) F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, NL-6500 HB Nijmegen, The Netherlands.
- 4) Department of Psychology, University of Turin, Italy .

* Corresponding author: Dr. B. de Gelder, Cognitive and Affective Neurosciences Laboratory, Department of Psychology, Tilburg University, P.O. 90153. 5000 LE TILBURG. Tel.: +31 13 466 24 95; Fax.: +31 13 466 2067; E-mail: B.deGelder@uvt.nl

Keywords: fear, body, scene, voice, P1, N170.

Abbreviations

EBA	extrastriate body area
FFA	fusiform face area
LGN	lateral geniculate nucleus
PHC	parahippocampal cortex
OFA	occipital face area
OFC	orbitofrontal cortex
RSC	retrosplenial cortex
SC	superior colliculus
STS	superior temporal sulcus

P1 = an event related potential component with positive deflection occurring at about 100 ms after stimulus onset.

N170 = an event related potential component with negative deflection occurring at about 170 ms after stimulus onset.

Abstract:

Humans optimize behavior by deriving context-based expectations. Contextual data that are important for survival are extracted rapidly, using coarse information, adaptive decision strategies and dedicated neural infrastructure. In the field of object perception, the influence of a surrounding context has been a major research theme, and it has generated a large literature. That visual context, as typically provided by natural scenes, facilitates object recognition has been convincingly demonstrated (Bar, 2004). Just like objects, faces are generally encountered as part of a natural scene. Thus far, the facial expression literature has neglected that context and treats facial expressions as if they stand on their own. This neglect most probably is a mistake. Facial expressions tend to appear in a context of head and body orientations, body movements, posture changes, and other object-related actions with a similar or at least a closely related meaning. For instance, one would expect a frightened face when confronted to an external danger to be at least accompanied by withdrawal movements of head and shoulders. Furthermore, some cues provided by the environment or the context in which a facial expression appears may have a direct relation with the emotion displayed by the face. The brain may even fill in the natural scene context typically associated with the facial expression. Recognition of the facial expression may also profit from processing the vocal emotion as well as the emotional body language that normally accompany it. Here we review the emerging evidence on how the immediate visual and auditory context influence the recognition of facial expressions.

Introduction:

It is surprising that, except for a few isolated studies, the literature on face recognition has not yet addressed the issue of context. So far, much of the face recognition literature has been dominated by the issue of face modularity, or the notion that our ability to process faces reflects a functional and neurobiological specialization. From the point of view of face specificity theorists, face processing may be immune to surrounding context recognition processes because faces are uniquely salient and attention-grabbing signals. If so, context influence may just be another dimension on which face and object processing differ considerably as face processing may not be sensitive to context.

Emotion researchers have predominantly used isolated facial expressions at the detriment of means like posture, movement, and voice or scenes by which emotions are expressed and communicated. This may or may not be justified. On the one hand, facial expression may be very special indeed. For highly social species like humans, facial expressions may be by far the most salient carriers of emotional information, dwarfing objects or natural scenes as cues of emotional significance. In that case, the meaning of facial expressions will be computed automatically irrespective of other cues present together with the face like emotional body language, arousing context or emotional voice expressions. On the other hand, the manner in which we process facial expressions and what behavioral consequences are associated with it is influenced by the context in which a facial expression appears. This may be because facial expressions are often ambiguous and additional signals from the body, the voice and the context are required to compute the meaning and the behavioral consequences of a facial expression.

In this chapter we review recent investigations of three very important contexts in which facial expressions are frequently encountered: whole bodies, natural scenes and emotional voices. Recent shifts in the theoretical perspective in cognitive and affective neurosciences converge on important notions like embodiment, affect programs and multisensory based perception integration and open a perspective in which context may play a crucial role, even for highly automated processes like the ones underlying recognition of facial expressions. We briefly sketch this background before presenting recent findings on the three contexts of face processing we deem essential for an ecologically valid theory of facial expressions.

Background

Since its reintroduction in experimental psychology, emotion research has mainly focused on *visual* processes associated with seeing emotional stimuli, *de facto* facial expressions. Recent findings point to close links between the visual and the sensorimotor system and to the role the body plays in perception as for example shown in research on embodied cognition (Barsalou, 1999). The leading intuition which can now be approached with the novel methods of cognitive and affective neuroscience is that individuals embody the emotional gestures of other people, including facial expressions, posture, and vocal affect. Imitative behavior produces a corresponding state in the perceiver, leading to the general suggestion that embodied knowledge produces corresponding emotional states.

In the early stages of processing core emotions (Ekman, 1992) bodily resonance is automatic and reflexlike while in the later, more cognitive and conscious processing stages it is under strategic control and influenced by higher order knowledge. The notion of embodiment in a more general meaning has also come to the foreground of emotion theories again with the proposals made by Damasio (1994; 1999).

From a more evolutionary inspired perspective, emotions and facial expressions are adaptations (Schmidt and Cohn, 2001) as already envisaged in the Darwinian view that emotions are closely related to actions and therefore likely to involve the whole body. Emotion provoking stimuli trigger affect programs (Darwin, 1872; Frijda, 1986; Panksepp, 1998; Russell and Feldman Barrett, 1999; Tomkins, 1963), which produce an ongoing stream of neurophysiological change (or change in a person's homeostatic state) and are associated with evolutionary-tuned behaviors for dealing with stimuli of significant value. Along with the orbitofrontal cortex (OFC) and amygdala, the insula and somatosensory cortex are involved in the modulation of emotional reactions involving the body via connections to brain stem structures (Damasio, 1994; 1999; LeDoux, 1996). This function of the insula and somatosensory cortex may underlie their important role in emotion perception (Adolphs et al., 2000; Heberlein and Adolphs, 2004; Winston et al., 2003). Processes engaging somatosensory cortex and insula may involve simulating the viewed emotional state via the generation of a somatosensory image of the associated body state.

Recognition of faces and facial expressions

A great deal of effort has been devoted in trying to establish that faces constitute a particular category of stimuli processed with dedicated behavioral skills, in specific cortical areas of the brain and possibly with shorter latencies than other stimuli. In the modular model proposed by Kanwisher (Kanwisher et al., 1997), a small region in the fusiform gyrus, the so-called fusiform face area (FFA) is specialized in face perception (cf. Gauthier and Nelson, 2001; Gauthier et al., 1999; Gauthier et al., 1998). This view seems only to concern the neural basis of personal identity cues as provided by the face and neither the facial expression nor the context in which faces typically appear (body, scene, voice).

These caveats are accounted for in the distributed models for face perception (de Gelder et al., 2003; de Gelder and Rouw, 2000; Haxby et al., 2000; 2002) which also consider other aspects of faces besides person identity (Adolphs, 2002; Adolphs et al., 2000; de Gelder et al., 2003; de Gelder and Rouw, 2000; Haxby et al., 2000; Haxby et al., 1994; Haxby et al., 1996; Hoffman and Haxby, 2000; Puce et al., 1996). In distributed models, different areas of the brain process each different attributes of the face, such as identity (FFA and the occipital face area (OFA)) gaze direction (superior temporal sulcus (STS)) and expression and/or emotion (OFC, amygdala, anterior cingulate cortex, premotor cortex, somatosensory cortex). Several of these structures, e.g. OFC, amygdala and somatosensory cortex have clearly direct and indirect connections with visceral, autonomic and muscular centers (Adolphs, 2002) and thereby influencing the affective homeostasis and making the body part of the perceptual process, i.e. embodiment.

Within this multitude of regions, there is a division of labor. The first route, a subcortical pathway to the amygdala via the superior colliculus (SC) and pulvinar, is concerned with fast and more coarse but subconscious processing (de Gelder et al., 2001; de Gelder et al., 1999b; Morris et al., 2001; Morris et al., 1998b; Pegna et al., 2005) in case of highly salient, especially threatening stimuli, while the second route, via the lateral geniculate nucleus (LGN) and striate cortex to cortical regions like STS, OFA and FFA, is more concerned with detailed and fine-grained processing in case stimuli are ambiguous and full blown awareness of the perceived face is necessary.

These regions of the parallel routes interact (Adolphs, 2002; de Gelder and Rouw, 2000) and modulate each other with feedforward and feedback projections in order to establish a fine-grained

percept composed of identity and emotional aspects of the face, which can be accessible to consciousness. Especially the amygdala has strong functional and structural connections with several cortical regions, such as FFA, STS and OFC (functional connectivity: e.g. Carmichael and Price, 1995; Lidaka et al., 2001; Morris et al., 1998a; Vuilleumier et al., 2004); structural connectivity; e.g. (Carmichael and Price, 1995) or with striate cortex (structural connectivity; Amaral and Price, 1984; Catani et al., 2003).

Electrophysiological studies have shed light on the temporal characteristics of neuronal processing of faces. Two early components that can be readily identified in the waveform of visual event-related potentials (ERP) or magnetic fields (ERF), i.e the P1 and N170, show sensitivity to faces, hinting at that dedicated systems are attuned to the processing of faces. The first component would point to the involvement in global encoding, i.e. categorizing a face as such (Liu et al., 2002), while the second deflection would reflect configural perceptual processing (Bentin et al., 1996).

The face-selective N170 waveform shows a robust face-selective “inversion” effect indicative of configural processing, i.e. it is enhanced and delayed to faces that are presented upside down, but not to inverted objects (Stekelenburg and de Gelder, 2004; Watanabe et al., 2003). Controversy exists about the underlying neuronal source of the N170. Some studies point to the STS as generator (Henson et al., 2003; Itier and Taylor, 2004b), while others propose the fusiform gyrus, where the FFA resides, as possible candidate (Halgren et al., 2000; Pizzagalli et al., 2002; Shibata et al., 2002). Whether the N170 is generated in the fusiform gyrus or STS may depend on the exact nature of the task and the stimuli being used. The N170 amplitude is affected by biological motion (Jokisch et al., 2005), eye gaze (Watanabe et al., 2002), facial motion (Puce et al., 2003), facial expressions (Batty and Taylor, 2003; Stekelenburg and de Gelder, 2004), expressional change (Miyoshi et al., 2004) and affective facial features (Pizzagalli et al., 2002).

Recent studies challenge the N170 as earliest marker of selective face processing and draw attention to an earlier component peaking between 100-130ms post-stimulus. The P1 ERP component (or its MEG equivalent) is mainly generated in “early” extrastriate visual areas (Di Russo et al., 2005; Linkenkaer-Hansen et al., 1998) and it is commonly thought to reflect processing of the low-level features of a stimulus. A few recent studies however suggest that higher-order visual processing can already occur at this early stage. Successful categorization of stimuli as faces was found to correlate with an early MEG component at 100-120ms after onset (Di Russo et al., 2005; Itier and Taylor, 2002;

2004a; Linkenkaer-Hansen et al., 1998; Liu et al., 2002. Both the MEG and ERP component show an inversion effect, suggesting that some configurational processing already takes place at this early stage. In addition, this component appears to be sensitive to facial likeability (Pizzagalli et al., 2002) and emotional facial expressions in contrast to neutral expression, but not between emotional expression (Batty and Taylor, 2003; Eger et al., 2003; Esslen et al., 2004; Halgren et al., 2000).

Facial expressions in the context of whole bodies.

Perception of bodies is a relatively new field as is perception of bodily expressions of emotion. Recent research on neutral and instrumental body postures and movements has set out to raise some of the familiar questions of face researchers. Are the perceptual characteristics of faces and bodies alike? Is one specific brain region dedicated to body perception (modularity hypothesis), or are multiple brain regions involved (distributed model hypothesis). Or does perception of face and body expression share an underlying common neural basis?

Evidence from single cell recordings suggests a degree of specialization for either face or neutral body images (Rizzolatti et al., 1996). This view is corroborated by studies reporting that neurons of monkey posterior STS react selectively to body posture and by the fMRI study of Downing and co-workers (Downing et al., 2001) in which a region near the middle occipital gyrus, the so-called extrastriate body area (EBA), reacted selectively to body form and body parts but showed little activation to isolated faces.

However, a recent electrophysiological investigation in humans lends support for common configural perceptual processing mechanisms for faces and bodies. A typical but slightly faster N170 component commonly obtained for faces was also found for the perception of human bodies (Stekelenburg & de Gelder, 2004). Most interestingly, the N170 showed an inversion effect for bodies, comparable to the inversion effect earlier found for faces (Stekelenburg & de Gelder, 2004).

In the studies of Tamietto and co-workers, the simultaneous presentation to both visual hemifields of two emotionally congruent faces (Tamietto et al., submitted) or two emotionally congruent bodies (Tamietto et al., 2005b) leads to shorter latencies for stimulus detection as compared to the unilateral presentation of the same stimuli to either the left or right hemifield. Additionally, patients with hemineglect and visual extinction, who typically fail to report the presence of a

contralesional stimulus under conditions of bilateral stimulation, could more easily detect a contralesional happy or angry facial expression than a neutral facial expression (Vuilleumier and Schwartz, 2001). This finding was replicated with emotional bodily expressions in the study of Tamietto and colleagues (Tamietto et al., 2005a), in which fearful bodily expression were more easily detected than neutral bodily expressions for the contralesional field. These findings indicate similarities in perceptual properties between faces and bodies, and the ability of emotional biological stimuli to attract attention in unattended visual space.

There appears to be also large similarities between emotional bodily and facial expression at the neural level. A striking finding (Hadjikhani and de Gelder, 2003) is that observing bodily expressions activates two well-known face areas, such as FFA and OFA, predominantly associated with processing faces but also linked with biological movement (Bonda et al., 1996). These activations in face related areas may result from mental imagery (O'Craven and Kanwisher, 2000) or alternatively – and more probably- from context driven high-level perceptual mechanisms filling in the face information missing from the input. However, this is unlikely to be the only explanation for similarities between fearful facial expressions and bodily expressions (cf. Cox et al., 2004)). The finding of Hadjikhani & de Gelder (2003) was supported by the studies of de Gelder and colleagues (de Gelder et al., 2004a), for bodily expressions, and by Peelen & Downing (2005) for neutral body postures.

As there is as of yet no literature on how recognition of facial expression is affected by emotional body contexts we have recently started to explore this critical issue. We used photographs of fearful and angry faces and bodies to create realistically looking face-body compound images, with either matched or mismatched emotional expressions. The emotions fear and anger were selected because they are both emotions with a negative valence and each is associated with evolutionary relevant threat situations. A short stimulus presentation time was used (200 ms), requiring observers to judge the faces on the basis of a “first impression” and to rely on global processing rather than on extensive analysis of separate facial features. Participants attended to the face and made judgments about the facial expression. The recognition of the emotion conveyed by the face was found to be systematically influenced by the emotion expressed by the body (Meeren et al., submitted). Observers made significantly better (81% correct) and faster (774 ms) decisions when faces were accompanied by a matching bodily expression than when a bodily expression did not match the facial expression (67% and 840 ms). The fact that a reliable influence was obtained in an implicit paradigm in which the

bodies were not task relevant nor explicitly attended to suggests that the influence they exercise is rapid and automatic. To further test the automatic processing hypothesis we recorded EEG while subjects performed the task. An enlargement of the occipital P1-component as early as 115 ms after presentation onset was found (Meeren et al., submitted). This points to the existence of an ultrarapid neural mechanism sensitive to the degree of agreement between simultaneously presented facial and bodily emotional expressions, even when the latter are unattended.

Facial expressions in the context of scenes.

Faces routinely appear as part of natural scenes. Hierarchical models of perception tend to assimilate scene effects with semantic effects occurring relatively late at higher cognitive centers (Bar, 2004). However, the processing of objects is influenced by the properties of a scene at an early level. It has been reported that the rapid extraction of the gist of a scene seems to be based on low spatial frequency coding (Oliva and Schyns, 1997) Brief exposure to a known scene activates a representation of its layout that contributes to subsequent processing of spatial relations across the scene (Sanocki, 2003). Segmentation of object from background scenes occurs rapidly, during the first 100 ms of processing (Lamme, 1995), and object detection is faster when presented in an implicitly learned context configurations (Olson et al., 2001). The results support the role of feedback modulations in an early level of processing in animal (Lamme and Roelfsema, 2000) and human studies (Foxy and Simpson, 2002) .

Semantic contexts effects on object processing occur on a much later stage of processing. Objects that are congruent with their context are identified better (Davenport and Potter, 2004) and faster (Ganis and Kutas, 2003) and the interaction occurs at about 390 ms after stimulus-onset (i.e., the N400 component), which is assumed to be a high level of semantic representation of object and scene (Ganis and Kutas, 2003). In an fMRI study it was found that the parahippocampal cortex (PHC) and retrosplenial cortex (RSC) are involved in a system that associates object with contexts (Bar, 2004; Bar and Aminoff, 2003).

The effects of emotional contexts may occur on a much earlier level than semantic effects and may involve different neural systems (Hariri et al., 2002). We recently investigated how emotional visual scenes influence face processing. Event-related potentials were recorded for faces

(fearful/neutral) embedded in scenes (fearful/neutral) while participants performed an orientation-decision task (face upright/inverted). Thus, the task condition was kept irrelevant to the emotion in context and face. Increased structural encoding, as indicated by the N170 response to faces, was found when faces were perceived in a fearful context as opposed to a neutral context (Righart and de Gelder, 2005). This N170 response was even more increased for fearful faces in a fearful context, possibly as a consequence of congruency. Preliminary behavioral data substantiate these congruency effects, as it was found that facial expressions (e.g., a disgust expression) were recognized faster when they were accompanied by a congruent emotional context (for instance, a rubbish dump). A control condition showed that the increased response on the N170 could not be attributed to the exclusive presence of the context, as the amplitudes did not differ between fearful and neutral contexts without a face.

The N170 of faces, particular fearful faces, in a threatening context, may be increased in order to enhance structural encoding. In a potentially dangerous situation it is important to analyze instantly what is happening. The results may be consistent with the model proposed by Haxby (Haxby et al., 2000). Source analysis studies suggest that an enlarged N170 may be indicative of increased activation in fusiform gyrus or STS (Henson et al., 2003; Itier and Taylor, 2004b; Pizzagalli et al., 2002; Shibata et al., 2002). Fearful faces and contexts may activate the amygdala and modulate activity in the fusiform gyrus (Lang et al., 1998; Morris et al., 1998a; Surguladze et al., 2003) and by this way face processing by enhancing the N170 amplitude. Alternatively, activity in the STS has been related to the perception of social cues (Allison et al., 2000). This functional interpretation accords with findings that the N170 amplitude is profoundly affected by biological motion (Jokisch et al., 2005), eye gaze (Watanabe et al., 2002), facial motion (Puce et al., 2003), facial expressions (Batty and Taylor, 2003; Stekelenburg and de Gelder, 2004), expressional change (Miyoshi et al., 2004) and affective facial features (Pizzagalli et al., 2002).

It is not clear yet whether congruent emotions engage the system that includes the PHC, similar to semantic associations between object and context (Bar and Aminoff, 2003). Future studies should determine whether processing of emotional relations in contexts should be distinguished from semantic relations in context. Our recent data show that the time-courses differ (Ganis and Kutas, 2003; Righart and de Gelder, 2005), but no fMRI data are as yet available as to what neural systems are involved in processing faces in emotional contexts.

Further, an interesting question is in what way fearful context do enhance perceptual analysis of faces. According to the model of Bar (Bar, 2004), the gist of the scene is extracted by perception of the low spatial frequencies (Oliva and Schyns, 1997), which provides a rough and quick image of the scene information, on which the high spatial frequencies provide the detailed fill-in of the object. Low spatial frequencies in the face increase amygdala responses for fearful faces (Vuilleumier et al., 2003). If low spatial frequencies are important for fear processing in general, then the model could explain why N170 amplitudes are increased for faces in fearful contexts. In such a model, fearful contexts may provide a first coarse template on which the perceptual analysis faces is interpreted.

Enhanced N170 amplitudes for faces in fearful contexts may be related to enhanced encoding of identity, which may improve recognition memory for faces. Data of prosopagnosia patients indicate that impaired structural encoding, as reflected in the N170, may disrupt facial identification (de Gelder and Stekelenburg, 2005; Eimer, 2000), and that facial expressions may improve their performance on face recognition (de Gelder et al., 2003). It has already been shown that object recognition memory (e.g., tools, furniture, clothing) is better for object that were presented in a positive valenced context background than in a negative or neutral valenced background (Smith et al., 2004). Similar increases in accuracy were obtained for words primed by positive background (Erk et al., 2003). An interesting question is whether recognition memory is also improved for faces that are presented in emotional contexts.

Facial expressions in the context of voices.

Human cognition and emotion researches tend to focus on how organisms process information from one sensory system at a time (usually the visual system) but information processing in everyday life is typically multisensory. In many higher species communication involves multiple sensory systems often in combination. Animal researchers are traditionally more interested in co-occurring behavioral signals and a number of studies have explored the close link between vocal and visual communication (Parr, 2004) and discovered synergies between the evolutionary history and the functionality of visual and auditory communication signals (Cooper and Goller, 2004). Audiovisual vocalizations are ethologically relevant and thus may tap into specialized neural mechanisms (Ghazanfar and Santos, 2004).

As stated above, the traditional emphasis is on visual processes, foremost facial expressions. In comparison with the processing of facial expressions, there have been only a few attempts to identify the specific neural sites for processing emotions in the voice (de Gelder et al., 2004b; George et al., 1996; Ross, 2000). Available research shows that listeners can readily recognize a speaker's emotion from his tone of voice. Rapid recognition of affect in auditory expressions happens within the first 100-150 ms of stimulus presentation (Bostanov and Kotchoubey, 2004; Goydke et al., 2004) and is based primarily on voice characteristics.

The ability to decode emotional cues in prosody and facial expressions may have a common processing and/or representational substrate in the human brain (Borod et al., 2000; de Gelder and Bertelson, 2003; Pourtois et al., 2002), facilitating processing and integration of these distinct but often calibrated sources of information. Most studies on multi-sensory emotion perception have focused on the integration of facial expression with information in the voice (de Gelder and Vroomen, 2000; de Gelder et al., 1999a; Massaro and Egan, 1996; Pourtois et al., 2000). Judging the emotional state of a speaker is possible via facial or vocal cues (Banse and Scherer, 1996; Scherer et al., 1991) alone but both judgment accuracy and speed seem to benefit from combining the modalities, e.g., response accuracy increases and response speed decreases when a face is paired with a voice expressing the same emotion. This improvement of performance occurs even when participants are instructed to ignore the voice and rate only the face, suggesting that extracting affective information from a voice may be automatic and/or mandatory (de Gelder and Vroomen, 2000). The fact that prosodic and facial expressions of emotion frequently correlate suggests that the underlying cognitive mechanisms are highly sensitive to shared associations activated by cues in each channel (de Gelder et al., 1999a; Massaro et al., 1996).

To assess how emotional judgments of the face are biased by prosody Massaro and Egan (1996) presented computer-generated faces expressing a *happy*, *angry*, or *neutral* emotion accompanied a word spoken in one of the three emotional tones. De Gelder and Vroomen (2000) presented photographs taken from the Ekman and Friesen (1976) series with facial expressions rendered emotionally-ambiguous by "morphing" the expressions between *happy* and *sad* as the two endpoints. The emotional prosody tended to facilitate how accurately and quickly subjects rate an emotionally *congruent as compared to an incongruent* face. These findings indicate that the emotional value of prosody-face events is registered and somehow integrated during perceptual tasks, affecting

behavioral responses according to the emotion congruity of the combined events. Moreover, these cross-modal influences appear to be resistant to increased attentional demands induced by a dual task implying that combining the two forms of input may be mandatory (Vroomen et al., 2001) . The conclusion of mandatory integration is now considerably strengthened in a study using patients who could recognize facial expressions without being aware of the visual stimuli presented (hemianopic patients suffering from loss of primary visual cortex exhibiting affective blindsight) (de Gelder et al., 2005; de Gelder et al., 2002) .

Our current knowledge of bimodal integration of visual and auditory primate vocal signals in the brain is derived almost exclusively from human neuroimaging studies of audiovisual speech. STS and superior temporal gyrus are consistently activated by bimodal speech signals and often show enhanced activity over unimodal induced signals (Callan et al., 2003; Calvert et al., 2000; Stein and Meredith, 1993) but audiovisual perception of ecologically valid stimuli may not follow the rules derived from firing patterns of cells with audiovisual receptive fields and superadditivity may not be the correct criterion (de Gelder and Bertelson, 2003; Ghazanfar and Santos, 2004) .

A few studies have explored brain areas involved in processing faces in the context of emotional voices. The classical candidate is multisensory convergence in heteromodal cortex (Mesulam, 1998) . Cortical areas like STS (Baylis et al., 1987) and ventral premotor cortex (Kohler et al., 2002) appear to play an important role. A recent study in rhesus monkeys has confirmed such integration in the STS at the level of single units for biologically meaningful actions (Barraclough et al., 2005) . In a PET-study, we found enhanced activity for bimodal stimuli compared to unimodal stimuli situated in the left lateral temporal cortex. Separate analysis for positive and negative emotions showed supplementary convergence area's anteriorly in the left and right hemisphere respectively (Pourtois et al., 2005) . Subcortical audiovisual emotion convergence sites have been found in the amygdala and SC in fMRI-studies (de Gelder et al., 2005; Dolan et al., 2001) . These subcortical nuclei might play a more important role than hitherto expected in part also because of their role in orienting to novel and highly significant stimuli in the environment.

Information about time course may be more critical than anything else to clarify processing properties. All our EEG studies so far (de Gelder, 2005; de Gelder et al., 1999b; Pourtois et al., 2000; Pourtois et al., 2002) point in the direction of early interaction between the facial expression and the emotion in the voice.

Conclusions

Recent data show that different types of context influence the recognition of facial expression. When a face is accompanied by a body or voice expressing the same emotion, or when it is presented in a congruent emotional scene, the recognition of facial expression typically improves, i.e. both the judgment accuracy and speed increase. Hence, both the immediate visual and auditory contexts function to disambiguate the signals of facial expression. Our behavioral and electrophysiological data suggest that this perceptual integration of information does not require high-level semantic analysis occurring relatively late at higher cognitive centers. Instead, the integration appears to be an automatic and mandatory process, which takes place very early in the processing stream, before full structural encoding of the stimulus and conscious awareness of the emotional expression is established.

References

- Adolphs, R. (2002) Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav Cognit Neurosci Rev* **1**, 21-61.
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G. and Damasio, A. R. (2000) A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *J Neurosci* **20**, 2683-90.
- Allison, T., Puce, A. and McCarthy, G. (2000) Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences* **4**, 267-278.
- Amaral, D. G. and Price, J. L. (1984) Amygdalo-cortical projections in the monkey (*Macaca fascicularis*). *J Comp Neurol* **230**, 465-96.
- Banse, R. and Scherer, K. R. (1996) Acoustic profiles in vocal emotion expression. *J Pers Soc Psychol* **70**, 614-36.
- Bar, M. (2004) Visual objects in context. *Nat Rev Neurosci* **5**, 617-29.
- Bar, M. and Aminoff, E. (2003) Cortical analysis of visual context. *Neuron* **38**, 347-58.

- Barraclough, N. E., Xiao, D., Baker, C. I., Oram, M. W. and Perrett, D. I. (2005) Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. *J Cogn Neurosci* **17**, 377-91.
- Barsalou, L. W. (1999) Perceptual symbol systems. *Behav Brain Sci* **22**, 577-609; discussion 610-60.
- Batty, M. and Taylor, M. J. (2003) Early processing of the six basic facial emotional expressions. *Cognitive Brain Research* **17**, 613-620.
- Baylis, G. C., Rolls, E. T. and Leonard, C. M. (1987) Functional subdivisions of the temporal lobe neocortex. *J Neurosci* **7**, 330-42.
- Bentin, S., Allison, T., Puce, A., Perez, E. and McCarthy, G. (1996) Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience* **8**, 551-565.
- Bonda, E., Petrides, M., Ostry, D. and Evans, A. (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J Neurosci* **16**, 3737-44.
- Borod, J. C., Rorie, K. D., Pick, L. H., Bloom, R. L., Andelman, F., Campbell, A. L., Obler, L. K., Tweedy, J. R., Welkowitz, J. and Sliwinski, M. (2000) Verbal pragmatics following unilateral stroke: emotional content and valence. *Neuropsychology* **14**, 112-24.
- Bostanov, V. and Kotchoubey, B. (2004) Recognition of affective prosody: continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology* **41**, 259-68.
- Callan, D. E., Jones, J. A., Munhall, K., Callan, A. M., Kroos, C. and Vatikiotis-Bateson, E. (2003) Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport* **14**, 2213-8.
- Calvert, G. A., Campbell, R. and Brammer, M. J. (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* **10**, 649-57.
- Carmichael, S. T. and Price, J. L. (1995) Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J Comp Neurol* **363**, 615-641.
- Catani, M., Jones, D. K., Donato, R. and Ffytche, D. H. (2003) Occipito-temporal connections in the human brain. *Brain* **126**, 2093-107.
- Cooper, B. G. and Goller, F. (2004) Multimodal signals: enhancement and constraint of song motor patterns by visual display. *Science* **303**, 544-6.

Cox, D., Meyers, E. and Sinha, P. (2004) Contextually evoked object-specific responses in human visual cortex. *Science* **304**, 115-7.

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

Damasio, A. R. (1999) *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Hartcourt Brace.

Darwin (1872) *The expression of the emotion in man and animals*. London: J. Murray.

Davenport, J. L. and Potter, M. C. (2004) Scene consistency in object and background perception. *Psychol Sci* **15**, 559-64.

de Gelder, B. (2005) Non-conscious emotions: new findings and novel perspectives. In: *Emotion: Conscious and Unconscious Processes*. In press. Eds. L. F. Barret, P. Niedenthal and P. Winkielman.

de Gelder, B. and Bertelson, P. (2003) Multisensory integration, perception and ecological validity. *Trends Cogn Sci* **7**, 460-7.

de Gelder, B., Bocker, K. B., Tuomainen, J., Hensen, M. and Vroomen, J. (1999a) The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci Lett* **260**, 133-6.

de Gelder, B., Frissen, I., Barton, J. and Hadjikhani, N. (2003) A modulatory role for facial expressions in prosopagnosia. *Proc Natl Acad Sci U S A*.

de Gelder, B., Morris, J. S. and Dolan, R. J. (2005) Two emotions in one brain: interactions between consciously and unconsciously processed facial expressions, emotional scenes and voices. *in preparation*.

de Gelder, B., Pourtois, G., van Raamsdonk, M., Vroomen, J. and Weiskrantz, L. (2001) Unseen stimuli modulate conscious visual experience: evidence from inter-hemispheric summation. *Neuroreport* **12**, 385-91.

de Gelder, B., Pourtois, G. and Weiskrantz, L. (2002) Fear recognition in the voice is modulated by unconsciously recognized facial expressions but not by unconsciously recognized affective pictures. *Proc Natl Acad Sci U S A* **99**, 4121-6.

de Gelder, B. and Rouw, R. (2000) Configural face processes in acquired and developmental prosopagnosia: evidence for two separate face systems? *Neuroreport* **11**, 3145-50.

de Gelder, B., Snyder, J., Greve, D., Gerard, G. and Hadjikhani, N. (2004a) Fear fosters flight: A mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proc Natl Acad Sci U S A* **101**, 16701-6.

de Gelder, B. and Stekelenburg, J. J. (2005) Naso-temporal asymmetry of the N170 for processing faces in normal viewers but not in developmental prosopagnosia. *Neuroscience letters* **376**, 40-45.

de Gelder, B. and Vroomen, J. (2000) The perception of emotions by ear and by eye. *Cognition & Emotion* **14**, 289-311.

de Gelder, B., Vroomen, J. and Pourtois, G. (2004b) Multisensory perception of affect, its time course and its neural basis. In: *Handbook of multisensory processes.*, pp. 581-596. Eds. G. Calvert, C. Spence and B. E. Stein. MIT: Cambridge, MA.

de Gelder, B., Vroomen, J., Pourtois, G. and Weiskrantz, L. (1999b) Non-conscious recognition of affect in the absence of striate cortex. *Neuroreport* **10**, 3759-63.

Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D. and Hillyard, S. A. (2005) Identification of the neural sources of the pattern-reversal VEP. *Neuroimage* **24**, 874-86.

Dolan, R. J., Morris, J. S. and de Gelder, B. (2001) Crossmodal binding of fear in voice and face. *Proc Natl Acad Sci U S A* **98**, 10006-10.

Downing, P. E., Jiang, Y., Shuman, M. and Kanwisher, N. (2001) A cortical area selective for visual processing of the human body. *Science* **293**, 2470-3.

Eger, E., Jedynak, A., Iwaki, T. and Skrandies, W. (2003) Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. *Neuropsychologia* **41**, 808-17.

Eimer, M. (2000) Event-related brain potentials distinguish processing stages involved in face perception and recognition. *Clin Neurophysiol* **111**, 694-705.

Ekman, P. (1992) An argument for basic emotions. *Cognition and Emotion* **6**, 169-200.

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

Erk, S., Kiefer, M., Grothe, J., Wunderlich, A. P., Spitzer, M. and Walter, H. (2003) Emotional context modulates subsequent memory effect. *Neuroimage* **18**, 439-47.

- Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K. and Lehmann, D. (2004) Brain areas and time course of emotional processing. *Neuroimage* **21**, 1189-203.
- Foxe, J. J. and Simpson, G. V. (2002) Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Exp Brain Res* **142**, 139-50.
- Frijda, N. H. (1986) *The emotions*, Cambridge, Cambridge University Press.
- Ganis, G. and Kutas, M. (2003) An electrophysiological study of scene effects on object identification. *Brain Res Cogn Brain Res* **16**, 123-44.
- Gauthier, I. and Nelson, C. A. (2001) The development of face expertise. *Curr Opin Neurobiol* **11**, 219-24.
- Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P. and Gore, J. C. (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci* **2**, 568-73.
- Gauthier, I., Williams, P., Tarr, M. J. and Tanaka, J. (1998) Training 'greeble' experts: a framework for studying expert object recognition processes. *Vision Res* **38**, 2401-28.
- George, M. S., Parekh, P. I., Rosinsky, N., Ketter, T. A., Kimbrell, T. A., Heilman, K. M., Herscovitch, P. and Post, R. M. (1996) Understanding emotional prosody activates right hemisphere regions. *Arch Neurol* **53**, 665-70.
- Ghazanfar, A. A. and Santos, L. R. (2004) Primate brains in the wild: the sensory bases for social interactions. *Nat Rev Neurosci* **5**, 603-16.
- Goydke, K. N., Altenmuller, E., Moller, J. and Munte, T. F. (2004) Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Brain Res Cogn Brain Res* **21**, 351-9.
- Hadjikhani, N. and de Gelder, B. (2003) Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr Biol* **13**, 2201-5.
- Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V. and Hari, R. (2000) Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb Cortex* **10**, 69-81.
- Hariri, A. R., Tessitore, A., Mattay, V. S., Fera, F. and Weinberger, D. R. (2002) The amygdala response to emotional stimuli: A comparison of faces and scenes. *Neuroimage* **17**, 317-323.

Haxby, J. V., Hoffman, E. A. and Gobbini, M. I. (2000) The distributed human neural system for face perception. *Trends Cogn Sci* **4**, 223-233.

Haxby, J. V., Hoffman, E. A. and Gobbini, M. I. (2002) Human neural systems for face recognition and social communication. *Biol Psychiatry* **51**, 59-67.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P. and Grady, C. L. (1994) The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J Neurosci* **14**, 6336-53.

Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I. and Grady, C. L. (1996) Face encoding and recognition in the human brain. *Proc Natl Acad Sci U S A* **93**, 922-7.

Heberlein, A. S. and Adolphs, R. (2004) Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 7487-7491.

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

Hoffman, E. A. and Haxby, J. V. (2000) Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nat Neurosci* **3**, 80-4.

Iidaka, T., Omori, M., Murata, T., Kosaka, H., Yonekura, Y., Okada, T. and Sadato, N. (2001) Neural interaction of the amygdala with the prefrontal and temporal cortices in the processing of facial expressions as revealed by fMRI. *J Cogn Neurosci* **13**, 1035-47.

Itier, R. J. and Taylor, M. J. (2002) Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: a repetition study using ERPs. *Neuroimage* **15**, 353-72.

Itier, R. J. and Taylor, M. J. (2004a) Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *Neuroimage* **21**, 1518-32.

Itier, R. J. and Taylor, M. J. (2004b) Source analysis of the N170 to faces and objects. *Neuroreport* **15**, 1261-5.

Jokisch, D., Daum, I., Suchan, B. and Troje, N. F. (2005) Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. *Behavioural Brain Research*.

- Kanwisher, N., McDermott, J. and Chun, M. M. (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* **17**, 4302-11.
- Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V. and Rizzolatti, G. (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science* **297**, 846-8.
- Lamme, V. A. (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* **15**, 1605-15.
- Lamme, V. A. and Roelfsema, P. R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* **23**, 571-9.
- Lang, P. J., Bradley, M. M., Fitzsimmons, J. R., Cuthbert, B. N., Scott, J. D., Moulder, B. and Nangia, V. (1998) Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology* **35**, 199-210.
- LeDoux (1996) *The emotional brain: The mysterious underpinnings of emotional life*. New York: Simon and Schuster.
- Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J. and Ilmoniemi, R. J. (1998) Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neurosci Lett* **253**, 147-50.
- Liu, J., Harris, A. and Kanwisher, N. (2002) Stages of processing in face perception: an MEG study. *Nat Neurosci* **5**, 910-6.
- Massaro, D. W., Cohen, M. M. and Smeele, P. M. (1996) Perception of asynchronous and conflicting visual and auditory speech. *J Acoust Soc Am* **100**, 1777-86.
- Massaro, D. W. and Egan, P. B. (1996) Perceiving Affect from the Voice and the Face. *Psychonomic Bulletin and Review* **3**, 215-221.
- Mesulam, M. M. (1998) From sensation to cognition. *Brain* **121 (Pt 6)**, 1013-52.
- Miyoshi, M., Katayama, J. and Morotomi, T. (2004) Face-specific N170 component is modulated by facial expressional change. *Neuroreport* **15**, 911-4.
- Morris, J. S., DeGelder, B., Weiskrantz, L. and Dolan, R. J. (2001) Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain* **124**, 1241-52.

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J. and Dolan, R. J. (1998a) A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain* **121** (Pt 1), 47-57.

Morris, J. S., Öhman, A. and Dolan, R. J. (1998b) Conscious and unconscious emotional learning in the human amygdala. *Nature* **393**, 467-70.

O'Craven, K. M. and Kanwisher, N. (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* **12**, 1013-23.

Oliva, A. and Schyns, P. G. (1997) Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognit Psychol* **34**, 72-107.

Olson, I. R., Chun, M. M. and Allison, T. (2001) Contextual guidance of attention: human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain* **124**, 1417-25.

Panksepp, J. (1998) *Affective neuroscience: The foundation of human and animal emotions*. New York, Oxford University Press.

Parr, L. A. (2004) Perceptual biases for multimodal cues in chimpanzee (*Pan troglodytes*) affect recognition. *Anim Cogn* **7**, 171-8.

Peelen, M. V. and Downing, P. E. (2005) Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* **93**, 603-8.

Pegna, A. J., Khateb, A., Lazeyras, F. and Seghier, M. L. (2005) Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat Neurosci* **8**, 24-5.

Pizzagalli, D. A., Lehmann, D., Hendrick, A. M., REGARD, M., Pascual-Marqui, R. D. and Davidson, R. J. (2002) Affective judgments of faces modulate early activity (not similar to 160 ms) within the fusiform gyri. *Neuroimage* **16**, 663-677.

Pourtois, G., de Gelder, B., Bol, A. and Crommelinck, M. (2005) Perception of facial expressions and voices and of their combination in the human brain. *Cortex* **41**, 49-59.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B. and Crommelinck, M. (2000) The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport* **11**, 1329-33.

- Pourtois, G., Debatisse, D., Despland, P. A. and de Gelder, B. (2002) Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Res Cogn Brain Res* **14**, 99-105.
- Puce, A., Allison, T., Asgari, M., Gore, J. C. and McCarthy, G. (1996) Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *J Neurosci* **16**, 5205-15.
- Puce, A., Syngeniotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J. and Castiello, U. (2003) The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. *Neuroimage* **19**, 861-869.
- Righart, R. and de Gelder, B. (2005) Contexts influence early perceptual analysis of faces. An electrophysiological study. *In preparation*.
- Rizzolatti, G., Fadiga, L., Gallese, V. and Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res* **3**, 131-41.
- Ross, E. D. (2000) Affective prosody and the aprosodias. In M. Mesulam (Ed.): *Principles of behavioral and cognitive neurology*. Oxford, England: Oxford University Press., 316-331.
- Russell, J. A. and Feldman Barrett, L. (1999) Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of Personality and Social Psychology* **76**, 805-819.
- Sanocki, T. (2003) Representation and perception of scenic layout. *Cognit Psychol* **47**, 43-86.
- Scherer, K. R., Banse, R., Wallbott, H. G. and Goldbeck, T. (1991) Vocal Cues in Emotion Encoding and Decoding. *Motivation and Emotion* **15**, 123-148.
- Schmidt, K. L. and Cohn, J. F. (2001) Human facial expressions as adaptations: Evolutionary questions in facial expression research. *Am J Phys Anthropol Suppl* **33**, 3-24.
- Shibata, T., Nishijo, H., Tamura, R., Miyamoto, K., Eifuku, S., Endo, S. and Ono, T. (2002) Generators of visual evoked potentials for faces and eyes in the human brain as determined by dipole localization. *Brain Topogr* **15**, 51-63.
- Smith, A. P. R., Henson, R. N. A., Dolan, R. J. and Rugg, M. D. (2004) fMRI correlates of the episodic retrieval of emotional contexts. *Neuroimage* **22**, 868-878.
- Stein, B. E. and Meredith, M. A. (1993) *The merging of the senses*. MIT Press: Cambridge, MA.

Stekelenburg, J. J. and de Gelder, B. (2004) The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport* **15**, 777-780.

Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C. and Phillips, M. L. (2003) A preferential increase in the extrastriate response to signals of danger. *Neuroimage* **19**, 1317-28.

Tamietto, M., de Gelder, B., Genero, R. and Geminiani, G. (2005a) Automatic encoding of emotional body language in spatial neglect and visual extinction. Paper presented at the 4th Dutch Endo-Neuro-Psycho Meeting, Doorwerth, The Netherlands, 31 May-3 June.

Tamietto, M., Geminiani, G. and de Gelder, B. (2005b) Inter-hemispheric interaction for bodily emotional expressions: Is the right hemisphere superiority related to facial rather than emotional processing? In: *Proceedings of the XVIII European Conference on Visual Perception Annual Meeting, A Coruña, Spain, , August 22-26*.

Tamietto, M., Latini Corazzini, L., de Gelder, B. and Geminiani, G. (submitted) Functional asymmetry and interhemispheric cooperation in the perception of emotions from facial expressions.

Tomkins, S. S. (1963) *Affect, imagery consciousness: Vol. 2. The negative effects*. New York: Springer verlag.

Vroomen, J., Driver, J. and de Gelder, B. (2001) Is cross-modal integration of emotional expressions independent of attentional resources? *Cogn Affect Behav Neurosci* **1**, 382-7.

Vuilleumier, P., Armony, J. L., Driver, J. and Dolan, R. J. (2003) Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nat Neurosci* **6**, 624-31.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J. and Dolan, R. J. (2004) Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat Neurosci* **7**, 1271-8.

Vuilleumier, P. and Schwartz, S. (2001) Emotional facial expressions capture attention. *Neurology* **56**, 153-8.

Watanabe, S., Kakigi, R. and Puce, A. (2003) The spatiotemporal dynamics of the face inversion effect: a magneto- and electro-encephalographic study. *Neuroscience* **116**, 879-95.

Watanabe, S., Miki, K. and Kakigi, R. (2002) Gaze direction affects face perception in humans. *Neuroscience Letters* **325**, 163-166.

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