

Chapter 11

Cross-Modal Modulation of Spatial Attention by Emotion

Tobias Brosch and Didier Grandjean

Abstract A huge amount of environmental stimulus input constantly enters the brain via the different sensory channels of the organism. Due to its limited capacity, the brain cannot process all the inputs exhaustively, and thus needs to select a subset of stimuli for further processing at the cost of others. Emotional stimuli, for example social signals such as angry faces or happy voices, are privileged in the competition for attentional processing resources: the neural representation of emotional stimuli is stronger and more robust compared to neutral stimuli; emotional stimuli are prioritized in perception, draw attention more quickly, and impede attentional disengagement longer than neutral stimuli. The representations of emotional stimuli are thus intensified at different stages of processing. This generates a vivid conscious percept allowing organisms to prepare and implement adequate responses. By modulating frontoparietal attention systems, emotional stimuli also impact on the perceptual processing of subsequent stimuli appearing at the same location as an emotional stimulus. Such neurocognitive selection mechanisms thus drastically reorganize our representation and perception of the environment by focusing on emotionally and motivationally relevant events and their immediate spatial and temporal periphery.

Until now, the effects of emotional stimuli on attentional processes have mainly been described within a sensory modality, most frequently using pictures of emotional stimuli to modulate visuospatial attention. However, in real-life situations humans typically encounter simultaneous input to several different senses, such as vision, audition, olfaction, and touch. Signals entering these different channels might originate from a common, emotionally relevant source. To receive maximal benefit

T. Brosch

New York University, Department of Psychology, 6 Washington Place, New York, NY 10003, USA
e-mail: tobias.brosch@nyu.edu

D. Grandjean (✉)

Department of Psychology, Centre Interfacultaire en Sciences Affectives (NCCR),
University of Geneva, 7 rue des Batoirs, 1205 Geneva, Switzerland
e-mail: didier.grandjean@unige.ch

from multimodal sensory input, the brain must coordinate the input appropriately so that signals from a relevant common source are rapidly processed and integrated across the different input channels to allow for the preparation and implementation of adaptive responses.

We review the current evidence for cross-modal modulation of spatial attention by emotional information. Presenting data from behavioral and electrophysiological investigations in human subjects we illustrate, for example, the effects of emotional voices on visual attention and the effects of emotional images on haptic attention. The data converge to show that emotion modulates attentional processing across sensory modalities by boosting early sensory stages of processing, potentially implemented by a large-scale neural network centered around the amygdala, providing direct and indirect top-down signals to sensory pathways and frontoparietal pathways involved in exogenous and endogenous attentional selection processes.

This rapid cross-modal integration at multiple stages of processing may reflect a fundamental principle of human brain organization: to prioritize the processing of emotionally relevant stimuli, even if they are outside the focus of spatial attention, thus facilitating the multimodal assessment of emotionally relevant stimuli in the environment.

1 Introduction

Our environment constantly confronts us with large amounts of information. Due to capacity limits of the brain, we cannot process all the information entering our senses thoroughly, but have to select important information and prioritize its processing at the cost of other, less relevant information. This competition for neural processing capacity is driven by attentional mechanisms (Driver, 2001) which are influenced by several factors, related to the current needs and goals of the observer (*endogenous attention*) as well as to basic physical properties of the stimulus (*exogenous attention*). In addition, the emotional relevance of a stimulus constitutes an important selection criterion for prioritized processing. Efficient processing of emotional stimuli is highly adaptive, as emotion highlights the relevance of a stimulus for the well-being and survival of the organism (Scherer, 2001). Emotional stimuli should thus be noticed readily and, once detected, become the focus of attention, evaluation, and action. It has been suggested that dedicated neural circuits may underlie the prioritization of emotional stimuli (*emotional attention*, Vuilleumier, 2005). The amygdala, a limbic region critically involved in the processing of emotional information (LeDoux, 2000; Phelps, 2006; Sander, Grafman, & Zalla, 2003), is thought to play a critical role by modulating the processing of incoming sensory stimuli through direct feedback projections to sensory cortex and subsequent biasing signals to frontoparietal attention regions.

Up to now, most studies investigating the preferential role of emotional stimuli in attention and perception have examined *within-modality* effects, most frequently using pictures of emotional stimuli to modulate visual attention. However, humans

typically encounter simultaneous input to several different senses, such as vision, audition, olfaction, and touch. Signals entering these different channels might originate from a common emotionally relevant source, requiring mechanisms for the integration of information conveyed by multiple sensory channels. This integration allows for a more detailed and efficient representation of the world than any single modality in isolation, as it may capitalize on the individual strengths of the different modalities. For example, audition covers a larger spatial area than vision. The rapid detection of an emotionally arousing sound may subsequently lead to an increased allocation of visual attention toward the spatial source of the sound, allowing for a more thorough analysis of the situation based on visual input.

In this chapter, we review the literature investigating cross-modal modulations of attention by emotional information. We first summarize research on the effects and mechanisms of exogenous and endogenous attention selection within and across modalities. We then highlight the special role of emotional information in attention and perception, reviewing both behavioral evidence and evidence from neuroimaging. We conclude by presenting a neurocognitive model describing the mechanisms underlying cross-modal emotional attention.

2 Mechanisms of Attentional Selection: Endogenous and Exogenous Attention

Not all incoming environmental stimulation can be processed in parallel and evaluated thoroughly due to capacity limits of the human brain (Marois & Ivanoff, 2005). To allow for a rapid and efficient analysis of behaviorally important information in the environment, dedicated attention systems therefore serve to select a subset of all incoming stimuli for more in-depth processing and preferential access to conscious awareness (Driver, 2001). Attentional prioritization leads to preferential processing via increases in sensory gain (Hillyard, Vogel, & Luck, 1998), as evidenced by perceptual enhancement such as faster stimulus detection (Posner, 1980) or increases in contrast sensitivity (Carrasco, Ling, & Read, 2004). Attentional selection can be guided by stimulus-related and by observer-dependent effects. Distinct functional subprocesses related to different selection criteria have been put forward, and their respective properties and contributions to attentional selection mechanisms have been isolated using both behavioral and brain-imaging methods. *Exogenous attention* refers to effects driven by the intrinsic physical salience of sensory inputs (Egeth & Yantis, 1997; Theeuwes, 1991; Wolfe & Horowitz, 2004). Low-level properties such as stimulus intensity, color, or size may trigger an involuntary, stimulus-driven, bottom-up attention process. Experimentally, this form of attentional selection has been demonstrated using the exogenous cueing paradigm (Posner, 1980), where participants have to indicate the location of a target that appears either at the same location as a previous exogenous cue (e.g., a bright flash) or at the opposite location. Importantly, the cue is nonpredictive of the target location, i.e., in 50% of the trials the target replaces the cue (valid trials), in 50% of the trials it appears at

the opposite location (invalid trials). Faster responses to targets in valid trials indicate exogenous attention capture by the cue. This effect has been demonstrated within the visual (Posner, 1980), the auditory (Spence & Driver, 1994), and the tactile modality (Miles, Poliakoff, & Brown, 2008). Furthermore, cross-modal cueing studies have demonstrated that directing exogenous attention to a stimulus in one modality (e.g., with a nonpredictive sound) facilitates the speed of responding of spatially coincident stimuli in another modality (e.g., towards a visual or a tactile target). This cross-modal facilitation has been observed for all combinations of visual, auditory, and tactile stimuli (see Driver & Spence, 1998; Koelewijn, Bronkhorst, & Theeuwes, 2010, for reviews). Some asymmetries have been observed related to the modality of the cue: visual cues lead to a narrower focusing of the attentional field in which facilitation is achieved compared to auditory cues, an effect that may be related to the different spatial resolutions of the different sensory modalities (Spence, 2010). In contrast to the reflexive exogenous attention mechanisms, *endogenous attention* refers to a voluntary top-down process, initiated by implicit or explicit expectations for a specific object or location (Desimone & Duncan, 1995; Posner, Snyder, & Davidson, 1980). This process selects stimuli important to the current behavior and goals of the organism. This form of attentional selection has been demonstrated using the endogenous cueing task (Posner et al., 1980), in which a centrally presented arrow indicates the location where a subsequent target stimulus will probably appear, thus creating an expectation in the participants. Faster responses to validly cued targets (i.e., targets that appear at the location indicated by the arrow) reflect voluntary endogenous attention shifts. Again, this effect has been demonstrated for the visual (Posner et al., 1980), auditory (Spence & Driver, 1994), and tactile modalities (Lloyd, Bolanowski, Howard, & McGlone, 1999). Furthermore, cross-modal cueing studies have demonstrated that directing endogenous attention to one modality (e.g., creating an expectation for a sound at a specific location) facilitates the speed of responding of spatially coincident stimuli in another modality (e.g., towards a visual or a tactile target). Again, this effect has been observed for all combinations of visual, auditory and tactile stimuli (see Koelewijn et al., 2010, for a review). Besides expectations for target locations, endogenous attention can be directed toward and improve detection of other features of potential target objects such as shape, color, or direction of motion (Rossi & Paradiso, 1995) or towards complete objects (Yantis, 1992). First evidence for cross-modal object-based attention has been presented recently (Turatto, Mazza, & Umiltà, 2005), demonstrating that auditory objects may affect the deployment of visual attention.

According to a recent neurocognitive model of attention, both endogenous and exogenous attention primarily implicate frontoparietal networks of cortical regions (Corbetta, Patel, & Shulman, 2008; Corbetta & Shulman, 2002; see also Peelen, Heslenfeld, & Theeuwes, 2004), with endogenous attention control being exerted by interactions of dorsal regions such as the intraparietal sulcus (IPS) and the frontal eye fields (FEF), and exogenous reorienting of the attentional focus mediated by more ventral regions in the right hemisphere such as the right ventral frontal cortex (VFC) and temporoparietal junction (TPJ). Even though most neuroimaging data

investigating these two attentional networks have been collected in the visual modality, the available evidence supports a supramodal function. The ventral network is sensitive to salient events in the visual, auditory, and tactile modality, and similar ventral and dorsal frontoparietal regions are modulated by reorienting in different modalities (Corbetta et al., 2008; Eimer & Driver, 2001).

ERP studies measuring the neural effects of cross-modal endogenous and exogenous attention suggest that attentional facilitation effects are operating at early perceptual stages. Cross-modal attentional modulations affect early modality-specific ERP components (up to 200 ms after target onset), but show smaller or no effects at later components linked to post-perceptual stages (later than 200 ms, see Eimer & Driver, 2001, for a review). Studies using fMRI data and source localization models of EEG data point to the involvement of heteromodal areas such as the superior temporal sulcus (STS) as well as early modality-specific sensory areas in cross-modal attention modulation (see Koelewijn et al., 2010, for a review). McDonald, Teder-Salejarvi, Di Russo, and Hillyard (2003) measured modulations of visually evoked brain activity by nonpredictive exogenous auditory cues using ERPs and observed a first modulation in the superior temporal cortex (120–140 ms after stimulus onset), followed by a second modulation in the ventral occipital cortex of the fusiform gyrus (150–170 ms after stimulus onset). This spatiotemporal sequence suggests that enhanced visual perception produced by cross-modal exogenous attention results from feedback from multimodal superior temporal cortex to early modality-specific visual areas. Cross-modal exogenous attention may thus first facilitate processing of spatially coincident visual stimuli in the posterior parts of superior temporal gyrus and superior temporal sulcus (STG/STS), regions of multisensory convergence and integration (Hein & Knight, 2008; Kreifelts, Ethofer, Shiozawa, Grodd, & Wildgruber, 2009). Reentrant feedback from STG/STS to early visual areas may then enhance activation in early modality-specific areas by increasing sensory gain.

3 The Special Role of Emotion in Attention and Perception

In addition to endogenous and exogenous attention mechanisms, the emotional relevance of a stimulus has been shown to constitute another important feature influencing selection by attention. Behavioral findings across many different tasks and paradigms indicate that perception is facilitated and attention prioritized for emotional information. Thus, emotion processing does not only enrich our experiences with affective flavor, but can directly shape the content of our percepts and awareness. Emotional stimuli may draw attention quicker and impede attentional disengagement longer than neutral stimuli. In visual search tasks, the detection of a target among distractors is faster when the target is emotional, as opposed to neutral (e.g., Ohman, Flykt, & Esteves, 2001). Conversely, emotional distractors may prolong the search for a nonemotional target (Rinck, Reinecke, Ellwart, Heuer, & Becker, 2005). In the attentional blink task, the detection of a target word in a rapid serial visual stream (items appearing successively at fixation at ~10 Hz) is impaired

when it occurs shortly after another target. However, this deficit is greatly attenuated for emotional stimuli (e.g., Anderson & Phelps, 2001). Conversely, the deficit may increase for a second neutral target following an emotional one, suggesting that the emotional meaning of items tend to grab or divert attention in situations where resources cannot be equally deployed to every stimulus (Smith, Most, Newsome, & Zald, 2006). In the visual prior-entry paradigm, two stimuli are presented simultaneously or almost simultaneously, and participants have to indicate which of the stimuli they perceived first. In this task, fearful faces are perceived earlier in time than neutral faces, reflecting accelerated perception due to attentional prioritization (West, Anderson, & Pratt, 2009). Attentional prioritization has been observed at very early cortical stages of processing, such as primary visual cortex (V1) for threatening visual stimuli (Pourtois, Grandjean, Sander, & Vuilleumier, 2004; West, Anderson, Ferber, & Pratt, 2011). Once attention has been drawn to and engaged by emotional stimuli, it may also dwell longer at their location and facilitate the processing of subsequent nonemotional target stimuli appearing at the same location. Such emotional orienting effects have been demonstrated using the dot probe task (MacLeod, Mathews, & Tata, 1986), where participants must respond to a target (a line or a dot) that replaces one of two simultaneously presented cues—one being emotionally significant (e.g., a fearful face) and the other neutral. Importantly, the cues are equated on basic physical properties such as brightness, contrast, color so that any observed preferential cueing effect is not due to exogenous attention based on low-level stimulus differences, but can be attributed to the perceived emotionality of the cues. Typical results show faster responses to targets replacing the emotional rather than the neutral cue. This effect has been demonstrated both for the visual (Brosch, Sander, & Scherer, 2007; Lipp & Derakshan, 2005) and for the auditory modality (Bertels, Kolinsky, & Morais, 2010). Emotional cueing may also increase contrast sensitivity for the subsequent target (Phelps et al., 2006). These cueing effects occur despite the fact that the cue is not predictive of target location and their emotional meaning is task-irrelevant. Modulation of attention by emotion has furthermore been observed in brain-damaged patients. The dorsal attentional network can be disrupted by stroke in the right parietal regions, resulting in neglect and/or extinction. Studies in patients with these symptoms have demonstrated that the extinction of visual and auditory stimuli can be modulated by emotional stimulus content. Pictures of spiders compared to flowers can decrease the amount of visual extinction in neglect patients (Vuilleumier & Schwartz, 2001). Similarly, emotional prosody can reduce auditory extinction in neglect patients, as demonstrated in a dichotic listening task (Grandjean, Sander, Lucas, Scherer, & Vuilleumier, 2008).

Until now, studies on the emotional modulation of spatial attention have mainly examined within-modality effects, most frequently using pictures of emotional stimuli to modulate visual attention. However, some studies have recently begun to investigate cross-modal emotional attention. In a series of studies, we adapted the emotional dot probe paradigm to investigate cross-modal bias of visual spatial attention by auditory emotion (Brosch, Grandjean, Sander, & Scherer, 2008, 2009). More specifically, we investigated whether emotional prosody (see Grandjean, Bänziger, & Scherer, 2006) influences the spatial deployment of visual attention

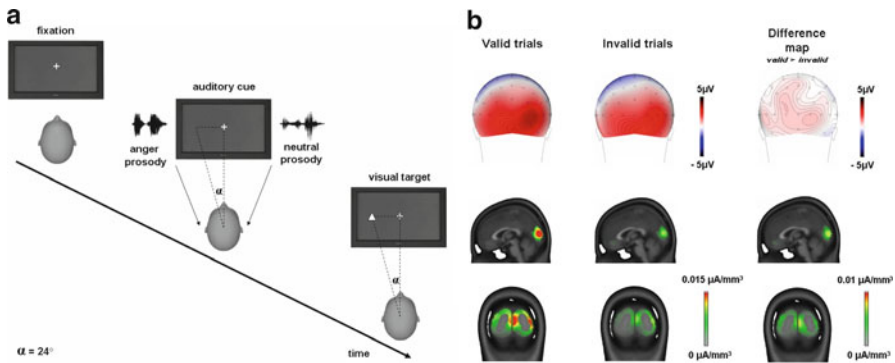


Fig. 11.1 The cross-modal emotional dot probe paradigm. **(a)** Experimental sequence of Brosch et al. (2008, 2009). Each trial started with a random time interval between 500 and 1,000 ms, after which the acoustic cue sound pair was presented. One of the sounds in the pair had emotional prosody, the other one neutral prosody. The target, a neutral geometric figure was presented with a variable cue–target stimulus onset asynchrony sound onset, on the left or right side. The angle between the target and the fixation cross was 24° , equivalent to the synthesized location of the audio stimulus pairs. In a *valid* trial, the target appeared on the side of the emotional sound, in an *invalid* trial, the target appeared on the side of the neutral sound. **(b)** Electrophysiological data confirm cross-modal effects of emotional prosody on early visual processing. *Top row*: Topographic maps for the P1 in valid and invalid trials and topographic difference map. *Middle and bottom rows*: Source localization revealed the intracranial generators of the P1 in striate and extrastriate visual cortex

when emotional and neutral utterances were presented simultaneously (see Fig. 11.1a). In order to give the subjective impression that the sounds originated from a specific location in space (at an angle of 24° to the left and to the right of the participants, corresponding to the locations where the visual target could appear on screen), we manipulated the interaural time difference of the sounds. We used spatially localized stimuli instead of the simpler dichotic presentation mode, as it is a closer approximation of real life contexts in which concomitant auditory and visual information can originate from a common source localized in space. We observed faster responses towards targets when they appeared at the location of the source of the emotional prosody. Importantly, this cross-modal emotional effect was not present when using synthesized control stimuli matched for the mean fundamental frequency and the amplitude envelope, two low-level acoustic parameters related to emotional prosody, of each vocal stimulus used in the experiment, ruling out the possibility that only low-level acoustic parameters trigger the cross-modal emotional effect.

Using a similar approach, Poliakoff and colleagues investigated the effect of threatening visual cues on tactile attention. In a modified cueing paradigm, visual cues were presented close to the participant's hands, which were hidden from view behind a computer screen. The cues consisted of one picture of either a threatening (snakes or spiders) or a nonthreatening stimulus (flowers or mushrooms) presented either close to the left hand or close to the right hand. Following the cue, a tactile stimulus was presented to one of the hands. Pictures of snakes led to faster responses

to the tactile stimulus than nonthreatening pictures. Remarkably, this facilitation effect was enhanced in participants with high fear of snakes, showing that the cross-modal attentional facilitation is driven by the individually perceived threat value (Poliakoff, Miles, Li, & Blanchette, 2007). Following up on these results, Van Damme and colleagues compared the impact of the presentation of threatening pictures on tactile and auditory attention using the prior-entry paradigm (Van Damme, Gallace, Spence, Crombez, & Moseley, 2009). In this paradigm two target stimuli are presented simultaneously or almost simultaneously, and participants have to indicate which target they perceived first. Attentional prioritization of a target leads to accelerated perception. In some trials, participants were presented two tactile targets (vibrations with minimal stimulus onset asynchronies, between 5 and 120 ms), one to the left hand and one to the right hand. In other trials, they were presented two auditory targets emanating from two loudspeakers. In each trial, participants had to indicate which target they perceived first. Before presentation of the target pair, one of the potential target sides was cued with a picture of either a threat to the hand (such as a knife), a general threat (such as an exploding truck), or a picture with emotionally neutral content. All responses were faster when cued by threatening compared to neutral pictures, confirming cross-modal attentional bias by threat. However, in trials with tactile targets, tactile attention was modulated more strongly by pictures showing threats to the hand than by pictures showing general threat. In trials with auditory target pairs, however, attention was biased more strongly by general threat than by threat to the hand. Thus, a visual emotional stimulus indicating imminent threat to a body part leads to attentional bias toward the input from that body part, suggesting some degree of specificity in cross-modal emotional attention. In a similar vein, Schirmer and colleagues investigated to what extent being touched by a friend can modulate early stages of visual processing. Early ERP components such as the N100 and the P200 were modulated by the touch of a friend during negative and neutral pictures viewing. Furthermore, the Late Positive Component (LPC) was increased during negative picture presentations when human touch occurred compared to negative pictures without human touch (Schirmer et al., [in press](#)).

Taken together, the behavioral data reviewed here indicate that perception is facilitated and attention prioritized for emotional information. Emotional stimuli capture attention quicker and may prolong attentional disengagement relative to neutral stimuli. Depending on the task, the prioritization of emotional material can improve behavioral performance (when the target of the task is emotional), but may also lead to interference (when an emotional stimulus competes with a nonemotional target for processing resources). Longer dwelling times of attention at the location of emotional stimuli may furthermore facilitate the processing of subsequent target stimuli that appear at the same location. Whereas most studies have looked at within-modality effects of emotional attention, first studies investigating cross-modal emotional attention demonstrate that emotional attention is not restricted to one modality, but operates across modalities. Here, we reviewed evidence for the modulation of visual attention by auditory emotional information (Brosch, Grandjean et al., 2008; Brosch et al., 2009), evidence for the modulation

of tactile and auditory attention by visual emotional information (Poliakoff et al., 2007; Van Damme et al., 2009), as well as evidence for the modulation of visual processing by tactile emotional information (Schirmer et al., [in press](#)).

4 Neural Mechanisms of Within-Modality Emotional Attention

Consistent with the behavioral findings reviewed above, brain imaging studies using fMRI have consistently revealed increased neural responses to many different emotional stimuli compared to emotionally neutral stimuli, both in early sensory areas like primary visual cortex, and in higher-level regions associated with object and face recognition. Enhanced responses have been observed for emotional pictures in the visual cortex (Whalen et al., 1998), emotional faces in the fusiform face area (Vuilleumier, Armony, Driver, & Dolan, 2001), and emotional body movements in the fusiform body area (Peelen, Atkinson, Andersson, & Vuilleumier, 2007). Similar results have been found in the auditory modality, in that emotional prosody increases activity in the associative auditory cortex (Ethofer, Anders, Wiethoff et al., 2006). Altogether, these findings suggest a selective modulation of brain regions involved in the processing of the specific stimulus categories by emotion. This emotional boosting of neural processing was observed even when the focus of endogenous attention was directed away from the emotional stimuli by secondary tasks, as observed both for the visual (Vuilleumier et al., 2001) and the auditory modality (Grandjean et al., 2005; Sander et al., 2005). Research using electroencephalography (EEG) has yielded similar results, revealing modulatory effects of emotion at several stages of cortical processing, including both early, sensory-related processes and later processes related to more elaborate evaluations of these stimuli, subsequent autonomic arousal, and/or memory formation (see, e.g., Eimer & Holmes, 2007; Olofsson, Nordin, Sequeira, & Polich, 2008; Vuilleumier & Pourtois, 2007, for reviews). Thus, brain imaging and electrophysiological data converge to show that emotional stimuli are represented by more robust neural signatures than neutral ones, and can consequently profit from preferential access to further cognitive processing, behavior control, and awareness.

It has been suggested that the prioritization of emotional information is driven by dedicated neural circuits (Brosch, Pourtois, Sander, & Vuilleumier, 2011; Vuilleumier, 2005; Vuilleumier & Brosch, 2009), separate from the frontoparietal networks involved in endogenous and exogenous attention allocation (Corbetta et al., 2008; Corbetta & Shulman, 2002; see also Peelen et al., 2004). In this model, the amygdala, a limbic region critically involved in the processing of emotional information (LeDoux, 2000; Phelps, 2006) is thought to play a critical role by modulating the processing of incoming sensory stimuli through direct feedback projections to visual cortex (Amaral, Behnia, & Kelly, 2003) and biasing signals to frontoparietal attention regions (Pourtois, Thut, Grave de Peralta, Michel, & Vuilleumier, 2005). Consistent with this suggestion, several PET and fMRI studies have reported that cortical increases to emotional stimuli were significantly correlated

with amygdala responses, i.e., the more the amygdala was sensitive to the emotional meaning, the more the modulation observed in sensory areas.

The boosting of emotional stimuli by the amygdala not only may directly impact on sensory cortices, thus augmenting the neural representation of the emotional stimulus, but it can also recruit the frontoparietal endogenous attention network toward the location of the stimulus, so that subsequent information arising at the same location as emotional cues will benefit from enhanced processing resources. This effect has been demonstrated using the emotional dot probe task where the processing of a nonemotional target is facilitated if it appears at the same location as a previous emotional cue. A series of studies recording event-related potentials (ERPs) during the emotional dot probe task (Brosch et al., 2011; Brosch, Sander, Pourtois, & Scherer, 2008; Pourtois et al., 2004) have shown that emotional stimuli lead to a rapid gain increase in sensory cortex by means of which attended locations or stimuli receive increased perceptual processing (Hillyard et al., 1998). This gain increase is preceded by an early posterior parietal negativity, suggesting a functional coupling between activation of the frontoparietal attention network and a gain increase in early sensory cortex (Pourtois et al., 2005). Using fMRI recordings during the emotional dot probe, greater activation was observed in the intraparietal sulcus (IPS) when targets were preceded by a fearful face than a neutral face, consistent with enhanced attentional orienting and faster detection of targets on valid trials. This contrasted with strongly reduced activation on invalid trials, suggesting that IPS may become unresponsive to targets subsequent to the enhanced focusing of attention on the contralateral emotional cue task (Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2006). A recent fMRI study investigating active search for threatening stimuli reported increased connectivity between amygdala and IPS, FEF and fusiform gyrus when participants were searching for threatening compared to neutral targets (Mohanty, Egner, Monti, & Mesulam, 2009). This finding suggests that actively searching for emotional information elicits amygdalar input into the frontoparietal attention network and inferotemporal visual areas, which may facilitate the rapid detection of emotional stimuli.

Taken together, within-modality work on emotional attention has demonstrated how emotional stimuli can induce a distinctive cascade of neural events which does not only boost the processing of the stimulus itself but also influences mechanisms responsible for orienting and shifting attention in space, such that subsequent information arising at the same location as an emotional cue will also benefit from enhanced processing resources.

5 A Neurocognitive Model of Cross-Modal Emotional Attention

To receive maximal benefit from multimodal input, the brain must coordinate and integrate the input appropriately so that signals from an emotionally relevant source are prioritized across the different input channels. Thus, for example, auditory information about an emotional stimulus should lead to increased neural processing of visual information originating at the same location. This integration and cross-modal

prioritization is a computational challenge, as the properties of the representation of information are highly modality-specific and differ greatly between the input channels: vision is represented retinotopically, touch somatotopically, audition first tonotopically and then head-centered (Driver & Spence, 1998). However, our attention mechanisms seem to be able to perform the necessary computations rapidly. ERP studies investigating nonemotional attention suggest that cross-modal attentional effects on early perceptual processing are based on an allocentric frame of reference reflecting common coordinates of external space (Eimer, Cockburn, Smedley, & Driver, 2001; Kennett, Eimer, Spence, & Driver, 2001). The spatial integration across modalities may be organized by convergence zones in posterior parietal areas, which have been shown to receive multimodal input and to code modality-specific coordinate frames into a common spatial representation (Andersen, Snyder, Bradley, & Xing, 1997). Additionally, single-cell recordings have confirmed the existence of heteromodal neurons with overlapping receptive fields for the different modalities, which are most sensitive to the location of an event, rather than to the modality it activates (Cerf et al., 2010; Stein & Stanford, 2008).

Most studies investigating the neural mechanisms underlying cross-modal attention have looked at the effects of nonemotional stimuli, whereas only few studies have investigated the neural correlates of cross-modal modulation of attention by emotion. Keil and colleagues used ERPs to measure resource allocation to a startle probe (a noise burst) while participants were watching emotional and neutral pictures or listening to emotional and neutral sounds. They observed a decreased amplitude of the P3 potential when startle probes were presented during emotional, as opposed to neutral, stimuli for both sound and picture foregrounds. These results indicate that emotional stimuli cross-modally attract processing resources, leading to optimized processing of the emotional stimulus and reduced processing capacity for concurrent stimuli (Keil et al., 2007). Dowman (2007) and Dowman and Ben-Avraham (2008) identified a network of brain areas involved in the detection and attentional reorienting toward the location of an unexpected painful somatosensory electrical stimulus, when endogenous attention is deployed not towards the tactile, but the visual modality. Using EEG measurements and source localization techniques, they concluded that the detection of the threatening tactile stimulus occurs in sensory cortex (somatosensory cortex and insula) during very early perceptual processing (as early as 70 ms), followed by increased activation in medial prefrontal cortex (130–300 ms), a structure sensitive to situations requiring changes in attentional control. Medial prefrontal cortex is then thought to signal to lateral prefrontal regions that endogenous attention needs to be redirected towards the threat (Bishop, Duncan, Brett, & Lawrence, 2004).

Whereas the work reviewed so far focused on the interruption of ongoing voluntary processing by emotional stimuli, another study has looked at the neural mechanisms underlying perceptual facilitation by cross-modal emotional attention. In our emotional dot probe paradigm investigating cross-modal bias of visual spatial attention by auditory emotion (Brosch, Grandjean et al., 2008; Brosch et al., 2009), we recorded ERPs to investigate at what stage of stimulus processing the deployment of visuospatial attention toward visual targets was affected by spatially congruent or

incongruent emotional information conveyed in affective prosody. Faster response times to visual targets appearing at the location of the source of emotional prosody were accompanied by increased P1 amplitudes towards the target. Source localization indicated that the P1 modulation originated from generators localized in visual cortex (see Fig. 11.1b), suggesting that the cross-modal modulation of spatial attention triggered by emotional prosody affected early sensory stages of visual processing. These early effects at the level of the P1 mirror within-modality effects using the emotional dot probe paradigm (Brosch et al., 2011; Brosch, Sander et al., 2008; Pourtois et al., 2004), and imply that emotionally relevant stimuli may lead to a gain increase in early sensory cortex even when perceived in a different sensory modality. In a similar vein, the specificity of the results by Van Damme et al. (2009) presented earlier, revealing increased tactile attentional bias to a hand when a visual stimulus indicates impending threat to this hand, indirectly suggest a gain effect in primary sensory cortex S1, where somatotopic maps of the body surface have been documented (Penfield & Rasmussen, 1950).

Thus, electrophysiological studies of cross-modal emotional attention reveal that emotional information may interfere with voluntary processing across sensory modalities to boost and optimize the processing of emotional stimuli, and may furthermore amplify the early perceptual processing of multimodal information originating at the location of the emotional stimulus.

We suggest that cross-modal emotional attention may operate via two complementary pathways modulating the neural representation of emotional events across modalities (see Fig. 11.2). Previous research has shown that the amygdala plays a key role in the cross-modal integration of visual and auditory emotional information (Dolan, Morris, & de Gelder, 2001). For example, emotional prosody has been shown to lead to increased activation of the amygdala (Grandjean et al., 2005; Sander & Scheich, 2001), but also to increased activation of visual cortex (Sander et al., 2005; see also von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005), probably reflecting a functional coupling between auditory and visual cortices. Functional connectivity analyses suggest that cross-modal effects of an emotional voice on visual processing are accompanied by increased connectivity between visual areas and the amygdala, but not directly between unimodal visual areas and auditory sensory areas (Ethofer, Anders, Erb et al., 2006). This suggests that cross-modal enhancements by emotion may not be mediated by direct coupling between modality-specific areas, but rather via supramodal relay areas. In addition to the amygdala, the superior temporal gyrus and sulcus may play an important role. Cross-modal exogenous cueing by nonemotional auditory signals has been shown to operate via reentrant feedback from STG/STS to early visual areas (McDonald et al., 2003). Posterior superior temporal sulcus acts as a convergence zone for the integration of emotional visual and auditory information and sends top-down feedback signals back to unimodal cortices (Campanella & Belin, 2007). Perceptual facilitation by cross-modal emotional attention thus may also operate by increased coupling between STG/STS and regions of unimodal cortex, potentially driven by the boosting of emotional information by the amygdala.

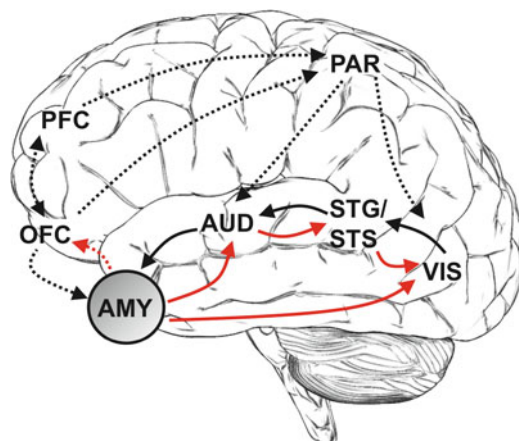


Fig. 11.2 Two neural pathways underlying cross-modal emotional attention, as illustrated here for the effects of emotional auditory information on visual perception. **(1)** Cross-modal boosting of emotional information (bold arrows): Emotional information originating in auditory cortex (AUD) is amplified by feedback signals from the amygdala (AMY). This amplification may reach visual cortex (VIS) via convergence zones such as superior temporal gyrus/sulcus (STG/STS). Additionally, the amygdala may directly mediate the functional coupling between auditory and visual unimodal cortices. **(2)** Reorienting of frontoparietal attention networks (dotted arrows): Amygdala signals may bias fronto-parietal attention regions (OFC, PFC, PAR) toward the location of emotional events to supramodally amplify information processing at this location red arrows indicate direct feedback signals originating from the amygdala

In addition to the direct enhancement of the neural representation of emotional information, the amygdala has been shown to reorient frontoparietal attention networks toward the location of an emotional stimulus (Pourtois et al., 2006; Vuilleumier & Brosch, 2009). Attentional facilitation effects of a frontoparietal reorienting have been shown to operate cross-modally for nonemotional information (Eimer & Driver, 2001). Thus, amygdala-driven recruitment of frontoparietal attention networks toward emotional stimuli will lead to benefits for subsequent information arising at the same location, independent of the modality of this information (Brosch et al., 2009). Conversely, this may lead to a reduction of processing capacities for ongoing voluntary processing in all modalities (Keil et al., 2007).

To conclude, the data reviewed here converge to show that emotion modulates attentional processing across sensory modalities by boosting early sensory stages of processing, potentially implemented by a large-scale neural network centered around the amygdala, providing direct and indirect top-down signals to sensory pathways and frontoparietal pathways involved in exogenous and endogenous attentional selection processes. This rapid cross-modal integration at multiple stages of processing may reflect a fundamental principle of human brain organization: to prioritize the processing of emotionally relevant stimuli, even if they are outside the focus of spatial attention, thus facilitating the multimodal assessment of emotionally relevant stimuli in the environment.

References

- 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*, 1099–1120.
- Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience*, *20*, 303–330.
- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*, 305–309.
- Bertels, J., Kolinsky, R., & Morais, J. (2010). Emotional valence of spoken words influences the spatial orienting of attention. *Acta Psychologica*, *134*, 264–278.
- Bishop, S. J., Duncan, J., Brett, M., & Lawrence, A. D. (2004). Prefrontal cortical function and anxiety: Controlling attention to threat-related stimuli. *Nature Neuroscience*, *7*, 184–188.
- Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition*, *106*, 1497–1503.
- Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2009). Cross-modal emotional attention: Emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience*, *21*, 1670–1679.
- Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008). Beyond fear: Rapid spatial orienting towards positive emotional stimuli. *Psychological Science*, *19*, 362–370.
- Brosch, T., Sander, D., & Scherer, K. R. (2007). That baby caught my eye... Attention capture by infant faces. *Emotion*, *7*, 685–689.
- 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.
- Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*, 535–543.
- Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience*, *7*, 308–313.
- Cerf, M., Thiruvengadam, N., Mormann, F., Kraskov, A., Quiroga, R. Q., Koch, C., et al. (2010). On-line, voluntary control of human temporal lobe neurons. *Nature*, *467*, 1104–1108.
- Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron*, *58*, 306–324.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201–215.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 10006–10010.
- Dowman, R. (2007). Neural mechanisms of detecting and orienting attention toward unattended threatening somatosensory targets. I. Intermodal effects. *Psychophysiology*, *44*, 407–419.
- Dowman, R., & Ben-Avraham, D. (2008). An artificial neural network model of orienting attention toward threatening somatosensory stimuli. *Psychophysiology*, *45*, 229–239.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, *92*, 53–78.
- Driver, J., & Spence, C. (1998). Crossmodal attention. *Current Opinion in Neurobiology*, *8*, 245–253.
- Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*, 269–297.
- Eimer, M., Cockburn, D., Smedley, B., & Driver, J. (2001). Cross-modal links in endogenous spatial attention are mediated by common external locations: Evidence from event-related brain potentials. *Experimental Brain Research*, *139*, 398–411.

- Eimer, M., & Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: Evidence from event-related brain potential studies. *Neuroscience and Biobehavioral Reviews*, 25, 497–511.
- Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, 45, 15–31.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human Brain Mapping*, 27, 707–714.
- Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., et al. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport*, 17, 249–253.
- Grandjean, D., Bänziger, T., & Scherer, K. R. (2006). Intonation as an interface between language and affect. *Progress in Brain Research*, 156, 235–247.
- Grandjean, D., Sander, D., Lucas, N., Scherer, K. R., & Vuilleumier, P. (2008). Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect. *Neuropsychologia*, 46, 487–496.
- Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8, 145–146.
- Hein, G., & Knight, R. T. (2008). Superior temporal sulcus—It's my area: Or is it? *Journal of Cognitive Neuroscience*, 20, 2125–2136.
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 353, 1257–1270.
- Keil, A., Bradley, M. M., Junghofer, M., Russmann, T., Lowenthal, W., & Lang, P. J. (2007). Cross-modal attention capture by affective stimuli: Evidence from event-related potentials. *Cognitive Affective and Behavioral Neuroscience*, 7, 18–24.
- Kennett, S., Eimer, M., Spence, C., & Driver, J. (2001). Tactile-visual links in exogenous spatial attention under different postures: Convergent evidence from psychophysics and ERPs. *Journal of Cognitive Neuroscience*, 13, 462–478.
- Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multi-sensory integration: A review of audiovisual studies. *Acta Psychologica*, 134, 372–384.
- Kreifelts, B., Ethofer, T., Shiozawa, T., Grodd, W., & Wildgruber, D. (2009). Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia*, 47, 3059–3066.
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23, 155–184.
- Lipp, O. V., & Derakshan, N. (2005). Attentional bias to pictures of fear-relevant animals in a dot probe task. *Emotion*, 5, 365–369.
- Lloyd, D. M., Bolanowski, S. J., Jr., Howard, L., & McGlone, F. (1999). Mechanisms of attention in touch. *Somatosensory and Motor Research*, 16, 3–10.
- MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95, 15–20.
- Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, 9, 296–305.
- McDonald, J. J., Teder-Salejarvi, W. A., Di Russo, F., & Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *Journal of Cognitive Neuroscience*, 15, 10–19.
- Miles, E., Poliakoff, E., & Brown, R. J. (2008). Investigating the time course of tactile reflexive attention using a non-spatial discrimination task. *Acta Psychologica*, 128, 210–215.
- Mohanty, A., Egner, T., Monti, J. M., & Mesulam, M. M. (2009). Search for a threatening target triggers limbic guidance of spatial attention. *Journal of Neuroscience*, 29, 10563–10572.
- Ohman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General*, 130, 466–478.
- Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, 77, 247–265.

- Peelen, M., Atkinson, A., Andersson, F., & Vuilleumier, P. (2007). Emotional modulation of body-selective visual areas. *Social Cognitive and Affective Neuroscience*, 2, 274–283.
- Peelen, M., Heslenfeld, D. J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *Neuroimage*, 22, 822–830.
- Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man*. New York, NY: Macmillan.
- Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, 57, 27–53.
- Phelps, E. A., Ling, S., & Carrasco, M. (2006). Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological Science*, 17, 292–299.
- Poliakoff, E., Miles, E., Li, X., & Blanchette, I. (2007). The effect of visual threat on spatial attention to touch. *Cognition*, 102, 405–414.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology*, 109, 160–174.
- Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex*, 14, 619–633.
- Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2006). Neural systems for orienting attention to the location of threat signals: An event-related fMRI study. *Neuroimage*, 31, 920–933.
- Pourtois, G., Thut, G., Grave de Peralta, R., Michel, C., & Vuilleumier, P. (2005). Two electrophysiological stages of spatial orienting towards fearful faces: Early temporo-parietal activation preceding gain control in extrastriate visual cortex. *Neuroimage*, 26, 149–163.
- Rinck, M., Reinecke, A., Ellwart, T., Heuer, K., & Becker, E. S. (2005). Speeded detection and increased distraction in fear of spiders: Evidence from eye movements. *Journal of Abnormal Psychology*, 114, 235–248.
- Rossi, A. F., & Paradiso, M. A. (1995). Feature-specific effects of selective visual attention. *Vision Research*, 35, 621–634.
- Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences*, 14, 303–316.
- Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *Neuroimage*, 28, 848–858.
- Sander, K., & Scheich, H. (2001). Auditory perception of laughing and crying activates human amygdala regardless of attentional state. *Cognitive Brain Research*, 12, 181–198.
- Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York, NY: Oxford University Press.
- Schirmer, A., Teh, K. S., Wang, S., Vijayakumar, R., Ching, A., Nithianantham, D., Escoffier, N., Cheok, A. D. (2011). Squeeze me, but don't tease me: human and mechanical touch enhance visual attention and emotion discrimination. *Social Neuroscience*, 6(3), 219–230.
- Smith, S. D., Most, S. B., Newsome, L. A., & Zald, D. H. (2006). An emotion-induced attentional blink elicited by aversively conditioned stimuli. *Emotion*, 6, 523–527.
- Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Science*, 1191, 182–200.
- Spence, C., & Driver, J. (1994). Covert spatial orienting in audition: Exogenous and endogenous mechanisms facilitate sound localization. *Journal of Experimental Psychology-Human Perception and Performance*, 20, 555–574.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: Current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, 9, 255–266.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception and Psychophysics*, 49, 83–90.
- Turatto, M., Mazza, V., & Umiltà, C. (2005). Crossmodal object-based attention: Auditory objects affect visual processing. *Cognition*, 96, B55–B64.

- Van Damme, S., Gallace, A., Spence, C., Crombez, G., & Moseley, G. L. (2009). Does the sight of physical threat induce a tactile processing bias? Modality-specific attentional facilitation induced by viewing threatening pictures. *Brain Research, 1253*, 100–106.
- von Kriegstein, K., Kleinschmidt, A., Sterzer, P., & Giraud, A. L. (2005). Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience, 17*, 367–376.
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences, 9*, 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*, 829–841.
- Vuilleumier, P., & Brosch, T. (2009). Interactions of emotion and attention. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences IV* (pp. 925–934). Cambridge, MA: MIT Press.
- Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia, 45*, 174–194.
- Vuilleumier, P., & Schwartz, S. (2001). Beware and be aware: Capture of spatial attention by fear-related stimuli in neglect. *Neuroreport, 12*, 1119–1122.
- West, G. L., Anderson, A. A., Ferber, S., & Pratt, J. (2011). Electrophysiological evidence for biased competition in V1 for fear expressions. *Journal of Cognitive Neuroscience, 23*(11), 3410–3418.
- West, G. L., Anderson, A. A., & Pratt, J. (2009). Motivationally significant stimuli show visual prior entry: Evidence for attentional capture. *Journal of Experimental Psychology-Human Perception and Performance, 35*, 1032–1042.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18*, 411–418.
- 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*, 495–501.
- Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24*, 295–340.