



Modulation of spatial attention to visual targets by emotional environmental sounds

Neil R. Harrison and Simon J. Davies

Liverpool Hope University, Liverpool, England

Abstract

Previous research has shown that visual spatial attention can be modulated by emotional prosody cues, but it is not known whether such crossmodal modulation of visual attention is associated with the engagement or disengagement of attentional resources. To test this, we employed a modified spatial cueing task, where participants indicated whether a visual target appeared either on the left or the right, after hearing a spatially non-predictive peripheral sound. Prior studies using prosody cues have found that modulation of visual attention by emotional auditory cues was lateralized, but this may have been due to the speech content of the stimuli; here instead we used non-speech environmental sounds. The sound was either emotional (pleasant, unpleasant) or neutral, and was presented either on the same side as the visual target ('valid' trial) or on the opposite side ('invalid' trial). For the cue validity index (RT to invalid cue minus RT to valid cue), we found differences between emotional and neutral cues, but only for visual targets presented in the right hemifield; here the cue validity index was lower for unpleasant compared to neutral and pleasant cues. Absolute RTs for targets on the right were faster for invalid trials following unpleasant cues, compared to pleasant and neutral cues, indicating that the reduced cue validity effect was due to faster disengagement from unpleasant auditory cues. Further, our results show that the laterality effect is related to the emotional nature of the cues, rather than the speech content of the stimuli. **Keywords:** emotion, crossmodal, attention, spatial, cueing.

Received 05 July 2013; received in revised form 08 August 2013; accepted 17 August 2013. Available online 23 December 2013.

Introduction

The perceptual system receives information from multiple modalities, and it is well-known that a stimulus in one modality can modify the processing of a cue presented in a different modality (for review see e.g., Driver & Noesselt, 2008). An important feature of multisensory processing is the ability of a cue in one modality to direct attention towards a particular region of space, which can result in enhanced processing of a signal presented at the same location in a separate modality (for review see Eimer & Driver, 2001). This ability of a stimulus in one modality to act as an involuntary spatial cue for a stimulus in a separate modality has generally been studied using non-affective cues such as beeps and light flashes (e.g., McDonald, Teder-Sälejärvi, & Hilyard, 2000; McDonald & Ward, 2000; Spence & Driver, 1997).

Cues in the environment often acquire salience and preferential processing due to their emotional nature (Pourtois, Schettino, & Vuilleumier, 2012; Vuilleumier, 2005; Yiend, 2010), and several studies

have shown that emotional cues in one modality can influence processing in a second modality. For example, it has been shown that emotional pictures can improve categorization of auditory cues (Tartar, de Almeida, McIntosh, Rosselli, & Nash, 2012), and that emotional pictures, compared to neutral pictures, can bias attention towards non-emotional tactile and auditory targets (Van Damme, Gallace, Spence, Crombez, & Moseley, 2009). In particular, facilitated processing of visual targets has been reported following presentation of spatially uninformative cues conveying negative (Brosch, Grandjean, Sander and Scherer, 2008a) and positive prosody (Brosch, Grandjean, Sander, & Scherer, 2008b), and this type of crossmodal exogenous spatial cueing by emotional stimuli is thought to occur at a relatively early (i.e., perceptual) stage of processing, as revealed by event-related potentials (Brosch, Grandjean, Sander, & Scherer, 2009).

Experiments on crossmodal attentional modulation by auditory emotion have mainly used dot-probe tasks (e.g., Brosch et al., 2008a; Brosch et al., 2009), where a visual probe is presented immediately following the presentation of two simultaneous auditory cues, one to each ear, for example an unpleasant cue to the left ear and a neutral cue to the right ear. The simultaneous presentation of two cues, however, limits the investigation into the components of attention that contributed to the crossmodal

modulatory effect. That is, differences in reaction times to visual probes presented on the side of the emotional cue compared to the side of the neutral cue could be caused by either faster or slower attentional engagement to the emotional cue, or by faster or slower attentional disengagement from the affective cue. By contrast, a modified spatial cueing paradigm, where a single auditory cue is presented on each trial, allows investigation into the effects of engagement and disengagement attentional components elicited by the emotional cues by comparing reaction times to valid (i.e., at the cued location) emotional trials compared to valid neutral trials, and by comparing reaction times to invalid (i.e., at the non-cued location) trials for emotional and neutral cues, respectively (see e.g., Yiend & Mathews, 2001). The present experiment used a modified exogenous spatial cueing design, based on the original spatial cueing task (e.g., Posner, 1980), to test the effects of emotional (pleasant, unpleasant, or neutral) auditory stimuli on the engagement and disengagement components of visual spatial attention. Participants were required to indicate whether a visual target appeared either in the left or the right visual field, after hearing a spatially non-predictive peripheral auditory cue.

The emotional modulation of spatial attention has been shown to be lateralized, with modulated processing of visual targets restricted to the right visual field following presentation of emotional compared to neutral word-like utterances (Brosch et al., 2008a). It is not clear, however, whether this lateralized effect was due to the emotional properties of the cues, or instead was driven by the speech content of the auditory cues, as meaningless speech sounds are preferentially processed by the left hemisphere (Vouloumanos, Kiehl, Werker, & Liddle, 2001). Consequently, speech cues may have primed the left hemisphere, leading to attentional bias towards the contralateral (i.e., right) side. To rule out whether the laterality effect was due to the speech content of the cues, we considered it important to test whether crossmodal spatial cueing effects are lateralized when employing non-speech environmental sounds, which, compared to speech sounds, are known to activate cortical regions more bilaterally (Meyer, Zysset, von Cramon, Alter, 2005).

Methods

Participants

Twenty-two participants took part in the experiment (age $M = 33.1$, $SD = 12.3$; 13 female), all right-handed. All participants reported normal hearing and normal or corrected-to-normal vision. One participant was excluded from the analysis for failure to adequately follow the task instructions ($> 15\%$ errors). The experiment was approved by the Ethics Committee of the Psychology Department at Liverpool Hope University.

Stimuli and apparatus

The auditory stimuli consisted of 30 sounds from the International Affective Digitalized Sounds database (IADS-2; Bradley & Lang, 2007) selected on the basis of their valence and arousal norms. All 30 stimuli were non-speech environmental sounds and did not contain content involving vocal-fold vibration. Ten stimuli were rated as unpleasant (valence < 4 ; e.g., explosion), ten as pleasant (valence > 6 ; e.g., applause), and ten as neutral (valence $> 4.0 < 6.0$; e.g. typewriter). Details of the selected IADS-2 sounds are listed in the Appendix.

The sounds from the IADS-2 database were approximately six seconds long. A 1000 ms segment was selected from each original sound (using Audacity) on the basis that it was representative of the content of the original sound (see also Scherer and Larsen, 2011, for a similar procedure). An independent group of participants rated the segments on the dimensions of valence and arousal. 12 participants (7 female) judged each sound using a computerised Self-Assessment Manikin (SAM) rating system (see Bradley & Lang, 2007, for a full description of the rating method). Pleasant stimuli had a higher valence rating than neutral $t(11) = 3.361$, $p < 0.01$ (Pleasant: $M = 6.1$, $SD = 0.9$; Neutral: $M = 4.5$, $SD = 1.1$), and unpleasant ratings ($M = 3.0$, $SD = 1.1$) had a lower valence than neutral stimuli, $t(11) = 6.277$, $p < .001$. For arousal ratings, the unpleasant sound segments had higher arousal ratings than neutral, $t(11) = 2.604$, $p < 0.03$ (Unpleasant: $M = 5.2$, $SD = 1.7$; Neutral: $M = 4.4$, $SD = 1.5$), and the pleasant sounds ($M = 5.8$, $SD = 1.2$) had higher ratings than neutral, $t(11) = 3.776$, $p < 0.005$, but the pleasant and unpleasant arousal ratings did not differ, $t < 1.5$.

The interaural level differences (ILDs) of the left and right channels were adjusted to create sounds that were approximately spatially matched to the locations of the visual targets, as this offers a closer resemblance to real-world environments (where correlated auditory and visual information often arise from a single spatial location (c.f., Brosch et al. 2009)), compared to simple dichotic presentation. A preliminary experiment was run with three participants to determine the ILDs which best matched the locations of the visual targets, following a similar procedure to Meienbrock, Naumer, Doehrmann, Singer, & Muckli (2007). 41 versions of the 'dog growl' stimulus were created that varied in volume in each ear in steps of 4.25% to simulate locations from 85° to the left of center to 85° to the right of center. The duration of each sound was 1000ms. The 41 sounds were presented sequentially, with an inter-stimulus interval of 500 ms, and participants judged which sound best fitted the spatial location (17° to the left and right of center) of the two visual targets, by pressing a key on the keyboard. The ILDs of the sound positions judged closest to the visual targets were applied to the 30 IADS-2 sound segments, producing 30 stimuli perceived as located at the position of the

left target, and 30 stimuli located at the position of the right target. The interaural time difference was the same in both ears. The intensities of the segments were equated by equalizing the root mean square power of all sound files, using MATLAB. In the preliminary study and in the main experiment, auditory stimuli were presented over Sennheiser HD201 headphones and the mean dB level for all sounds was 65 dBA. Participants' heads were not stabilized during the experiment, but participants were instructed to refrain from moving their heads during the experiment.

Visual stimuli were presented on a 19" Samsung SyncMaster monitor. The visual target was a small dot presented to the left or right side of the screen, 17° from fixation, at the same height as fixation. The experiment was controlled using E-Prime 2.0.

Procedure and data analysis

Participants were seated at a distance of 60 cm from the screen. The experiment began with a practice block of 10 trials using sounds that were not included in the main experiment. The main experiment consisted of three blocks of trials. In 'valid' trials the sound location was on the same side as the target; in 'invalid' trials the sound location was on the opposite side. The position (left/ right) refers to the location of the visual target. Each of the 30 sounds was presented four times during each experimental block (valid left, valid right, invalid left, invalid right), in randomized order. In addition to the 120 target trials, there were 6 'no-target' trials (two for pleasant, two for unpleasant, two for neutral) per block where no visual target appeared after the sound, to prevent response strategies (Brosch, 2008a). In total the experiment consisted of 378 trials.

Each trial began with a central fixation cross lasting 500 ms, immediately followed by presentation of the auditory stimulus. Directly after offset of the auditory cue, the visual target appeared on the screen for 30 ms (except in no-target trials). Participants had 1500 ms to respond after onset of the target. After response, or after 1500 ms in the event of no response, there followed a random inter-trial interval of between 500 – 1500 ms.

Results

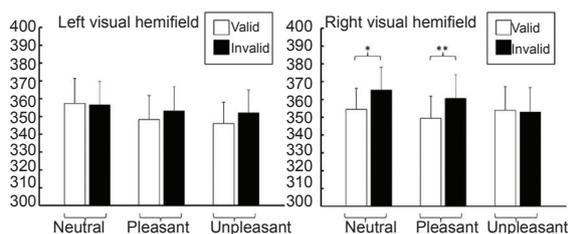


Figure 1. Mean reaction times for targets presented to the left and right visual hemifields. Responses were faster for valid compared to invalid trials for targets presented to the right visual hemifield following neutral and positive auditory cues. Error bars indicate standard error of the mean. * $p < .05$; ** $p < .01$.

Mean error rate was 4.1% and mean false alarm rate was less than 0.5%. These results were not analysed further. Only reaction times on trials with correct responses were included in the analysis. Trials with RTs less than 100 ms (too fast) or greater than 1200 ms (too slow) were discarded (< 1% of data). A three-way repeated-measures ANOVA was conducted on response times with factors cue validity (valid, invalid), emotion (neutral, pleasant, unpleasant) and visual hemifield of presentation (left, right). Analysis revealed a significant main effect of cue validity, $F(1, 20) = 5.97$, $p < .03$, with faster reaction times for targets presented on the same side as the auditory stimulus, and a significant main effect of emotion, $F(2, 40) = 4.51$, $p < .04$. These main effects were qualified by a significant interaction between cue validity, emotion, and hemifield, $F(2, 40) = 3.52$, $p < .04$. To follow up the significant three-way interaction, the simple interaction effect of validity and emotion was analysed at each level of the factor hemifield of presentation. For reaction times in the left hemifield, there was no interaction between emotion and cue validity, $F < 1$. For responses to right sided targets, there was a significant interaction, $F(2,40) = 4.10$, $p < .03$; follow-up paired t-tests showed faster responses for valid ($M = 354.4$ ms, $SD = 54.2$) compared to invalid ($M = 365.3$ ms, $SD = 59.2$) trials for neutral cues, $t(20) = 2.69$, $p < .02$, and faster responses for pleasant cues for valid ($M = 349.4$ ms, $SD = 56.7$) compared to invalid trials ($M = 360.7$ ms, $SD = 60.1$), $t(20) = 3.57$, $p < .002$, indicating cue validity effects for neutral and pleasant cues when targets were presented to the right hemifield.

A second analysis step assessed differences in magnitude of the cue validity effect between different emotional categories, by calculating a cue validity index (i.e., subtracting RTs on valid trials from RTs on invalid trials) for each emotional category, separately for targets in the left and right hemifields (see Figure 2). This measure summarises the interaction term of cue validity and emotional category for reaction times. A one-way repeated-measures ANOVA confirmed no differences between emotions for targets presented to the left hemifield, $F < 1$, but differences between emotions were observed in the right hemifield, $F(2, 40) = 4.10$, $p < .025$. Paired t-tests found greater cue validity for neutral ($M = 10.9$ ms, $SD = 18.5$) compared to unpleasant cues ($M = -9.6$ ms, $SD = 12.3$), $t(20) = 2.41$, $p < .03$; and increased cue validity for pleasant ($M = 11.3$ ms, $SD = 14.4$) compared to unpleasant cues $t(20) = 2.98$, $p < .01$.

To investigate the attentional processes of engagement and disengagement underlying the cue validity effect for right-sided targets, reaction times were compared across emotional conditions, separately for valid and invalid trials. There were no differences between emotional categories for valid cues on the right ($F < .6$), but there was a significant effect for invalid cues, $F(2, 40) = 3.82$, $p < .030$, where invalid trials were faster for unpleasant cues ($M = 352.9$ ms,

$SD = 62.9$) compared to neutral cues ($M = 365.3$ ms, $SD = 59.2$), $t(20) = 3.02$, $p < .01$.

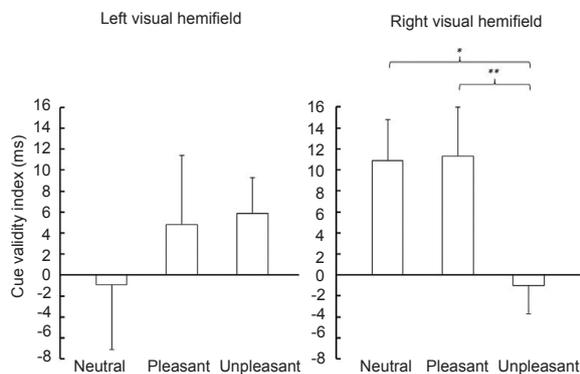


Figure 2. Mean cue validity effects (RT invalid - RT valid) for targets presented to the left and right visual hemifields. There was a reduced cue validity effect for unpleasant auditory cues compared to pleasant and neutral cues, for visual targets presented to the right. Error bars indicate standard error of the mean. * $p < .05$, ** $p < .01$.

Discussion

The current study investigated the effects of emotional auditory cues on the allocation of visual spatial attention. A modified spatial cueing paradigm was used that permitted enquiry into attentional engagement and disengagement attentional elicited by the emotional cues (see e.g., Yiend & Mathews, 2001). While previous research has found lateralized effects for multisensory emotional modulation of attention (Brosch et al., 2008a), using speech as the affective cue, the current experiment also tested whether a laterality effect was present with emotional auditory cues that were non-speech environmental sounds, that are known to activate cortical areas more bilaterally compared to speech sounds (Meyer et al., 2005). We found a crossmodal effect of spatial attention by emotion only for visual targets presented on the right. For right-sided targets, neutral and pleasant auditory cues led to faster responses to valid trials (i.e., auditory and visual stimuli on the same side) compared to invalid trials (i.e., auditory and visual stimuli on opposite sides). Importantly, the degree of cue validity (RTs to invalid cues minus RTs to valid cues) was significantly *lower* for unpleasant compared to pleasant and neutral cues. For right sided targets, we found faster attentional disengagement from unpleasant cues compared to neutral and pleasant cues, showing that attention was directed away from unpleasant auditory stimuli on the left.

The current results extend the findings of previous research into crossmodal modulation of attention by auditory emotional cues by elucidating the processes of engagement and disengagement of attentional resources evoked by the affective cue. Further, the findings confirm that lateralized modulation of

visual attention by auditory emotion is not confined to auditory cues containing speech (Brosch et al., 2008a), as the current experiment employed non-speech environmental sounds as the acoustic stimulus. For visual targets presented to the right, unpleasant auditory cues resulted in a reduced cue validity effect, compared to neutral and positive cues (Figure 2), indicating less efficient processing of the visual target following unpleasant auditory cues at the same location, compared to neutral or positive cues. It appears that the reduced cue validity effect for unpleasant stimuli with targets in the right hemifield was driven by faster disengagement of attention from unpleasant auditory cues presented to the left side, as responses to invalid trials in the unpleasant condition were faster compared to the neutral and pleasant conditions. More rapid disengagement from an unpleasant auditory stimulus on the left led to increased attention to the opposite (i.e., right) side in the visual modality, consistent with prior (purely visual) studies showing faster RTs on invalid trials following negative emotional cues compared to neutral cues (Lee, Franklin, Turkel, Goetz, & Woods, 2012; Koster, Crombez, Verschuere, Van Damme, & Wiersema, 2006). Facilitated attentional disengagement from unpleasant cues is assumed to reflect attentional avoidance of the unpleasant emotional stimulus (Lee, Franklin, Turkel, Goetz, & Woods, 2012; Koster et al., 2006). Enhanced processing of unpleasant auditory cues presented to the left is in line with theoretical models that attribute increased right hemisphere activation for unpleasant sounds (see Demaree, Everhart, Youngstrom, & Harrison, 2005, for a review), given that lateralized auditory stimuli are processed preferentially by the contralateral auditory cortex (e.g., Woldorff et al., 1999).

Differences in the duration of the emotional stimuli may explain the apparent discrepancy between the reduced cue validity effect in the current experiment and the facilitated processing of visual targets following negative prosody cues reported by Brosch et al. (2008a). In the current study, visual targets appeared 1000 ms after onset of the auditory emotional cues, whereas Brosch and colleagues presented visual targets 500 ms after onset of the prosody signals. Further studies should therefore assess whether the magnitude of cue validity effects for visual targets following emotional auditory cues critically depends on inter-stimulus timing. Moreover, auditory stimuli were presented over headphones in the current study, but a more realistic situation could be created by presenting auditory stimuli from loudspeakers located next to the visual targets, as it is known that precise spatial alignment between visual and auditory cues is necessary for optimal multisensory binding (Meyer, Wuerger, Röhrbein, & Zetzsche, 2005). Finally, in the current experiment visual targets were non-emotional, but future experiments could usefully investigate the effect of emotional sounds on spatial attention to emotionally congruent visual targets, as multisensory

processes are modulated by semantic congruence (see e.g., Meyer, Harrison, & Wuerger, 2013).

Appendix A. Stimulus list

Table 1: Stimulus List

Neutral		Pleasant		Unpleasant	
IADS-2 Number	Description	IADS-2 Number	Description	IADS-2 Number	Description
320	Office	725	SodaFizz	115	Bees
322	Typewriter	254	Videogame	116	Buzzing
373	Paint	270	Whistling	380	Jackhammer
403	Helicopter	378	Doorbell	420	Car horns
602	Thunderstorm	815	Rock n roll	502	Engine-failure
700	Toilet flush	820	Funk music	730	Glass-break
708	Clock	808	Bugle	422	Tire squeeze
722	Walking	717	Slotmachine	424	Car-Wreck
728	Paper	351	Applause	626	Explosion
358	Writing	817	Bongos	711	Siren

References

- Bradley, M. M., & Lang, P. J. (2007). *The International Affective Digitized Sounds (2nd ed.; IADS-2): Affective ratings of sounds and instruction manual (Tech. Rep. No. B-3)*. Gainesville, FL: University of Florida.
- Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2008a). Behold the voice of wrath: Cross-modal modulation of visual attention by anger prosody. *Cognition*, *106*, 1497–1503.
- Brosch, T., Sander, D., Pourtois, G., & Scherer, K. R. (2008b). Beyond fear: Rapid spatial orienting towards positive emotional stimuli. *Psychological Science*, *19*, 362–370.
- Brosch, T., Grandjean, D., Sander, D., & Scherer, K. R. (2009). Crossmodal emotional attention: emotional voices modulate early stages of visual processing. *Journal of Cognitive Neuroscience*, *21*, 1670–1679.
- Demaree, H. A., Everhart, D. E., Youngstrom, E. A., & Harrison, D. W. (2005). Brain lateralization of emotional processing: Historical roots and a future incorporating “dominance”. *Behavioral and Cognitive Neuroscience Reviews*, *4*, 3–20.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron*, *57*, 11–23.
- 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.
- Koster, E. H. W., Crombez, G., Verschuere, B., Van Damme, S., & Wiersema, J. R. (2006). Components of attentional bias to threat in high trait anxiety: Facilitated engagement, impaired disengagement, and attentional avoidance. *Behaviour Research and Therapy*, *44*, 1757–1771.
- Lee, H.-J., Franklin, S. A., Turkel, J. E., Goetz, A. R., & Woods, D. W. (2012). Facilitated attentional disengagement from hair-related cues among individuals diagnosed with trichotillomania: An investigation based on the exogenous cueing paradigm. *Journal of Obsessive-Compulsive and Related Disorders*, *1*, 8–15.
- McDonald, J. J., Teder-Sälejärvi, W. A., & Hilyard, S. A. (2000). Involuntary orienting to sound improves visual perception. *Nature*, *407*, 906–908.
- McDonald, J. J., & Ward, L. M. (2000). Involuntary listening aids seeing: Evidence from human electrophysiology. *Psychological Science*, *11*, 167–171.
- Meienbrock, A., Naumer, M. J., Doehrmann, O., Singer, W., & Muckli, L. (2007). Retinotopic effects during spatial audio-visual integration. *Neuropsychologia*, *45*, 531–539.
- Meyer, G. F., Wuerger, S. M., Röhrbein, F., & Zetzsche, C. (2005). Low-level integration of auditory and visual motion signals requires spatial co-localisation. *Experimental Brain Research*, *166*, 538–547.
- Meyer, G., Harrison, N., & Wuerger, S. (2013). The time course of auditory–visual processing of speech and body actions: Evidence for the simultaneous activation of an extended neural network for semantic processing. *Neuropsychologia*, *51*, 1716–1725.
- Meyer, M., Zysset, S., von Cramon, D. Y., & Alte, K. (2005). Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. *Cognitive Brain Research*, *24*, 291–306.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.
- Pourtois, G., Schettino, A., & Vuilleumier, P. (2012). Brain mechanisms for emotional influences on perception and attention: What is magic and what is not. *Biological Psychology*, *92*, 492–512.
- Scherer, L. D., & Larsen, R. J. (2011). Cross-modal evaluative priming: Emotional sounds influence the processing of emotion words. *Emotion*, *11*, 203–208.
- Spence, C., & Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Perception & Psychophysics*, *59*, 1–22.
- Tartar, J. L., de Almeida, K., McIntosh, R. C., Rosselli, M., & Nash, A. J. (2012). Emotionally negative pictures increase attention to a subsequent auditory stimulus. *International Journal of Psychophysiology*, *83*, 36–44.
- 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.
- Vouloumanos, A., Kiehl, K. A., Werker, J. F., & Liddle, P. F. (2001). Detecting sounds in the auditory stream: Event-related fMRI evidence for differential activation to speech and non-speech. *Journal of Cognitive Neuroscience*, *13*, 994–1005.
- Vuilleumier, P. (2005). How brains beware: Neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*, 585–594.
- Woldorff, M. G., Tempelmann, C., Fell, J., Tegeler, C., Gaschler-Markefski, B., Hinrichs, H., Heinze, H. J., & Scheich, H. (1999). Lateralized auditory spatial perception and the contralaterality of cortical processing as studied with functional magnetic resonance imaging and magnetoencephalography. *Human Brain Mapping*, *7*, 49–66.
- Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, *24*, 3–47.
- Yiend, J., & Mathews, A. (2001). Anxiety and attention to threatening pictures. *Quarterly Journal of Experimental Psychology A*, *54*, 665–681.