



# Emotional valence and spatial congruency differentially modulate crossmodal processing: an fMRI study

Dhana Wolf<sup>1,2,3,\*†</sup>, Lisa Schock<sup>1,2,3†</sup>, Saurabh Bhavsar<sup>1,2,3</sup>, Liliana R. Demenescu<sup>1,2,3</sup>, Walter Sturm<sup>2,4</sup> and Klaus Mathiak<sup>1,2,3</sup>

<sup>1</sup> Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Aachen, Germany

<sup>2</sup> Interdisciplinary Centre for Clinical Research, Medical School, RWTH Aachen University, Aachen, Germany

<sup>3</sup> JARA–Translational Brain Medicine, Research Centre Jülich, Jülich, Aachen, Germany

<sup>4</sup> Department of Neurology, Clinical Neuropsychology, Medical School, RWTH Aachen University, Aachen, Germany

## Edited by:

Benjamin Kreifelts, University of Tübingen, Germany

## Reviewed by:

Marianne Latinus, Aix-Marseille Université, France

Ralf Veit, Institute of Medical Psychology, Germany

## \*Correspondence:

Dhana Wolf, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical School, RWTH Aachen University, Pauwelsstr. 30, D-52074 Aachen, Germany  
e-mail: dhwolf@ukaachen.de

<sup>†</sup> These authors have contributed equally to this work.

Salient exogenous stimuli modulate attentional processes and lead to attention shifts—even across modalities and at a pre-attentive level. Stimulus properties such as hemispheric laterality and emotional valence influence processing, but their specific interaction in audio-visual attention paradigms remains ambiguous. We conducted an fMRI experiment to investigate the interaction of supramodal spatial congruency, emotional salience, and stimulus presentation side on neural processes of attention modulation. Emotionally neutral auditory deviants were presented in a dichotic listening oddball design. Simultaneously, visual target stimuli (schematic faces) were presented, which displayed either a negative or a positive emotion. These targets were presented in the left or in the right visual field and were either spatially congruent (valid) or incongruent (invalid) with the concurrent deviant auditory stimuli. According to our expectation we observed that deviant stimuli serve as attention-directing cues for visual target stimuli. Region-of-interest (ROI) analyses suggested differential effects of stimulus valence and spatial presentation on the hemodynamic response in bilateral auditory cortices. These results underline the importance of valence and presentation side for attention guidance by deviant sound events and may hint at a hemispheric specialization for valence and attention processing.

**Keywords: hemodynamic mismatch response, spatial congruency, crossmodal spatial cueing, attention, valence, auditory cortex**

## INTRODUCTION

For an efficient interaction with the environment, information from various sensory modalities is integrated into a unified spatial representation. Since the processing capacity of perceptual input is limited, spatial attention needs to be selectively allocated to relevant stimuli. The importance of such a mechanism is reflected, for instance, in the huge impact of impairments of attention in the left visual field—unilateral neglect—on simple activities of daily living in patients with right-hemisphere damage (for a review see Danckert and Ferber, 2006). The allocation of attention is guided by the processing and evaluation of information from various sensory modalities as well as top-down control mechanisms. The underlying neural mechanisms of attention allocation are modulated by different stimulus properties such as presentation side and emotional content.

## SPATIAL CUEING AND SPATIAL CONGRUENCY

Across modalities, attention networks interact at a supramodal level to guide attention to the most important stimulus in the environment. Attention distribution mechanisms modulate processing in the target-modality sensory cortices and also in other, usually to be ignored modalities. This was demonstrated for audio-visual, tactile-visual, and tactile-auditory interactions (for

a review, see Eimer and Driver, 2001). Thus, attention shifts following exogenous cues influence target detection even across modalities (Eimer and Driver, 2001; McDonald et al., 2003; Menning et al., 2005).

When a sensory stimulation is presented on the same side as the target stimulus, this valid cue reduces reaction time. When the stimulation is presented on the opposite side (invalid cue), attention is directed away from the target location, leading to increased reaction time (Posner, 1980). The auditory oddball paradigm is a powerful mean to investigate this cueing effect at a pre-attentive level. The paradigm comprises deviant auditory stimuli presented in a rapid sequence of frequent standard stimuli. Deviants of any type implemented in a series of standard events constitute a violation of the established pattern and elicit a mismatch response in the auditory cortex, triggering attentional shifts to the side of the deviant (Schröger, 1996; Schock et al., 2013). Schröger et al. combined an auditory oddball design presented on a to-be-ignored ear with an auditory GoNogo-task on the other ear. They found a prolonged reaction time to the task when it was preceded by a deviant sound on the ignored ear. This was interpreted as a shift of attention induced by the deviant sound.

This oddball-induced mismatch response is a well-established tool in the investigation of neural responses (EEG: Näätänen et al.,

1989; Schröger, 1996; Näätänen et al., 2004; MEG: Mathiak et al., 2005; fMRI: Mathiak et al., 2002; Schock et al., 2012). These studies reported that even non-attended changes in the auditory stream induce increased activation of the auditory cortex. Further, this effect is modulated by attention (Alho et al., 1999; Zvyagintsev et al., 2011), cross-modal information (Calvert and Campbell, 2003; Kayser et al., 2005, 2007; Zvyagintsev et al., 2009) and top-down regulation (van Atteveldt et al., 2007; Zvyagintsev et al., 2009; Hsieh et al., 2012).

### THE INFLUENCE OF VALENCE ON ATTENTION DISTRIBUTION

Emotional valence, an important factor contributing to the salience of stimuli, influences attention-related processing. Negative valence draws attention effectively and enhances stimulus processing (Eastwood et al., 2001; Fenske and Eastwood, 2003; Rowe et al., 2007). A negative stimulus can narrow the attentional focus accompanied by enhanced processing. In visual search, negative stimuli are found faster and with fewer errors than positive stimuli (Eastwood et al., 2001). In a preceding behavioral study, Schock et al. (2013) combined schematic faces—representing emotionally salient target stimuli—with dichotic syllables presented in an oddball paradigm. Visual and auditory stimuli were presented on the same side (spatially congruent) or on opposite sides (spatially incongruent). Prolonged response time was found to left-lateralized spatially incongruent stimuli if they were of positive valence, while emotionally negative targets were not affected by spatial incongruency. This behavioral effect may be induced by pre-attentive auditory processing.

### AIMS AND HYPOTHESES

Although neural mechanisms of attention have been extensively studied, the influence of emotional valence, hemispheric lateralization, and their specific interactions on early auditory processing, remain ambiguous. The goal of the present study was to investigate the interaction of emotional valence, visual field, and spatial congruency on auditory cortex activation. For this aim, we conducted an fMRI experiment combining a dichotic listening oddball design with emotional visual targets. We hypothesized that deviant sounds would trigger a shift of attention towards the side of the deviant and thereby accelerate response time to spatially congruent visual stimuli, while response time to spatially incongruent stimuli would be increased. Emotional salience was expected to interact with spatial congruency, in that negative salience would reduce this reaction time costs in spatially incongruent trials. Regarding the neurophysiological data, we hypothesized that auditory cortex responses to deviant sounds would be modulated by spatial congruency with, and emotion of, the target faces. Furthermore, attention and salience-related networks were expected to respond stronger to spatially incongruent than to congruent stimuli.

### MATERIALS AND METHODS

We conducted an fMRI experiment combining a dichotic listening oddball design with emotional visual targets. A deviant stimulus would induce an attention shift to the side of the deviant. Visual target stimuli were schematic faces that exclusively displayed the feature of positive and negative face expressions. They were

presented in the left or right visual field either spatially congruent (valid) or incongruent (invalid) with the auditory deviants. Participants were instructed to ignore the auditory modality and to indicate the detection of a visual target stimulus via button press. We have successfully applied this paradigm in a previous behavioral study and demonstrated significant interactions of valence and visual field and of valence and spatial congruency on reaction time (Schock et al., 2013).

### SUBJECTS

Sixteen healthy volunteers (8 females, age 21–36 years, mean age  $24.9 \pm 4.5$  years) participated in the study. Fifteen subjects were right-handed, as indicated by the laterality quotient of the Edinburgh Inventory (Oldfield, 1971) and one subject was ambidextrous. Subjects were screened with the Structured Clinical Interview (SCID-I; Wittchen et al., 1997) for the diagnostic and Statistical Manual of Mental Disorders (DSM-IV) to exclude subjects with psychiatric disorders. The study was approved by the local Ethics Committee of the Medical School of the RWTH Aachen University and was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained from all subjects.

### STIMULI

#### *Auditory stimuli*

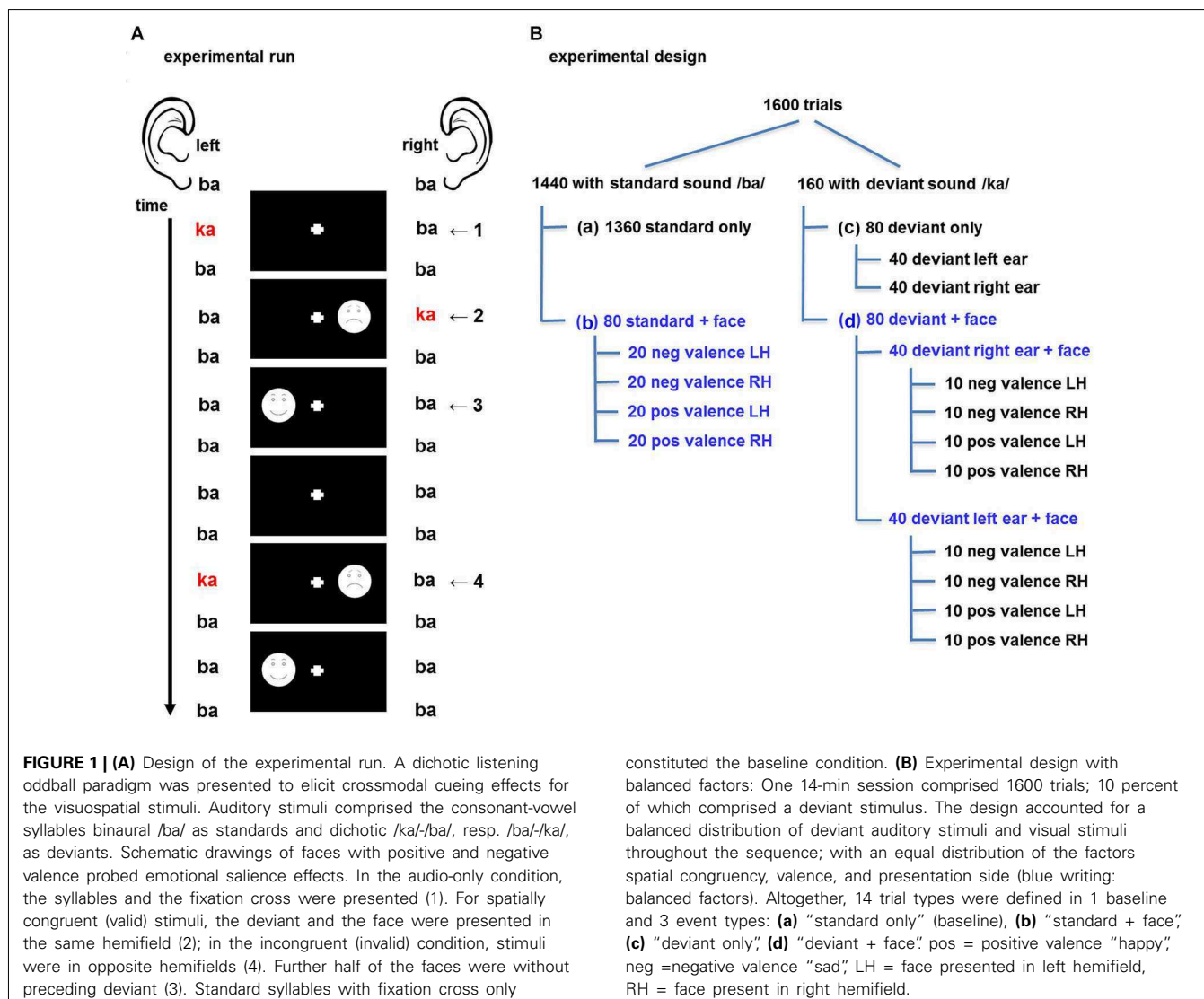
The stimuli comprised two consonant-vowel syllables (/ba/ and /ka/) presented in a dichotic listening oddball design. In 90% of the presentations the standard /ba/ was presented to both ears. In the remaining 10% the deviant /ka/ was presented to one ear (5% to the left ear and 5% to the right ear) and the standard /ba/ to the other ear (left/ba/-right/ka/ and left/ka/-right/ba/). The syllables were delivered synchronous over both ears, with equal sound level and duration (300 ms), and with a constant stimulus onset asynchrony (SOA) of 500 ms.

#### *Visual stimuli*

Visual stimuli were composed of schematic drawings of faces with either a positive or negative emotional expression. The positive expression included a half-elliptic mouth shape and round-shaped eyebrows. The negative expression was achieved by inverting the mouth downwards and orienting the inner eyebrow ends upwards. These expressions were validated beforehand and were rated as either happy or sad (see Schock et al., 2013). For trials including a face, the visual stimuli were presented 175 ms after auditory stimulus onset and lasted for 150 ms.

### DESIGN

The experiment comprised three sessions of 14 min each. In each session, 1600 trials were presented in a rapid, event-related design. A trial was defined as the presentation of one auditory stimulus with or without a face stimulus, so that the length of each trial corresponded to the auditory stimuli onset asynchrony of 500 ms (Figure 1A). Audio-visual trial types were specified by the factors *valence* (positive vs. negative facial expression), *visual field* (left vs. right visual field), and *ear* (left- vs. right-ear deviant). In total 14 events of interest emerged: 4 times



**FIGURE 1 | (A)** Design of the experimental run. A dichotic listening oddball paradigm was presented to elicit crossmodal cueing effects for the visuospatial stimuli. Auditory stimuli comprised the consonant-vowel syllables binaural /ba/ as standards and dichotic /ka-/ba/, resp. /ba-/ka/, as deviants. Schematic drawings of faces with positive and negative valence probed emotional salience effects. In the audio-only condition, the syllables and the fixation cross were presented (1). For spatially congruent (valid) stimuli, the deviant and the face were presented in the same hemifield (2); in the incongruent (invalid) condition, stimuli were in opposite hemifields (4). Further half of the faces were without preceding deviant (3). Standard syllables with fixation cross only

constituted the baseline condition. **(B)** Experimental design with balanced factors: One 14-min session comprised 1600 trials; 10 percent of which comprised a deviant stimulus. The design accounted for a balanced distribution of deviant auditory stimuli and visual stimuli throughout the sequence; with an equal distribution of the factors spatial congruency, valence, and presentation side (blue writing: balanced factors). Altogether, 14 trial types were defined in 1 baseline and 3 event types: **(a)** “standard only” (baseline), **(b)** “standard + face”, **(c)** “deviant only”, **(d)** “deviant + face”: pos = positive valence “happy”, neg = negative valence “sad”, LH = face presented in left hemifield, RH = face present in right hemifield.

face and standard sound (face left/right, positive/negative), 2 distinct deviants (left/right), and 8 events comprising a face and a deviant (combination of the previous factors). These trial types are displayed in **Figure 1B**. Each session comprised the baseline condition with 1440 trials with “standard only” event types (standard auditory event without visual stimulus; **Figure 1B-a**), 80 trials with the four “standard + face” event types (20 each, **Figure 1B-b**), 80 trials with two “deviant only” event types (40 each, **Figure 1B-c**) and 80 trials with eight “deviant + face” event types (10 each, **Figure 1B-d**). In 50% of the cases the auditory and visual stimulus were presented on the same side (spatially congruent, “valid”) and in the other 50% the stimuli were presented on the opposite side (spatially incongruent, “invalid”; see **Figure 1A**).

The design accounted for a balanced distribution of auditory deviants and visual stimuli throughout the oddball sequence. A minimum of two standards (left/ba/-right/ba/) were placed between two deviants (left/ba/-right/ka/ or left/ka/-right/ba/).

Likewise, at least two trials without visual stimulation were presented between two trials with face presentation to prevent a button press overlap. Each session was individually randomized: first the deviant distribution was randomized. Then, left-ear and right-ear deviant events were randomly selected and assigned to the “deviant + face” events and in a third step the standard events were randomly selected and assigned to the “standard + face” events.

Participants were instructed to ignore the sound and report detection of the face stimulus via button pressing. Visual stimuli were displayed via MR-compatible video goggles (Visua-StimDigital, Resonance Technology, RT, Northridge, CA, USA) and sounds were delivered with MR-compatible headphones. Sound levels were adjusted to comfortable hearing level and good audibility during scanning. Stimulus presentation and response time recording was performed using the software *Presentation* (Version 10.0; Neurobehavioral Systems, Inc., Albany, CA).

## MR IMAGING

Functional imaging was conducted on a 3T Magnetom Trio MR scanner (Siemens Medical Systems, Erlangen, Germany) in the department of Psychiatry, Psychotherapy and Psychosomatics at the Medical School of RWTH Aachen University. Functional images were collected with echo planar imaging (EPI) sensitive to blood oxygenation level dependent (BOLD) contrast (interleaved acquisition of 34 slices; repetition time [TR] = 2000 ms; echo time [TE] = 28 ms; flip angle [FA] = 77°; slice thickness = 3 mm; gap 0.75 mm; matrix size = 64 × 64; field of view [FOV] = 192 × 192 mm<sup>2</sup>; voxel size = 3 × 3 × 3 mm<sup>3</sup>). Slices covered the entire cerebral cortex and were positioned oblique-transversally to achieve maximal brain coverage. 420 volumes were collected per session, of which the first seven were discarded to remove the influence of T1 saturation effects. Head movement was minimized with the use of foam wedges to securely hold the head in the 12-channel head coil. Structural images were obtained using a high-resolution T1-weighted 3D-sequence (TR = 1900 ms; inversion time TI = 900 ms; TE = 2.52 ms; FA = 9°; FOV = 256 × 256 mm<sup>2</sup>; 176 3D-partitions with an isotropic resolution of 1 mm<sup>2</sup>).

## BEHAVIORAL DATA ANALYSIS

Button presses in response to visual stimuli were analyzed by subtracting the onset of the visual stimulus from the onset of button press. The mean reaction time (in ms) was calculated for each event type and each participant. A 2 × 2 × 2 repeated-measures ANOVA with the factors *valence* (positive, negative emotion), *visual field* (left, right), and *spatial congruency* (auditory and visual presentation on the same/opposite side) was conducted. Significance level was set at  $p < 0.05$  after Bonferroni correction. Paired  $t$ -tests disentangled the effects *post-hoc*.

## IMAGING DATA ANALYSIS

Functional MRI data analysis was conducted with the Statistical Parametric Mapping software (SPM8<sup>1</sup>; implemented in MATLAB) (MathWorks, Natick, MA, USA). After discarding the first seven volumes, 413 volumes per session from each participant were spatially realigned to the mean image to correct for head movement, normalized to the stereotaxic anatomical MNI (Montreal Neurological Institute) space with 2 mm isotropic voxels, and spatially smoothed with an 8 mm (FWHM) isotropic Gaussian kernel to account for inter-subject variability in brain anatomy and to increase signal-to-noise ratio. A rapid event-related design was chosen to model the experimental conditions (14 audio-visual events, see Section Design) with a general linear model (GLM).

### Whole brain analysis

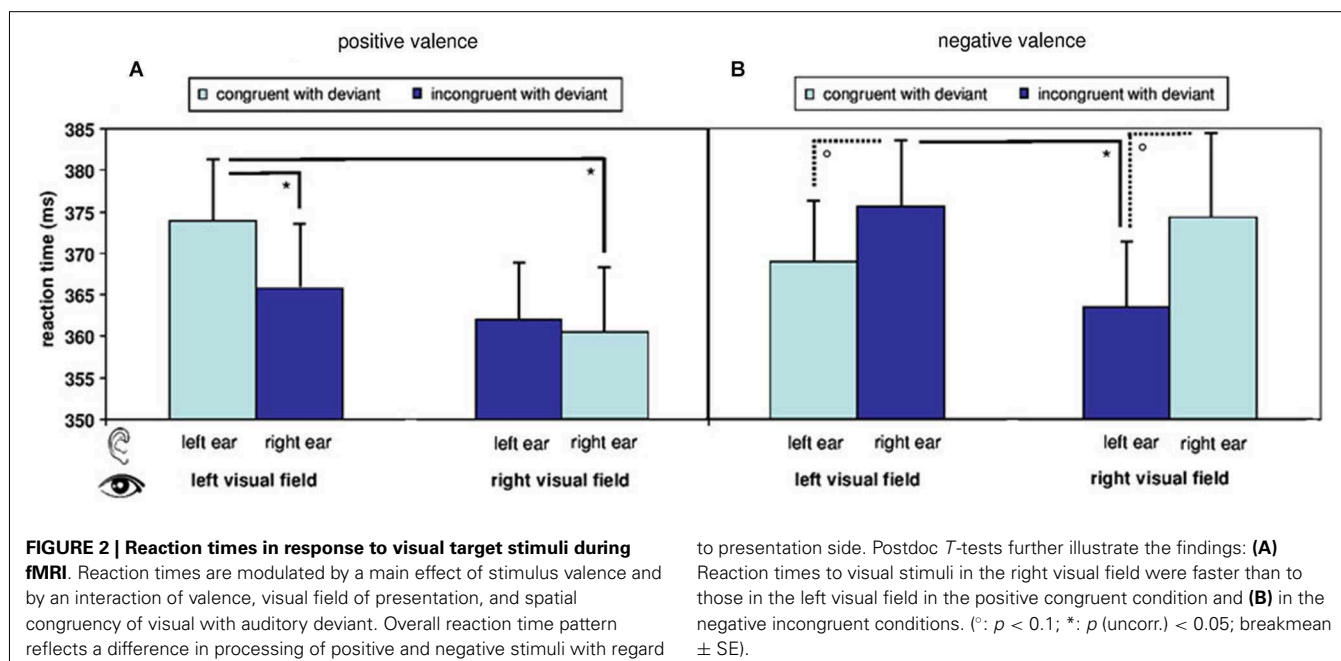
The contrast “deviant without face” vs. “standard without face” (baseline) was built to document auditory cortex activation in response to deviant stimuli with a  $t$ -test. Results were corrected with a family-wise-error (FWE) rate of  $p < 0.05$  to account for multiple testing.

To investigate the neural mechanisms underlying spatial congruency effects, a  $t$ -contrast was designed, which compared spatially incongruent trials (“left deviant + right face”, and “right face + left deviant”) with spatially congruent trials (“left deviant + left face”, and “right deviant + right face”).

### Region-of-interest (ROI) analysis

The contrast estimates for each subject were extracted from bilateral auditory activation cluster peak voxel of the contrast “deviant without face” vs. “standard without face” (baseline; **Figure 2**,

<sup>1</sup>www.fil.ion.ucl.ac.uk





**Table 1 | Deviance responses.**

Contrast	Anatomical area	Hemisphere	BA	MNI coordinates			Peak T-values	Cluster size [voxel]
				X	Y	Z		
"deviant only" vs. "standard only"	Superior temporal gyrus	R	42	68	-26	8	6.80	229
	Superior temporal gyrus	L	22	-62	-24	4	5.77	56

Threshold  $p < 0.05$ , FWE-corr. and cluster extent  $\geq 50$  voxels, BA: Brodmann Area, MNI: Montreal Neurological Institute.

**Table 1).** The response amplitudes at bilateral Auditory cortex (AC) were assessed by a  $2 \times 2 \times 2$  repeated-measures ANOVA, with the factors stimulus *valence* (positive, negative), *visual field* (left, right) and *ear* of deviant presentation (left, right). For this analysis, only trials comprising a deviant and a face were considered.

## RESULTS

### BEHAVIORAL RESULTS

The here applied, uninformative cueing paradigm yielded no speeding of reaction times after auditory deviants ( $368.10 \pm 28.98$  ms) as compared to standards ( $367.13 \pm 25.39$  ms;  $T_{15} = 0.575$ ,  $p = 0.574$ , n.s.). A repeated-measure ANOVA including the factors *spatial congruency*, *visual field*, and *valence* of the visual stimulus analyzed the reaction times in response to stimuli after deviants. A significant main effect emerged for *valence* ( $F_{[1,15]} = 7.820$ ,  $p = 0.014$ ). Mean differences ( $\pm$ SE) yielded significantly faster reactions to positive as compared to negative targets ( $-5.010 \pm 1.792$ ). The factor *spatial congruency* yielded a trend-level effect ( $F_{[1,15]} = 3.476$ ,  $p = 0.082$ ), while *visual field* did not yield significant effects ( $F_{[1,15]} = 2.305$ ,  $p = 0.150$ , n.s.). However, a significant result emerged for the triple interaction between *visual field*, *valence*, and *spatial congruency* ( $F_{[1,15]} = 9.262$ ,  $p = 0.008$ ). In the *post-hoc* *t*-tests with Bonferroni-correction for multiple comparisons, no contrast between the eight audio-visual conditions yielded significance (see descriptive analysis in **Figure 2**).

### fMRI RESULTS

#### Whole brain analysis

Deviant auditory stimuli yielded large cluster in bilateral AC ( $p < 0.05$ , FWE corr., **Figure 3**, **Table 1**). The incongruency contrast, comparing spatially incongruent with congruent trials, did not reveal significant effects in a whole-brain contrast.

#### ROI analysis

The contrast estimates of the contrast "deviant" vs. "standard" were extracted from the cluster peak voxel in bilateral AC (MNI coordinates [68, -26, 8] and [-62, -23, 4]; **Figure 4**). *Visual field* yielded a significant main effect (left AC:  $F_{[1,15]} = 6.64$ ,  $p = 0.0112$ ; right AC:  $F_{[1,15]} = 10.04$ ,  $p = 0.0019$ ), and *ear* (deviant presentation side) yielded a main effect in the right AC ( $F_{[1,15]} = 9.94$ ,  $p = 0.002$ ). *Valence* trended toward significance in the right AC ( $F_{[1,15]} = 3.16$ ,  $p = 0.078$ ). Furthermore, a significant interaction of *visual field* and *valence* emerged in both ACs (left AC:  $F_{[1,15]} = 4.29$ ,  $p = 0.0405$ ; right AC:  $F_{[1,15]} = 16.4$ ,  $p < 0.001$ ). *Spatial congruency* (interaction of *ear* and *visual field*) failed a significant effect (left AC:  $F_{[1,15]} = 0.01$ ,  $p = 0.918$ , n.s.; right AC:  $F_{[1,15]} = 0.13$ ,  $p = 0.717$ , n.s.). Likewise, significance failed for the

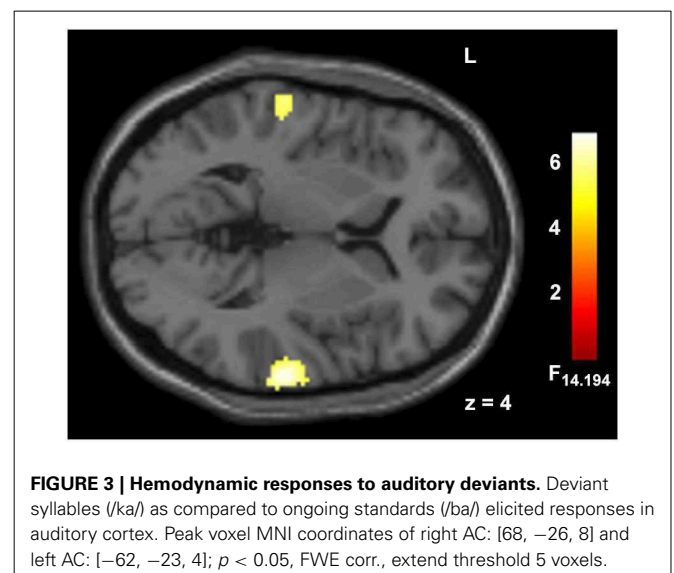
interaction of *ear* and *valence* (left AC:  $F_{[1,15]} = 0.57$ ,  $p = 0.451$ ; right AC:  $F_{[1,15]} = 0.82$ ,  $p = 0.367$ ), as well as a triple interaction of *ear*, *visual field*, and *valence* (left AC:  $F_{[1,15]} = 0.01$ ,  $p = 0.908$ ; right AC:  $F_{[1,15]} = 0.45$ ,  $p = 0.503$ ).

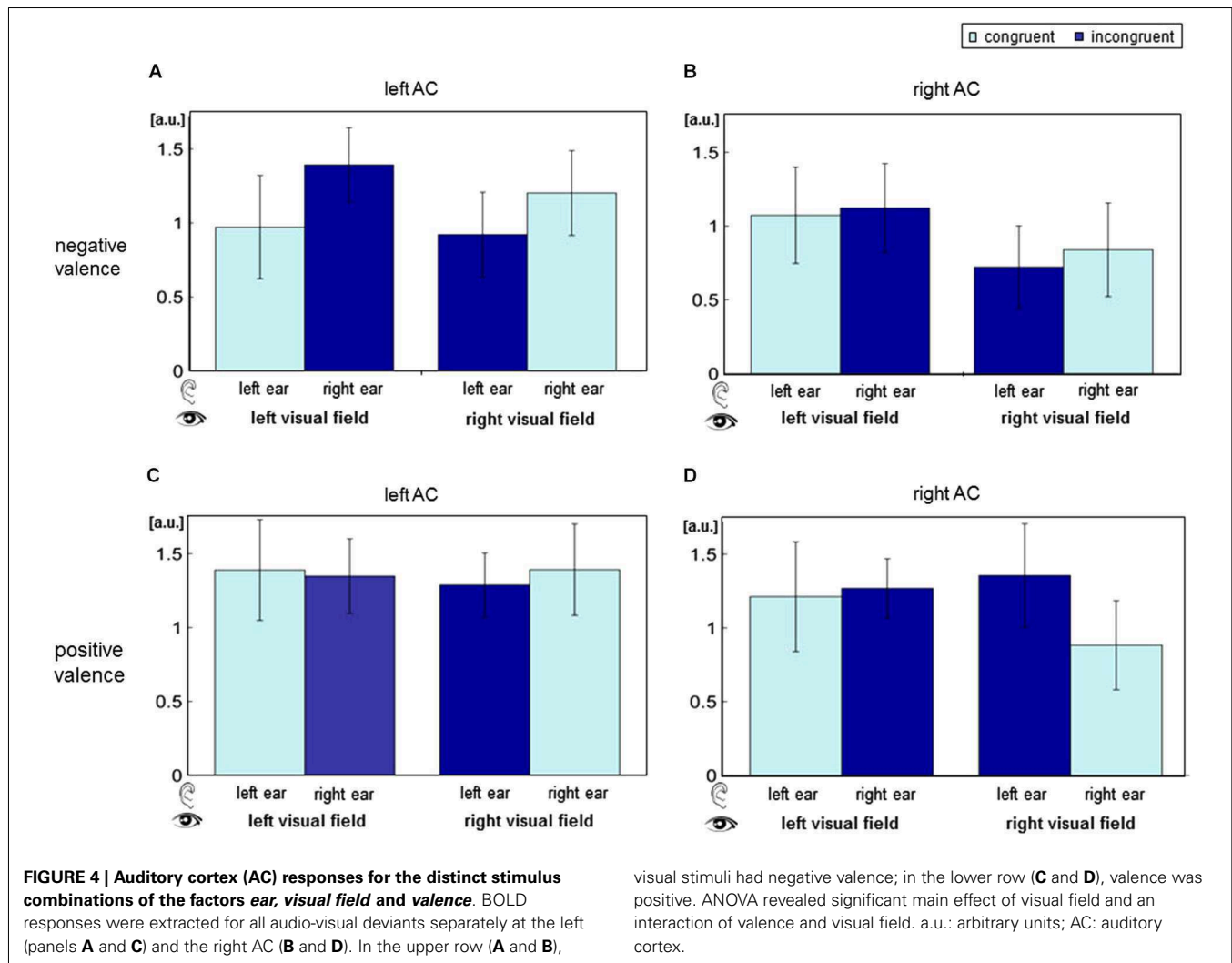
## DISCUSSION

In the present study neural correlates of attention processes in a dichotic listening oddball paradigm were investigated. An auditory stimulus (standard or deviant) preceded visual target presentation (positive and negative schematic faces in the left or right visual field). Events comprising a deviant auditory stimulus were equally distributed between spatially congruent and incongruent trials, thus constituting uninformative spatial cues with a 50% validity ratio. In analogy to the cueing effect of Posner (1980), but applying uninformative crossmodal cues (Spence and Driver, 1997; Ward et al., 2000; Mazza et al., 2007), the auditory deviants were expected to serve as crossmodal spatial cues for visual target stimuli. Indeed, presentation side and valence of visual stimuli significantly affected AC response.

### BEHAVIORAL DATA

To our knowledge, this is the first study to combine schematic faces as emotional stimuli with a spatial cueing (mismatch auditory oddball) design in fMRI. The effectiveness of schematic faces for realistic emotion display (Dyck et al., 2008) and for affective facial stimuli in modulating the allocation of attention in behavioral studies has been demonstrated previously.





Fox et al. (2001) used schematic faces as spatially incongruent cues in a visual paradigm and demonstrated that negative faces draw and hold attention effectively. In a similar vein, other studies have demonstrated comparable valence effects for the presentation of schematic, or cartoon faces (Eastwood et al., 2001; Fenske and Eastwood, 2003; Santos et al., 2011).

In our previous behavioral study employing the here reported paradigm, we observed that reaction time in response to a visual stimulus following a deviant auditory stimulus was facilitated. Furthermore, reaction time costs emerged in response to spatially incongruent stimuli only in the positive valence condition, whereas negative valence overcame this incongruity effect (Schock et al., 2013). The interaction between spatial presentation and emotional valence of target stimuli is a well-documented phenomenon (Eastwood et al., 2001; Fenske and Eastwood, 2003). The current findings partially reflect the previously observed effects (Schock et al., 2013). Deviant sounds preceding a face did not reduce reaction times compared to faces preceded by a standard sound, but a main effect of *valence* and an interaction

of spatial *congruency*, *visual field* and *valence* was replicated. In summary, alerting may not have been efficient in the noisy MR-environment, but the overall reaction time pattern reflects a difference in processing of positive and negative stimuli with spatial congruency constituting a modulating factor.

#### EFFECT OF DEVIANTS AND SPATIAL INCONGRUENCY

Contrasting the presentation of a deviant sound with a standard, deviance yielded increased hemodynamic responses in bilateral AC. Increased AC activation in response to non-attended deviants within an oddball listening design is well established (EEG: Sams et al., 1985; Näätänen et al., 1989, 2004; MEG: Alho et al., 1998; Mathiak et al., 2000; Phillips et al., 2000; fMRI: Mathiak et al., 2002; Schock et al., 2012). This effect on very early processing steps may underlie the redirection of attention toward changing features in the environment (Näätänen, 1995). Deviant stimuli were presented either spatially congruent or incongruent with a visual target. Contrasting our hypothesis, no spatial congruency effect emerged at the whole-brain level. Our design, comprising uninformative cues (equal probability

of presentation with or without a face) and a wide range of stimulus combinations (valence, visual field, and spatial congruency) may have been inadequate for this question. Until now, only few fMRI studies have investigated whole-brain activation changes in response to spatially incongruent vs. congruent audiovisual stimuli. Sestieri et al. (2006) reported increased response of superior temporal sulcus to spatially congruent audiovisual stimuli, as compared to spatially incongruent stimuli. However, in contrast to our paradigm, participants were explicitly asked to pay attention to both the visual and the auditory stimuli.

### EFFECTS OF DEVIANCE AND VISUAL FIELD ON AC PROCESSING

ANOVA analyses of AC peak voxel responses revealed a main effect of *visual field* and an interaction effect of *visual field* and *valence* in both AC. Furthermore, a trend-level effect of *valence* emerged in the right AC.

Up until now, the direct comparison of AC responses to spatially incongruent vs. congruent audiovisual stimuli has not been investigated with an auditory oddball design and fMRI. In an EEG study, Teder-Sälejärvi et al. (2005) compared the ERP of bimodal audiovisual stimuli in spatially congruent and incongruent conditions. For incongruent stimulus pairings, the authors reported a shift in phase and amplitude of activity at 100–400 ms, which were allocated to ventral occipito-temporal cortex. Spatially congruent pairings yielded an amplitude modulation of activity at 260–280 ms, localized to superior temporal regions. Though this result hints at an effect of spatial congruency on early auditory processing, we did not find an interaction of *ear* of presentation and *visual field* (i.e., spatial congruency). However, the significant effect of deviance and of visual field of face presentation on AC response suggests a similar notion.

In line with our results, several studies reported that visual input can modulate auditory processing (Kayser et al., 2007; Zvyagintsev et al., 2009; Hsieh et al., 2012). Direct connections between primary auditory and visual cortex have additionally been identified by anatomic studies in primates and fMRI connectivity analysis (Eckert et al., 2008).

### EFFECTS OF EMOTIONAL VALENCE ON AC PROCESSING

The influence of emotion and valence of visual stimuli on early auditory processing has been reported in a wide range of electrophysiological studies. In an EEG-experiment (Alexandrov et al., 2007) emotional context (monetary reward or punishment) led to significantly larger auditory cortex event-related potentials in response to negative as compared to positive trials. Visually induced emotional states with positive or negative pictures have also been reported to modulate event-related potentials of auditory stimuli (Surakka et al., 1998; Yamashita et al., 2005; Sugimoto et al., 2007; Domínguez-Borràs et al., 2009; Wang et al., 2009). Functional MRI studies complement this picture: Schock et al. (2012) observed right lateralized prefrontal cortex (PFC) activation and enhanced processing of right-ear deviants in the sad mood condition. Fear conditioning, as conveyed by visual stimuli, also affected AC activity (Armony and Dolan, 2001). AC response to fear-conditioned stimuli was modulated by the

presence of a visual context for the likelihood of aversive stimuli appearance.

Complementing these studies, we were the first to investigate face valence effect on AC response in the context of spatial congruency (induced by deviating syllables in an auditory oddball design) and visual field of presentation. The significant interaction of valence and visual field indicates that emotional valence of visual stimuli may modulate auditory processing at an early stage. This finding may support the hypotheses of hemispheric specialization for multimodal integration and emotion processing. According to the valence hypothesis, emotional contexts, especially negative stimuli, are primarily lateralized to the right hemisphere, while positive stimuli are dominantly processed in the left hemisphere (for a review of the valence hypothesis, see Demaree et al., 2005). Similarly, Schock et al. reported enhanced ipsilateral processing of right-ear deviants in the right AC during sad mood, but not during happy or neutral mood. Further, Petit et al. report a right hemispheric enhancement in frontal and temporal areas in response to unattended deviant tones (Petit et al., 2007). Although recent meta-analyses of neuroimaging studies investigating laterality effects on emotional face processing were not able to confirm the valence hypothesis (Wager et al., 2003; Fusar-Poli et al., 2009), our results suggest that within multimodal settings, processing of deviant tones may be affected by laterality and valence—at least at an early auditory stage.

### LIMITATIONS

To our knowledge, this study is the first to combine fMRI measurement with a dichotic listening oddball design and emotional faces as target stimuli. The resulting high number of conditions and parameters are challenging to take into account. Due to limitations in paradigm length, visual stimuli with neutral emotion were not implemented. Furthermore, deviant oddball paradigms yield comparably small number of events for subsequent analyses since a subdivision into the large number of varying conditions is necessary. However, increasing the experiment duration may induce unwanted effects on attention and arousal. This design-intrinsic small event number as well as a relatively low number of participants may account for the missing results of the whole brain contrasts. The neural response to deviants was clearly detectable in the AC, yet this observation was not fully reflected in the behavioral responses. Our inability to fully reproduce the behavioral effect of spatially congruent stimuli may be explained by the fMRI measurement. The difficulty to reproduce behavioral data within the scanner environment and concurrent reaction time slowing is frequently reported (Plank et al., 2012). Participants in our fMRI experiment responded on average 40 ms slower and with a larger deviation than participants in the behavioral study utilizing the same paradigm (Schock et al., 2013). The fMRI environment may create increased arousal and ongoing distraction due to scanner noise. Indeed, Skouras et al. (2013) recently reported an effect of scanner noise on affective brain processes. Therefore, the fMRI environment may have prevented any salience-driven response acceleration in our study. Nevertheless, although the visual target stimuli exhibit clear valence, the stimuli's arousal value is rather low. Future studies may address the question of stimulus arousal and valence

differences. Higher arousal values may increase the observed effects of valence and visual presentation side. Furthermore, a neutral condition could be implemented to further disentangle emotion and attention effects and laterality of hemispheric activation.

## CONCLUSION

Deviant stimuli of a dichotic listening oddball design were shown to serve as attention-directing cues for visual target stimuli during fMRI measurement. Deviant events and visual targets of different valence induced supramodal modulation of auditory cortices. This modulation was significantly dependent on visual field of presentation and on the interaction of valence and visual field. The present results underline the importance of valence and presentation side for attention guidance by deviant sound events and support the notion of intermodal effects on auditory processing, which may be modulated by emotions.

## ACKNOWLEDGMENTS

This research project was supported by the “Interdisciplinary Centre for Clinical Research (IZKF) Aachen” (N4.2) of the Faculty of Medicine at RWTH Aachen University and the German Research Foundation (Deutsche Forschungsgemeinschaft DFG, IRTG 1328 and MA2631/4-1).

## REFERENCES

- Alexandrov, Y. I., Klucharev, V., and Sams, M. (2007). Effect of emotional context in auditory-cortex processing. *Int. J. Psychophysiol.* 65, 261–271. doi: 10.1016/j.ijpsycho.2007.05.004
- Alho, K., Medvedev, S. V., Pakhomov, S. V., Roudas, M. S., Tervaniemi, M., Reinikainen, K., et al. (1999). Selective tuning of the left and right auditory cortices during spatially directed attention. *Brain Res. Cogn. Brain Res.* 7, 335–341. doi: 10.1016/s0926-6410(98)00036-6
- Alho, K., Winkler, I., Escera, C., Huottilainen, M., Virtanen, J., Jääskeläinen, I. P., et al. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: magnetoencephalographic recordings. *Psychophysiology* 35, 211–224. doi: 10.1111/1469-8986.3520211
- Armory, J. L., and Dolan, R. J. (2001). Modulation of auditory neural responses by a visual context in human fear conditioning. *Neuroreport* 12, 3407–3411. doi: 10.1097/00001756-200110290-00051
- Calvert, G. A., and Campbell, R. (2003). Reading speech from still and moving faces: the neural substrates of visible speech. *J. Cogn. Neurosci.* 15, 57–70. doi: 10.1162/089892903321107828
- Danckert, J., and Ferber, S. (2006). Revisiting unilateral neglect. *Neuropsychologia* 44, 987–1006. doi: 10.1016/j.neuropsychologia.2005.09.004
- Demaree, H. A., Everhart, D. E., Youngstrom, E. A., and Harrison, D. W. (2005). Brain lateralization of emotional processing: historical roots and a future incorporating “dominance”. *Behav. Cogn. Neurosci. Rev.* 4, 3–20. doi: 10.1177/1534582305276837
- Domínguez-Borràs, J., Trautmann, S.-A., Erhard, P., Fehr, T., Herrmann, M., and Escera, C. (2009). Emotional context enhances auditory novelty processing in superior temporal gyrus. *Cereb. Cortex* 19, 1521–1529. doi: 10.1093/cercor/bhn188
- Dyck, M., Winbeck, M., Leiberg, S., Chen, Y., Gur, R. C., and Mathiak, K. (2008). Recognition profile of emotions in natural and virtual faces. *PLoS One* 3:e3628. doi: 10.1371/journal.pone.0003628
- Eastwood, J. D., Smilek, D., and Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Percept. Psychophys.* 63, 1004–1013. doi: 10.3758/bf03194519
- Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, D., and Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: evidence from intrinsic fMRI connectivity analysis. *Hum. Brain Mapp.* 29, 848–857. doi: 10.1002/hbm.20560
- Eimer, M., and Driver, J. (2001). Crossmodal links in endogenous and exogenous spatial attention: evidence from event-related brain potential studies. *Neurosci. Biobehav. Rev.* 25, 497–511. doi: 10.1016/s0149-7634(01)00029-x
- Fenske, M. J., and Eastwood, J. D. (2003). Modulation of focused attention by faces expressing emotion: evidence from flanker tasks. *Emotion* 3, 327–343. doi: 10.1037/1528-3542.3.4.327
- Fox, E., Russo, R., Bowles, R., and Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *J. Exp. Psychol. Gen.* 130, 681–700. doi: 10.1037/0096-3445.130.4.681
- Fusar-Poli, P., Placentino, A., Carletti, F., Allen, P., Landi, P., Abbamonte, M., et al. (2009). Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neurosci. Lett.* 452, 262–267. doi: 10.1016/j.neulet.2009.01.065
- Hsieh, P.-J., Colas, J. T., and Kanwisher, N. (2012). Spatial pattern of BOLD fMRI activation reveals cross-modal information in auditory cortex. *J. Neurophysiol.* 107, 3428–3432. doi: 10.1152/jn.01094.2010
- Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron* 48, 373–384. doi: 10.1016/j.neuron.2005.09.018
- Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *J. Neurosci.* 27, 1824–1835. doi: 10.1523/jneurosci.4737-06.2007
- Mathiak, K., Hertrich, I., Lutzenberger, W., and Ackermann, H. (2000). Encoding of temporal speech features (formant transients) during binaural and dichotic stimulus application: a whole-head magnetoencephalography study. *Brain Res. Cogn. Brain Res.* 10, 125–131. doi: 10.1016/s0926-6410(00)00035-5
- Mathiak, K., Hertrich, I., Zvyagintsev, M., Lutzenberger, W., and Ackermann, H. (2005). Selective influences of cross-modal spatial-cues on preattentive auditory processing: a whole-head magnetoencephalography study. *Neuroimage* 28, 627–634. doi: 10.1016/j.neuroimage.2005.06.030
- Mathiak, K., Rapp, A., Kircher, T. T. J., Grodd, W., Hertrich, I., Weiskopf, N., et al. (2002). Mismatch responses to randomized gradient switching noise as reflected by fMRI and whole-head magnetoencephalography. *Hum. Brain Mapp.* 16, 190–195. doi: 10.1002/hbm.10041
- Mazza, V., Turatto, M., Rossi, M., and Umiltà, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia* 45, 514–522. doi: 10.1016/j.neuropsychologia.2006.02.010
- McDonald, J. J., Teder-Sälejärvi, W. A., Di Russo, F., and Hillyard, S. A. (2003). Neural substrates of perceptual enhancement by cross-modal spatial attention. *J. Cogn. Neurosci.* 15, 10–19. doi: 10.1162/089892903321107783
- Menning, H., Ackermann, H., Hertrich, I., and Mathiak, K. (2005). Spatial auditory attention is modulated by tactile priming. *Exp. Brain Res.* 164, 41–47. doi: 10.1007/s00221-004-2212-7
- Näätänen, R. (1995). The mismatch negativity: a powerful tool for cognitive neuroscience. *Ear Hear.* 16, 6–18. doi: 10.1097/00003446-199502000-00002
- Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., and Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neurosci. Lett.* 98, 217–221. doi: 10.1016/0304-3940(89)90513-2
- Näätänen, R., Pakarinen, S., Teemu, R., and Rika, T. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clin. Neurophysiol.* 115, 140–144. doi: 10.1016/j.clinph.2003.04.001
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Petit, L., Simon, G., Joliet, M., Andersson, F., Bertin, T., Zago, L., et al. (2007). Right hemisphere dominance for auditory attention and its modulation by eye position: an event related fMRI study. *Restor. Neurol. Neurosci.* 25, 211–225.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: an MEG mismatch study. *J. Cogn. Neurosci.* 12, 1038–1055. doi: 10.1162/08989290051137567
- Plank, T., Rosengarth, K., Song, W., Ellermeier, W., and Greenlee, M. W. (2012). Neural correlates of audio-visual object recognition: effects of implicit spatial congruency. *Hum. Brain Mapp.* 33, 797–811. doi: 10.1002/hbm.21254
- Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231



- Rowe, G., Hirsh, J. B., and Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proc. Natl. Acad. Sci. U S A* 104, 383–388. doi: 10.1073/pnas.0605198104
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. doi: 10.1016/0168-5597(85)90054-1
- Santos, A., Mier, D., Kirsch, P., and Meyer-Lindenberg, A. (2011). Evidence for a general face salience signal in human amygdala. *Neuroimage* 54, 3111–3116. doi: 10.1016/j.neuroimage.2010.11.024
- Schock, L., Bhavsar, S., Demenescu, L. R., Sturm, W., and Mathiak, K. (2013). Does valence in the visual domain influence the spatial attention after auditory deviants? Exploratory data. *Front. Behav. Neurosci.* 7:6. doi: 10.3389/fnbeh.2013.00006
- Schock, L., Dyck, M., Demenescu, L. R., Edgar, J. C., Hertrich, I., Sturm, W., et al. (2012). Mood modulates auditory laterality of hemodynamic mismatch responses during dichotic listening. *PLoS One* 7:e31936. doi: 10.1371/journal.pone.0031936
- Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *J. Cogn. Neurosci.* 8, 527–539. doi: 10.1162/jocn.1996.8.6.527
- Sestieri, C., Di Matteo, R., Ferretti, A., Del Gratta, C., Caulo, M., Tartaro, A., et al. (2006). “What” versus “where” in the audiovisual domain: an fMRI study. *Neuroimage* 33, 672–680. doi: 10.1016/j.neuroimage.2006.06.045
- Skouras, S., Gray, M., Critchley, H., and Koelsch, S. (2013). fMRI scanner noise interaction with affective neural processes. *PLoS One* 8:e80564. doi: 10.1371/journal.pone.0080564
- Spence, C., and Driver, J. (1997). Audiovisual links in exogenous covert spatial orienting. *Percept. Psychophys.* 59, 1–22. doi: 10.3758/bf03206843
- Sugimoto, S., Nittono, H., and Hori, T. (2007). Visual emotional context modulates brain potentials elicited by unattended tones. *Int. J. Psychophysiol.* 66, 1–9. doi: 10.1016/j.ijpsycho.2007.05.007
- Surakka, V., Tenhunen-Eskelinen, M., Hietanen, J. K., and Sams, M. (1998). Modulation of human auditory information processing by emotional visual stimuli. *Brain Res. Cogn. Brain Res.* 7, 159–163. doi: 10.1016/S0926-6410(98)00021-4
- Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., and Hillyard, S. A. (2005). Effects of spatial congruity on audio-visual multimodal integration. *J. Cogn. Neurosci.* 17, 1396–1409. doi: 10.1162/0898929054985383
- van Atteveldt, N. M., Formisano, E., Goebel, R., and Blomert, L. (2007). Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. *Neuroimage* 36, 1345–1360. doi: 10.1016/j.neuroimage.2007.03.065
- Wager, T. D., Phan, K. L., Liberzon, I., and Taylor, S. F. (2003). Valence, gender and lateralization of functional brain anatomy in emotion: a meta-analysis of findings from neuroimaging. *Neuroimage* 19, 513–531. doi: 10.1016/S1053-8119(03)00078-8
- Wang, J., Nicol, T., Skoe, E., Sams, M., and Kraus, N. (2009). Emotion modulates early auditory response to speech. *J. Cogn. Neurosci.* 21, 2121–2128. doi: 10.1162/jocn.2008.21147
- Ward, L. M., McDonald, J. J., and Lin, D. (2000). On asymmetries in cross-modal spatial attention orienting. *Percept. Psychophys.* 62, 1258–1264. doi: 10.3758/bf03212127
- Wittchen, H. U., Wunderlich, U., Gruschwitz, S., and Zaudig, M. (1997). *SKID-I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft*. Göttingen: Hogrefe.
- Yamashita, H., Okamoto, Y., Morinobu, S., Yamawaki, S., and Kähkönen, S. (2005). Visual emotional stimuli modulation of auditory sensory gating studied by magnetic P50 suppression. *Eur. Arch. Psychiatry Clin. Neurosci.* 255, 99–103. doi: 10.1007/s00406-004-0538-6
- Zvyagintsev, M., Nikolaev, A. R., Sachs, O., and Mathiak, K. (2011). Early attention modulates perceptual interpretation of multisensory stimuli. *Neuroreport* 22, 586–591. doi: 10.1097/WNR.0b013e32834901f3
- Zvyagintsev, M., Nikolaev, A. R., Thönnessen, H., Sachs, O., Dammers, J., and Mathiak, K. (2009). Spatially congruent visual motion modulates activity of the primary auditory cortex. *Exp. Brain Res.* 198, 391–402. doi: 10.1007/s00221-009-1830-5

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 June 2013; accepted: 08 August 2014; published online: 27 August 2014.

Citation: Wolf D, Schock L, Bhavsar S, Demenescu LR, Sturm W and Mathiak K (2014) Emotional valence and spatial congruency differentially modulate crossmodal processing: an fMRI study. *Front. Hum. Neurosci.* 8:659. doi: 10.3389/fnhum.2014.00659

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2014 Wolf, Schock, Bhavsar, Demenescu, Sturm and Mathiak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.