

# Modulation of affective face processing deficits in Schizophrenia by congruent emotional sounds

Veronika I. Müller,<sup>1,2,3,4</sup> Tanja S. Kellermann,<sup>2,3,4</sup> Sarah C. Seligman,<sup>5</sup> Bruce I. Turetsky,<sup>6</sup> and Simon B. Eickhoff<sup>1,2,3,4</sup>

<sup>1</sup>Institute of Clinical Neuroscience and Medical Psychology, Heinrich Heine University, D-40225 Düsseldorf, Germany, <sup>2</sup>Department of Neuroscience und Medicine, INM-1, Research Center Jülich, D-52428 Jülich, Germany, <sup>3</sup>Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, D-52074 Aachen, Germany, <sup>4</sup>JARA-BRAIN Translational Brain Medicine, Jülich/Aachen, Germany, <sup>5</sup>Department of Psychology, Temple University, PA 19122 Philadelphia, USA, and <sup>6</sup>Neuropsychiatry Division, Department of Psychiatry, University of Pennsylvania School of Medicine, PA 19104 Philadelphia, USA

**Schizophrenia is a psychiatric disorder resulting in prominent impairments in social functioning. Thus, clinical research has focused on underlying deficits of emotion processing and their linkage to specific symptoms and neurobiological dysfunctions. Although there is substantial research investigating impairments in unimodal affect recognition, studies in schizophrenia exploring crossmodal emotion processing are rare. Therefore, event-related potentials were measured in 15 patients with schizophrenia and 15 healthy controls while rating the expression of happy, fearful and neutral faces and concurrently being distracted by emotional or neutral sounds. Compared with controls, patients with schizophrenia revealed significantly decreased P1 and increased P2 amplitudes in response to all faces, independent of emotion or concurrent sound. Analyzing these effects with regard to audiovisual (in)congruence revealed that P1 amplitudes in patients were only reduced in response to emotionally incongruent stimulus pairs, whereas similar amplitudes between groups could be observed for congruent conditions. Correlation analyses revealed a significant negative correlation between general symptom severity (Brief Psychiatric Rating Scale-V4) and P1 amplitudes in response to congruent audiovisual stimulus pairs. These results indicate that early visual processing deficits in schizophrenia are apparent during emotion processing but, depending on symptom severity, these deficits can be restored by presenting concurrent emotionally congruent sounds.**

**Keywords:** ERP; audiovisual; emotion; congruence; schizophrenia

## INTRODUCTION

Emotion recognition is a crucial component of social functioning and interpersonal relationships. Importantly, affect recognition in daily life is rarely based on emotional information in solely one sensory modality but is rather the result of the evaluation of stimuli coming from multiple sensory channels. Studies on crossmodal emotional integration have demonstrated greater accuracy and faster reaction times in congruent bimodal compared with unimodal emotional conditions. If, in contrast, the emotional content in the two modalities is incongruent, interference effects on reaction time and performance can occur (de Gelder and Vroomen, 2000; Dolan *et al.*, 2001; Kreifelts *et al.*, 2007; Collignon *et al.*, 2008). On the neural level, studies using event-related potentials (ERPs) found facilitated auditory N1 amplitudes (Pourtois *et al.*, 2000) when comparing emotional congruent to incongruent conditions, as well as longer auditory P2b latencies in affective incongruent pairings (Pourtois *et al.*, 2002). These, although few, results indicate that early sensory processing can be modulated depending on the congruence of concurrent stimuli from other sensory channels.

Schizophrenia is a psychiatric disorder characterized not only by symptoms of hallucinations, delusions and thought disorder but also by affective flattening and social impairments (American Psychiatric Association, 2000). Clinical research addressing the underlying deficits of the latter symptoms has provided accumulating evidence that schizophrenia patients exhibit impairments in facial affect recognition

(Kucharska-Pietura *et al.*, 2005; Turetsky *et al.*, 2007; Bach *et al.*, 2009a; Kohler *et al.*, 2010) as well as in categorization of prosody (Kucharska-Pietura *et al.*, 2005; Bozikas *et al.*, 2006; Bach *et al.*, 2009a; Leitman *et al.*, 2010). This impairment could be found for both emotional (positive and negative) and neutral stimuli. Therefore, for the visual domain a general deficit of early sensory processing was suggested, possibly related to dysfunctions in the magnocellular pathways (Butler *et al.*, 2001; Doniger *et al.*, 2002). This dysfunctional processing then in turn may further contribute to higher sensory processing deficits like, for example, problems in emotional perception. In line with these deficits, patients reveal aberrant ERPs in response to faces such as decreased N170 (Herrmann *et al.*, 2004; Johnston *et al.*, 2005; Campanella *et al.*, 2006; Bediou *et al.*, 2007; Caharel *et al.*, 2007; Turetsky *et al.*, 2007; Lynn and Salisbury, 2008) and P1 amplitudes (Campanella *et al.*, 2006; Caharel *et al.*, 2007). However, these studies have focused primarily on unimodal emotion processing, whereas, as mentioned earlier, emotional perception in daily life is rarely based on information from only one sensory channel. Thus, it is questionable whether one can generalize from unimodal deficits to impairments of emotion processing in real life. Some behavioral studies have already taken this into account. For example, de Gelder *et al.* (2005) used an audiovisual emotion recognition paradigm and found that patients are less influenced by emotional voices when rating a face compared with healthy controls but, conversely, the effect of faces on the rating of voices is increased. In contrast, de Jong *et al.* (2009) reported decreased influence of faces on voice perception, and Van den Stock *et al.* (2011) found an increased effect of voices on the rating of bodily expressions. Another phenomenon often assessed in the context of audiovisual integration is the McGurk fusion (McGurk and MacDonald, 1976), which denotes an illusion where the auditory perception of vowels is altered by concurrently seeing a face pronouncing a different vowel. In this paradigm, reduced audiovisual

Received 27 April 2012; Accepted 5 September 2012

Advance Access publication 13 September 2012

B.I.T. and S.B.E. contributed equally to this work.

This study was supported by Deutsche Forschungsgemeinschaft (DFG, IRTG 1328), by the Human Brain Project (R01-MH074457-01A1), by National Institute of Mental Health (grant no. 5P50MH064045) and the Helmholtz Initiative on systems biology (The Human Brain Model). Funding for the Open Access publication charges of this article was provided by the Central library of the Research Centre Jülich.

Correspondence should be addressed to Veronika Müller, Department of Neuroscience and Medicine, INM-1, Research Center Jülich, Leo-Brandt-Strasse, D-52428 Jülich, Germany. E-mail: v.mueller@fz-juelich.de

integration in patients was demonstrated by fewer McGurk fusion reports than controls (de Gelder *et al.*, 2003; Ross *et al.*, 2007). While across these studies the specific findings regarding the relative influence of one sensory modality on another are thus inconsistent, they all clearly suggest that patients with schizophrenia are not only impaired in unimodal emotion processing but also in the integration of emotional information from different modalities.

Here, we use ERPs to explore potentially aberrant neural correlates of audiovisual emotion integration in schizophrenia. In particular, we are interested in examining differences between patients and controls in the modulation of face-selective ERPs by concurrently presented congruent and incongruent sounds.

Given the unimodal emotion processing deficits in schizophrenia as well as aberrant audiovisual integration at the behavioral level, it is hypothesized that in response to audiovisual stimulus pairs, patients show reduced visual P1 and N1 amplitudes compared with healthy controls. Furthermore, based on behavioral results in audiovisual integration, we would expect that visual ERPs are differently affected by congruent and incongruent sounds in patients and controls. Decreased difference between the ERP response to incongruent and congruent conditions in patients compared with controls would indicate reduced audiovisual integration. In contrast, increased divergence between the two congruence conditions would point to increased integration in the patient group.

## METHODS

### Subjects

A total of 17 patients with schizophrenia and 16 healthy controls were recruited for the study through the Schizophrenia Research Center in the Neuropsychiatry Division of the Hospital of the University of Pennsylvania. Two patients and one control subject were subsequently excluded due to excessive electroencephalography (EEG) artifacts or an inability to perform the task. The final sample included 15 patients (four females; mean age  $35.1 \pm 9.26$  years, mean education  $14.1 \pm 2.23$  years) and 15 control participants (three females; mean age  $40.8 \pm 10.67$  years, mean education  $13.9 \pm 1.98$  years). All participants were right handed, reported normal or corrected-to-normal vision and were tested negative on a urine drug screen before the experiment. The groups were comparable in gender distribution ( $X^2_1 = 0.186$ , ns), mean age ( $t_{28} = -1.55$ , ns) and years of education ( $t_{28} = 0.26$ , ns).

All patients were stable outpatients at the time of testing. Consensus best-estimate Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) diagnoses were established using data gathered through the Structured Clinical Interview for DSM-IV (First, 1997) and any additional information available from medical record review, family and care providers. Ten patients met DSM-IV criteria for paranoid subtype (295.30), four were undifferentiated subtype (295.90) and one was diagnosed with the residual subtype (295.60). Only patients without comorbid psychiatric or neurological illness, substance abuse or addiction in the last 6 months were included in the study. Patients' scores on the Brief Psychiatric Rating Scale (BPRS-V4; Ventura *et al.*, 1993), the Scale for Assessment of Positive Symptoms (SAPS; Andreasen, 1984) and the Scale for Assessment of Negative Symptoms (SANS; Andreasen, 1983) were obtained by assessors trained to >85% inter-rater reliability. Nine patients were treated with atypical antipsychotics, three with typical antipsychotics and three were unmedicated. Five medicated patients were concurrently taking antidepressant agents and one was taking anticholinergic drugs. Four medicated patients were also being treated with benzodiazepines. Table 1 summarizes the demographic and clinical characteristics of the patient group.

**Table 1** Clinical profile of the patient group

Gender	11 males/4 females
Age	$35.10 \pm 9.26$
Age of onset	$20.8 \pm 4.46$
Duration of illness	$14.33 \pm 9.12$
Medicated/unmedicated	12/3
Brief Psychiatric Rating Scale (BPRS-V4)	$46.93 \pm 9.68$
Positive Symptoms Scale (SAPS)—total score	$29.80 \pm 21.6$
Negative Symptoms Scale (SANS)—total score	$37.80 \pm 13.3$
Number of patients with past alcohol abuse	2
Number of patients with past cannabis and phencyclidine abuse	1

The healthy control subjects were all without any history of neurological or psychiatric disorder, including substance abuse, or any history of an Axis I psychotic disorder in a first-degree relative. None was taking any psychoactive medication.

Following a full explanation of study procedures, written informed consent was obtained in compliance with guidelines established by the University of Pennsylvania Institutional Review Board and in accordance with The Code of Ethics of the World Medical Association (1964 Declaration of Helsinki).

### Stimuli

A detailed description of stimuli development is provided in Müller *et al.* (2011). In short, the visual stimuli obtained from the facial emotions for brain activation inventory (Gur *et al.*, 2002) consisted of 30 color pictures of five male and five female faces, each showing three different expressions (happy, neutral or fearful). Furthermore, 10 masks (neutral faces blurred with a mosaic filter) were used. For the auditory stimuli, 10 happy (laughs), 10 fearful (screams) and 10 neutral (yawning) sounds (five male and five female in each case) were employed.

### Procedure

In total, 180 audiovisual stimulus pairs were presented. Every face condition (happy, fear and neutral) was paired with every sound condition (happy, fear and neutral), resulting in a  $3 \times 3$  design with nine different conditions. Every face and every sound was presented twice per condition leading to 20 audiovisual pairs per condition. The combination was pseudo-random and matched with regard to gender. Furthermore, as every stimulus was presented twice per condition, pairing of the same sound and face stimulus was possible and allowed. The pairing was pseudo-randomized individually for every subject, so that every subject was presented with different pairs in varying order.

All stimulus pairs were presented for 1500 ms with a jittered inter-stimulus interval between 2500 and 4500 ms during which a blank black screen was shown. Every trial started with presentation of a sound concurrently with a blurred neutral face which was displaced by a clear face after 1000 ms and presented with the continuing sound for another 500 ms. Immediately after the audiovisual stimulus pairs, a rating scale was presented visually. Participants were instructed to ignore the sound and blurred face and to just rate the expression of the clear face. Response was not allowed during presentation of the audiovisual stimulus pair. Rather, the task was to wait until the face stimulus goes off and to respond as fast and accurately as possible as soon as the eight-point rating scale was displayed. The rating scale, ranging from extremely fearful to extremely happy, was visualized by eight buttons on the screen. Endpoints of the scale (button 1 and button 8) were labelled as extremely fearful (button 1) and extremely happy (button 8). Stimuli were presented with the software Presentation 14.8 (<http://www.neurobs.com/>) and responses were given manually using eight response buttons [extremely

fearful—left little finger (button 1) to extremely happy—right little finger (button 8)].

### Additional tests and questionnaires

After the EEG session, participants rated the valence and arousal of all sounds and pictures individually on a nine-point rating scale. The scales ranged from very fearful to very happy and not at all arousing to very arousing, respectively. The valence scale included the eight EEG test ratings plus an additional neutral category. Due to technical problems, the off-line rating of one control subject could not be used.

### EEG data acquisition and processing

EEG data were recorded using a 64-channel BioSemi ActiveTwo amplifier system with 24-bit A/D conversion (Amsterdam, The Netherlands) and customized electrodes with an integrated first stage amplifier. Electrodes were situated on the scalp using a 64-channel headcap. Horizontal and vertical eye movements were recorded from bipolar electrodes located lateral to the left and right epicanthi and above and below the right eye, respectively. Data were filtered online with a 0.16–100 Hz bandpass filter and sampled at 512 Hz. The EEG data were re-referenced off-line to the average of all channels, bandpass filtered between 0.5 and 30 Hz (12 dB/octave) and corrected for eye movements using the algorithm of Gratton *et al.* (1983). The data then underwent an automatic artifact rejection (minimal allowed amplitude:  $-100 \mu\text{V}$ , maximal allowed amplitude:  $+100 \mu\text{V}$ ; removal 250 ms before and after the event) and the continuous EEG was split into segments time-locked for three different sound conditions (fearful sounds, neutral sounds, and happy sounds) and 10 different face conditions [all fearful faces, all neutral faces, all happy faces, fearful face with congruent (fearful) sound, fearful face with incongruent (happy) sound, happy face with congruent (happy) sound, happy face with incongruent (fearful) sound, neutral face with scream, neutral face with yawn, and neutral face with laugh] with a duration of 600 ms (starting 100 ms before the onset of the sound or face). Sound conditions were time-locked to the onset of the sounds, whereas face conditions were time-locked to the onset of the faces. The segments were then baseline corrected ( $-100$  ms until stimulus onset) and average evoked potentials were computed for each condition. Number of averaged segments did not differ between patients and controls ( $t_{28} = -0.37$ , ns).

### Peak detection

For each condition, grand average waveforms were constructed and the peak latencies of the first positive (P1), first negative (N1) and second positive (P2) peaks were taken as references to establish an appropriate time interval for each peak (peak  $\pm 30$  ms) for the individual subject peak detection (Supplementary Table S1). For each component, peak latency was defined at those electrode(s) exhibiting the maximum component activity in the grand average waveforms. Thus, for the auditory ERPs (sound conditions), the peaks were defined at Cz. For the visual ERPs (face conditions), the positive peak latencies were defined at PO7 and PO8; N1 latency was defined at P9 and P10. Figure 1 illustrates the voltage distribution of auditory and visual ERPs across the scalp. For every subject, peak search was first done automatically in the established time interval and then, if applicable, manually adjusted. The mean area under the curve (peak  $\pm 14$  ms) was then computed for each channel, condition and subject and exported into Statistical Package for Social Science (SPSS).

### Statistical analysis

Behavioral and ERP data were analyzed off-line using IBM SPSS 19.0.0. All data were confirmed to be normally distributed and multivariate analyses of variance (MANOVAs)/analyses of variance (ANOVAs) were calculated. For *post hoc* analyses *t*-tests, Bonferroni corrected for multiple comparisons, were calculated. Pearson correlations of ERP measures with the total scores of SAPS, SANS and BPRS-V4 were computed (Bonferroni corrected for multiple comparisons). As benzodiazepines might affect evoked potentials (Rockstroh *et al.*, 1991), all tests were recalculated excluding patients taking benzodiazepines to confirm that the observed effects were not due to benzodiazepine intake in the patient group. Finally, antipsychotic medication was converted into chlorpromazine equivalent dosages (Andreasen *et al.*, 2010) and correlated with ERP measures as well as SAPS, SANS and BPRS-V4 scores (Bonferroni corrected for multiple comparisons).

## RESULTS

### Behavioral data

An ANOVA with the factors sound (scream/yawn/laugh), face (fearful/neutral/happy) and group (patient/control) and the dependent variable 'on-line rating during EEG measurement' revealed a significant main effect of face ( $F_{2,56} = 193.36$ ,  $P < 0.05$ ). However, there was no significant main effect of group ( $F_{1,28} = 0.007$ , ns) or interaction with group (sound  $\times$  group:  $F_{2,56} = 0.60$ , ns; face  $\times$  group:  $F_{2,56} = 0.44$ , ns; sound  $\times$  face  $\times$  group:  $F_{4,112} = 1.96$ , ns).

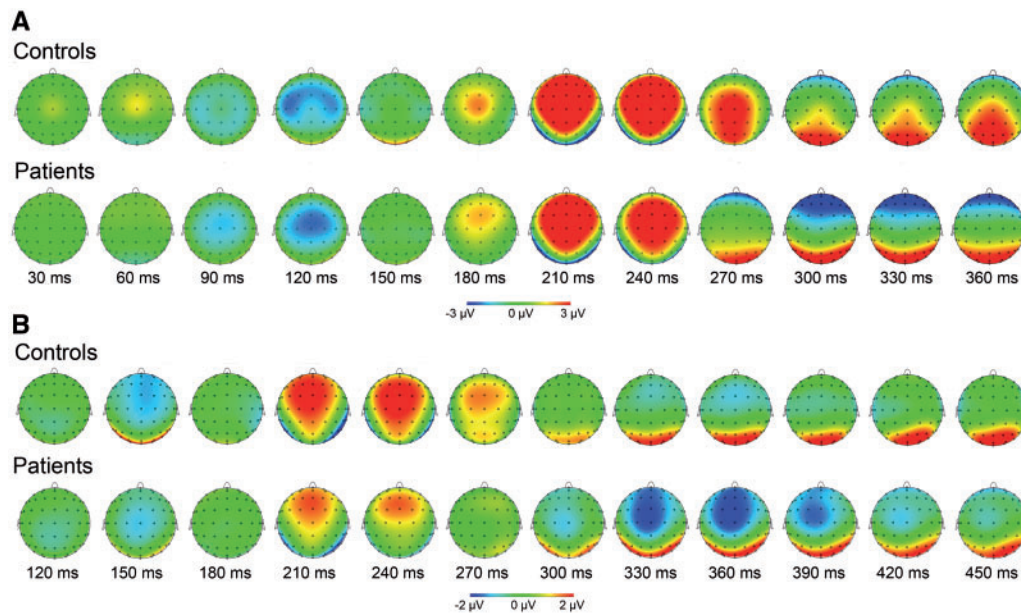
Two MANOVAs, one analyzing the off-line rating of valence and the other that of arousal for the (isolated) faces and sounds, resulted in significant effects of emotion (valence:  $F_{4,24} = 198.27$ ,  $P < 0.05$ ; arousal:  $F_{4,24} = 11.89$ ,  $P < 0.05$ ), which were significant for both faces (valence:  $F_{2,54} = 150.44$ ,  $P < 0.05$ ; arousal:  $F_{2,54} = 6.21$ ,  $P < 0.05$ ) and sounds (valence:  $F_{2,54} = 303.37$ ,  $P < 0.05$ ; arousal:  $F_{2,54} = 15.24$ ,  $P < 0.05$ ) in the univariate comparisons. *Post hoc* analyses revealed that all types of faces and sounds differed from each other in valence rating (faces: fearful *vs* neutral:  $t_{28} = -5.68$ ,  $P < 0.05$ ; happy *vs* neutral:  $t_{28} = -13.72$ ,  $P < 0.05$ ; fearful *vs* happy:  $t_{28} = -14.11$ ,  $P < 0.05$ ; all Bonferroni corrected; sounds: scream *vs* yawn:  $t_{28} = -16.00$ ,  $P < 0.05$ ; laugh *vs* yawn:  $t_{28} = -9.51$ ,  $P < 0.05$ ; scream *vs* laugh:  $t_{28} = -22.43$ ,  $P < 0.05$ ; all Bonferroni corrected) and that emotional faces and sounds were rated as more arousing than neutral ones (faces: fearful *vs* neutral:  $t_{28} = 3.40$ ,  $P < 0.05$ ; happy *vs* neutral:  $t_{28} = -3.13$ ,  $P < 0.05$ ; fearful *vs* happy:  $t_{28} = -0.11$ , ns; all Bonferroni corrected; sounds: scream *vs* yawn:  $t_{28} = 4.61$ ,  $P < 0.05$ ; laugh *vs* yawn:  $t_{28} = -5.69$ ,  $P < 0.05$ ; scream *vs* laugh:  $t_{28} = 2.13$ , ns; all Bonferroni corrected). However, no group effects could be found (valence: group:  $F_{2,26} = 1.92$ , ns; emotion  $\times$  group:  $F_{4,24} = 0.77$ , ns; arousal: group:  $F_{2,26} = 0.37$ , ns; emotion  $\times$  group:  $F_{4,24} = 2.13$ , ns). For all analyses of the behavioral data, the results were completely confirmed when excluding patients taking benzodiazepines (Supplementary Table S2).

### ERP data

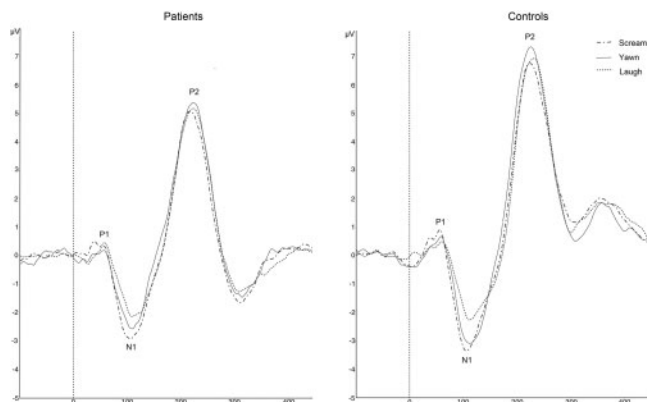
#### ERPs in response to sounds

Figure 2 displays the grand averages of the ERPs in response to screams, yawns and laughs at Cz in patients and healthy controls. A MANOVA with the independent variables emotional expression of the sound (scream/yawn/laugh) and group (patient/control) and the dependent variables amplitudes of P1, N1 and P2 at Cz was calculated to test for group differences. The data of one control subject were excluded due to excessive noise at Cz. The main effect of emotion was significant ( $F_{6,22} = 4.23$ ,  $P < 0.05$ ). However, there was no significant difference between ( $F_{3,25} = 1.52$ , ns) or interaction ( $F_{6,22} = 0.45$ , ns) with group.





**Fig. 1** ERP voltage maps. Voltage distribution of auditory (A) and visual (B) grand average ERPs, separately for patients and controls.



**Fig. 2** ERPs in response to screams, laughs and yawns in patients with schizophrenia and healthy controls.

Additionally, the individual ANOVAs for P1, N1 and P2 did only reveal a significant effect of emotion for N1 ( $F_{2,54} = 5.94, P < 0.05$ ) and P2 ( $F_{2,54} = 3.27, P < 0.05$ ), but not for P1 ( $F_{2,54} = 0.61, ns$ ). Furthermore, no significant main effect of group (P1:  $F_{1,27} = 2.02, ns$ ; N1:  $F_{1,27} = 0.54, ns$ ; P2:  $F_{1,27} = 3.41, ns$ ) or interaction with group (P1:  $F_{2,54} = 0.16, ns$ ; N1:  $F_{2,54} = 0.52, ns$ ; P2:  $F_{2,54} = 0.02, ns$ ) was found. *Post hoc* tests for the main effect of emotion revealed that the amplitude of N1 was higher when screams were presented compared with laughs ( $t_{28} = -2.91, P < 0.05$ , Bonferroni corrected). In contrast, the amplitude of N1 did not differ for scream vs yawn ( $t_{28} = -1.85, ns$ , Bonferroni corrected) and yawn vs laugh ( $t_{28} = -2.01, ns$ , Bonferroni corrected). Furthermore, the P2 amplitude did not significantly differ between any of the emotion conditions (scream vs laugh:  $t_{28} = -1.09, ns$ ; scream vs yawn:  $t_{28} = -2.41, ns$ ; laugh vs laugh:  $t_{28} = -1.67, ns$ ; all Bonferroni corrected).

**ERPs in response to faces**

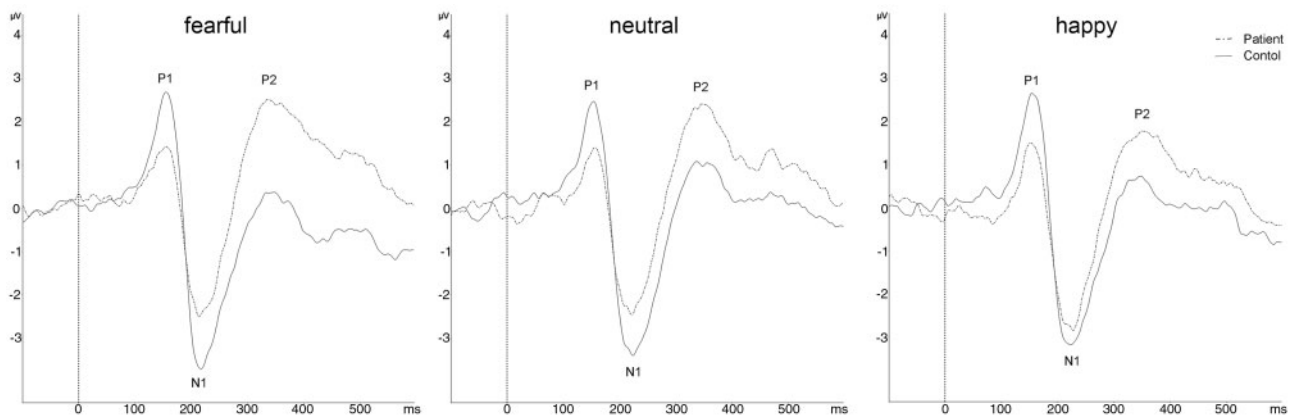
With regard to ERPs in response to faces (independent of the concurrently presented sounds), a MANOVA with the independent variables

facial emotional expression (fearful/neutral/happy), side (left/right), caudality (P9, P10/PO7, PO8) and group (patient/control) and the dependent variables amplitudes of P1, N1, and P2 was calculated. This analysis revealed a significant main effect of group ( $F_{3,26} = 3.34, P < 0.05$ ), emotional expression ( $F_{6,23} = 2.58, P < 0.05$ ) as well as significant interaction of group  $\times$  emotion ( $F_{6,23} = 2.58, P < 0.05$ ). While the effect of emotional expression as well as the interaction of group  $\times$  emotion was not significant for any of the dependent variables in the univariate comparisons (emotion: P1:  $F_{2,56} = 0.83, ns$ ; N1:  $F_{2,56} = 0.36, ns$ ; P2:  $F_{2,56} = 2.71, ns$ ; group  $\times$  emotion: P1:  $F_{2,56} = 0.41, ns$ ; N1:  $F_{2,56} = 2.52, ns$ ; P2:  $F_{2,56} = 2.78, ns$ ), the main effect of group was significant for P1 ( $F_{2,28} = 4.32, P < 0.05$ ) and P2 ( $F_{2,28} = 5.18, P < 0.05$ ) with patients showing smaller P1 and higher P2 responses compared with controls (Figure 3). These group differences in P1 ( $F_{1,24} = 9.19, P < 0.05$ ) and P2 ( $F_{1,24} = 6.37, P < 0.05$ ) persisted after excluding those patients taking benzodiazepines.

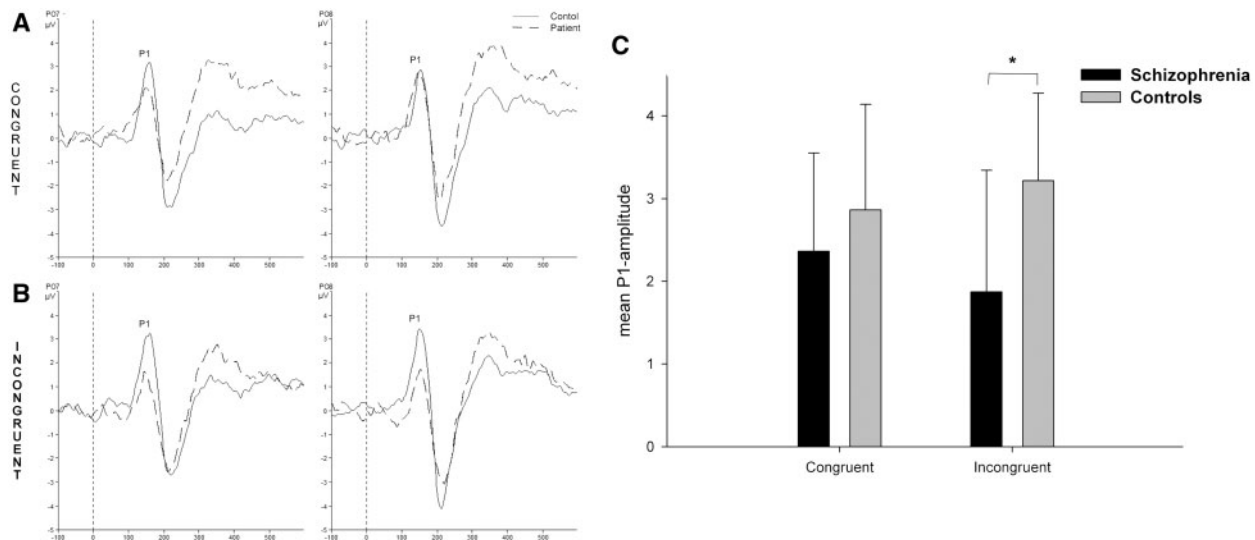
**ERPs in response to emotional congruent and incongruent stimulus pairs (incongruence of emotional valence)**

To investigate whether these general effects found over all face conditions differed for incongruent and congruent conditions, two further ANOVAs were conducted for the dependent variables P1 and P2 amplitudes, respectively, and the independent variables facial emotional expression (fearful/happy), side (left/right), caudality (P9, P10/PO7, PO8), auditory congruency (congruent/incongruent) and group (patient/control). Only conditions in which emotional faces (fearful, happy) and sounds (scream, laugh) were presented were included as conditions containing neutral stimuli lack incongruence of emotional valence.

Analysis of the P1 response revealed a significant main effect of group ( $F_{1,28} = 4.65, P < 0.05$ ) and a significant interaction between group  $\times$  incongruence ( $F_{1,28} = 6.50, P < 0.05$ ). *Post hoc* analysis of the interaction demonstrated that the main effect of group was mainly due to the incongruent conditions. While patients showed a reduced P1 amplitude compared with controls in response to incongruent audiovisual stimuli ( $t_{28} = -2.87, P < 0.05$ , Bonferroni



**Fig. 3** ERPs in response to fearful, neutral and happy faces in patients with schizophrenia (dashed line) and healthy controls (solid line). Amplitudes are pooled over electrodes P9, P10, P07 and P08. P1 and P2 amplitudes differ between patients and controls across facial conditions.



**Fig. 4** ERPs in response to faces in congruent (A) and incongruent (B) audiovisual conditions in patients with schizophrenia (dashed lines) and healthy controls (solid line) at P07 and P08. (C) Mean P1 amplitudes to faces in congruent and incongruent audiovisual conditions in patients with schizophrenia (black) and healthy controls (gray). Values are collapsed across electrodes (P9, P10, P07 and P08) and emotion (fearful and happy face), bars represent standard deviations. \* indicates significant differences between patients and controls at  $P < 0.05$ .

corrected; Figure 4), they did not differ in the amplitudes of congruent conditions ( $t_{28} = -1.11$ , ns; Figure 4). Neither controls ( $t_{14} = -1.561$ , ns) nor patients ( $t_{14} = 2.03$ , ns) showed a significant different P1 response to congruent compared with incongruent conditions. Furthermore, the main effect of emotion and the other interactions with group and incongruence were not significant for the analysis of the P1 amplitude in the ANOVA. Analyses excluding patients taking benzodiazepines revealed the same results (group:  $F_{1,24} = 12.813$ ,  $P < 0.05$ ; group  $\times$  incongruence:  $F_{1,24} = 8.048$ ,  $P < 0.05$ ).

Analysis of the P2 response revealed only a significant main effect of group ( $F_{1,28} = 6.24$ ,  $P < 0.05$ ) but no interaction between group  $\times$  incongruence ( $F_{1,28} = 0.57$ , ns) and no main effect of emotion ( $F_{1,28} = 1.23$ , ns). These effects persisted when excluding patients taking benzodiazepines (group:  $F_{1,24} = 7.38$ ,  $P < 0.05$ ; group  $\times$  incongruence:  $F_{1,24} = 0.98$ , ns; emotion:  $F_{1,24} = 2.67$ , ns).

**ERPs in response to congruent and incongruent stimulus pairs for neutral faces**

In addition to incongruence of emotion valence, incongruence effects were also analyzed for the neutral face condition (incongruence of

emotion presence). Due to lack of homogeneity at some electrodes for the neutral face conditions, mean P1 and P2 amplitudes were calculated across electrodes (P9, P10, P07 and P08). The ANOVAs with the factors sound (scream/yawn/laugh) and group (patient/control) and the dependent variables mean P1 and P2 amplitudes, respectively, did not reveal any significant effects, neither for P1 nor for P2 (Supplementary Table S3 and Figure S4).

**Correlations**

For the patient group, mean amplitudes for P1 and P2 in response to all faces were calculated across electrodes (P9, P10, P07 and P08) and emotion (fearful, neutral and happy face). Similarly, means were computed separately for congruent and incongruent stimulus pairs. These mean amplitudes were then correlated with SANS, SAPS and BPRS-V4 scores. Mean ERPs across emotion and electrodes were used to reduce the number of correlation analyses and therefore ameliorate the need to control for a high number of multiple comparisons (using Bonferroni correction). Moreover, we did not find any significant emotion  $\times$  group interactions in the performed MANOVA/ANOVAs that could have been taken as a hint toward emotion-specific impairments.

**Table 2** Correlations between SANS, SAPS, BPRS-V4 scores and mean P1 and P2 amplitudes in response to all, congruent (C) and incongruent (IC) faces

	SANS	SAPS	BPRS-V4
P1_all faces			
<i>r</i>	0.022	−0.345	−0.394
<i>P</i>	0.938	0.207	0.147
P2_all faces			
<i>r</i>	0.265	0.103	0.262
<i>P</i>	0.339	0.714	0.345
P1_C			
<i>r</i>	−0.206	−0.265	<b>−0.704</b>
<i>P</i>	0.461	0.340	<b>0.003</b>
P1_IC			
<i>r</i>	−0.051	−0.120	−0.333
<i>P</i>	0.857	0.670	0.225
P2_C			
<i>r</i>	0.099	0.151	0.330
<i>P</i>	0.726	0.390	0.229
P2_IC			
<i>r</i>	0.250	0.390	0.276
<i>P</i>	0.368	0.151	0.319

Bold numbers indicate a significant correlation (Bonferroni corrected for multiple comparisons).

Nevertheless, for the sake of completeness, correlations separately for every emotion are provided in Supplementary Tables S5 and S6. There was no significant correlation between P1 and P2 amplitudes over all faces and any of the symptomology scores (Table 2).

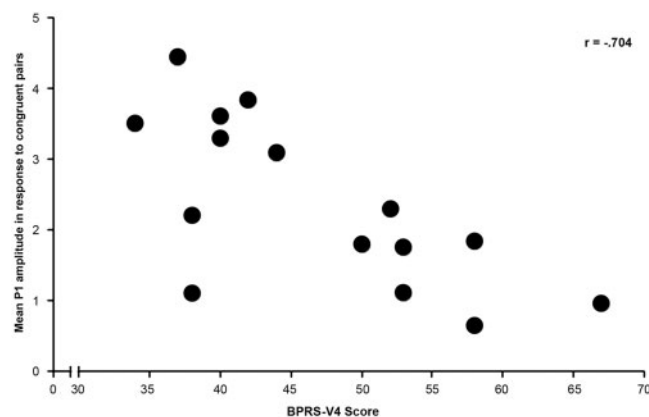
In contrast, correlations of mean P1 amplitudes of congruent and incongruent conditions (separately) with SANS, SAPS and BPRS-V4 scores revealed a significant negative correlation between P1 in response to congruent audiovisual stimulus pairs and BPRS-V4 scores ( $r = -0.704, P < 0.05$ , Bonferroni corrected; Table 2, Figure 5). This correlation persisted when correlating the scores with the ERPs of only those patients who were not taking benzodiazepines ( $r = -0.653, P < 0.05$ ). All other correlations were not significant (Table 2), both when including all patients and when including only patients who were not taking benzodiazepines. Furthermore, there was no significant correlation between chlorpromazine equivalent dosages and any ERP measure or SAPS, SANS and BPRS-V4 scores.

**DISCUSSION**

Here, we used an emotional audiovisual paradigm to investigate ERPs in response to emotionally congruent and incongruent audiovisual stimuli in patients with schizophrenia. Despite no significant differences between groups in behavioral performance or ERPs in response to sounds, this study demonstrated reduced visual P1 and increased P2 responses in patients compared with controls. Analyzing these effects in detail showed that the effect in P1 was mainly due to differences in response to faces presented with a concurrent incongruent sound. In contrast, P1 response to faces presented in a congruent auditory context was similar in patients and controls. Moreover, a negative correlation between mean P1 amplitude in response to congruent audiovisual stimulus pairs and general psychiatric symptoms was found in the patient group, suggesting a P1 response similar to that of healthy controls when fewer symptoms are present. These results indicate that congruent sounds can modify early visual responses in patients, leading to similar neural response as in healthy individuals but importantly, this modulation is highly dependent on symptom severity.

**Behavioral data**

Patients did not show any differences, neither in on-line ratings of faces while instructed to ignore the sounds nor in off-line ratings in which valence and arousal were rated independently for faces and



**Fig. 5** Negative correlation between BPRS-V4 scores and mean P1 amplitude across electrodes (P9, P10, P07 and P08) and emotion (fearful and happy face) in response to congruent audiovisual stimulus pairs.

sounds. Studies to date concerning multimodal emotional integration in schizophrenia are rare and results rather inconsistent. Although all studies report differences between schizophrenia and controls, some find reduced (de Gelder *et al.*, 2005; de Jong *et al.*, 2009) and others excessive (de Gelder *et al.*, 2005; Van den Stock *et al.*, 2011) influence of one modality on another. In contrast, the results of our study indicate that both groups are similarly affected by voices when rating the valence of a face. When considering this apparent discrepancy, it should be noted that the current paradigm differs from previous experiments in several crucial aspects. While previous studies used emotional decision paradigms and calculated accuracy scores for each condition, we used a (subjective) eight-point rating scale, which mainly captured valence intensity. Furthermore, we used human vocalizations as auditory stimuli, whereas former studies investigated audiovisual emotional integration with emotional prosody. These differences in the experimental design may have produced these contradictory results.

The individual rating of faces and sounds following the EEG paradigm also did not reveal any group differences, indicating that perception and subjective experience of sounds and faces were not impaired in our schizophrenia group—a finding that is in line with some previous reports (Tüscher *et al.*, 2005; Lynn and Salisbury, 2008; Wynn *et al.*, 2008). Nevertheless, it contradicts results of other studies of emotional face and prosody perception (Bozikas *et al.*, 2006; Turetsky *et al.*, 2007; Bach *et al.*, 2009a; Kohler *et al.*, 2010; Leitman *et al.*, 2010). Again, our study was different in that we did not calculate objective accuracy scores. Tüscher *et al.* (2005) also did not find any differences in the valence and arousal ratings of emotional environmental sounds between patients and controls indicating that, in contrast to prosody, perception of environmental and human sounds seems to be preserved in schizophrenia.

**Electrophysiological data**

**Auditory processing**

Although previous literature reports deficits in auditory processing in schizophrenia (Boutros *et al.*, 2004; Bramon *et al.*, 2004; Turetsky *et al.*, 2009), to our knowledge no study has investigated auditory ERPs in response to human (emotional) vocalizations in this disorder. However, two functional magnetic resonance imaging studies that focus on dysfunctions of auditory emotion processing in schizophrenia report altered lateralization patterns (Mitchell *et al.*, 2004; Bach *et al.*, 2009b) during emotion prosody processing. By comparing ERPs between schizophrenia patients and controls in response to screams, laughs and yawns, this study demonstrates that patients have similar



auditory ERP responses to healthy controls, suggesting relatively intact auditory processing of human vocalizations, as opposed to simple tones.

### Visual processing

The waveform morphology time-locked to the face presentation in our study appeared similar to the typical waveform elicited by unimodal face presentation, but with specific ERP components appearing at somewhat prolonged latencies. This might be due to the experimental design, in which a blurred face was presented before the target face and the visual stimulus was paired with a concurrent auditory stimulus. In line with this view, Isoglu-Alkac *et al.* (2007) report longer latencies in response to bimodal compared with unimodal conditions.

Deficits in early visual processing of faces in schizophrenia have been reported in a number of studies, with the majority reporting N170 reductions (Herrmann *et al.*, 2004; Johnston *et al.*, 2005; Campanella *et al.*, 2006; Bediou *et al.*, 2007; Caharel *et al.*, 2007; Turetsky *et al.*, 2007; Lynn and Salisbury, 2008), but some also found decreased P1 (Campanella *et al.*, 2006; Caharel *et al.*, 2007) and increased P2 (Herrmann *et al.*, 2004; Ramos-Loyo *et al.*, 2009) amplitudes. In contrast, other studies did not find any differences in the P1 (Herrmann *et al.*, 2004; Johnston *et al.*, 2005; Bediou *et al.*, 2007; Turetsky *et al.*, 2007; Wynn *et al.*, 2008) or N170 response (Streit *et al.*, 2001; Wynn *et al.*, 2008; Ramos-Loyo *et al.*, 2009) between patients and controls.

Our results, demonstrating decreased P1 amplitudes during face processing in bimodal conditions, are in line with Caharel *et al.* (2007) and Campanella *et al.* (2006), but are also consistent with studies using unimodal non-face stimuli (Doniger *et al.*, 2002; Yeap *et al.*, 2008), which indicate a deficit occurring at early stages of visual processing in patients. The P1 response represents general primary visual processing (Campanella *et al.*, 2006) and Butler *et al.* (2001) as well as Doniger *et al.* (2002) suggest that this impairment in schizophrenia reflects deficits in magnocellular pathways. These deficits, in turn, may contribute to higher visual processing deficits, resulting in abnormal facial visual scanning strategies (Schwartz *et al.*, 1999; Loughland *et al.*, 2002), which then impede facial emotion recognition.

In contrast to reduced P1 amplitude, we did not find any N170 differences between patients and controls. As the majority of previous research supports a N170 decrease (Herrmann *et al.*, 2004; Johnston *et al.*, 2005; Campanella *et al.*, 2006; Bediou *et al.*, 2007; Caharel *et al.*, 2007; Turetsky *et al.*, 2007; Lynn and Salisbury, 2008) in schizophrenia, our finding might indicate that this deficit is more readily apparent when faces are processed without any context, namely when a face alone is processed without any accompanying stimulation from other modalities. Our results therefore indicate that the N170 response following face presentation is preserved in schizophrenia when coupled with auditory information.

In accordance with Herrmann *et al.* (2004) and Ramos-Loyo *et al.* (2009), we additionally observed increased P2 amplitudes in the patient group. We suggest that higher P2 activity in schizophrenia may illustrate compensation for early visual deficits, which then leads to intact perception of the emotional expression of faces demonstrated by the lack of behavioral differences between groups. In accordance with this view, Ramos-Loyo *et al.* (2009) also did not find any behavioral differences between groups.

### Incongruence effects

De Gelder *et al.* (1999) and Pourtois *et al.* (2000) report that audiovisual integration takes place as early as 110 ms after stimulus presentation and occurs automatically. Therefore, stimuli from one modality may influence processing in another before having been fully structurally encoded. The interaction of audiovisual congruence and group in

P1 amplitude in this study supports this view by showing that in patients the earliest positive component in response to faces can be modulated by concurrent presentation of a congruent sound. Importantly, this effect can only be found for (in)congruence of emotion valence and not when pairing neutral faces with emotional or neutral sounds (incongruence of emotion presence).

It has to be noted that in both groups we did not find a difference between ERPs to congruent and incongruent conditions. Therefore, our results neither support our primary hypothesis of reduced nor that of increased audiovisual integration in the patient group. Rather, only a between-group difference in the incongruent condition was found. This result indicates that additional presentation of sounds that convey the same emotional expression as the target face leads to similar P1 responses in patients and controls. Therefore, deficits in early visual processing in schizophrenia seem to be restored by congruent auditory emotional information.

This finding extends previous research investigating exploratory eye movements in schizophrenia when confronted with audiovisual information, demonstrating that patients showed an increase in total number of gaze points when looking at a smiling face of a baby accompanied by laughter compared with processing the face in isolation (Ishii *et al.*, 2010). Therefore, presenting a congruent sound may increase visual attention and hence cortical processing of visual stimuli in patients. In particular, due to top-down information extracted from the sound, attention may be directed toward specific emotional characteristics of faces, which then leads to more appropriate scanning strategies in congruent audiovisual conditions. In contrast, incongruent sounds misdirect attention so that deficits in visual scanning (Loughland *et al.*, 2002) may persist or even be reinforced.

Furthermore, a negative correlation of P1 amplitudes in response to congruent stimulus pairs with symptomology was found. More precisely, the fewer symptoms patients show, the more similar they are to healthy controls in their P1 response to congruent audiovisual information. This correlation therefore indicates that the beneficial effects of congruent auditory context processing are diminished in patients with more severe symptomatology. Therefore, contextual modulation of P1 deficits might only be possible in patients with relatively mild symptoms, whereas more severely affected patients might not benefit from additional auditory stimulation.

These findings generate a new perspective on views of emotional impairments in schizophrenia, indicating that when confronted with congruent multimodal emotional information, early visual processing as well as subjective perception of emotions seem to be intact. Studies reporting deficits in emotion processing in schizophrenia have mainly used unimodal emotional paradigms, but emotion recognition in daily life is hardly based on processing of a face or a voice alone. Therefore, symptoms such as affect flattening and social impairments might be less associated with deficits found in unimodal face processing than previously thought. In particular, social deficits might not necessarily be the result of impaired sensory processing but rather reflect problems in higher executive functioning. Therefore, based on these results, we suggest that future studies should apply more naturalistic experimental designs to investigate emotional and social deficits in schizophrenia.

It is interesting to consider the possibility that this crossmodal modulation, which had a positive effect on early visual processing in our paradigm, might reflect the same mechanism that underlies hallucinations in schizophrenia. It has been hypothesized that hallucinations may reflect an imbalance between imagery and perception, arising from an over-interpretation of top-down information in patients (Behrendt, 1998; Grossberg, 2000; Aleman *et al.*, 2003). Increased influence of top-down auditory expectation might therefore have led to modulation of the P1 amplitude in response to congruent audiovisual pairs in our schizophrenia group.

## CONCLUSION

In summary, this study demonstrates deficits in early visual face processing in schizophrenia, which may be counteracted or attenuated by concurrently presented emotionally congruent sounds. These results expand previous research in unimodal emotion processing by showing that to a certain degree of symptom severity, patients' deficits in emotion processing are less apparent in (congruent) audiovisual emotion processing.

## SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

## Conflict of Interest

None declared.

## REFERENCES

- Aleman, A., Bocker, K.B., Hijman, R., de Haan, E.H., Kahn, R.S. (2003). Cognitive basis of hallucinations in schizophrenia: role of top-down information processing. *Schizophrenia Research*, 64, 175–85.
- American Psychiatric Association. (2000). *Diagnostic Criteria from DSM-IV-TR*. Washington, D.C.: The American Psychiatric Association.
- Andreasen, N.C. (1983). *The Scale for the Assessment of Negative Symptoms (SANS)*. Iowa City, IA: The University of Iowa.
- Andreasen, N.C. (1984). *The Scale for the Assessment of Positive Symptoms (SAPS)*. Iowa City, IA: The University of Iowa.
- Andreasen, N.C., Pressler, M., Nopoulos, P., Miller, D., Ho, B.C. (2010). Antipsychotic dose equivalents and dose-years: a standardized method for comparing exposure to different drugs. *Biological Psychiatry*, 67, 255–62.
- Bach, D.R., Buxtorf, K., Grandjean, D., Strik, W.K. (2009a). The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. *Psychological Medicine*, 39, 927–38.
- Bach, D.R., Herdener, M., Grandjean, D., Sander, D., Seifritz, E., Strik, W.K. (2009b). Altered lateralisation of emotional prosody processing in schizophrenia. *Schizophrenia Research*, 110, 180–7.
- Bediou, B., Henaff, M.A., Bertrand, O., et al. (2007). Impaired fronto-temporal processing of emotion in schizophrenia. *Neurophysiologie Clinique*, 37, 77–87.
- Behrendt, R.P. (1998). Underconstrained perception: a theoretical approach to the nature and function of verbal hallucinations. *Comprehensive Psychiatry*, 39, 236–48.
- Boutros, N.N., Korzyuko, O., Oliwa, G., et al. (2004). Morphological and latency abnormalities of the mid-latency auditory evoked responses in schizophrenia: a preliminary report. *Schizophrenia Research*, 70, 303–13.
- Bozikas, V.P., Kosmidis, M.H., Anezoulaki, D., Giannakou, M., Andreou, C., Karavatos, A. (2006). Impaired perception of affective prosody in schizophrenia. *Journal of Neuropsychiatry and Clinical Neuroscience*, 18, 81–5.
- Bramon, E., Rabe-Hesketh, S., Sham, P., Murray, R.M., Frangou, S. (2004). Meta-analysis of the P300 and P50 waveforms in schizophrenia. *Schizophrenia Research*, 70, 315–29.
- Butler, P.D., Schechter, L., Zemon, V., et al. (2001). Dysfunction of early-stage visual processing in schizophrenia. *American Journal of Psychiatry*, 158, 1126–33.
- Caharel, S., Bernard, C., Thibaut, F., et al. (2007). The effects of familiarity and emotional expression on face processing examined by ERPs in patients with schizophrenia. *Schizophrenia Research*, 95, 186–96.
- Campanella, S., Montedoro, C., Streeb, E., Verbanck, P., Rosier, V. (2006). Early visual components (P100, N170) are disrupted in chronic schizophrenic patients: an event-related potentials study. *Neurophysiologie Clinique*, 36, 71–8.
- Collignon, O., Girard, S., Gosselin, F., et al. (2008). Audio-visual integration of emotion expression. *Brain Research*, 1242, 126–35.
- de Gelder, B., Bocker, K.B., Tuomainen, J., Hensen, M., Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neuroscience Letters*, 260, 133–6.
- de Gelder, B., Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, 14, 289–311.
- de Gelder, B., Vroomen, J., Annen, L., Masthof, E., Hodiomont, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, 59, 211–8.
- de Gelder, B., Vroomen, J., de Jong, S.J., Masthoff, E.D., Trompenaars, F.J., Hodiomont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia Research*, 72, 195–203.
- de Jong, J.J., Hodiomont, P.P., Van den Stock, J., de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophrenia Research*, 107, 286–93.
- 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, 10006–10.
- Doniger, G.M., Foxe, J.J., Murray, M.M., Higgins, B.A., Javitt, D.C. (2002). Impaired visual object recognition and dorsal/ventral stream interaction in schizophrenia. *Archives of General Psychiatry*, 59, 1011–20.
- First, M.B. (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders SCID-I: Clinician Version, Administrative Booklet*. Washington, DC: American Psychiatric Press.
- Gratton, G., Coles, M.G., Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, 55, 468–84.
- Grossberg, S. (2000). How hallucinations may arise from brain mechanisms of learning, attention, and volition. *Journal of the International Neuropsychological Society*, 6, 583–92.
- Gur, R.C., Sara, R., Hagendoorn, M., et al. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115, 137–43.
- Herrmann, M.J., Ellgring, H., Fallgatter, A.J. (2004). Early-stage face processing dysfunction in patients with schizophrenia. *American Journal of Psychiatry*, 161, 915–7.
- Ishii, Y., Morita, K., Shouji, Y., Nakashima, Y., Uchimura, N. (2010). Effects of emotionally charged sounds in schizophrenia patients using exploratory eye movements: comparison with healthy subjects. *Psychiatry and Clinical Neuroscience*, 64, 10–8.
- Isoglu-Alkac, U., Kedzior, K., Keskindemirci, G., Ermutlu, N., Karamursel, S. (2007). Event-related potentials to visual, auditory, and bimodal (combined auditory-visual) stimuli. *International Journal of Neuroscience*, 117, 259–73.
- Johnston, P.J., Stojanov, W., Devir, H., Schall, U. (2005). Functional MRI of facial emotion recognition deficits in schizophrenia and their electrophysiological correlates. *European Journal of Neuroscience*, 22, 1221–32.
- Kohler, C.G., Walker, J.B., Martin, E.A., Healey, K.M., Moberg, P.J. (2010). Facial emotion perception in schizophrenia: a meta-analytic review. *Schizophrenia Bulletin*, 36, 1009–19.
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: an event-related fMRI study. *Neuroimage*, 37, 1445–56.
- Kucharska-Pietura, K., David, A.S., Masiak, M., Phillips, M.L. (2005). Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. *British Journal of Psychiatry*, 187, 523–8.
- Leitman, D.I., Laukka, P., Juslin, P.N., Saccente, E., Butler, P., Javitt, D.C. (2010). Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. *Schizophrenia Bulletin*, 36, 545–56.
- Loughland, C.M., Williams, L.M., Gordon, E. (2002). Visual scanpaths to positive and negative facial emotions in an outpatient schizophrenia sample. *Schizophrenia Research*, 55, 159–70.
- Lynn, S.K., Salisbury, D.F. (2008). Attenuated modulation of the N170 ERP by facial expressions in schizophrenia. *Clinical EEG and Neuroscience*, 39, 108–11.
- McGurk, H., MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–8.
- Mitchell, R.L., Elliott, R., Barry, M., Crutten, A., Woodruff, P.W. (2004). Neural response to emotional prosody in schizophrenia and in bipolar affective disorder. *British Journal of Psychiatry*, 184, 223–30.
- Müller, V.I., Habel, U., Derntl, B., et al. (2011). Incongruence effects in crossmodal emotional integration. *Neuroimage*, 54, 2257–66.
- Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11, 1329–33.
- Pourtois, G., Debatisse, D., Despland, P.A., de Gelder, B. (2002). Facial expressions modulate the time course of long latency auditory brain potentials. *Brain Research Cognitive Brain Research*, 14, 99–105.
- Ramos-Loyo, J., Gonzalez-Garrido, A.A., Sanchez-Loyo, L.M., Medina, V., Basar-Eroglu, C. (2009). Event-related potentials and event-related oscillations during identity and facial emotional processing in schizophrenia. *International Journal of Psychophysiology*, 71, 84–90.
- Rockstroh, B., Elbert, T., Lutzenberger, W., Altenmüller, E. (1991). Effects of the anticonvulsant benzodiazepine clonazepam on event-related brain potentials in humans. *Electroencephalography and Clinical Neurophysiology*, 78, 142–9.
- Ross, L.A., Saint-Amour, D., Leavitt, V.M., Molholm, S., Javitt, D.C., Foxe, J.J. (2007). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophrenia Research*, 97, 173–83.
- Schwartz, B.L., Rosse, R.B., Johri, S., Deutsch, S.I. (1999). Visual scanning of facial expressions in schizophrenia. *Journal of Neuropsychiatry and Clinical Neuroscience*, 11, 103–6.
- Streit, M., Wolwer, W., Brinkmeyer, J., Ihl, R., Gaebel, W. (2001). EEG-correlates of facial affect recognition and categorisation of blurred faces in schizophrenic patients and healthy volunteers. *Schizophrenia Research*, 49, 145–55.
- Turetsky, B.I., Bilker, W.B., Siegel, S.J., Kohler, C.G., Gur, R.E. (2009). Profile of auditory information-processing deficits in schizophrenia. *Psychiatry Research*, 165, 27–37.
- Turetsky, B.I., Kohler, C.G., Indersmitten, T., Bhati, M.T., Charbonnier, D., Gur, R.C. (2007). Facial emotion recognition in schizophrenia: when and why does it go awry? *Schizophrenia Research*, 94, 253–63.
- Tüscher, O., Silbersweig, D., Pan, H., et al. (2005). Processing of environmental sounds in schizophrenic patients: disordered recognition and lack of semantic specificity. *Schizophrenia Research*, 73, 291–5.



- Van den Stock, J., de Jong, S.J., Hodiament, P.P., de Gelder, B. (2011). Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Social Neuroscience*, 6, 537–47.
- Ventura, J., Lukoff, D., Nuechterlein, K.H., Liberman, R.P., Green, M., Shaner, A. (1993). Appendix 1: Brief Psychiatric Rating Scale (BPRS) Expanded Version (4.0) scales, anchor points and administration manual. *International Journal of Methods in Psychiatric Research*, 3, 227–43.
- Wynn, J.K., Lee, J., Horan, W.P., Green, M.F. (2008). Using event related potentials to explore stages of facial affect recognition deficits in schizophrenia. *Schizophrenia Bulletin*, 34, 679–87.
- Yeap, S., Kelly, S.P., Sehatpour, P., et al. (2008). Visual sensory processing deficits in Schizophrenia and their relationship to disease state. *European Archives of Psychiatry and Clinical Neuroscience*, 258, 305–16.