Incongruence effects in crossmodal emotional integration

Veronika I. Müller[a,b,c,*], **Ute Habel**[a,c], **Birgit Derntl**[a,f], **Frank Schneider**[a,c], **Karl Zilles**[b,c,d], **Bruce I. Turetsky**[e], **Simon B. Eickhoff**[a,b,c]

[a]Department of Psychiatry und Psychotherapy, RWTH Aachen University, Germany

[b]Department of Neuroscience und Medicine, INM-2, Research Centre Jülich, Germany

[c]JARA-BRAIN Translational Brain Medicine, Jülich/Aachen, Germany

[d]C. and O. Vogt Brain Research Institute, University of Düsseldorf, Germany

[e]Neuropsychiatry Division, Department of Psychiatry, University of Pennsylvania School of Medicine, Philadelphia, PA, USA

[f]Institute for Clinical, Biological and Differential Psychology, Faculty of Psychology, University of Vienna, Austria

Abstract

Emotions are often encountered in a multimodal fashion. Consequently, contextual framing by other modalities can alter the way that an emotional facial expression is perceived and lead to emotional conflict. Whole brain fMRI data was collected when 35 healthy subjects judged emotional expressions in faces while concurrently being exposed to emotional (scream, laughter) or neutral (yawning) sounds. The behavioral results showed that subjects rated fearful and neutral faces as being more fearful when accompanied by screams than compared to yawns (and laughs for fearful faces). Moreover, the imaging data revealed that incongruence of emotional valence between faces and sounds led to increased activation in the middle cingulate cortex, right superior frontal cortex, right supplementary motor area as well as the right temporoparietal junction. Against expectations no incongruence effects could be found in the amygdala. Further analyses revealed that, independent of emotional valence congruency, the left amygdala was consistently activated when the information from both modalities was emotional. If a neutral stimulus was present in one modality and emotional in the other, activation in the left amygdala was significantly attenuated. These results indicate that incongruence of emotional valence in audiovisual integration activates a cingulate-fronto-parietal network involved in conflict monitoring and resolution. Furthermore in audiovisual pairing amygdala responses seem to signal also the absence of any neutral feature rather than only the presence of an emotionally charged one.

Keywords

fMRI; Emotional conflict; Incongruence; Amygdala; Audiovisual

Introduction

Affect processing and the ability to recognize emotions in facial expressions is a crucial part of social cognition and inter-personal relationships. In everyday life, however, the evaluation of facial emotions is rarely based on the expression of a face alone. Rather, most of the time a face is perceived in a particular context, for instance, we see a face in a specific scene or accompanied by a particular sound. Thus, facial affect recognition usually involves the interpretation of a face in a specific context, even if the latter is not processed consciously. Importantly, this context may not always be congruent with the emotional expression displayed by the face, which implies that concurrently receiving two or more different emotional inputs can lead to emotional conflict.

Non-emotional conflict has been studied extensively using behavioural and neuroimaging experiments (Botvinick et al., 2004; Carter et al., 1998; Carter and van Veen, 2007; Durston et al., 2003; Egner and Hirsch, 2005; Kerns et al., 2004; MacDonald et al., 2000; Ochsner et al., 2009; Weissman et al., 2004). The results of these studies have led to the conflict monitoring hypothesis, which suggests that conflict is detected by the dorsal anterior cingulate cortex (ACC), which in turn recruits prefrontal regions to increase cognitive control (Botvinick et al., 2004; Carter and van Veen, 2007; Kerns et al., 2004).

In contrast, rather few studies have focused on the effects of affective conflict, reporting increased reaction times (Collignon et al., 2008; de Gelder and Vroomen, 2000; Dolan et al., 2001; Haas et al., 2006; Ochsner et al., 2009; Wittfoth et al., 2010) in incongruent compared to congruent conditions for both unimodal and crossmodal affective conflict.

On the neurobiological level affective conflict has been linked to activation in ACC for (within-modality) conflicts in the visual and auditory domain (Haas et al., 2006; Ochsner et al., 2009; Wittfoth et al., 2010), as well as right dorsolateral prefrontal cortex (DLPFC) and bilateral posterior medial frontal cortex for visual (Ochsner et al.,2009) and superior temporal gyrus (STG) for auditory conflict processing (Wittfoth et al., 2010).

As outlined above emotional conflicts may arise not only from incompatible information encountered in the same modality but in particular also from receiving non-matching cues from different sensory channels, such as when we see a frightened face and concurrently hear laughter. In spite of their behavioral relevance, however, the neurobiology of multi-modal emotional conflicts has rarely been studied. Studies focusing on the neural correlates of audiovisual integration independent of a potential conflict, consistently report increased activation of posterior superior temporal sulcus (pSTS) (Beauchamp, 2005; Kreifelts et al., 2007; Pourtois et al., 2005; Robins et al., 2009), for both emotional and non-emotional integration. Furthermore Dolan et al. (2001) contrasted emotional congruent to emotional incongruent conditions in an audiovisual paradigm and found greater activation of the left amygdala and right fusiform gyrus (FFG) in congruent as compared to incongruent conditions, but did not report a significant effect of the reverse contrast. Nevertheless studies which focus on the neural correlates of audiovisual emotional incongruence processing are still sparse. Furthermore, most of the existing studies investigating emotional incongruence

effects either did not use neutral stimuli or did not differentiate between different ways how incongruence could arise in such settings. In particular, it may be argued that two different types of incongruence can occur in audiovisual pairing when, in addition to emotional, neutral stimuli are also presented. These two types may be called incongruence of emotional valence and incongruence of emotion-presence. On one hand, incongruence of emotional valence arises when in both modalities an emotional stimulus is presented but these two stimuli differ in emotional valence (i.e. happy face paired with scream). That is, the information from the two sensory channels disagrees in the nature or valence of the portrayed emotion. On the other hand incongruence of emotion-presence is characterized by pairing a neutral stimulus with an emotional one (i. e. neutral face paired with a scream). In other words, an emotional stimulus may be present in one modality but absent in the other (because a neutral stimulus is presented). Given that emotion processing is supposed to be centered about filtering behaviorally relevant information, in this case a conflict arises between the two sensory channels, with one indicating a (potentially) salient event, while the other is not.

Finally, clinical studies in patients with depression show that this disorder goes along with a negativity bias (Drevets et al., 2008) and problems inhibiting negative information (Goeleven et al., 2006; Joormann, 2010). These problems have repeatedly been linked to dysregulation in ACC, dorsolateral prefrontal cortex and the amygdala (Abler et al., 2007; Drevets, 2000; Fales et al., 2008). Furthermore symptom severity of depression correlates with amygdala activation in emotion regulation (Abler et al., 2007) as well as inhibitory-related activity in ACC during a cognitive inhibition task (Matthews et al., 2009). Lastly, Abler et al. (2010) demonstrated that activation in the amygdala does even correlate with subclinical depression scores. Whether a similar correlation of depressivity scores with regional brain activation can also be found in crossmodal emotional integration is yet not clarified. Moreover, revealing and understanding the relationship between subclinical traits and neural function in the healthy brain may help to provide information about neurobiological vulnerability markers for depression.

Given the open questions outlined above, the primary goal of the current study is to investigate the influence of emotional sounds on facial affect perception and the underlying neural substrates involved in the mediation of both types of audiovisual (in-) congruence effects. We hypothesize that behaviorally, emotional sounds influence the perception of facial expressions while incongruence effects should elicit increased activation in dACC and prefrontal regions. In contrast, based on previous work, we would assume that emotional congruence enhances activation in amygdala and FFG.

As a secondary aim the correlation of (subclinical) depressive traits as quantified by the Beck Depression Inventory (BDI) with the neural correlates of audiovisual inhibition processes is investigated. Given previous reports in depressed patients (Abler et al., 2007; Drevets, 2000; Fales et al., 2008), we would hypothesize that activity in the amygdala, ACC and DLPFC over all task conditions relative to baseline correlates with BDI scores.

Methods

Subjects

In total 40 healthy subjects were examined. Five subjects were excluded from the analysis due to BDI or alexithymia scores exceeding the cut-off for clinically relevant depression (> 15) and alexithymia (>60), respectively, or excessive movement during scanning. Hence, 35 healthy right-handed volunteers (20 females; mean age 29.4 ± 8.1 years, mean education 14.1 ± 3.2, 15 males; mean age 34.0 ± 13.3, mean education 14.5 ± 3.2) were included in the analysis. Men and women did not significantly differ in age (t_{33} = 1.29, ns.) or education (t_{33} = 0.34, ns.). All participants had normal or corrected-to-normal vision and were right handed as confirmed by the Edinburgh Inventory (Oldfield, 1971). The structured clinical interview of DSM IV (SCID; Wittchen et al., 1997) was used to screen the subjects revealing no history of neurological or psychiatric disorders including substance abuse. No subject took mood- or cognition-altering medication. Subjects gave informed consent into this study which was approved by the ethics committee of the School of Medicine of the RWTH Aachen University.

Stimuli

The visual stimuli, obtained from the FEBA inventory (Gur et al., 2002a), consisted of color pictures of male and female faces, showing either a happy, neutral or fearful expression. In total 10 (5 males and 5 females) faces of different actors each showing every expression (happy, neutral, fear) were used. To optimize the design, a behavioral pre-study with 48 subjects was conducted where different versions of the paradigm and varying presentation times were tested. These results indicated that emotional faces were too clear to allow any consistent context framing effects. As a consequence, happy and fearful faces were degraded in emotionality in order to make them more ambiguous. This was done by merging every emotional facial image with the neutral mouths of the same actor.

Furthermore, as a longer auditory and a shorter visual stimulus had to be combined, each trial also employed (visual) masks. This was necessary because a long auditory presentation is needed to contain clear emotional information, whereas a short visual presentation is important to allow context effects. As visual input should be held constant for the whole trial 10 different neutral faces blurred with a mosaic filter were included as masks.

The auditory stimuli were either neutral (yawning) or socio-emotional salient sounds like laughing (happy) or screaming (fear). 210 different sounds were obtained from various sound libraries and equalized with respect to length and volume. In a pre-experiment, these were then rated by a different sample of 10 subjects for validation and selection of the stimuli used in the experiment. In total 10 laughs, 10 yawns and 10 screams (5 males and 5 females in each case) were selected.

Procedure

In the experiment a total of 180 stimulus pairs, each consisting of a visual and an auditory stimulus were presented. Every face condition (happy, fear, neutral) was combined with every sound condition (happy, fear, neutral) resulting in a 3×3 design with nine different

conditions and 20 audiovisual pairs per condition. The pairing was pseudo-random and matched with regard to gender, so that a male face was always paired with a male voice and vice versa.

Audiovisual stimulus pairs were presented for 1500 ms with a jittered inter-stimulus-interval between 4000 and 6000 ms during which a blank black screen was shown (see Fig. 1). Every trial started with the presentation of a sound to establish a context. Concurrently with the sound a blurred neutral face was shown. After 1000 ms the blurred face was displaced by the target stimulus (neutral or emotional face) and presented with the continuing sound for another 500 ms. The subjects were asked to ignore the sound and to rate as fast and accurate as possible the presented facial expression on an eight-point rating scale from extremely fearful to extremely happy. Subjects were instructed that the study focuses on attention processes in order to avoid effects of expectation due to knowledge about the aims of the study. Stimuli were presented with the software Presentation 14.2 (http://www.neurobs.com/) and response was given manually with the right or left hand on an MR-compatible response pad (from extremely fearful—left little finger (1) to extremely happy—right little finger (8)).

Additional tests and questionnaires

After the scanning session participants rated the valence and arousal of sounds and pictures individually on a nine-point valence scale ranging from very fearful to very happy and not at all arousing to very arousing, respectively. Contrary to the rating in the scanner where subjects had to make a forced choice on the emotional content due to the absence of a "neutral" response, the rating of faces and sounds after scanning included a neutral category. Moreover the TAS-D (Bagby et al., 1994) was carried out to control for alexithymia and the Beck Depression Inventory (BDI-II; Hautzinger et al., 2006) to check for subclinical depressive symptoms.

fMRI data acquisition

Images were acquired on a Siemens Trio 3 T whole-body scanner (Erlangen, Germany) in the Research Center Jülich using blood-oxygen-level-dependent (BOLD) contrast (Gradient-echo EPI pulse sequence, TR = 2.2 s, in plane resolution = 3.1×3.1 mm, 36 axial slices (3.1 mm thickness)) covering the entire brain. Image acquisition was preceded by 4 dummy images allowing for magnetic field saturation. These were discharged prior to further processing. Images were analysed using SPM5 (www.fil.ion.ucl.ac.uk/spm). First, the EPI images were corrected for head movement by affine registration using a two-pass procedure, by which images were initially realigned to the first image and subsequently to the mean of the realigned images. After realignment, the mean EPI image for each subject was spatially normalized to the MNI single subject template using the "unified segmentation" approach (Ashburner and Friston, 2005). The resulting parameters of a discrete cosine transform, which define the deformation field necessary to move the subjects data into the space of the MNI tissue probability maps, were then combined with the deformation field transforming between the latter and the MNI single subject template. The ensuing deformation was subsequently applied to the individual EPI volumes that were hereby transformed into the MNI single subject space and resampled at $2 \times 2 \times 2$ mm^3 voxel size. The normalized images

were spatially smoothed using an 8 mm FWHM Gaussian kernel to meet the statistical requirements of the General Linear Model and to compensate for residual macro-anatomical variations across subjects.

Statistical analysis

Behavioral data were analyzed off-line using SPSS 18.0.0 (SPSS Inc., Chicago, IL). All data were confirmed to be normally distributed and MANOVAs / ANOVAs were calculated. Violations of sphericity were corrected by the Greenhouse–Geisser or Huynh–Feldt correction. Post-hoc analyses, corrected for multiple comparisons, were calculated for significant main effects and interactions.

The imaging data were analyzed using a General Linear Model as implemented in SPM5. Each experimental condition (fearful face/scream, fearful face/yawn, fearful face/laugh, neutral face/scream, neutral face/yawn, neutral face/laugh, happy face/scream, happy face/ yawn, happy face/laugh) as well as the response event were separately modelled by a boxcar reference vector convolved with a canonical hemodynamic response function and its first-order temporal derivative. Additionally, low-frequency signal drifts were filtered using a cutoff period of 128 s. Parameter estimates were subsequently calculated for each voxel using weighted least squares to provide maximum likelihood estimators based on the temporal autocorrelation of the data (Kiebel et al., 2003). No global scaling was applied. For each subject, simple main effects for each of the nine experimental conditions and the response were computed by applying appropriate baseline contrasts. These individual first-level contrasts were then fed to a second-level group-analysis using an ANOVA (factor: condition, blocking factor subject) employing a random-effects model. In the modelling of variance components, violations of sphericity have been allowed by modelling non-independence across images from the same subject and allowing unequal variances between conditions and subjects using the standard implementation in SPM 5.

In the ANOVA mean (weighted by non-sphericity correction) parameter estimates over all subjects were calculated for every condition (fearful face/scream, fearful face/yawn, fearful face/laugh, neutral face/scream, neutral face/yawn, neutral face/laugh, happy face/scream, happy face/yawn, happy face/laugh). Based on these estimates of the second level analysis, separate t-contrasts were calculated for every differential effect (effect of fearful faces, effect of happy faces, effect of fearful sounds, effect of happy sounds, incongruence effects of emotional valence, incongruence effects of emotion-presence) by applying the respective contrast to the 2nd level parameter estimates. In other words, the ANOVA was used to calculate the group level parameter estimates for each condition whereas *t*-contrasts to the parameter estimates were employed to test for differential effects between these. The resulting SPM (T) maps were then thresholded at *p* <0.05, correcting for multiple comparisons by controlling the family-wise error (FWE) rate according to the theory of Gaussian random fields (Worsley et al., 1996). FWE corrects for accumulation of type I error in multiple testing with non-compact support, i.e., correlated statistics. By regarding the statistical image as a lattice representation of a random Gaussian field (whose smoothness is estimated using the residual field of the statistical analysis), the probability that one or more false positive voxels are found in the entire search volume can be estimated

for each threshold applied to the statistic. The threshold is then chosen as to control the family-wise error rate at $p < 0.05$, i.e., in only one of twenty analyses of random fields of the same type, size and smoothness one would expect to find a single voxel above the chosen threshold.

Additionally, another model was estimated where the nine conditions and responses and BDI-scores as a covariate were modeled.

Ensuing activations were anatomically localised using version 1.6b of the SPM Anatomy toolbox (Eickhoff et al., 2005, 2006, 2007; www.fz-juelich.de/ime/spm_anatomy_toolbox).

Results

Valence and arousal of faces and sounds in the off-line rating

An one-factorial MANOVA with the factor emotion (fear/neutral/happy) and the dependent variables "off-line valence rating of faces" and "off-line valence rating of sounds" was conducted revealing a significant main effect of emotion ($F_{2,4} = 318.11$, $p < 0.05$). This effect was significant for both dependent variables, faces (Greenhouse–Geisser corrected: $F_{2,66} = 271.2$; $\varepsilon = 0.74$, $p < 0.05$) and sounds (Greenhouse–Geisser corrected: $F_{2,66} = 983.42$, $\varepsilon = 0.76$, $p < 0.05$). Fearful stimuli got the lowest ratings, followed by neutral faces and sounds, while the most positive ratings were given for happy stimuli (all post-hoc comparisons were significant at $p < 0.05$, cf. Table 1).

An additional MANOVA with ratings of arousal also led to a significant main effect of emotion ($F_{2,4} = 53.14$, $p < 0.05$), significant for both stimulus modalities (faces: $F_{2,68} = 13.72$; $p < 0.05$; sounds: $F_{2,68} = 95.95$; $p < 0.05$). Neutral faces and sounds were rated as least arousing, followed by happy stimuli, whereas fearful stimuli were perceived as most arousing (all post-hoc comparisons were significant at $p < 0.05$ except for comparison of happy to neutral faces which revealed a tendency towards significance, cf. Table 1).

On-line behavioral data

An ANOVA with the factors face (fear/neutral/happy) and sound (scream/yawn/laugh) and the dependent variable "rating obtained in the scanner" revealed significant main effects of face (Huynh–Feldt corrected: $F_{2,68} = 253.96$, $\varepsilon = 0.66$, $p < 0.01$) and sound ($F_{2,68} = 7.96$, $p < 0.01$), as well as a significant interaction between face x sound ($F_{4,136} = 3.86$, $p < 0.01$).

Post-hoc analyses of the main effect face showed the same pattern as observed in the off-line rating, with lowest ratings for fearful, highest for happy and intermediate values for neutral faces (all post-hoc comparisons were significant at $p < 0.05$).

As the task was to ignore the sound and to rate the face, the main effect of sound reflects the influence of a sound on a face. Post-hoc analyses of this effect did show that screams led, independently of facial expression, to a more fearful rating of the faces compared to yawns and laughs (post-hoc comparisons were significant at $p < 0.05$), whereas subjects did not rate faces differently when exposed to laughs compared to yawns.

Post-hoc analyses of the interaction revealed that subjects rated fearful faces as being more fearful when accompanied by screams compared to yawns or laughs. Neutral faces were also perceived as more fearful when paired with screams compared to yawns but not to laughs (Fig. 2). No context effect could be found for happy faces (Tukey's HSD=0.096).

Imaging data

To identify emotion specific areas all happy and fearful faces were separately contrasted to neutral faces. For happy relative to neutral faces, this contrast revealed increased activation in bilateral calcarine sulcus, left middle occipital gyrus, left inferior parietal cortex, left supplementary motor area, primary motor and postcentral gyrus as well as left insula and right thalamus (p<0.05, FWE corrected, Table 2). In contrast, there was no significant greater activation for fearful compared to neutral faces at the corrected level. Comparison of screams and laughs, respectively, to yawns showed activation in bilateral STG for either emotional sound. Furthermore screams compared to yawns activated bilateral inferior parietal cortex (IPC), right supplementary motor area (SMA) and left middle cingulate cortex (p<0.05, FWE corrected, Table 3).

Incongruence effects of emotional valence

To identify brain regions involved in detecting and processing emotional conflict of valence, emotionally incongruent conditions (fearful face paired with laugh and happy face paired with scream) were contrasted to emotionally congruent conditions (fearful face paired with scream and happy face paired with laugh). Significant incongruence effects could be found in the middle cingulate cortex, the right superior frontal gyrus, the right SMA and the right temporoparietal junction (TPJ) (p<0.05, FWE corrected, Table 4, Fig. 3).

In contrast, congruent compared to incongruent conditions did not evoke significantly increased activation in any brain region. As neither incongruence nor congruence effects were found in the amygdala, a small volume correction (based on the coordinates by Dolan et al., 2001; x=−20, y=−8, z=−14 Talairach coordinates converted into MNI space x=−20, y=−8, z=−17; search volume 6 mm) was carried out for both congruence and incongruence effects. Again, no significant effect was detected, neither FWE corrected, nor uncorrected.

Incongruence effects of emotion-presence

To elucidate the neural correlates of incongruence concerning the presence and absence of emotion, we compared pure emotional conditions (fearful face with laugh or scream, happy face with laugh or scream) with conditions where a neutral stimulus was presented in one of the two modalities and an emotional one in the other (fearful or happy face with yawn, neutral face with scream or laugh).

This contrast revealed significant activation (cluster peak at x=−20, y=−3, z=−21, p<0.05, FWE corrected, Fig. 4) of the left laterobasal/superficial nuclei of the amygdala (LB/SF; Amunts et al., 2005). The overlap of the activated cluster with LB was 15.2%, while 78.6 % of its volume was allocated to SF; the maximum probabilities were 60 % and 70 %, respectively.

Furthermore this contrast led to significant greater bilateral occipital and temporal activation (Table 5).

Correlation with BDi-scores

As it is known that depression goes along with impairments in emotional processing, we were interested if even subclinical depressive symptoms correlate with neural activation in an emotion-related task.

An analysis with BDI-Scores as covariate revealed a negative correlation of BDI with activation (across all conditions) bilaterally in middle occipital gyrus (x=−38; y=−86; z=9 and x=33; y=−89; z=8) and in right pSTS (x=45; y=−44; z=8) and FFG (x=29; y=−60; z=−9, all cluster-level corrected $p<0.05$) (Fig. 5). There was no significant interaction between BDI and condition.

Discussion

By using fMRI we investigated the influence of emotional sounds on the perception of facial expressions and the neural correlates of incongruence processing. The results revealed an influence of screams on the perception of fearful and neutral faces and incongruence effects of emotional valence were found in middle cingulate cortex, right superior frontal gyrus, right supplementary motor area (SMA) and right temporoparietal junction (TPJ). Furthermore activation in the left amygdala was attenuated when an emotional stimulus was paired with a neutral one compared to when in both modalities the stimulus was emotional.

Off-line rating of faces and sounds and neural correlates of face and sound processing

The off-line rating of faces and sounds showed that both, faces and sounds were recognized correctly by the subjects with lowest ratings for fearful faces and sounds, followed by neutral stimuli and the highest (happiest) ratings for happy faces and sounds. Additionally the rating of faces reflect that the perceived emotion was more intense in happy as compared to fearful faces, evident by a larger difference to the neutral category.

This pattern is reflected in the imaging data. While happy compared to neutral faces activate a widespread cortical network, no significantly (at the corrected level) increased activity was observed for fearful as compared to neutral faces. Increased activation in occipital areas (Britton et al., 2006), insula (Britton et al., 2006; Chen et al., 2009) and thalamus (Dolan et al., 1996) for happy faces are well in accordance with previous results.

In contrast, on the FWE correction level, fearful faces did not recruit more activation in any areas compared to neutral faces. This reflects the relative moderate expression of fear in the fearful face condition and goes along with the less intense rating of these faces after scanning.

Comparing laughs and screams to yawns activated bilateral superior temporal gyrus (STG), which is in line with previous findings on emotional vs. neutral vocalizations (Fecteau et al., 2007; Meyer et al., 2005; Phillips et al., 1998). As Fecteau et al. (2007) pointed out this activation may either reflect differences in acoustic characteristics or indicate that the STG

plays a role in extracting emotional information from socially relevant sounds. Moreover, screams compared to yawns led to greater activation in bilateral inferior parietal cortex (1PC), right SMA and left middle cingulate cortex. This pattern may reflect that the higher arousal evoked by the screams (as reflected by the off-line ratings) has led to activation of parietal areas to augment attention (Corbetta and Shulman, 2002), cingulate cortex for monitoring (Botvinick et al., 2004) and SMA for increased motor control (for review see Nachev et al., 2008).

Context effects on ratings of facial affect

Only two studies (de Gelder and Vroomen, 2000; Ethofer et al., 2006) previously investigated how emotional sounds influence the perception of (emotional) facial expressions. Our behavioral data obtained in the scanner demonstrated that context, established by concurrent social–emotional sounds, influences the perception of facial expressions. Fearful and neutral faces were perceived more fearful when accompanied by screams compared to yawns (or laughs in case of fearful faces). In contrast, these context effects were not found for happy faces. This is in line with the results of Ethofer et al. (2006) who also report context effects for fearful and neutral but not for happy faces. One explanation for the lack of effect on happy faces may be differences in ambiguity, as indicated by the off-line ratings. This difference occurred even though happy and fearful faces were matched in intensity based on the classification of Gur et al. (2002a), were degraded by the neutral mouth and showed equally clear separation from neutral faces in a prestudy conducted in a different sample of 18 subjects. These results thus support earlier notions that happiness is the easiest and fear one of the hardest recognizable emotions in faces (Gur et al., 2002a; Montagne et al., 2007).

Compared to yawns only screams, not laughs led to a significantly more emotional rating, potentially relating to the higher arousal evoked by these sounds which makes them much harder to ignore. This higher saliency fits well with the biological perspective that screams, as threat signals, are far more important for survival than laughs (Calder et al., 2001; Green and Phillips, 2004) and is further-more reflected by the more widespread activation evoked by these (as described above).

Neuronal effects of incongruence of emotional valence

Emotional incongruence or conflict (pairing of fearful faces with laughs and happy faces with screams) evoked greater activation in middle cingulate cortex, right superior frontal gyrus, right SMA and right TPJ compared to emotional congruent trials. Crossmodal emotional conflict thus recruits a network similar to those previously described to be involved in unimodal emotional and cognitive conflict and attentional control (Botvinick et al., 2004; Corbetta et al., 2008; Ochsner et al., 2009; Wittfoth et al., 2010).

The (anterior) cingulate cortex (ACC) has repeatedly been linked to conflict monitoring (Botvinick et al., 2004; Carter et al., 1998; Carter and van Veen, 2007) and has been found by Ochsner et al. (2009) and Egner et al. (2008) for cognitive and affective conflict in the visual domain and by Wittfoth et al. (2010) for auditory emotional conflict. It should be noted, that the location of these findings, although reported as "ACC" corresponds well to

the increased activation observed in the middle cingulate cortex in our study. This discrepancy should be attributable to the still widespread use of the outdated two-region classification of the cingulate cortex introduced by Brodmann in spite of convincing evidence for at least four major regions of this structure (Palomero-Gallagher et al., 2009; Vogt et al., 2004). As this region has already been demonstrated to respond to unimodal (Egner et al., 2008; Ochsner et al., 2009; Wittfoth et al., 2010) and non-emotional audiovisual (Weissman et al., 2004) conflict, current findings in emotional audiovisual conflict provide further evidence for a generalized role of the dorsal middle cingulate cortex in conflict monitoring and attentional control, irrespectively of modality, integration demands or affective content. Similarly, several studies have already linked the superior frontal / dorsolateral prefrontal cortices to the resolution of (non-emotional) conflict (Corbetta and Shulman, 2002; Durston et al., 2003; Egner and Hirsch, 2005; MacDonald et al., 2000) mainly by top-down control of attentional resources. Again, our results extend these previous findings to emotional conflict, pointing to an overlapping neuronal framework.

The SMA plays a major role in voluntary action, cognitive control and initiation/inhibition of motor responses (Grefkes et al., 2008; Kasess et al., 2008; Nachev et al., 2008; Sumner et al., 2007). SMA activation was also found for emotional conflict in the study of Ochsner et al. (2009) in an affective flanker task. Like these authors we would conclude that higher SMA activity may reflect increased executive control needed to select an adequate response in the presence of conflicting (emotional) stimuli.

The TPJ, finally, has been implicated in predictive coding (Jakobs et al., 2009), attentional reorienting, theory of mind and perspective taking tasks (Corbetta et al., 2008; Derntl et al., 2010; Mitchell, 2008). In our current data, we would side with the idea of predictive coding postulating that the brain predicts upcoming events based on previous information in order to reduce computational load and to deal with ambiguous information (Creutzig and Sprekeler, 2008). The auditory stimulus, starting 1000 ms before the face, could have induced expectation of a visual stimulus congruent with the emotion expressed in the sound. The presentation of an emotionally incongruent face should then have led to violation of these expectations. Consequently, if the TPJ integrates sensory signals and against the expectations on these, the prediction error between bottom-up facial information and expectation built on the sound may underlie the greater TPJ activation in the incongruent condition.

In summary, audiovisual emotional incongruence leads to activation of a cingulate-fronto-parietal network previously shown to be involved in conflict monitoring and resolution (Botvinick et al., 2004; Carter et al., 1998; Ochsner et al., 2009; Weissman et al., 2004; Wittfoth et al., 2010). Conflict caused by the convergence of two incongruent information streams is detected by the dACC/MCC (Botvinick et al., 2004; Carter et al., 1998; Carter and van Veen, 2007; Kerns et al., 2004), which recruits the dorsal prefrontal cortex and the SMA to increase attentional and motor control to enable selection of an adequate response. The current results expand this conflict monitoring hypothesis in two ways. Firstly, we would argue for a role of the TPJ in processing temporally asynchronous conflicting information by integrating predictions and incoming stimuli (Jakobs et al., 2009). Second,

our results indicate that the same conflict monitoring and resolution network demonstrated for cognitive and action related tasks, is likewise involved in emotional conflict processing.

This suggestion is in line with the biased competition model of attention introduced by Desimone and Duncan (1995), postulating that competition for neural representation occurs when two or more stimuli with different task relevance are presented. This competition can in turn be modulated (biased) in two different ways, namely either based on bottom-up (salience) or by top-down (cognitive control) influences. The increased activity of cingulate, frontal and parietal areas in incongruent compared to congruent emotional conditions observed in our study may reflect such top-down mechanisms. In particular, given that subjects were instructed to attend to and rate the faces rather than the sounds, these regions may be the origin of top-down signals biasing competition towards the task relevant stimuli (faces) in incongruent conditions where competition and attentional demands are higher. In line with this suggestion, other studies also demonstrate greater activity in ACC (Bishop et al., 2004a) and parietal and frontal (Mitchell et al., 2007) areas going along with higher attentional demands when a cognitive task is distracted by emotional stimuli. Furthermore it is reported that attention towards a non-emotional task suppresses activity in the amygdala (Amting et al., 2009; Bishop et al., 2004b; Mitchell et al., 2008; Pessoa et al., 2005). The absence of a (valence) congruency effect in the amygdala could therefore just be the result of biased competition, as the salience of emotional sounds is likewise suppressed in favor of the (attended) face.

Role of the amygdala

In contrast to the finding of Dolan et al. (2001) we did not observe any (in-) congruence effects of emotional valence in the amygdala when explicitly testing for these. However, the testing of incongruence effects with respect to the presence and absence of emotional stimuli (incongruence of emotion-presence) revealed a significant effect in the left basolateral/ superficial amygdala complex.

It should be noted that our result should be interpreted with caution given the current resolution of fMRI and the susceptibility of the medial temporal lobe to imaging artifacts. Nevertheless, the applied histological atlas allows anatomical allocation in a probabilistic fashion, hereby accommodating some of the uncertainty in establishment of structure-function relationships (cf. Amunts et al., 2005; Eickhoff et al., 2005). Moreover, this approach has already been repeatedly shown to allow the attribution of functional activation to specific parts of the amygdala (Ball et al., 2007; Goossens et al., 2009; Roy et al., 2009).

It should also be taken into account that we used a somewhat larger smoothing kernel as previous studies which targeted particularly at amygdala responses (Ball et al., 2007; Goossens et al., 2009; Roy et al., 2009). In this context, however, it is important to note, that while smoothing blurs the data, it should per se not affect localisation accuracy (Fox et al., 2001). Rather spatial uncertainty in neuroimaging results mainly derives from inter-individual variability and imprecision in spatial normalisation (Eickhoff et al., 2009; Grachev et al., 1999; Hellier et al., 2003). To minimise the latter potential confound, we used the individual subjects' mean EPI image directly as the source image for spatial normalisation rather than co-registering T1. This approach has the advantage of avoiding

influences of non-linear distortions between EPI and T1 images and yields a precise registration of the functional data into MNI space (see supplementary Fig. 1).

Though, our result indicates that in audiovisual pairing the left amygdala is active independent of crossmodal valence congruency when emotional information is not accompanied by concurrent neutral information. Conversely, it may be hypothesized, that a neutral stimulus in one sensory channel may counteract the emotional saliency provided by the stimulus in the other modality by attenuating amygdala responses. In other words, in audiovisual integration it is not only the presence of an emotional stimulus but rather the absence of a concurrent, mitigating neutral stimulus that may be crucial for left amygdala responses. Could this effect be related to priming, so that the presentation of a neutral stimulus influences amygdala response to the subsequent facial stimulus? Importantly this is not the case, as the pattern of activation is the same independently of whether the neutral stimulus (yawn) was presented first or as a neutral face 1000 ms after the onset of an (emotional) sound.

The current findings thus place a new twist on the current views about the function and role of the left amygdala. It has previously been demonstrated that this structure is activated by attended and unattended emotional stimuli in the visual or auditory modality (Belin et al., 2004; Fecteau et al., 2007; Gur et al., 2002b; Habel et al., 2007; Phillips et al., 1998; Sander and Scheich, 2001). We could show that the left laterobasal/superficial amygdala nuclei also integrate neutral stimuli across modalities, which probably leads to an "all-clear" signal. From studies in macaque monkeys it is known that the laterobasal amygdala receives direct input from auditory and visual cortices (McDonald, 1998; Yukie, 2002). Moreover feedback projections from the basal nucleus of the amygdala to sensory-associative areas have been demonstrated (Amaral et al., 2003; Yukie, 2002). This connectivity pattern in combination with the observed response pattern may be interpreted as follows. The left amygdala receives socially salient information from all sensory domains which are then integrated and evaluated. Information processing in visual and auditory areas is then modulated via the feedback projections from the laterobasal amygdala to early sensory cortices providing a route for enhancing behaviourally relevant and suppressing less important information. Ultimately, integrating a neutral stimulus into the evaluation by the amygdala should then lead to top-down inhibition in associative visual and auditory areas whereas emotional stimuli in both modalities induce top-down facilitation by the amygdala as evident from significantly higher activation in the purely emotional conditions in lingual gyrus, calcarine sulcus, STG and middle temporal gyrus (MTG).

Correlation with BDI

BDI scores were negatively correlated with activation in left and right occipital gyrus as well as right FFG and STS, i.e., regions contributing to the neural system for facial perception (Haxby et al., 2000). Moreover, the posterior superior temporal sulcus plays also a crucial role in audiovisual integration (Beauchamp, 2005; Campanella and Belin, 2007). As we only investigated healthy subjects with BDI scores in the normal range, these scores may be regarded as subclinical depressive symptoms or traits which nevertheless were shown to correlate with neuronal activation in an emotion-related task. The observed correlation in

sensory cortices may point to a dysregulation of neuronal activity by depressive traits spreading beyond areas involved in emotion regulation (for a review on findings in depression see Leppanen, 2006) to early stages of face processing. Alternatively, the negative correlation of BDI scores and activity in STS could indicate decreased crossmodal integration associated with higher depressivity, which has yet to be tested in patients with major depression. In particular, it remains to be seen, how evaluation in the amygdala and integration in the STS may be differentially affected in these patients.

Conclusion

Our audiovisual integration paradigm demonstrated that crossmodal emotional conflict recruits a cingulate-fronto-parietal network previously shown for cognitive, motor and (unimodal) emotional conflict monitoring and resolution. We propose that emotional conflict is detected by the TPJ integrating bottom-up input and expectation while the middle cingulate cortex recruits the dorsal prefrontal cortex and SMA to increase cognitive and motor control for selecting the adequate response.

Additionally, our results reveal that irrespective of congruency the left laterobasal/superficial amygdala is attenuated as soon as one sensory channel carries neutral but ecologically valid information. This indicates that crossmodal evaluation in the amygdala integrates emotional as well as neutral information with the latter providing a "clear-off" for implicit arousal regulation.

At last, BDI scores correlated negatively with activation in occipital gyrus, FFG and pSTS. These results emphasize the importance of taking even subclinical traits into account when investigating emotional processing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abler B, Erk S, Herwig U, Walter H, 2007. Anticipation of aversive stimuli activates extended amygdala in unipolar depression. J. Psychiatr. Res 41, 511–522. [PubMed: 17010993]

Abler B, Hofer C, Walter H, Erk S, Hoffmann H, Traue HC, Kessler H, 2010. Habitual emotion regulation strategies and depressive symptoms in healthy subjects predict fMRI brain activation patterns related to major depression. Psychiatry Res. 183, 105–113. [PubMed: 20630713]

Amaral DG, Behniea H, Kelly JL, 2003. Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. Neuroscience 118, 1099–1120. [PubMed: 12732254]

Amting JM, Miller JE, Chow M, Mitchell DG, 2009. Getting mixed messages: the impact of conflicting social signals on the brain's target emotional response. Neuroimage 47, 1950–1959. [PubMed: 19446638]

Amunts K, Kedo O, Kindler M, Pieperhoff P, Mohlberg H, Shah NJ, Habel U, Schneider F, Zilles K, 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal

cortex: intersubject variability and probability maps. Anat. Embryol. (Berl) 210, 343–352. [PubMed: 16208455]

Amunts K, Malikovic A, Mohlberg H, Schormann T, Zilles K, 2000. Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? Neuroimage 11, 66–84. [PubMed: 10686118]

Ashburner J, Friston KJ, 2005. Unified segmentation. Neuroimage 26, 839–851. [PubMed: 15955494]

Bagby RM, Taylor GJ, Parker JD, 1994. The Twenty-item Toronto Alexithymia Scale-II. Convergent, discriminant, and concurrent validity. J. Psychosom. Res 38, 33–40. [PubMed: 8126688]

Ball T, Rahm B, Eickhoff SB, Schulze-Bonhage A, Speck O, Mutschler I, 2007. Response properties of human amygdala subregions: evidence based on functional MRI combined with probabilistic anatomical maps. PLoS ONE 2, e307. [PubMed: 17375193]

Beauchamp MS, 2005. See me, hear me, touch me: multisensory integration in lateral occipital– temporal cortex. Curr. Opin. Neurobiol 15, 145–153. [PubMed: 15831395]

Belin P, Fecteau S, Bedard C, 2004. Thinking the voice: neural correlates of voice perception. Trends Cogn. Sci 8, 129–135. [PubMed: 15301753]

Bishop S, Duncan J, Brett M, Lawrence AD, 2004a. Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. Nat. Neurosci 7, 184–188. [PubMed: 14703573]

Bishop SJ, Duncan J, Lawrence AD, 2004b. State anxiety modulation of the amygdala response to unattended threat-related stimuli. J. Neurosci 24, 10364–10368. [PubMed: 15548650]

Botvinick MM, Cohen JD, Carter CS, 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci 8, 539–546. [PubMed: 15556023]

Britton JC, Taylor SF, Sudheimer KD, Liberzon I, 2006. Facial expressions and complex IAPS pictures: common and differential networks. Neuroimage 31, 906–919. [PubMed: 16488159]

Calder AJ, Lawrence AD, Young AW, 2001. Neuropsychology of fear and loathing. Nat. Rev. Neurosci 2, 352–363. [PubMed: 11331919]

Campanella S, Belin P, 2007. Integrating face and voice in person perception. Trends Cogn. Sci 11, 535–543. [PubMed: 17997124]

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD, 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science 280, 747–749. [PubMed: 9563953]

Carter CS, van Veen V, 2007. Anterior cingulate cortex and conflict detection: an update of theory and data. Cogn. Affect. Behav. Neurosci 7, 367–379. [PubMed: 18189010]

Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K, 2006. The human inferior parietal cortex: cytoarchitectonic parcellation and interindividual variability. Neuroimage 33, 430–448. [PubMed: 16949304]

Chen YH, Dammers J, Boers F, Leiberg S, Edgar JC, Roberts TP, Mathiak K, 2009. The temporal dynamics of insula activity to disgust and happy facial expressions: a magnetoencephalography study. Neuroimage 47, 1921–1928. [PubMed: 19442746]

Collignon O, Girard S, Gosselin F, Roy S, Saint-Amour D, Lassonde M, Lepore F, 2008. Audio-visual integration of emotion expression. Brain Res. 1242, 126–135. [PubMed: 18495094]

Corbetta M, Patel G, Shulman GL, 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron 58, 306–324. [PubMed: 18466742]

Corbetta M, Shulman GL, 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci 3, 201–215. [PubMed: 11994752]

Creutzig F, Sprekeler H, 2008. Predictive coding and the slowness principle: an information-theoretic approach. Neural Comput. 20, 1026–1041. [PubMed: 18085988]

de Gelder B, Vroomen J, 2000. The perception of emotions by ear and by eye. Cogn. Emotion 14, 289–311.

Derntl B, Finkelmeyer A, Eickhoff S, Kellermann T, Falkenberg DI, Schneider F, Habel U, 2010. Multidimensional assessment of empathic abilities: neural correlates and gender differences. Psychoneuroendocrinology 35, 67–82. [PubMed: 19914001]

Desimone R, Duncan J, 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci 18, 193–222. [PubMed: 7605061]

Dolan RJ, Fletcher P, Morris J, Kapur N, Deakin JF, Frith CD, 1996. Neural activation during covert processing of positive emotional facial expressions. Neuroimage 4, 194–200. [PubMed: 9345509]

Dolan RJ, Morris JS, de Gelder B, 2001. Crossmodal binding of fear in voice and face. Proc. Natl. Acad. Sci. U. S. A 98, 10006–10010. [PubMed: 11493699]

Drevets WC, 2000. Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. Prog. Brain Res 126, 413–431. [PubMed: 11105660]

Drevets WC, Price JL, Furey ML, 2008. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. Brain Struct. Funct 213, 93–118. [PubMed: 18704495]

Durston S, Davidson MC, Thomas KM, Worden MS, Tottenham N Martinez A, Watts R, Ulug AM, Casey BJ, 2003. Parametric manipulation of conflict and response competition using rapid mixed-trial event-related fMRI. Neuroimage 20, 2135–2141. [PubMed: 14683717]

Egner T, Etkin A, Gale S, Hirsch J, 2008. Dissociable neural systems resolve conflict from emotional versus nonemotional distracters. Cereb. Cortex 18, 1475–1484. [PubMed: 17940084]

Egner T, Hirsch J, 2005. The neural correlates and functional integration of cognitive control in a Stroop task. Neuroimage 24, 539–547. [PubMed: 15627596]

Eickhoff SB, Heim S, Zilles K, Amunts K, 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage 32, 570–582. [PubMed: 16781166]

Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT, 2009. Coordinate-based ALE meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum. Brain Mapp 30, 2907–2926. [PubMed: 19172646]

Eickhoff SB, Paus T, Caspers S, Grosbras MH, Evans AC, Zilles K, Amunts K, 2007. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage 36, 511–521. [PubMed: 17499520]

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K, 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25, 1325–1335. [PubMed: 15850749]

Ethofer T, Pourtois G, Wildgruber D, 2006. Investigating audiovisual integration of emotional signals in the human brain. Prog. Brain Res 156, 345–361. [PubMed: 17015090]

Fales CL, Barch DM, Rundle MM, Mintun MA, Snyder AZ, Cohen JD, Mathews J, Sheline YI, 2008. Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression. Biol. Psychiatry 63, 377–384. [PubMed: 17719567]

Fecteau S, Belin P, Joanette Y, Armony JL, 2007. Amygdala responses to nonlinguistic emotional vocalizations. Neuroimage 36, 480–487. [PubMed: 17442593]

Fox PT, Huang A, Parsons LM, Xiong JH, Zamarippa F, Rainey L, Lancaster JL, 2001. Location-probability profiles for the mouth region of human primary motor–sensory cortex: model and validation. Neuroimage 13, 196–209. [PubMed: 11133322]

Geyer A, Sebata T, Peenze I, Steele AD, 1996. Group B and C porcine rotaviruses identified for the first time in South Africa. J. S. Afr. Vet. Assoc 67, 115–116. [PubMed: 9120852]

Geyer S, 2004. The microstructural border between the motor and the cognitive domain in the human cerebral cortex. Adv. Anat. Embryol. Cell Biol 174, 1–89.

Goeleven E, De RR, Baert S, Koster EH, 2006. Deficient inhibition of emotional information in depression. J. Affect. Disord 93, 149–157. [PubMed: 16647141]

Goossens L, Kukolja J, Onur OA, Fink GR, Maier W, Griez E, Schruers K, Hurlemann R, 2009. Selective processing of social stimuli in the superficial amygdala. Hum. Brain Mapp 30, 3332–3338. [PubMed: 19347877]

Grachev ID, Berdichevski D, Rauch SL, Heckers S, Kennedy DN, Caviness VS, Alpert NM, 1999. A method for assessing the accuracy of intersubject registration of the human brain using anatomic landmarks. Neuroimage 9, 250–268. [PubMed: 9927554]

Green MJ, Phillips ML, 2004. Social threat perception and the evolution of paranoia. Neurosci. Biobehav. Rev 28, 333–342. [PubMed: 15225975]

Grefkes C, Eickhoff SB, Nowak DA, Dafotakis M, Fink GR, 2008. Dynamic intra- and interhemispheric interactions during unilateral and bilateral hand movements assessed with fMRI and DCM. Neuroimage 41, 1382–1394. [PubMed: 18486490]

Gur RC, Sara R, Hagendoorn M, Marom O, Hughett P, Macy L, Turner T, Bajcsy R, Posner A, Gur RE, 2002a. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. J. Neurosci. Meth 115, 137–143.

Gur RC, Schroeder L, Turner T, McGrath C, Chan RM, Turetsky BI, Alsop D, Maldjian J, Gur RE, 2002b. Brain activation during facial emotion processing. Neuroimage 16, 651–662. [PubMed: 12169250]

Haas BW, Omura K, Constable RT, Canli T, 2006. Interference produced by emotional conflict associated with anterior cingulate activation. Cogn. Affect. Behav. Neurosci 6, 152–156. [PubMed: 17007235]

Habel U, Windischberger C, Derntl B, Robinson S, Kryspin-Exner I, Gur RC, Moser E, 2007. Amygdala activation and facial expressions: explicit emotion discrimination versus implicit emotion processing. Neuropsychologia 45, 2369–2377. [PubMed: 17408704]

Hautzinger M, Keller F, Kühner C, 2006. Beck-Depressions-Inventar (BDI-II). Revision. Harcourt Test Services, Frankfurt/Main.

Haxby JV, Hoffman EA, Gobbini MI, 2000. The distributed human neural system for face perception. Trends Cogn. Sci 4, 223–233. [PubMed: 10827445]

Hellier P, Barillot C, Corouge I, Gibaut B, Le GG, Collins DL, Evans A, Malandain G, Ayache N, Christensen GE, Johnson HJ, 2003. Retrospective evaluation of intersubject brain registration. IEEE Trans. Med. Imaging 22, 1120–1130. [PubMed: 12956267]

Jakobs O, Wang LE, Dafotakis M, Grefkes C, Zilles K, Eickhoff SB, 2009. Effects of timing and movement uncertainty implicate the temporo-parietal junction in the prediction of forthcoming motor actions. Neuroimage 47, 667–677. [PubMed: 19398017]

Joormann J, 2010. Cognitive inhibition and emotion regulation in depression. Curr. Dir. Psychol. Sci 19, 161–166.

Kasess CH, Windischberger C, Cunnington R, Lanzenberger R, Pezawas L, Moser E, 2008. The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. Neuroimage 40, 828–837. [PubMed: 18234512]

Kerns JG, Cohen JD, MacDonald AW III, Cho RY, Stenger VA, Carter CS, 2004. Anterior cingulate conflict monitoring and adjustments in control. Science 303, 1023–1026. [PubMed: 14963333]

Kiebel SJ, Glaser DE, Friston KJ, 2003. A heuristic for the degrees of freedom of statistics based on multiple variance parameters. Neuroimage 20, 591–600. [PubMed: 14527620]

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, 1445–1456. [PubMed: 17659885]

Kurth F, Eickhoff SB, Schleicher A, Hoemke L, Zilles K, Amunts K, 2010. Cytoarchitecture and probabilistic maps of the human posterior insular cortex. Cereb. Cortex 20, 1448–1461. [PubMed: 19822572]

Leppanen JM, 2006. Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. Curr. Opin. Psychiatry 19, 34–39. [PubMed: 16612176]

MacDonald AWI, Cohen JD, Stenger VA, Carter CS, 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science 288, 1835–1838. [PubMed: 10846167]

Matthews S, Simmons A, Strigo I, Gianaros P, Yang T, Paulus M, 2009. Inhibition-related activity in subgenual cingulate is associated with symptom severity in major depression. Psychiatry Res. 172, 1–6. [PubMed: 19239982]

McDonald AJ, 1998. Cortical pathways to the mammalian amygdala. Prog. Neurobiol 55, 257–332. [PubMed: 9643556]

Meyer M, Zysset S, von Cramon DY, Alter K, 2005. Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. Brain Res. Cogn. Brain Res 24, 291–306. [PubMed: 15993767]

Mitchell DG, Luo Q, Mondillo K, Vythilingam M, Finger EC, Blair RJ, 2008. The interference of operant task performance by emotional distracters: an antagonistic relationship between the amygdala and frontoparietal cortices. Neuroimage 40, 859–868. [PubMed: 18234519]

Mitchell DG, Nakic M, Fridberg D, Kamel N, Pine DS, Blair RJ, 2007. The impact of processing load on emotion. Neuroimage 34, 1299–1309. [PubMed: 17161627]

Mitchell JP, 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. Cereb. Cortex 18, 262–271. [PubMed: 17551089]

Montagne B, Kessels RP, de Haan EH, Perrett DI, 2007. The Emotion Recognition Task: a paradigm to measure the perception of facial emotional expressions at different intensities. Percept. Mot. Skills 104, 589–598. [PubMed: 17566449]

Nachev P, Kennard C, Husain M, 2008. Functional role of the supplementary and pre-supplementary motor areas. Nat. Rev. Neurosci 9, 856–869. [PubMed: 18843271]

Ochsner KN, Hughes B, Robertson ER, Cooper JC, Gabrieli JD, 2009. Neural systems supporting the control of affective and cognitive conflicts. J. Cogn. Neurosci 21, 1842–1855. [PubMed: 18823233]

Oldfield RC, 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9, 97–113. [PubMed: 5146491]

Palomero-Gallagher N, Vogt BA, Schleicher A, Mayberg HS, Zilles K, 2009. Receptor architecture of human cingulate cortex: evaluation of the four-region neurobiological model. Hum. Brain Mapp 30, 2336–2355. [PubMed: 19034899]

Pessoa L, Padmala S, Morland T, 2005. Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. Neuroimage 28, 249–255. [PubMed: 15993624]

Phillips ML, Young AW, Scott SK, Calder AJ, Andrew C, Giampietro V, Williams SC, Bullmore ET, Brammer M, Gray JA, 1998. Neural responses to facial and vocal expressions of fear and disgust. Proc. Biol. Sci 265, 1809–1817. [PubMed: 9802236]

Pourtois G, de GB, Bol A, Crommelinck M, 2005. Perception of facial expressions and voices and of their combination in the human brain. Cortex 41, 49–59. [PubMed: 15633706]

Robins DL, Hunyadi E, Schultz RT, 2009. Superior temporal activation in response to dynamic audio-visual emotional cues. Brain Cogn. 69, 269–278. [PubMed: 18809234]

Roy AK, Shehzad Z, Margulies DS, Kelly AM, Uddin LQ, Gotimer K, Biswal BB, Castellanos FX, Milham MP, 2009. Functional connectivity of the human amygdala using resting state fMRI. Neuroimage 45, 614–626. [PubMed: 19110061]

Sander K, Scheich H, 2001. Auditory perception of laughing and crying activates human amygdala regardless of attentional state. Brain Res. Cogn. Brain Res 12, 181–198. [PubMed: 11587889]

Schleicher A, Palomero-Gallagher N, Morosan P, Eickhoff SB, Kowalski T, de VK, Amunts K, Zilles K, 2005. Quantitative architectural analysis: a new approach to cortical mapping. Anat. Embryol. (Berl) 210, 373–386. [PubMed: 16249867]

Sumner P, Nachev P, Morris P, Peters AM, Jackson SR, Kennard C, Husain M, 2007. Human medial frontal cortex mediates unconscious inhibition of voluntary action. Neuron 54, 697–711. [PubMed: 17553420]

Vogt BA, Hof PR, Vogt LJ, 2004. Cingulate Gyrus. In: Paxinos G, Mai JK (Eds.), The Human Nervous System. Amsterdam, Elsevier, pp. 915–949.

Weissman DH, Warner LM, Woldorff MG, 2004. The neural mechanisms for minimizing cross-modal distraction. J. Neurosci 24, 10941–10949. [PubMed: 15574744]

Wittchen HU, Zaudig M, Fydrich T, 1997. Strukturiertes Klinisches Interview für DSM-IV. Hogrefe, Göttingen.

Wittfoth M, Schroder C, Schardt DM, Dengler R, Heinze HJ, Kotz SA, 2010. On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. Cereb. Cortex 20, 383–392. [PubMed: 19505993]

Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC, 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp 4, 58–73. [PubMed: 20408186]

Yukie M, 2002. Connections between the amygdala and auditory cortical areas in the macaque monkey. Neurosci. Res 42, 219–229. [PubMed: 11900831]
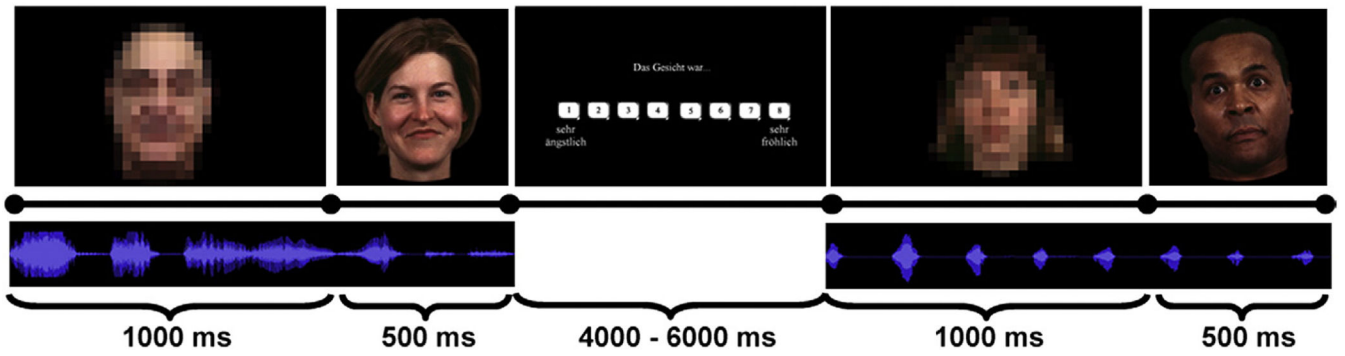
Fig. 1.
Experimental procedure: Every trial started with the presentation of a sound concurrently with a blurred neutral face. After 1000 ms the blurred face was displaced by an emotional or neutral face and presented with the continuing sound for another 500 ms.
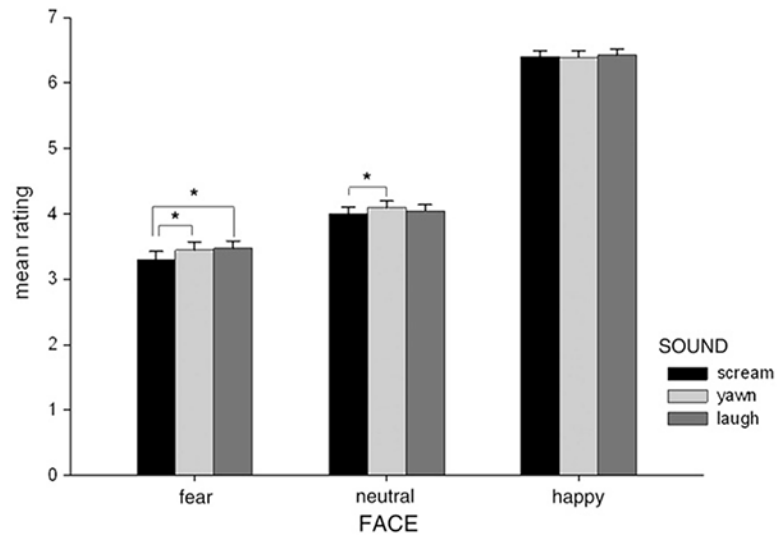
Fig. 2.
Mean ratings of fearful, neutral and happy faces while distracted with screams, yawns or laughs (0—very fearful, 8—very happy). Screams led significantly to a more fearful rating of fearful and neutral faces compared to yawns (and compared to laughs in case of fearful faces).
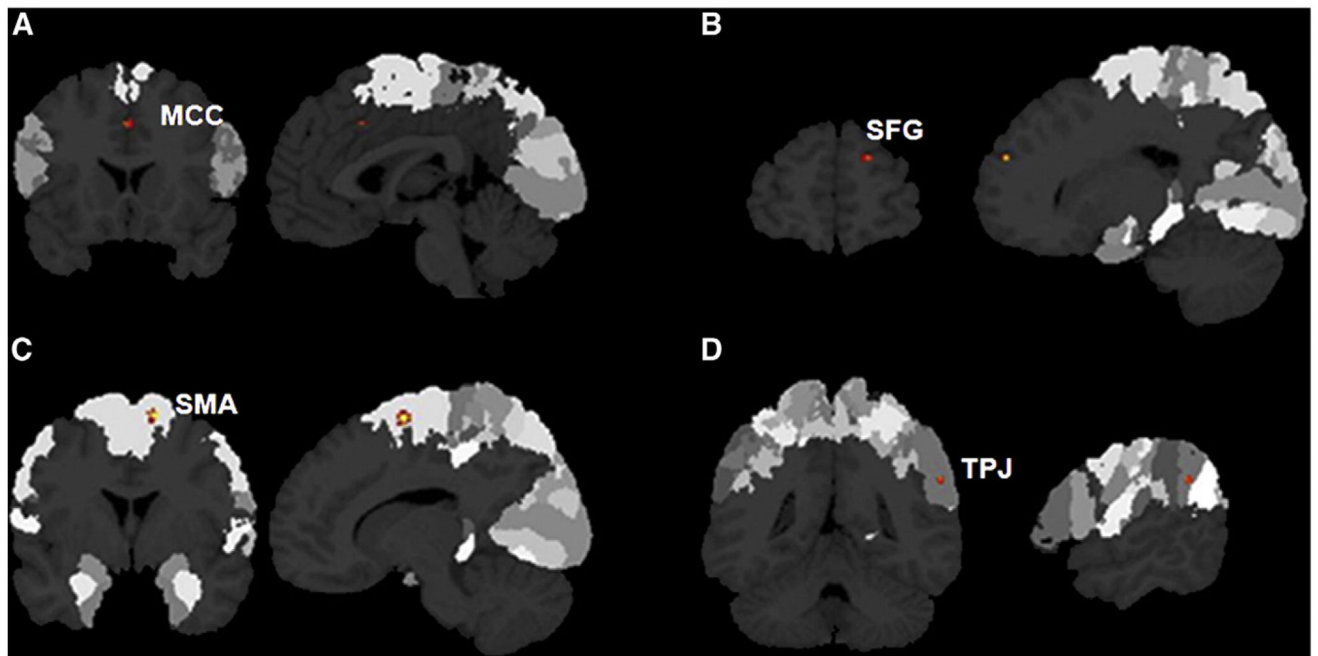
Fig. 3.
Activation in (A) middle cingulate cortex (MCC), (B) right superior frontal gyrus (SFG), (C) right supplementary motor area (SMA) and (D) right temporoparietal junction (TPJ) in emotional incongruent compared to congruent conditions. Dark grey brain regions represent regions that, until now, have not been mapped using observer-independent cytoarchitectonic analysis (Schleicher et al., 2005).
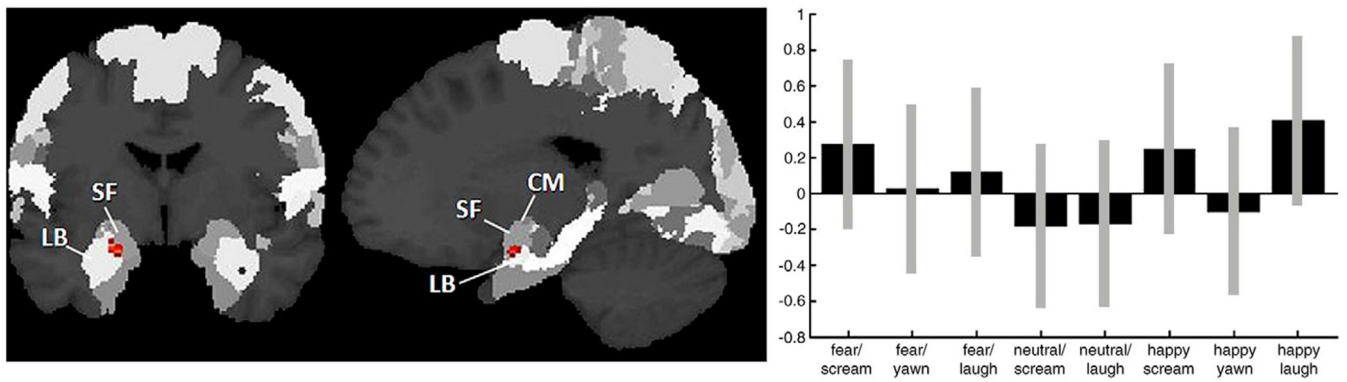
Fig. 4.
Amygdala activation when conditions where in both modalities an emotional stimulus was presented (fearful/scream, fearful/laugh, happy/scream, happy/laugh) were compared to conditions with a neutral stimulus in one and an emotional stimulus in the other modality (fearful/yawn, neutral/scream, neutral/laugh, happy/yawn). LB laterobasal, SF superficial, CM centromedial nucleus of the amygdala (Amunts et al., 2005).
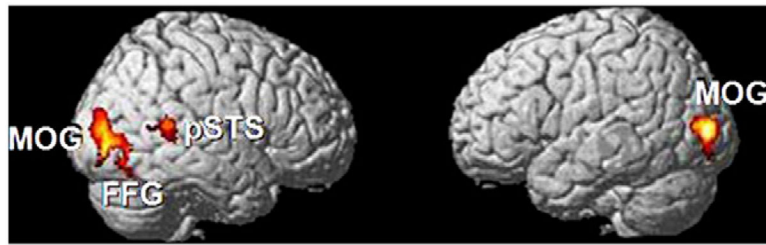
Fig. 5.
Negative correlation of BDI-scores with activation in left and right middle occipital gyrus (MOG) and right fusiform gyrus (FFG) and posterior superior temporal sulcus (pSTS).

Author Manuscript   Author Manuscript   Author Manuscript   Author Manuscript

Table 1

Mean ratings (standard deviation) of emotion and arousal of fearful, neutral and happy faces and sounds obtained outside the scanner.

	Face			Sound		
	Fearful	**Neutral**	**Happy**	**Scream**	**Yawn**	**Laugh**
Emotion	3.49 (0.7)	4.74 (0.43)	7.12 (0.67)	1.55 (0.65)	5.04 (0.40)	7.71 (0.71)
Arousal	5.46 (1.21)	4.23 (0.99)	4.74 (1.49)	7.11 (1.46)	2.57 (1.27)	4.80 (1.47)

Table 2

Activation happy versus neutral faces.

Region	Cytoarchitectonic area	x	y	z	Z-score	No. of voxels
Left calcarine	17	−18	−64	9	5.30	116
Left middle occipital gyrus	18	−14	−99	6	5.25	13
Left inferior parietal cortex	IPC (PFm)	−46	−54	46	5.27	21
Left postcentral gyrus	IPC (PFt)	−57	−16	26	5.06	12
Left primary motor area (M1)	4a	−36	−29	62	Inf	1251
Left supplementary motor area (SMA)	6	−4	−12	54	5.48	99
Left insula		−33	20	0	5.12	12
Right calcarine	17	15	−78	8	7.17	1522
Right thalamus		6	−4	−2	5.18	11

Coordinates are in MNI space; mapping of cytoarchitectonic areas: 17,18—(Amunts et al., 2000); PFm, PFt—(Caspers et al., 2006); 4a—(Geyer et al., 1996; Geyer, 2004).

Table 3

Activation scream versus yawn and laugh versus yawn.

Region	Cytoarchitectonic area	x	y	z	Z-score	No. of voxels
Scream vs. yawn						
Left middle cingulate cortex		−4	−20	42	6.58	203
Left IPC	IPC (PF)	−58	−39	30	5.98	286
	IPC (PFcm)	−46	−42	28	5.52	39
Left superior temporal gyrus (STG)	Id1	−45	−18	−6	Inf	1964
Right IPC	IPC (PFm)	56	−46	26	5.47	65
Right STG		45	−32	9	7.73	1732
Right SMA	6	6	10	58	5.39	25
Laugh vs. Yawn						
Left STG		−58	−18	2	Inf	1260
Right STG		63	−18	0	Inf	2201

Coordinates are in MNI space; mapping of cytoarchitectonic areas: PF, PFcm—(Caspers et al., 2006); 6—(Geyer, 2004); Id1— (Kurth et al., 2010).

Table 4

Regions showing incongruence effects of emotional valence.

Region	Cytoarchitectonic area	*x*	*y*	*z*	*Z*-score	No. of voxels
Middle cingulate cortex		−2	18	36	5.25	17
Right SMA	6	12	0	64	5.81	64
Right TPJ	IPC (PFm)	58	−45	28	5.26	16
Right superior frontal gyrus		18	56	22	5.46	16

Coordinates are in MNI space; mapping of cytoarchitectonic areas: PFm—(Caspers et al., 2006); 6—(Geyer, 2004).

Table 5

Additional regions showing incongruence effects of emotion-presence.

Region	Cytoarchitectonic area	x	y	z	Z-score	No. of voxels
Left lingual gyrus	17	−6	−75	2	5.33	48
Left middle temporal gyrus (MTG)		−48	−21	−6	5.23	23
Right calcarine	17	20	−68	4	6.18	330
Right STG		62	−24	12	5.24	16

Coordinates are in MNI space; mapping of cytoarchitectonic areas: 17—(Amunts et al., 2000).