available at www.sciencedirect.com

SCIENCE DIRECT®

www.elsevier.com/locate/brainres

ELSEVIER

BRAIN RESEARCH

Research Report

The emotional power of music: How music enhances the feeling of affective pictures

Thomas Baumgartner[a,*], *Kai Lutz*[b], *Conny F. Schmidt*[b,c], *Lutz Jäncke*[b]

[a]*Institute for Empirical Research in Economics and Neuroeconomics, University of Zurich, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland*
[b]*Institute of Psychology, Department of Neuropsychology, University of Zurich, Switzerland*
[c]*Institute for Biomedical Engineering, University of Zurich and Swiss Federal Institute of Technology Zurich, Switzerland*

ARTICLE INFO

Article history:
Accepted 8 December 2005
Available online 3 February 2006

Keywords:
fMRI
Emotion
Cognition
Picture
Music
Multimodal stimuli
Perception
Feelings

ABSTRACT

Music is an intriguing stimulus widely used in movies to increase the emotional experience. However, no brain imaging study has to date examined this enhancement effect using emotional pictures (the modality mostly used in emotion research) and musical excerpts. Therefore, we designed this functional magnetic resonance imaging study to explore how musical stimuli enhance the feeling of affective pictures. In a classical block design carefully controlling for habituation and order effects, we presented fearful and sad pictures (mostly taken from the IAPS) either alone or combined with congruent emotional musical excerpts (classical pieces). Subjective ratings clearly indicated that the emotional experience was markedly increased in the combined relative to the picture condition. Furthermore, using a second-level analysis and regions of interest approach, we observed a clear functional and structural dissociation between the combined and the picture condition. Besides increased activation in brain areas known to be involved in auditory as well as in neutral and emotional visual–auditory integration processes, the combined condition showed increased activation in many structures known to be involved in emotion processing (including for example amygdala, hippocampus, parahippocampus, insula, striatum, medial ventral frontal cortex, cerebellum, fusiform gyrus). In contrast, the picture condition only showed an activation increase in the cognitive part of the prefrontal cortex, mainly in the right dorsolateral prefrontal cortex. Based on these findings, we suggest that emotional pictures evoke a more cognitive mode of emotion perception, whereas congruent presentations of emotional visual and musical stimuli rather automatically evoke strong emotional feelings and experiences.

1. Introduction

Music is widely used to enhance the emotional impact of movies. For example, in thrilling horror movies, music and sound effects enhance the feeling of fear and anxiety. For many spectators, this congruent visual and auditory experience could become even so emotional that they could no longer bear it, and therefore, they would turn off the TV, in particular when watching scary movies. Another strategy to reduce the emotional experience to a tolerable level often reported by spectators is simply to turn off the sound of the movie. Surprisingly, this emotional "on–off switch" effect is

* *Corresponding author*. Fax: +41 44 634 49 07.
E-mail address: t.baumgartner@iew.unizh.ch (T. Baumgartner).

0006-8993/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2005.12.065

completely understudied in the brain research literature to date. To our best knowledge, there exists no neuroimaging study which explored this emotional enhancement effect of musical excerpts on emotional pictures.

However, some recent studies have demonstrated the emotional power of music. In these studies, musical excerpts presented without concomitant visual stimuli elicited strong emotional responses activating brain regions known to be involved in reward/motivation, approach/withdrawal, or arousal, including ventral striatum, midbrain, orbitofrontal cortex, the insula and ventral medial prefrontal cortex (e.g., Blood and Zatorre, 2001; Blood et al., 1999; Brown et al., 2004). In addition, there are few behavioral and brain activation studies which addressed the question of the integrated perception of emotion regarding voice and face stimuli. Behavioral studies have revealed that congruent crossmodal perception of emotional faces and voices leads to (1) a facilitation in facial emotion recognition (de Gelder and Vroomen, 2000) and (2) to enhanced affective judgements of emotional faces (for example, a fearful face is more likely to be perceived as fearful if accompanied by a fearful voice (Massaro and Egan, 1996)). The first study which directly addressed this integration question using fMRI revealed that congruent fearful face–voice pairs compared with incongruent pairs (happy voice + fearful face) elicited increased activation in the amygdala and the fusiform gyrus (Dolan et al., 2001). In a recent similar study using both single and combined modality conditions, activation in the right extended amygdala was obtained for fearful faces and fearful audio-visual pairs but not if fearful voices were presented alone, thus demonstrating no enhanced brain processing in the amygdala in the combined congruent emotional condition (Pourtois et al., 2005). However, the perception of audio-visual emotions (positive and negative) relative to unimodal conditions was associated with activation in the anterior fusiform gyrus and the medial temporal gyrus known to be involved in multisensory integration of neutral and emotional stimuli (Adolphs et al., 2002; Mesulam, 1998).

As demonstrated above, the influence of combined presentation of auditory and visual stimuli in the context of emotional experience has substantially been understudied. However, in a combined EEG and psychophysiology study of our laboratory using visual and musical stimuli, we recently demonstrated that congruent combined presentation of affective stimuli irrespective of valence enhances (1) the global cortical brain activation (measured by Alpha power) and (2) increases skin conductance responses, heart rate, respiration as well as psychometrical arousal measures compared to the picture condition (Baumgartner et al., 2005). However, electrical fields measured at particular electrode sites are too inaccurate to infer the underlying intracerebral cortical activations. Therefore, we designed the present neuroimaging study in order to explore the effect of crossmodal presentation of congruent emotional stimuli on emotional experience and concomitant cortical and subcortical activation pattern. In this study, we used the same emotional musical excerpts and affective pictures as in the EEG study. Fearful, happy, and sad pictures of the International Affective Picture System (Lang et al., 1995) were presented in a classical block design either alone or combined with congruent classical musical excerpts known to evoke the same emotional experience as the pictures (Peretz et al., 1998).

Due to the explorative nature of the study and to increase the statistical power, we focused our interest on two neural systems, proposed in a recent excellent review about the neurobiology of emotion perception: a ventral and a dorsal system (Phillips et al., 2003). Based on recent animal, human lesion, and functional neuroimaging studies, the authors suggest that the ventral system, including the amygdala, insula, striatum, thalamus, brainstem nuclei, and ventral regions of the anterior cingulate cortex and prefrontal cortex, is important for the identification of the emotional significance of the stimuli, the production of affective states (including autonomic, neuroendocrine, and somatomotor responses, as well as conscious emotional feeling), and automatic and autonomic regulation of emotional responses. The dorsal system, including the hippocampus and dorsal regions of the anterior cingulate gyrus and prefrontal cortex, is more important for cognitive processes, including selective attention, planning, and effortful rather than automatic regulation of affective states. We expected that the emotional musical stimuli mainly increase the emotional significance of the visual stimuli and, moreover, enhance the conscious emotional feeling and somatic body reactions (as we recently confirmed in the mentioned combined EEG and psychophysiology study of our laboratory (Baumgartner et al., 2005)). Therefore, we hypothesized an activation increase in the combined compared to the picture condition mainly in the automatic ventral system of emotion processing. Furthermore, because the subjects had no explicit cognitive task (they only had the instruction to feel the presented emotions), we hypothesized no change in activation between the two conditions or rather an activation decrease in the cognitive dorsal system in the combined condition. However, the medial temporal lobe memory system (including the hippocampus and the parahippocampus) which is part of the dorsal system has been shown to be strongly activated during the encoding phase of highly arousing emotional stimuli. This activation was present although the subjects were not explicitly instructed to memorize the presented stimulus material, indicating that the memory system is automatically activated by emotional stimuli (Dolcos et al., 2004). Therefore, we would also expect increased activation in the memory system of the brain in the more arousing combined condition of our experiment. In addition to the regions of the dorsal and ventral system of emotion processing, we included, based on the findings by Dolan et al. (2001) and Pourtois et al. (2005) mentioned above, the extrastriate cortex, and in particular the fusiform gyrus and the crossmodal integration areas (the medial temporal gyrus and the temporal pole) in our a priori regions of interest, expecting increased activation in these areas in the combined relative to the picture condition.

2. Results

2.1. Psychometrical results

Repeated measures ANOVA regarding the valence scale demonstrated a highly significant main effect of emotion

Fig. 1 – Mean valence ratings for all three emotional categories and two modalities (mean ± SD). Depicted is the trend for a significant interaction effect between emotion and modality ($P < 0.1$), indicating that the subjects experienced the happy conditions more positive and the fear and sad conditions more negative when congruent emotional music was simultaneously presented. *C* = Combined, *P* = Picture.

($F_{(2,7)}$=132.6, $P < 0.001$, ETA^2 = 0.974) and a trend for a significant interaction between emotion and modality ($F_{(2,7)}$ = 3.6, P = 0.083, ETA^2 = 0.509), indicating that the subjects clearly differentiated on a 5-point scale between the negative (fear condition: mean ± SD: 1.3 ± 0.3; sad condition: 1.4 ± 0.6) and positive emotions (happy condition: 4.7 ± 0.4). In addition and more interesting, the interaction showed that the subjects experienced the different emotions more extreme in the combined condition, that is they rated the happy condition more positive and the sad and fear condition more negative when congruent emotional music was simultaneously presented (see Fig. 1). Repeated measures ANOVA regarding the involvement scale revealed a highly significant main effect of modality ($F_{(1,8)}$=11.6, $P < 0.01$, ETA^2 = 0.594) which was qualified by an increased involvement and arousal experience in the combined (mean = 4.32, SD = 0.35) compared to the picture conditions (mean = 3.88, SD = 0.55). In addition, the subjects clearly indicated at the end of the experiment (separately for each emotion) that, despite of the noisy scanner environment, (1) they apprehended the musical excerpts (fear music: 4.22 ± 0.83; sad music: 4.33 ± 0.9; happy music: 4.33 ± 0.7); and (2) they recognized the emotional tune of the musical excerpts (fear music: 4.22 ± 1.1; sad music: 4.67 ± 0.5; happy music: 4.56 ± 0.7). Furthermore, all subjects also confirmed at the end of the experiment that they had experienced the combined conditions as more intense than the picture conditions (fear: 4.33 ± 0.7; sad: 4.67 ± 0.7; happy: 4.33 ± 0.8). One sample *t* tests revealed that these ratings significantly differed from the value 3 (all $P < 0.001$) which would have been chosen by the subjects if the emotional intensity in the combined conditions would have been equal as the intensity in the picture conditions. Summing up, the psychometrical ratings clearly indicated that the emotional experiences during the combined conditions were significantly increased compared to the picture conditions, and that the subjects were able to apprehend and recognize the emotional musical excerpts despite of the noisy scanner environment.

2.2. *Brain activation data*

2.2.1. *Combined versus fixation (C > F) and picture versus fixation (P > F)*

Results for these two contrasts are listed in Tables 1 and 2. As expected, both contrasts revealed widespread bilateral hemodynamic responses in the occipital lobe (inferior and middle occipital gyrus, cuneus, fusiform gyrus) and in the bilateral cerebellum. In both contrasts, activation was also found in the limbic system (parahippocampus and hippocampus) but with the distinction that the activation was more distributed and bilateral in the combined condition. Left amygdala activation, activation in the right precentral gyrus and in the pons as well as bilateral activation in the temporal lobe (superior, middle and inferior temporal gyrus) were only observed in the contrast $C > F$. Finally, only the contrast $P > F$ revealed strong hemodynamic responses in the right lateral prefrontal cortex (inferior frontal gyrus) as well as small hemodynamic responses in the left middle frontal gyrus.

Table 1 – Peak activations observed for the contrast combined versus fixation ($C > F$)

Brain regions	BA	Side	x	y	z	Max t score	Voxels
Frontal lobe							
Precentral gyrus	6	R	48	0	40	8.16**	14
Temporal lobe							
Superior temporal gyrus	22/38	L	−44	12	−28	6.71**	22
	22	R	52	−20	−4	5.60*	60
Middle temporal gyrus	21/38	R	56	−4	−20	6.26*	22
	21/38	L	−32	4	−40	6.57**	44
Inferior temporal gyrus	20	L	−40	−8	−36	5.11*	38
Parahippocampus/ Uncus	35/36	L	−20	−4	−28	14.16***	59
	35/36	R	20	−24	−24	6.73**	64
Hippocampus		L	−24	−8	−24	8.58**	36
Hippocampus		R	32	−28	−12	5.19*	21
Amygdala		L	−20	−4	−24	8.08**	9
Occipital lobe							
Middle occipital gyrus	18/19	L	−28	−96	8	6.11*	128
		R	28	−92	16	6.74**	83
Inferior occipital gyrus	17/18/19	L/R	12	−92	−16	8.42**	113
Cuneus	18	L	−16	−100	0	7.30**	
Fusiform gyrus	18/19/20	L	−24	−88	−24	12.18***	31
	18/19/20	R	24	−87	−24	8.21**	43
Subcortical structures							
Cerebellum		L	−4	−36	−8	11.61***	105
		R	40	−56	−36	7.83**	233
Pons		R	8	−32	−32	4.70*	15

The coordinates are given according to the MNI space together with its *T* scores and significant thresholds (*$P < 0.001$, **$P < 0.0001$ ***$P < 0.00001$ (all uncorrected for multiple comparisons)). Minimum cluster size 5 voxels. For every structure, only the highest maxima for the left and right hemisphere are reported.

Table 2 – Peak activations observed for the contrast picture versus fixation (P > F)

Brain regions	BA	Side	x	y	z	Max t score	Voxels
Frontal lobe							
Inferior frontal gyrus	44/45	R	52	8	32	7.11**	56
Middle frontal gyrus	45	L	−52	32	20	4.38*	6
Temporal lobe							
Parahippocampus/	27	L	−20	−32	−8	4.89*	12
Hippocampus	28	R	20	−28	−12	5.42*	9
Occipital lobe							
Middle occipital	18/19	L	−40	−80	0	11.90***	130
gyrus	18/19	R	28	−92	16	12.05***	65
Inferior occipital gyrus	17/18/19	L	−12	−96	−16	7.59**	134
	17/18/19	R	12	−96	−16	9.03***	140
Cuneus	18	L	−20	−100	0	11.34***	18
Fusiform gyrus	18/19	L	−24	−88	−24	10.55***	23
	18/19	R	22	−86	−20	4.82*	33
Subcortical structures							
Cerebellum		L	−48	−64	−28	6.29*	175
		R	24	−88	−28	9.12***	202

The coordinates are given according to the MNI space together with its T scores and significant thresholds (*$P < 0.001$, **$P < 0.0001$, ***$P < 0.00001$ (all uncorrected for multiple comparisons)). Minimum cluster size 5 voxels. For every structure, only the highest maxima for the left and the right hemisphere are reported.

2.2.2. *Combined versus picture (C > P) and picture versus combined (P > C)*

Results for these two contrasts are shown in Figs. 2 and 3 and listed in Table 3. Contrasting the combined condition with the picture condition (C > P) revealed as expected widespread bilateral activations in the superior and middle temporal gyrus. In addition, this contrast was associated with increased activations in the ventral part of the medial frontal lobe, in the parietal cortex (posterior cingulate/precuneus), in the bilateral fusiform gyrus, in the limbic system (bilateral amygdala, hippocampus, and parahippocampus), in the motor system (bilateral cerebellum, right precentral gyrus), in subcortical structures (left nucleus caudatus, right pons), and the left insula.

Contrasting the picture condition with the combined condition (P > C) mainly revealed activations in the dorsal part of the right prefrontal cortex (right inferior frontal gyrus, right middle frontal gyrus, left superior frontal gyrus). No further significant activations were observed despite of the relatively relaxed statistical threshold of $P < 0.005$ (uncorrected).

2.3. *Regions of interest analysis*

According to our main hypothesis, we conducted regions of interest analysis using unsmoothed data to examine the activation in the medial temporal lobe system, including the amygdala, the hippocampus and the parahippocampus. Two-way repeated measures ANOVAs regarding the extracted mean t values of the contrast images combined versus fixation (C > F) and picture versus fixation (P > F) revealed for all three examined structures of the medial temporal lobe system a significant main effect of modality (Amygdala: $F_{(1,8)} = 15.6$, $P < 0.01$, $ETA^2 = 0.661$; Hippocampus: $F_{(1,8)}=7.8$, $P < 0.05$, $ETA^2 = 0.494$; Parahippocampus: $F_{(1,8)} = 10.5$, $P < 0.05$, $ETA^2 = 0.567$). This main effect clearly indicated increased hemodynamic responses bilateral in the amygdala, the hippocampus, and the parahippocampus in the combined condition in comparison with the picture condition (see Fig. 4). One sample t tests further confirmed these findings showing that only the combined condition led to an activation increase in these structures which was significantly different from zero (all $P < 0.05$). No main effect of hemisphere and no interaction effect of hemisphere*modality was observed. We obtained similar ROI results using the smoothed data (12 mm) of the whole brain analysis (not presented here).

2.4. *Correlation analysis*

Because the region of interest analysis revealed strong bilateral activation in the amygdala, the parahippocampus, and hippocampus in the combined compared to the picture condition, we conducted the correlation analysis using the mean activation of both hemispheres. We found a strong positive correlation between the amygdala and the arousal ratings obtained at the end of the experiment ($r = 0.73$, $P = 0.024$, 2 tailed, see Fig. 5). No correlation between the arousal ratings and the hemodynamic responses in the hippocampus and parahippocampus were found (all $P > 0.35$).

3. Discussion

The observed psychometrical results and brain activation pattern confirmed our hypothesis of enhanced emotional processing in the combined condition compared to the picture condition (note that we only explored the negative emotional conditions). Subjects reported that the emotional experience was more intense during the combined conditions along with increased activation in most of the brain areas proposed to compose the automatic ventral system of emotion perception (Phillips et al., 2003), including bilateral amygdala, ventral frontal cortex, left striatum, left insula, and brainstem nuclei (right pons). Furthermore, we also observed in the combined condition compared to the picture condition a stronger activation in the medial temporal lobe memory system (including the hippocampus and parahippocampus) which is part of the more cognitive dorsal system for emotion perception. Finally, the results confirmed our hypothesis of enhanced brain activation in the combined condition in extrastriate visual processing areas (fusiform gyrus) and in brain areas known to be involved in neutral and emotional visual–auditory integration processes (bilateral temporal pole, bilateral medial temporal gyrus; (Mesulam, 1998). Apart from these hypothesized structures, we found in the combined condition an activation increase in brain areas involved in emotion and memory processes (posterior cingulate/precuneus (Maddock et al., 2003)) as well as in motor control

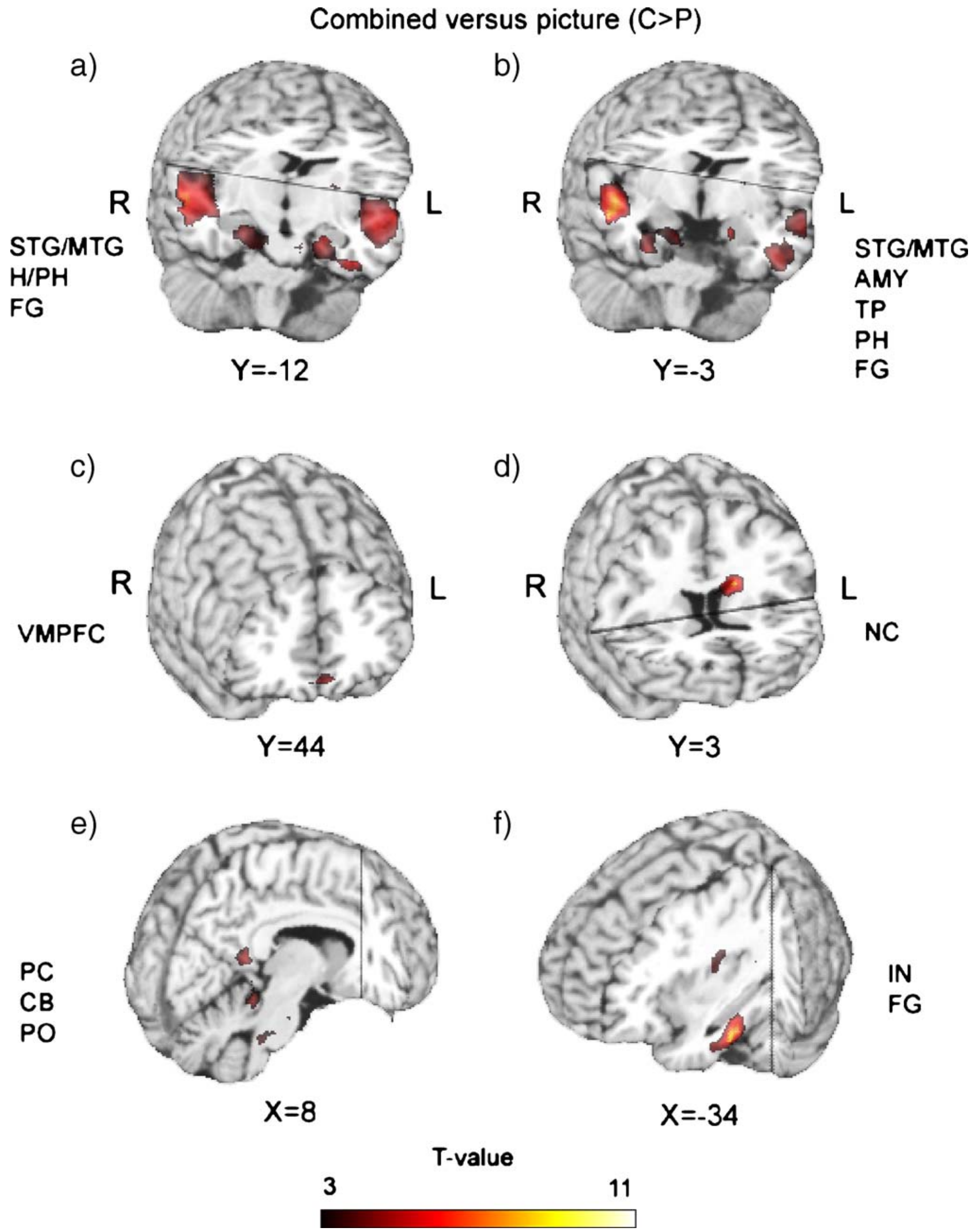

Fig. 2 – Depicted is the increased activity in the combined compared to the picture condition (all $P < 0.005$, uncorrected). SPM(t) maps are overlaid on a structural MRI brain. (a) Bilateral superior and middle temporal gyrus (STG, MTG), bilateral hippocampus/parahippocampus (H/PH) and left fusiform gyrus (FG); (b) bilateral superior and middle temporal gyrus (STG/MTG), left amygdala (AMY), left temporal pole (TP), right fusiform gyrus/inferior temporal gyrus (FG), right parahippocampus (PH); (c) ventral medial prefrontal cortex (VMPFC); (d) nucleus caudatus (NC); (e) right pons (PO), right cerebellum (CB), right posterior cingulate/precuneus (PC); (f) left fusiform gyrus (FG), left insula (IN).

(bilateral cerebellum, right precentral gyrus). Summing up, most of the hypothesized brain areas (as expected mainly in the ventral system of emotion perception) showed increased activation in the combined compared to the picture condition.

In contrast, the picture condition (composed only of negative pictures) only showed an activation increase in the more cognitive dorsal system for emotion perception, mainly in the right dorsolateral prefrontal cortex (right inferior frontal gyrus, right middle frontal gyrus) as well as a smaller activation increase in the left superior frontal gyrus. This frontal activation pattern is compatible with current theories of emotional processes (Davidson, 2003; Davidson and Irwin, 1999), suggesting increased left frontal brain processing for approach-related positive emotions and increased right frontal processing for withdrawal-related negative emotions. No other activation increase in the picture condition relative to the combined condition was observed despite of the relatively relaxed statistical threshold of $P < 0.005$ (uncorrected, cluster threshold: 5 voxels).

In the following, we will first discuss the increased activation in the combined condition ($C > P$) and afterwards the increased activation in the picture condition ($P > C$).

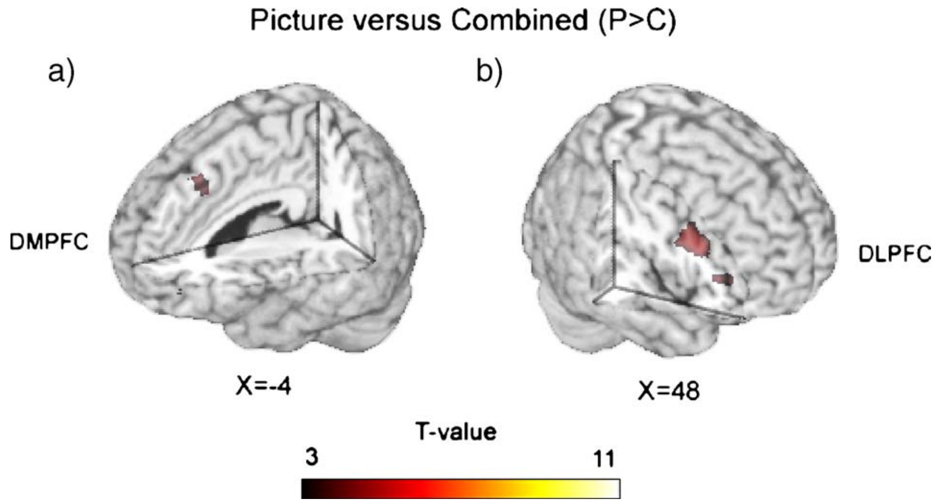

Fig. 3 – Depicted is the increased activity in the picture compared to the combined condition (all $P < 0.005$, uncorrected). SPM(t) maps are overlaid on a structural MRI brain. Note that increased activity was only found in the dorsal prefrontal cortex, mainly in the right dorsolateral prefrontal cortex (DLPFC). (a) left dorsomedial prefrontal cortex (DMPFC); (b) right dorsolateral prefrontal cortex (BA 44/45, DLPFC).

3.1. *Discussion of the increased activation in the combined condition (C > P)*

3.1.1. *Activations in auditory cortex and crossmodal integrations areas*

As expected, the combined condition revealed hemodynamic responses bilaterally in the entire primary and secondary auditory cortex (comprising BA 22, 41, 42) and the activation was more widespread in the right hemisphere, in agreement with studies of passive listening to music of non-musicians (Ohnishi et al., 2001; Zatorre et al., 1994). In addition, we found bilateral strong activation in the medial temporal gyrus (BA 21) and the right (BA 38) and left temporal polar region (BA 21). The medial temporal gyrus plays a role in multisensory integration (Mesulam, 1998), and brain damage to the temporal polar cortex impairs emotional recognition of both facial and vocal stimuli (Adolphs et al., 2002). In a recent crossmodal study using emotional voices and faces (Pourtois et al., 2005), it was found that the bimodal condition compared to the two unimodal conditions was associated with increased activity in the left medial temporal gyrus for positive voice–face pairs and in the right medial temporal gyrus for negative voice–face pairs. With our study, we could not confirm this lateralization pattern for positive and negative emotions because we found bilateral activation in this brain region for the combined condition (composed only of negative stimuli). Nevertheless, the results of our study suggest that the medial temporal gyrus and the temporal pole are not only involved in the crossmodal binding of voice and face stimuli but also in the crossmodal integration of congruent emotional pictures and musical excerpts.

3.1.2. *Activation in the amygdala*

The amygdala is a key structure in emotion processing involved in both the identification of the emotional significance of a stimulus and the production of an affective state (Phillips et al., 2003). We found activation within the amygdala only during the combined condition but not during the picture condition. Previous studies have reported amygdala activation to a wide variety of emotional stimuli, including fear faces (Morris et al., 1996; Phillips et al., 1997), aversive pictures (Simpson et al., 2000; Taylor et al., 1998), sad faces (Blair et al., 1999), happy faces (Breiter et al., 1996; Pessoa et al., 2002), positive pictures (Garavan et al., 2001; Hamann et al., 1999), film excerpts (Aalto et al., 2002; Reiman et al., 1997), as well as positive (laughing) and negative (crying) sound stimuli (Sander et al., 2003; Sander and Scheich, 2001). Our finding of an absence of amygdala activation to the picture condition is therefore surprising. However, in a recent meta-analysis of 55 neuroimaging studies, Phan et al. (2002) showed that only 50% of experiments using visual induction methods activated the amygdala. Furthermore, studies have demonstrated that an increasing intensity of sad (Blair et al., 1999) and fearful (Morris et al., 1996) facial expression is associated with enhanced activity in the amygdala. A recent study complemented these findings by demonstrating that the amygdala activation was dependent on both the intensity and type of displayed emotion, with significant responses to fearful facial expressions and to a lesser degree to happy expressions (Glascher et al., 2004). Therefore, the absence of amygdala activation in the picture condition as well as the presence of amygdala activation in the combined condition can be explained (at least to a large portion) by differences in arousal levels which are apparently in the picture condition to low to reach the threshold for amygdala activation. This interpretation is further strengthened by an observed strong positive correlation between psychometrical arousal measures and amygdala activation in the combined condition and is in line with current theories of amygdala function postulating for this structure a more general role for vigilance or for processing salience or attributes that make stimuli meaningful (Davis and Whalen, 2001).

3.1.3. *Activation in both memory systems*

In addition to the increased activation in the amygdala, we observed as expected in the combined relative to the picture

Table 3 – Peak activations observed for the contrasts picture versus combined (*P* > *C*) and combined versus picture (*C* > *P*)

	Brain regions	BA	Side	x	y	z	Max *t* score	Voxels
Picture versus combined (*P* > *C*)	Frontal lobe							
	Inferior frontal gyrus	44	R	52	24	28	5.37**	50
	Middle frontal gyrus	45/46	R	48	44	8	4.88*	17
	Superior frontal gyrus	8	L	−4	28	52	4.03*	8
Combined versus picture (*C* > *P*)	Frontal lobe							
	Medial frontal gyrus	11	L/R	0	44	−16	3.98*	15
	Precentral gyrus	6	R	52	0	44	6.05**	8
	Temporal lobe							
	Superior temporal gyrus	21	R	56	−20	−4	8.54***	235
		22	R	60	−24	4	8.00***	
		38	R	36	12	−32	4.36*	
		41	R	56	−28	8	7.82***	
		42	L	−64	−28	8	8.43***	162
		22	L	−56	−16	0	6.57**	
	Middle temporal gyrus	21	R	56	0	−20	10.34***	156
		21	R	64	−16	−12	7.65***	
		38	R	40	12	−44	5.52**	
		21	L	−52	8	−28	6.70**	142
	Parahippocampus	28	R	16	−8	−32	7.81***	27
		28	R	24	−16	−20	4.68*	
		36	L	−26	−20	−28	5.10**	36
	Amygdala		L	−17	−4	−20	5.89**	7
	Hippocampus/Amygdala		R	24	−8	−24	6.60**	27
	Hippocampus		L	−24	−20	−16	6.17**	18
	Fusiform gyrus	20	L	−32	−20	−32	9.39***	38
		20	L	−40	−8	−32	6.48**	
		20	R	35	−2	−36	4.04*	17
	Insula	13	L	−36	−8	12	3.68*	14
	Parietal lobe							
	Posterior cingulate/precuneus	29	L/R	8	−44	16	4.62*	19
	Subcortical structures							
	Nucleus caudatus		L	−8	16	0	5.95**	19
			L	−16	4	24	8.51***	
	Pons		L/R	4	−24	−36	4.67*	12
	Cerebellum		L/R	8	−40	−12	4.33*	17

The coordinates are given according to the MNI space together with its *T* scores and significant thresholds (*$P < 0.005$, **$P < 0.0005$, ***$P < 0.00005$ (all uncorrected for multiple comparisons)). Minimum cluster size 5 voxels. All observed maxima are reported. Note that all emotional conditions were only composed of negative stimuli (fear and sad, see methods part for an explanation).

condition an increased activity bilateral in the medial temporal lobe memory system supporting "declarative" memory (including the hippocampus and parahippocampus) as well as in the left nucleus caudatus supporting "procedural" memory (for a review of the two memory systems (Devan and White, 1999; Packard, 1999)). Recent evidence strongly support the idea that the amygdala exerts a modulatory influence on the hippocampus-dependent as well as caudate-dependent memory systems (McGaugh, 2004; Packard and Cahill, 2001). For example, memory for emotionally arousing material is not enhanced in human subjects with bilateral and selective lesions of the amygdala, as it is in normal subjects (Adolphs et al., 1997; Cahill et al., 1995). Furthermore, various recent fMRI studies using incidental learning paradigms (i.e., the subjects were not aware during the encoding phase of the subsequent memory tasks) demonstrated enhanced activity in the amygdala and the medial temporal lobe memory system to emotional arousing scenes during the encoding phase of the experiment. In addition, these studies also found significant correlations between episodic memory performance and the level of amygdala activation during encoding of negative and positive stimuli (Canli et al., 2000; Dolcos et al., 2004; Hamann et al., 1999). These studies clearly indicate that the medial temporal lobe memory system can be activated in a rather automatic way, modulated by strong activation of the amygdala similar as has been observed in the combined condition in our study. Therefore, we suggest to adapt the model by Phillips et al. (2003) by including the medial temporal lobe system not only in the cognitive dorsal system but also in the automatic ventral system for emotion processing as a transition zone between emotion and cognition.

3.1.4. *Activation in extrastriate visual processing areas*

Another interesting finding of this study (never reported before in neuroimaging papers) revealed that emotional musical stimuli can enhance the activity in visual processing areas (fusiform gyrus). Thus, our study confirms and extends the findings of studies demonstrating that crossmodal integration processes of congruent faces and voices increase the activity in the fusiform gyrus (Dolan et al., 2001; Pourtois et al., 2005). We suggest that this enhanced processing in the bilateral fusiform gyrus is evoked by markedly increased

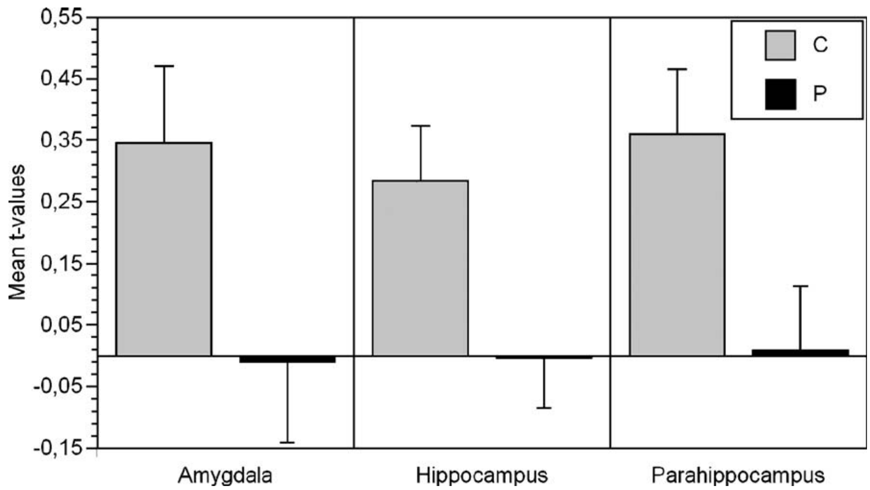

Fig. 4 – Region of interest analysis of the medial temporal lobe system (including amygdala, hippocampus, parahippocampus) using unsmoothed *t* values of the contrasts combined versus fixation ($C > F$) and picture versus fixation ($P > F$), respectively. Depicted are the significantly increased activations (mean ± SE) in the amygdala, the hippocampus, and parahippocampus in the combined (*C*) relative to the picture (*P*) condition (main effect of modality, all $P < 0.05$). No significant main effect of hemisphere or interaction effect of modality*hemisphere were observed.

arousal levels in the combined relative to the picture condition. In line with this interpretation, it has been shown that maximally intense expressions of emotions compared with neutral or modestly intense expressions were associated with enhanced activity in extrastriate visual processing area, and in particular the fusiform gyrus (Glascher et al., 2004; Lang et al., 1998; Surguladze et al., 2003; Taylor et al., 2000). Comparable with the influence of the amygdala on the medial temporal lobe memory system, it has been postulated (e.g., Morris et al., 1998; Rotshtein et al., 2001) that this enhanced activity in visual cortical areas is modulated by the amygdala. Consistent with this proposal, we found in the combined but not in the picture condition bilateral activation of the amygdala. Therefore, the increased activation in the fusiform gyrus in our study could indeed be modulated by the amygdala.

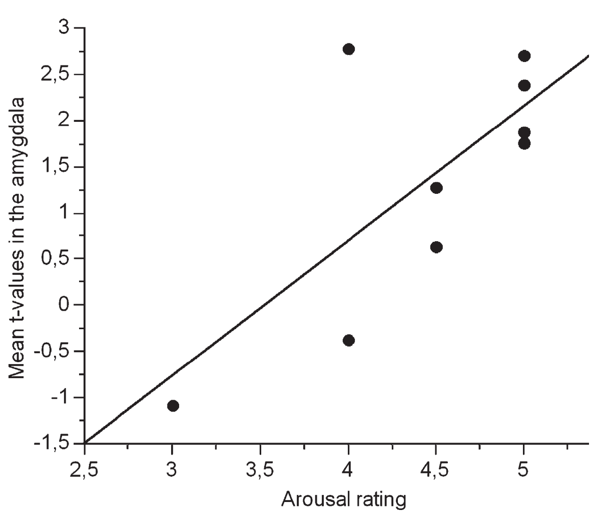

Fig. 5 – Strong positive correlation between the activation of the left and right amygdala in the combined condition and the psychometrical arousal measure obtained at the end of the experiment ($r = 0.73$, $P < 0.05$). Note that “3” on the rating scale means that the subjects experienced no emotional enhancement in the combined relative to the picture condition, whereas “4” and “5” on this scale mean that the subjects experienced the combined condition more arousing than the picture condition.

3.2. *Further activations*

Besides the increased activation of the amygdala, the medial temporal lobe memory system, and the fusiform gyrus in the combined condition, further structures showed increased activation in the combined compared to the picture condition, including hypothesized structures of the ventral system for emotion processing as the ventral medial frontal gyrus, the left insula, and brainstem nuclei (right pons) as well as not hypothesized structures as the posterior cingulate/precuneus, bilateral cerebellum, and right precentral gyrus. The activations of these structures are completely in line with the proposal of enhanced emotion processing in the combined compared to the picture condition. For example, ventral frontal regions (including the orbitofrontal cortex and the ventromedial frontal cortex) have been associated with the representation of the emotionally salient states of reward and punishment (O’Doherty et al., 2001; Rolls, 2000, 2004) and decision-making tasks requiring the manipulation of emotionally salient information (Bechara et al., 1998).

The insula, another structure showing increased activity in the combined condition, has been implicated in the mapping of body-related sensations, including temperature, pain, proprioception, and viscera (Craig, 2002), and consistent with this mapping hypothesis, insula activations were mainly found during aversive emotional stimulation that also evoked visceral/somatic sensation, including disgust (Phillips et al., 1997), sadness (George et al., 1996), fear conditioning (Buchel et al., 1999), and processing of fearful faces (Anderson et al., 2003). Moreover and in line with our finding of increased insula activation in the combined condition, Morris et al. (1998) showed that the insula was responsive to increasing intensity of fear.

The nuclei of the pons have been implicated in negative emotions of sadness and anger (Damasio et al., 2000) and especially acoustic startle responses to negative stimuli (Yeomans and Frankland, 1995). These nuclei receive projections from the cingulate cortex and the insula and send projections to the cerebellum. Therefore, Damasio et al. (2000) hypothesized that these projections guide the cerebellum in modulating and coordinating varied emotional action programs. Consistently, we also found bilateral activation of the cerebellum and the precentral gyrus in line with other emotion studies which have reported cerebellar activations both for negative and positive emotional stimuli (Damasio et al., 2000; Lane et al., 1997; Reiman et al., 1997; Taylor et al., 2000, 2003).

Finally, we found in the combined compared to the picture condition stronger activations of the posterior cingulate/precuneus. This brain area is involved in various functions including episodic memory (Maddock et al., 2003), spatial attention (Mesulam et al., 2001), and emotional stimulus processing (Damasio et al., 2000; Esslen et al., 2004; Maddock and Buonocore, 1997) and is heavily connected with frontal and temporal cortical areas (Allison et al., 2000; Morris et al., 1999). Therefore, it is suggested that this region mediates interactions of emotional and memory-related processes (Maddock et al., 2003), two processes which were strongly activated in the combined condition relative to the picture condition.

3.3. Discussion of the increased activation in the picture condition (P > C)

3.3.1. Activation in the prefrontal cortex

Most interestingly, we observed a clear functional and structural dissociation between the combined and picture condition. Whereas the combined condition demonstrated additional or increased activation in most structures of the ventral system known to be involved in emotion processing, the picture condition only showed one strong activation increase in the cognitive dorsal system, mainly in the right dorsolateral prefrontal cortex (BA 44/45). The dorsolateral prefrontal cortex is known to be involved in effortful regulation of attention, affective states, and autonomic responses (Phillips et al., 2003). Moreover, studies specifically observing the influence of different task instructions observed a similar functional and structural dissociation between more implicit (passive viewing) and more explicit emotion processing tasks (emotion-labeling task). Whereas the more implicit emotional processing tasks were associated with increased activity in the amygdala/hippocampal region, the more explicit emotion processing tasks were associated with increased activity in the prefrontal cortex, and in particular in the dorsolateral prefrontal cortex (BA 44/45 (Hariri et al., 2000; Nakamura et al., 1999)). In addition, consistent with these findings, Phan et al. (2002) demonstrated in a recent meta-analysis of emotional neuroimaging studies that experiments using an implicit emotion processing tasks were approximately 10 to 15% more numerous in demonstrating activation of subcortical regions (including the amygdala and the hippocampus), whereas emotional studies using a more explicit cognitive task were 15 to 20% more numerous in demonstrating activation in the dorsal anterior cingulate and prefrontal cortex. In our study, we did not vary the task instructions; the subjects had to follow in the combined as well as in the picture condition the same instruction, namely to place themselves into the same mood as expressed by the presented emotional stimuli. However, the distinctive brain activation patterns and subjective ratings clearly indicate that following the instruction during the combined condition must have been much easier and more automatic than during the picture condition. We hypothesize that the musical excerpts in the combined condition helped the subjects in a rather automatic and implicit way to feel the presented emotional experiences, whereas in the picture condition, the subjects cognitively evaluated the emotional stimuli or have even effortful and explicitly tried to up-regulate their affective states (as is known with moderate success), leading to activation of the dorsolateral prefrontal cortex. These findings strongly suggest a functional and structural dissociation between processes of cognitively evaluating emotions on the one hand and strongly feeling and experiencing emotions on the other hand. We postulate that different induction methods either activate more the cognitive mode or the feeling mode of emotion processing. Accordingly, these findings can give an explanation for the often discrepant findings in emotion research regarding for example amygdala activation or lateralization patterns.

3.4. Limitations

The finding of increased activation in the combined condition in most structures known to be involved in various emotional processes is striking (for excellent reviews, see Phan et al., 2002; Phillips et al., 2003). However, we are aware that this study has a major limitation. Other than the study of Pourtois et al. (2005), we have not included in our study a single musical condition. Therefore, we cannot not exclude that the activation pattern in the combined condition could also have been produced by the musical stimuli alone. In fact, most of the structures observed in the combined condition are reported in the literature to be activated by musical and sound stimuli (vocal and non-vocal) alone, including the orbitofrontal gyrus, the amygdala, the insula, the striatum, the cerebellum, the hippocampus, and the parahippocampus (Blood and Zatorre, 2001; Blood et al., 1999; Brown et al., 2004; Phillips et al., 1998; Sander and Scheich, 2001; Zatorre et al., 1994). Whereas most of these studies only showed activation in a few structures of the dorsal and ventral system for emotion perception, mainly one study (Blood and Zatorre, 2001) exploring the neural mechanism underlying intensely pleasant emotional responses to music using positron emissions tomography (PET) found a similar complex pattern of activation in this emotion network. However, no activation increase in the amygdala and the medial temporal lobe memory system was observed in this study, despite of high arousal levels. Interestingly, to date, only vocal (Phillips et al., 1998) and non-vocal sound stimuli (Sander et al., 2003) were reported to activate the amygdala. Similar to positive and negative visual stimuli which preferentially activated the amygdala (Phan et al., 2002), sound stimuli are important to evaluate changes in the environment, whereas musical stimuli alone have no direct evolutionary connection to the external world and therefore presumably activate a more internal mode of brain function (Baumgartner et al., 2005; Gusnard et al., 2001; Raichle et al., 2001). Consistent with these findings, it is postulated that the amygdala may be more responsible for processing of externally cued perceptual emotional stimuli (Reiman et al., 1997; Teasdale et al., 1999).

Nevertheless, despite of the interesting finding that no study to date has ever reported activation in the amygdala triggered by musical stimuli, we cannot rule out that the observed activation pattern in the combined condition could have been produced (at least partly) by musical stimuli alone. However, we want to emphasize the fact that the main goal of this study was not to explore specifically the effect of multimodal and unimodal brain processing, but rather to compare two different emotion induction methods which were never compared using neuroimaging methods. Therefore, this is the first neuroimaging study showing how musical stimuli can markedly enhance the emotional experience evoked by affective picture, namely by stronger or additional activation of most structures known to be involved in emotion perception and experience. Or in other words, although we used highly arousing emotional pictures of the International Affective Picture System, emotional musical excerpts still had the power to significantly increase the activation in most structures of the proposed emotional network for identification of the emotional significance of a stimulus and production of affective states. This is all the more remarkable because, in spite of using earphones and earplugs, the noisy scanner environment never can be completely suppressed. However, this noisy scanner environment might be the reason that two structures involved in emotion processing and arousal, including the thalamus (Anders et al., 2004) and the ventral part of the anterior cingulate (Elliott et al., 2000; Esslen et al., 2004), showed no activation increase in the combined condition in our study, whereas the study of Blood and Zatorre (2001) using similar emotional musical excerpts in the silent PET environment was able to observe an activation increase in these structures of the ventral system for emotion perception. Thus, we concluded that not surprisingly, the emotional power of music is attenuated in the noisy fMRI compared to a silent PET environment, yet still impressively effective demonstrated in this study by highly distinctive psychometrical ratings and brain activation patterns between the combined and the picture condition.

4. Conclusion

Summing up, in this study, we showed that classical musical stimuli strongly enhanced emotional processing of highly arousing affective pictures by increasing the activation in most structures of the ventral system for emotion processing, including the amygdala, the ventral medial frontal gyrus, the striatum, the insula, the brainstem, and the medial temporal lobe memory system (including the hippocampus and parahippocampus, proposed by this article to be also part of the ventral system). Moreover, this study also showed that congruent emotional musical stimuli can increase the activity in extrastriate visual processing areas, an area which has never before shown to be activated by musical stimuli alone. Finally, we found a clear structural and functional dissociation between the combined and the picture condition. Whereas the combined condition was associated with increased activation in a large and distributed network involved in emotion processing, the picture condition only showed one strong activation increase in the dorsolateral prefrontal cortex which is part of the cognitive dorsal system for emotion processing. We argue that these findings indicate that emotional pictures evoke a more cognitive emotional perception process, whereas combined presentations of congruent visual and musical emotional stimuli rather automatically evoke (strong) emotional feelings and experiences.

5. Experimental procedures

5.1. *Subjects*

Nine right-handed (tested with standard handedness tests revealing consistent right handedness in all subjects according to the criterion proposed by Annett (1970)), healthy females (mean ± SD age, 24.78 ± 2.9, range: 21–30) were examined in the study, all of them students at the University of Zurich, Switzerland. Female participants were chosen because previous studies showed that, compared to men, women are more likely to report intense emotional experiences along with stronger physiological reactions of the body (Lang et al., 1993). All subjects underwent a physical evaluation to screen out chronic diseases, mental disorders, medication, and drug or alcohol abuse. Furthermore, depression and anxiety were assessed by the German versions of the Self-Rating Depression Scale (SDS (Zung, 1965)) and the State-Trait Anxiety Inventory (STAI (Laux et al., 1981)). Four of the original 13 subjects had to be excluded. Two of them because their score in the STAI and SDS questionnaires were not within the normal range for the general population and the other two because of technical sound problems during the experiment. Each subject received payment for the participation. The study was carried out in accordance with the Declaration of Helsinki principles, approved by the ethics committee of the University of Zurich. All subjects gave written, informed consent and were informed of their right to discontinue participation at any time.

5.2. *Stimuli*

Fifty-four fear-, sad-, and happy-inducing pictures were taken from the International Affective Picture System (IAPS) or had been collected by the author (about 70% of the pictures were taken from the IAPS). All pictures contained humans or human faces, were matched for complexity, and rated for emotional content in a pilot experiment by 48 subjects on 9-point scales for valence and arousal ("9" indicating that the subjects felt very happy and aroused, respectively). The results of the pilot study clearly indicated that the presented stimuli were perceived as highly arousing negative or positive pictures. The mean ratings (±standard deviations) for the three picture categories were as follows: valence: 2.1 ± 0.71 (fear picture), 3.1 ± 0.73 (sadness picture), 7.7 ± 0.70 (happy picture); arousal: 6.8 ± 0.90 (fear picture), 5.5 ± 0.83 (sadness picture), 6.0 ± 0.77 (happy picture). The emotional musical stimuli were taken from classical orchestral pieces and consisted of excerpts of exactly 44-s duration. One excerpt of Gustav Holst (Mars—the Bringer of War from The Planets) was chosen to evoke fear, one excerpt of Samuel Barber (Adagio for

Strings) was chosen to evoke sadness, and one excerpt of Beethoven (Symphony no.6) was chosen to evoke the emotion of happiness. Various psychological and psychophysiological experiments have shown that these excerpts are capable of evoking the mentioned three basic emotions (e.g., Krumhansl, 1997; Peretz et al., 1998). Musical stimuli were presented binaurally through MRI-compatible headphones with an individually adjusted comfortable listening level (between 70 and 80 dB). The visual stimuli were presented via a video projector onto a translucent screen that subjects viewed inside the scanner via mirror. All stimuli were presented under computer control (Presentation, Neurobehavioral systems, version 0.79, 2003).

5.3. *Study design*

The fMRI measurements were conducted in the context of a classical block-design experiment during which stimulation (ON period) alternated with rest (OFF period). The whole experiment consisted of one run of 18-min and 28.8-s duration resulting in a total of 504 Scans (TR: 2.2 s). During this run, 6 different emotional blocks were repeated two times (3 emotions: fear, happiness, sadness; and two modalities: picture or combined) resulting in a total of 12 emotional stimulation blocks (6 picture conditions and 6 combined conditions). The emotional stimulation blocks were 20 TRs (44 s) in duration. A block consisted either of 9 different emotional pictures of the same emotional category alone (happy, fear, or sad pictures) or combined with the congruent musical excerpt (happy picture together with happy music, fear picture together with fear music, and sad picture together with sad music). The pictures were presented in pseudo-random order for 4.8 s with a gap of 100 ms between each picture. The same musical stimuli were presented during the whole emotional condition, but in order to avoid startling the participants, the beginning (2 s) and the end (2 s) of each stimulus were faded in and out, respectively. Both the visual as well as the musical stimuli were presented twice during the experiment. All visual stimuli were presented with and without music. The musical excerpts were never presented alone. Before every emotional condition, a short instruction screen was presented for 4.4 s (2 TRs) which instructed the subjects to place themselves into the same mood as expressed by the presented emotional stimuli (similar mood induction methods were also used by Esslen et al., 2004; Gur et al., 2002; Kimbrell et al., 1999). After each experimental condition, subjects gave two affective ratings via button-press on a computer-based 5-point scale (valence and involvement ratings, ranking from "1 = negative or weak" to "5 = positive or strong"). The involvement rating measured how strong the subjects were involved or engaged in the different emotional experiences. Each of these two affective questions was presented for 11 s resulting in 22 s for both questions (10 TRs). Finally, a fixation cross was presented for 22 s (10 TRs) between every emotional condition. Because some emotion papers have shown strong emotional habituation effects (e.g., Stark et al., 2004), we carefully controlled for these habituation and order effects. Thus, half of the subjects (n = 5) first experienced the picture conditions and afterwards the combined conditions, whereas the other half of the subjects (n = 4) first experienced the combined conditions and afterwards the picture conditions. The presentation of three picture conditions of every emotional category (fear, sad, and happy) was always followed by the presentation of three combined conditions of every emotional category or vice versa. The different emotions were further presented in a counterbalanced and pseudo-random order. At the end of the experiment, the following two questions regarding the noisy scanner environment had to be answered separately for each emotion on a 5-point scale, ranking from "1 = I agree not at all" to "5 = I agree very strongly": "Despite the noise scanner environment, (1) I was able to apprehend the musical excerpts and (2) I was able to recognize the emotional tune of the musical excerpts". Furthermore, the subjects had to rate their emotional intensity experience on a 5-point scale (separately for each emotion), where "3" indicated that the emotional experience in the combined conditions was equal intense as in the picture conditions, "1" and "2" indicated that the picture conditions were more intense than the combined conditions and, finally, where "4" and "5" indicated that the emotional experience in the combined conditions was stronger than in the picture conditions.

5.4. *Data acquisition*

The Experiment was conducted on a 3-T Philips Intera whole body MR Scanner (Philips Medical Systems, Best, The Netherlands) equipped with an 8-channel Philips SENSE head coil. Structural image acquisition consisted of 160 T_1-weighted transversal images (1 mm slice thickness). For functional imaging, a total of 504 volumes were obtained using a SENSitivity Encoded (SENSE (Pruessmann et al., 1999)) T_2^*-weighted echo-planar imaging sequence using an acceleration factor of 2.0. 36 axial slices were acquired covering the whole brain with a slice thickness of 3 mm; no inter-slice gap; interleaved acquisition; TR = 2200 ms; TE = 35 ms; flip angle=77°, field of view=220 mm; matrix size=80*80. The orientation of the axial slices was parallel to the AC–PC line.

5.5. *General image analysis*

For the preprocessing and statistical analyses, the statistical parametric mapping software package (SPM99, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab (Version 6.5) was used. For analysis, all images were realigned to the first volume, corrected for motion artefacts and normalized (4 mm^3) into standard stereotaxic space (template provided by the Montreal Neurological Institute), and smoothed using a 12 mm full-width-at-half-maximum Gaussian kernel. The six experimental conditions, fear, sadness, and happiness (picture and combined) were modeled with a box-car function convolved with a hemodynamic response function in the General Linear Model of SPM. A band-pass filter, which was composed of a discrete cosine-basis function with a cut-off period of 264 s for the high-pass filter and a canonical hemodynamic response function for the low-pass filter, was applied. In order to increase signal to noise ratio, global intensity changes were minimized by scaling each image to the grand mean.

5.6. *General statistical analysis*

For the different analyses, different t contrast images were calculated at the individual level (first level). To increase the statistical power due to the small number of repetition of the different conditions (only two per condition), we decided to combine the negative emotional conditions (fear and sadness) for statistical analysis due to the following reason. Several recent studies on brain activity have demonstrated that negative and aversive events elicit more prominent emotional responses than do positive or neutral ones (e.g., Carretie et al., 2001, 2004; Northoff et al., 2000), including our combined EEG and psychophysiology study in which we used the same emotional visual and musical stimuli as in the current fMRI study (Baumgartner et al., 2005). Thus, we focused our analysis in this study on the negative emotional conditions and the brain activation pattern of the happy conditions will not be analyzed and reported. However, we will report the psychometrical ratings for all 3 emotional categories. Therefore, 3 different conditions were used in this paper for the analysis of the different brain activation patterns: (1) combined condition (composed of sad and fearful musical and visual stimuli), (2) picture condition (composed of sad and fearful visual stimuli), and (3) fixation baseline (fixation cross in the middle of the screen). Accordingly, we calculated four t contrast images at the individual level: combined versus picture ($C > P$); picture versus combined ($P > C$); combined versus fixation ($C > F$);

picture versus fixation ($P > F$). The resulting contrast images were used for a subsequent random effects second-level analyses.

For the a priori regions of interest, we used for these second-level analyses voxel-wise intensity threshold of $P < 0.005$ (uncorrected for multiple comparison) with a spatial extent threshold of five contiguous voxels. The statistical criterion was selected on the basis of similar criteria used for the examination of a priori regions of interest in other neuroimaging studies employing random effects statistical models (e.g., Hamann and Mao, 2002). Based on previous studies (Dolan et al., 2001; Pourtois et al., 2005) and the two neural systems for emotion perception proposed by Phillips et al. (2003), the medial temporal lobe system (including the amygdala, the hippocampus, and the parahippocampus), the ventral and dorsal prefrontal cortex, the anterior cingulate, the insula, the striatum, the brainstem nuclei, the thalamus, the fusiform gyrus, the medial temporal gyrus, and the temporal pole were defined as the regions of interest. For all regions not belonging to these regions of interest, we did not use different statistical thresholds. Thus, voxels passing these lenient thresholds will be interpreted more carefully.

In addition to the second-level analysis, we conducted regions of interest (ROI) analysis for medial temporal lobe structures. ROIs were marked using an anatomic atlas for automatic labeling of structural brain images normalized into MNI space (Tzourio-Mazoyer et al., 2002) and included (1) amygdala; (2) hippocampus; (3) parahippocampal gyrus. For each ROI mean T values were computed for each subject and each condition based on the contrasts combined versus fixation baseline ($C > F$) and picture versus fixation baseline ($P > F$). To be sure that the activation in the amygdala is not influenced by the smoothing of neighbouring regions (we used a rather large smoothing filter for the whole brain analysis), we used unsmoothed data to calculate our ROI analysis. Finally, we used these mean T values of the ROI analysis to correlate the brain activation pattern in the amygdala, the hippocampus, and parahippocampal gyrus with the psychometrical arousal ratings obtained at the end of the experiment. For this purpose, we calculated mean psychometrical arousal ratings for rating sad and fear conditions.

We applied three different strategies to analyze the similarities and differences between the combined and picture conditions. First, we descriptively examined the brain regions displaying significant activation during the combined and the picture conditions separately for each condition compared to fixation ($C > F$ and $P > F$). Then we analyzed (by means of a second-level model) which of the brain regions showed statistically significant changes in activation between the combined and the picture conditions ($C > P$ and $P > C$). Finally, we explored the differences between combined and picture conditions using an ROI approach which focused specifically on the medial temporal lobe system and in particular on the amygdala.

Psychometrical and ROI analyses were performed using the statistical software package SPSS for PC (version 11.5). For the two psychometrical measures (valence and involvement rating scales), two-way repeated measures ANOVAs were performed with the following factors: "emotion" (fear, happiness, sadness) and "modality" (combined, picture). For every region of interest, we calculated a two-way repeated measure ANOVA with the following factors: "modality" (combined, picture) and "hemisphere" (left, right). Results were considered significant at the level of $P < 0.05$. In case of a significant multivariate effect, post hoc paired t tests were computed using the Bonferroni correction according to Holm (1979). As effect size measure ETA^2 is reported.

REFERENCES

Aalto, S., Naatanen, P., Wallius, E., Metsahonkala, L., Stenman, H., Niem, P.M., Karlsson, H., 2002. Neuroanatomical substrata of amusement and sadness: a PET activation study using film stimuli. NeuroReport 13, 67–73.

Adolphs, R., Cahill, L., Schul, R., Babinsky, R., 1997. Impaired declarative memory for emotional material following bilateral amygdala damage in humans. Learn. Mem. 4, 291–300.

Adolphs, R., Damasio, H., Tranel, D., 2002. Neural systems for recognition of emotional prosody: a 3-D lesion study. Emotion 2, 23–51.

Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues: role of the STS region. Trends Cogn. Sci. 4, 267–278.

Anders, S., Lotze, M., Erb, M., Grodd, W., Birbaumer, N., 2004. Brain activity underlying emotional valence and arousal: a response-related fMRI study. Hum. Brain Mapp. 23, 200.

Anderson, A.K., Christoff, K., Panitz, D., De, R.E., Gabrieli, J.D., 2003. Neural correlates of the automatic processing of threat facial signals. J. Neurosci. 23, 5627–5633.

Annett, M., 1970. A classification of hand preference by association analysis. Br. J. Psychol. 61, 303–321.

Baumgartner, T., Esslen, M., Jancke, L., 2005. From emotion perception to emotion experience: emotions evoked by pictures and classical music. Int. J. Psychophysiol.

Bechara, A., Damasio, H., Tranel, D., Anderson, S.W., 1998. Dissociation of working memory from decision making within the human prefrontal cortex. J. Neurosci. 18, 428–437.

Blair, R.J., Morris, J.S., Frith, C.D., Perrett, D.I., Dolan, R.J., 1999. Dissociable neural responses to facial expressions of sadness and anger. Brain 122 (Pt. 5), 883–893.

Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. Proc. Natl. Acad. Sci. U. S. A. 98, 11818–11823.

Blood, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C., 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. Nat. Neurosci. 2, 382–387.

Breiter, H.C., Etcoff, N.L., Whalen, P.J., Kennedy, W.A., Rauch, S.L., Buckner, R.L., Strauss, M.M., Hyman, S.E., Rosen, B.R., 1996. Response and habituation of the human amygdala during visual processing of facial expression. Neuron 17, 875–887.

Brown, S., Martinez, M.J., Parsons, L.M., 2004. Passive music listening spontaneously engages limbic and paralimbic systems. NeuroReport 15, 2033–2037.

Buchel, C., Dolan, R.J., Armony, J.L., Friston, K.J., 1999. Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. J. Neurosci. 19, 10869–10876.

Cahill, L., Babinsky, R., Markowitsch, H.J., McGaugh, J.L., 1995. The amygdala and emotional memory. Nature 377, 295–296.

Canli, T., Zhao, Z., Brewer, J., Gabrieli, J.D., Cahill, L., 2000. Event-related activation in the human amygdala associates with later memory for individual emotional experience. J. Neurosci. 20, RC99.

Carretie, L., Mercado, F., Tapia, M., Hinojosa, J.A., 2001. Emotion, attention, and the 'negativity bias', studied through event-related potentials. Int. J. Psychophysiol. 41, 75–85.

Carretie, L., Hinojosa, J.A., Martin-Loeches, M., Mercado, F., Tapia, M., 2004. Automatic attention to emotional stimuli: neural correlates. Hum. Brain Mapp. 22, 290–299.

Craig, A.D., 2002. How do you feel? Interoception: the sense of the physiological condition of the body. Nat. Rev., Neurosci. 3, 655–666.

Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J., Hichwa, R.D., 2000. Subcortical and cortical brain activity during the feeling of self-generated emotions. Nat. Neurosci. 3, 1049–1056.

Davidson, R.J., 2003. Affective neuroscience and psychophysiology: toward a synthesis. Psychophysiology 40, 655–665.

Davidson, R.J., Irwin, W., 1999. The functional neuroanatomy of emotion and affective style. Trends Cogn. Sci. 3, 11–21.
Davis, M., Whalen, P.J., 2001. The amygdala: vigilance and emotion. Mol. Psychiatry 6, 13–34.
de Gelder, B., Vroomen, J., 2000. The perception of emotions by ear and by eye. Cogn. Emot. 14, 289–311.
Devan, B.D., White, N.M., 1999. Parallel information processing in the dorsal striatum: relation to hippocampal function. J. Neurosci. 19, 2789–2798.
Dolan, R.J., Morris, J.S., de, G.B., 2001. Crossmodal binding of fear in voice and face. Proc. Natl. Acad. Sci. U. S. A. 98, 10006–10010.
Dolcos, F., LaBar, K.S., Cabeza, R., 2004. Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. Neuron 42, 855–863.
Elliott, R., Friston, K.J., Dolan, R.J., 2000. Dissociable neural responses in human reward systems. J. Neurosci. 20, 6159–6165.
Esslen, M., Pascual-Marqui, R.D., Hell, D., Kochi, K., Lehmann, D., 2004. Brain areas and time course of emotional processing. NeuroImage 21, 1189–1203.
Garavan, H., Pendergrass, J.C., Ross, T.J., Stein, E.A., Risinger, R.C., 2001. Amygdala response to both positively and negatively valenced stimuli. NeuroReport 12, 2779–2783.
George, M.S., Ketter, T.A., Parekh, P.I., Herscovitch, P., Post, R.M., 1996. Gender differences in regional cerebral blood flow during transient self-induced sadness or happiness. Biol. Psychiatry 40, 859–871.
Glascher, J., Tuscher, O., Weiller, C., Buchel, C., 2004. Elevated responses to constant facial emotions in different faces in the human amygdala: an fMRI study of facial identity and expression. BMC Neurosci. 5, 45.
Gur, R.C., Schroeder, L., Turner, T., McGrath, C., Chan, R.M., Turetsky, B.I., Alsop, D., Maldjian, J., Gur, R.E., 2002. Brain activation during facial emotion processing. NeuroImage 16, 651–662.
Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E., 2001. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 4259–4264.
Hamann, S., Mao, H., 2002. Positive and negative emotional verbal stimuli elicit activity in the left amygdala. NeuroReport 13, 15–19.
Hamann, S.B., Ely, T.D., Grafton, S.T., Kilts, C.D., 1999. Amygdala activity related to enhanced memory for pleasant and aversive stimuli. Nat. Neurosci. 2, 289–293.
Hariri, A.R., Bookheimer, S.Y., Mazziotta, J.C., 2000. Modulating emotional responses: effects of a neocortical network on the limbic system. NeuroReport 11, 43–48.
Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6, 65–70.
Kimbrell, T.A., George, M.S., Parekh, P.I., Ketter, T.A., Podell, D.M., Danielson, A.L., Repella, J.D., Benson, B.E., Willis, M.W., Herscovitch, P., Post, R.M., 1999. Regional brain activity during transient self-induced anxiety and anger in healthy adults. Biol. Psychiatry 46, 454–465.
Krumhansl, C.L., 1997. An exploratory study of musical emotions and psychophysiology. Can. J. Exp. Psychol. 51, 336–353.
Lane, R.D., Reiman, E.M., Ahern, G.L., Schwartz, G.E., Davidson, R.J., 1997. Neuroanatomical correlates of happiness, sadness, and disgust. Am. J. Psychiatry 154, 926–933.
Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O., 1993. Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 30, 261–273.
Lang, P.J., Bradley, M., Cuthbert, B., 1995. International Affective Picture System. Center for Research in Psychophysiology. Ref Type: Audiovisual Material, Gainsville, Florida.
Lang, P.J., Bradley, M.M., Fitzsimmons, J.R., Cuthbert, B.N., Scott, J.D., Moulder, B., Nangia, V., 1998. Emotional arousal and activation of the visual cortex: an fMRI analysis. Psychophysiology 35, 199–210.
Laux, L., Glanzmann, P., Schaffner, P., Spielberger, C.D., 1981. Das State-Trait-Angstinventar: Theoretische Grundlagen und Handanweisungen. Beltz, Weinheim.
Maddock, R.J., Buonocore, M.H., 1997. Activation of left posterior cingulate gyrus by the auditory presentation of threat-related words: an fMRI study. Psychiatry Res. 75, 1–14.
Maddock, R.J., Garrett, A.S., Buonocore, M.H., 2003. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. Hum. Brain Mapp. 18, 30–41.
Massaro, D.W., Egan, P.B., 1996. Perceiving affect from the voice and the face. Psychon. Bull. Rev. 3, 215–221.
McGaugh, J.L., 2004. The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu. Rev. Neurosci. 27, 1–28.
Mesulam, M.M., 1998. From sensation to cognition. Brain 121 (Pt. 6), 1013–1052.
Mesulam, M.M., Nobre, A.C., Kim, Y.H., Parrish, T.B., Gitelman, D.R., 2001. Heterogeneity of cingulate contributions to spatial attention. NeuroImage 13, 1065–1072.
Morris, J.S., Frith, C.D., Perrett, D.I., Rowland, D., Young, A.W., Calder, A.J., Dolan, R.J., 1996. A differential neural response in the human amygdala to fearful and happy facial expressions. Nature 383, 812–815.
Morris, J.S., Friston, K.J., Buchel, C., Frith, C.D., Young, A.W., Calder, A.J., Dolan, R.J., 1998. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 121 (Pt. 1), 47–57.
Morris, R., Pandya, D.N., Petrides, M., 1999. Fiber system linking the mid-dorsolateral frontal cortex with the retrosplenial/presubicular region in the rhesus monkey. J. Comp. Neurol. 407, 183–192.
Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., Kojima, S., 1999. Activation of the right inferior frontal cortex during assessment of facial emotion. J. Neurophysiol. 82, 1610–1614.
Northoff, G., Richter, A., Gessner, M., Schlagenhauf, F., Fell, J., Baumgart, F., Kaulisch, T., Kotter, R., Stephan, K.E., Leschinger, A., Hagner, T., Bargel, B., Witzel, T., Hinrichs, H., Bogerts, B., Scheich, H., Heinze, H.J., 2000. Functional dissociation between medial and lateral prefrontal cortical spatiotemporal activation in negative and positive emotions: a combined fMRI/MEG study. Cereb. Cortex 10, 93–107.
O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. Nat. Neurosci. 4, 95–102.
Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., Imabayashi, E., 2001. Functional anatomy of musical perception in musicians. Cereb. Cortex 11, 754–760.
Packard, M.G., 1999. Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning. Proc. Natl. Acad. Sci. U. S. A. 96, 12881–12886.
Packard, M.G., Cahill, L., 2001. Affective modulation of multiple memory systems. Curr. Opin. Neurobiol. 11, 752–756.
Peretz, I., Gagnon, L., Bouchard, B., 1998. Music and emotion: perceptual determinants, immediacy, and isolation after brain damage. Cognition 68, 111–141.
Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. Proc. Natl. Acad. Sci. U. S. A. 99, 11458–11463.
Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I., 2002. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. NeuroImage 16, 331–348.
Phillips, M.L., Young, A.W., Senior, C., Brammer, M., Andrew, C., Calder, A.J., Bullmore, E.T., Perrett, D.I., Rowland, D., Williams,

S.C., Gray, J.A., David, A.S., 1997. A specific neural substrate for perceiving facial expressions of disgust. Nature 389, 495–498.

Phillips, M.L., Young, A.W., Scott, S.K., Calder, A.J., Andrew, C., Giampietro, V., Williams, S.C., Bullmore, E.T., Brammer, M., Gray, J.A., 1998. Neural responses to facial and vocal expressions of fear and disgust. Proc. R. Soc. London, Ser. B Biol. Sci. 265, 1809–1817.

Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R., 2003. Neurobiology of emotion perception: I. The neural basis of normal emotion perception. Biol. Psychiatry 54, 504–514.

Pourtois, G., de, G.B., Bol, A., Crommelinck, M., 2005. Perception of facial expressions and voices and of their combination in the human brain. Cortex 41, 49–59.

Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P., 1999. SENSE: sensitivity encoding for fast MRI. Magn. Reson. Med. 42, 952–962.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676–682.

Reiman, E.M., Lane, R.D., Ahern, G.L., Schwartz, G.E., Davidson, R.J., Friston, K.J., 1997. Neuroanatomical correlates of externally and internally generated human emotion. Am. J. Psychiatry 154, 918–925.

Rolls, E.T., 2000. The orbitofrontal cortex and reward. Cereb. Cortex 10, 284–294.

Rolls, E.T., 2004. The functions of the orbitofrontal cortex. Brain Cogn. 55, 11–29.

Rotshtein, P., Malach, R., Hadar, U., Graif, M., Hendler, T., 2001. Feeling or features: different sensitivity to emotion in high-order visual cortex and amygdala. Neuron 32, 747–757.

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.

Sander, K., Brechmann, A., Scheich, H., 2003. Audition of laughing and crying leads to right amygdala activation in a low-noise fMRI setting. Brain Res. Brain Res. Protoc. 11, 81–91.

Simpson, J.R., Ongur, D., Akbudak, E., Conturo, T.E., Ollinger, J.M., Snyder, A.Z., Gusnard, D.A., Raichle, M.E., 2000. The emotional modulation of cognitive processing: an fMRI study. J. Cogn. Neurosci. 12 (Suppl. 2), 157–170.

Stark, R., Schienle, A., Walter, B., Kirsch, P., Blecker, C., Ott, U., Schafer, A., Sammer, G., Zimmermann, M., Vaitl, D., 2004. Hemodynamic effects of negative emotional pictures—A test–retest analysis. Neuropsychobiology 50, 108–118.

Surguladze, S.A., Brammer, M.J., Young, A.W., Andrew, C., Travis, M.J., Williams, S.C., Phillips, M.L., 2003. A preferential increase in the extrastriate response to signals of danger. NeuroImage 19, 1317–1328.

Taylor, S.F., Liberzon, I., Fig, L.M., Decker, L.R., Minoshima, S., Koeppe, R.A., 1998. The effect of emotional content on visual recognition memory: a PET activation study. NeuroImage 8, 188–197.

Taylor, S.F., Liberzon, I., Koeppe, R.A., 2000. The effect of graded aversive stimuli on limbic and visual activation. Neuropsychologia 38, 1415–1425.

Taylor, S.F., Phan, K.L., Decker, L.R., Liberzon, I., 2003. Subjective rating of emotionally salient stimuli modulates neural activity. NeuroImage 18, 650–659.

Teasdale, J.D., Howard, R.J., Cox, S.G., Ha, Y., Brammer, M.J., Williams, S.C., Checkley, S.A., 1999. Functional MRI study of the cognitive generation of affect. Am. J. Psychiatry 156, 209–215.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15, 273–289.

Yeomans, J.S., Frankland, P.W., 1995. The acoustic startle reflex: neurons and connections. Brain Res. Rev. 21, 301–314.

Zatorre, R.J., Evans, A.C., Meyer, E., 1994. Neural mechanisms underlying melodic perception and memory for pitch. J. Neurosci. 14, 1908–1919.

Zung, W.W.K., 1965. A self-rating depression scale. Arch. Gen. Psychiatry 12, 63–70.