Publication II

Viinikainen, M., Jääskeläinen, I.P., Balk, M.H., Autti, T., Sams, M. (2012) Neural processing of emotional valence of facial expressions, *Open Journal of Neuroscience,* 2-3.

Open Journal of Neuroscience

ROSS Science Publishers

Research Article

OPEN ACCESS

Neural processing of emotional valence of facial expressions

Mikko Viinikainen[1*], *Iiro P. Jääskeläinen*[1], *Marja H. Balk*[1,2], *Taina Autti*[2], *Mikko Sams*[1]

[1] Brain and Mind Laboratory, Department of Biomedical Engineering and Computational Science (BECS), School of Science, Aalto University, Espoo, Finland

[2] Department of Radiology, University of Helsinki and HUS Radiology (Medical Imaging Center), Helsinki, Finland

Corresponding Author & Address:

Mikko Viinikainen

Brain and Mind Laboratory, Department of Biomedical Engineering and Computational Science, School of Science, Aalto University, P.O. Box 12200, 00076 AALTO, Finland. Email: mikko.viinikainen@aalto.fi Tel: +358 40 7562053.

Published: 11th June, 2012 Accepted: 11th June, 2012

Received: 26th April, 2012

Open Journal of Neuroscience, 2012, **2**-3

Keywords: Emotion, Valence, Arousal, Face, fMRI, Correlation

ABSTRACT

Degree of emotional valence and arousal have been shown to covary with blood oxygen level dependent (BOLD) signal levels in several brain structures. Here we studied brain activity in 17 healthy subjects during perception of facial expressions varying in valence and arousal using functional magnetic resonance imaging (fMRI). Our results revealed correlations with the perceived valence in dorsolateral and ventrolateral prefrontal cortex, dorsomedial prefrontal cortex, and anterior insula. These findings corroborate results of our previous study where we used pictures of varying valence taken from the International Affective Picture System (IAPS). Together, the results of these two studies suggest existence of common brain areas processing valence of both emotional pictures and facial expressions. Additionally, BOLD signal exhibited distinctive dependency on perceived valence in intraparietal sulcus and supramarginal gyrus in the present study. BOLD activity correlated with negative and positive valence in separate cortical areas, and some areas demonstrated either a U-shaped or an inverted U-shaped relationship with valence (i.e., either minimal or maximal activation was observed to neutral expressions). This nonlinear dependency suggests that brain mechanisms underlying perception of negative and positive valence are at least to some extent independent. Perceived arousal correlated positively with the strength of the BOLD signal only in the left inferior frontal gyrus, which is an important node of the mirror neuron system.

INTRODUCTION

Smooth social interaction requires ability to perceive facial expressions that convey important information about emotional states of others. We quickly recognize pleasure and displeasure from facial expressions, and can also usually infer more specifically how another person is feeling. We can, for instance, perceive when someone is angry or frightened, and the magnitude of the expression informs us about the person's arousal.

Four to five year old children tend to divide facial expressions into good and bad instead of specifying the precise emotional category [1]. When asked to freely label facial expressions, children mainly use the words happy, angry, and sad, and less often words for other basic emotions such as scared, surprised, and disgusted [2]. Furthermore, happiness and sadness are the emotions that are most easily recognized by young children [3]. This lack of specificity at early stages of human development is one piece of evidence that has been taken to suggest that emotional expressions are processed in a rudimentary negative to positive dimension.

Naturally, on a cautious note, these findings could be explained by children's limited emotion vocabulary [4], however, based on factor analysis and multidimensional scaling, it has been suggested that valence (running from negative through neutral to positive emotions) and arousal (running from low to high) constitute orthogonal dimensions that represent the scope of human emotions [5, 6]. Although also other dimensional emotion models do exist (e.g. [7, 8]), the model with valence and arousal as fundamental dimensions is currently the most widely accepted one. The model posits that each emotion can be characterized by a vector in the two-dimensional valence-arousal space (e.g, sadness can be characterized as negative valence with low arousal and angriness, on the other hand, as negative valence and high arousal). The theory has been a very popular one since it gives conceptually separate building blocks of emotions, which are rooted in psychological reality [9]. It has even been suggested that discrete basic emotions (happiness, sadness, surprise, disgust, angriness, and fearfulness) exist only in the perceptual realm and that at the neural level valence is the fundamental processing unit [10].

Supporting the valence-arousal dimensional theory, many organisms, including humans, judge the degree of the positiveness or negativeness of various stimuli. Miller [11, 12] showed that approach-withdrawal tendencies of rats could be manipulated by modulating the closeness of reward and punishment. Emotion literature speaks of positivity offset and negativity bias (for a review, see [13, 14]). The former means that organisms regard their environment generally as slightly positive, and the latter that the gradient for experienced negativity is steeper. Concepts of valence and arousal also provide a useful framework for description of self-experienced subjective affective states [15]. When contemplating the emotional states of others, facial expressions are a key element conveying emotional information. Expressions of unpleasantness (e.g., sad, fearful) are perceived to be of negative valence and expressions of pleasantness (e.g., happy, positively surprised) are perceived to be of positive valence.

The neural basis of evaluation of valence and arousal can be studied using functional magnetic resonance imaging (fMRI). By parametrically varying the valence and arousal of stimuli, it is possible to correlate valence and arousal with blood oxygen level dependent (BOLD) signals of the brain. In previous studies, linear correlations have been calculated between BOLD signal and valence/arousal of olfactory [16, 17], gustatory [18], and visual stimuli [19, 20]. In these studies subjects evaluated valence or arousal of the stimuli and not the emotional feelings induced by the stimuli. In addition, social concepts [21], emotion-denoting words [22], and emotionally evocative sentences [23] have been used as stimuli in previous studies. In these studies, using verbal stimuli, subjects have been instructed to evaluate their own emotional feelings. This difference in study design does not necessarily prevent comparison of the results, since emotional stimuli likely induce at least weak emotions in subjects. These studies have demonstrated linear positive and/or negative correlations between BOLD signal and valence in several prefrontal cortical (PFC) regions and insula. Positive linear correlations between BOLD signal and arousal were further observed in the PFC and amygdala.

Lewis and co-workers, using emotion words [24], and Viinikainen and co-workers, using emotional visual stimuli [25] from the International Affective Picture System (IAPS; [26]), applied non-linear correlation models to study the relationship between BOLD signals and emotional valence and arousal of the stimuli. Their findings suggest that valence is not represented as a single continuum in the human brain, but rather that there are different neural mechanisms sensitive to negative and positive valence. Notably, these findings challenge the concept of unitary bipolar

valence (i.e., one running from negative through neutral to positive) that has been advocated by several researchers [15, 26]. In other words, unpleasantness and pleasantness would not seem to be represented along a single continuum from negative through neutral to positive at the neural level, but rather unpleasantness and its neural representation can vary from neutral to very intense and pleasantness and its neural representation can vary, at least partly independently, from neutral to very intense. Lewis et al. reported that there was a quadratic (U-shaped) dependency between BOLD signal and valence in the posterolateral orbitofrontal cortex (OFC), subgenual cingulate cortex, and anterior cingulate cortex (ACC) [24]. Support for these results was obtained using linear models separately for neutral-to-negative and neutral-to-positive valence.

Viinikainen et al. found inverted U-shaped correlations between valence and BOLD signals bilaterally in the dorsomedial prefrontal cortex (DMPFC), dorsolateral prefrontal cortex (DLPFC), and insular cortex [25]. A multitude of brain regions showed significant positive correlations between BOLD signal and negative-to-neutral valence and/or significant negative correlations between BOLD signal and neutral-to-positive valence (i.e., the activity was the strongest for neutral stimuli). Hence, the results support a fundamental division of brain mechanisms to those evaluating stimuli as pleasant vs. unpleasant. This is consistent with results of a previous study by Grabenhorst et al. [17] where it was observed that components of a mixture of pleasant and unpleasant odor, experienced as pleasant, were processed differently in areas sensitive to negative odors than in areas sensitive to positive odors. Furthermore, neurons responding specifically to either negative valence or positive valence stimuli have been documented in monkey amygdala [27]. Cacioppo and co-workers have suggested bivariate division of valence based on behavioral studies [13, 28]. Results of factor analytic approaches have also suggested that “Negative Affect” and “Positive Affect” are independent [29]. Taken together, these lines of research suggest that there might be distinct negative and positive valence dimensions (i.e. unpleasantness and pleasantness can vary at least partly independently). This contrasts with the view of bipolar valence, where positivity and negativity are considered to be reciprocally dependent [30].

In the present study, we aimed at finding out whether our results obtained with emotional pictures generalize to emotional facial expressions; whether we can see additional evidence suggesting that there are separate positive and negative valence representations in specific brain areas as suggested by our previous study. The paradigm that we used was therefore highly similar with that of Viinikainen et al. [25], but IAPS pictures were replaced by pictures of facial expressions of varying valence and arousal. Valence here refers to the perceived unpleasantness vs. pleasantness of facial expression. Arousal refers to the perceived level of calmness vs. alertness of facial expressions. In other words, the variables of interest were valence and arousal expressed by the facial expressions.

The stimulus presentation time was very brief in our study (100 ms). With short presentation times we tried to emphasize fast emotion-related perceptual processes [31] over cognitive evaluative processes. Short presentation times also enhanced the ecological validity of the experimental setting given that emotional expressions are often highly transient in real life. Our study used a block design paradigm in order to increase statistical power of the experiment. The perceived valence and arousal were modulated between blocks, and these modulations were covaried with the BOLD signal. Importantly, we tested bipolar, quadratic, as well as bivariate models of perceived valence in our study. In particular, we examined the perceived negative and positive valences of the facial expressions separately, which had not been done earlier. By using multiple parametric correlation models it was possible to explore the type of representation of perceived valence in different brain regions. Since earlier studies had shown several types of valence responses, we did not want to restrict our analyses a priori to any specific model.

Based on previous studies using emotionally evocative visual stimuli, we hypothesized that we will find valence-dependent correlations in the dorsal PFC and insula [19, 20, 25]. We further

hypothesized that negative and positive valence are represented by different neural populations, as suggested by findings in our previous IAPS study [25]. Importantly, the experimental paradigm was similar with our previous IAPS study, making the results directly comparable. As an alternative hypothesis, we used a linear bipolar model to test whether our findings would support the bipolar model of emotional valence. Furthermore, given that neural responses to observed categorical facial expressions have been in many studies linked to the functioning of the mirror neuron system [32, 33, 34, 35], we expected that the perceived valence and arousal would correlate with activity in the mirror neuron system (see [36], for a review), in particular the inferior frontal gyrus (IFG) and anterior insular cortex (AIC).

MATERIALS AND METHODS

Subjects

18 subjects (4 females, 14 males, 22–26 years of age) participated in the fMRI study. One male subject was excluded from the analysis due to extensive motion artifacts in the echo planar imaging (EPI) data. All subjects reported being right handed, having normal or corrected-to-normal visual acuity, and having neither neurological nor psychiatric disorders. They were naive to the experiment and were recruited amongst the students of Aalto University. Additionally, one male and one female participated in a pilot study designed for stimulus selection and grouping. The ethical committee of Helsinki and Uusimaa Hospital District for healthy subjects and basic healthcare approved the study, and a voluntary informed consent was obtained from each subject prior to participation.

Stimuli and their evaluation

The stimuli were still photographs of emotional facial expressions of eight professional actors and seven members of the laboratory staff. These persons were instructed on how to make stereotyped expressions and they privately rehearsed making them before the shoot. Detailed characterization of the photographs as well as Facial Action Coding System (FACS) and subject pool evaluations of part of the photographs can be found in [37]. The faces were masked with an oval mask to reduce the availability of other visual cues besides the facial expression and were displayed against a white background. The stimuli included neutral expressions and expressions of six basic emotions – happiness, surprise, anger, disgust, sadness, and fear – as well as combination expressions of happiness and disgust, and happiness and surprise (Fig. 1).

The combination expressions were natural expressions performed by the professional actors, not chimerical faces. Expressions of happiness and disgust as well as happiness and surprise were naturally combined, as such expressions do occur during social interaction. For example, combined happiness and disgust could be expressed when someone sits on a whoopee cushion. Combined happiness and surprise could be expressed, for instance, when one arrives at a surprise birthday party. The combination expressions helped to cover the positive side of valence-arousal space more thoroughly. In addition, they added diversity to the range of positive expressions. As a result there were five different classes of emotional expressions represented on the negative side and three on the positive side. In addition, the expressions of happiness included smiles both with and without teeth revealed. It was more difficult to reach heterogeneity on the positive than negative side of the valence scale. Intensities of the emotions varied, since the still photographs were acquired from different stages of the video clips of expressions. Although personal appearance certainly can have an effect on the perceived valence and arousal, our actors were normal in their appearance (see Fig. 1 for examples) and the main emotional ingredient comes from the facial expression.

Fig. (1). Examples of facial expressions used in the study. From top left to bottom right: happiness, happiness and disgust combined, happiness and surprise combined, surprise, anger, disgust, sadness, and fear. In addition, there were a few neutral expressions.

In a separate behavioral pilot experiment, two subjects rated the valence and arousal of 573 facial expressions on a rating scale 1–9. For the actual experiment 197 of these stimuli were discarded, based on large inter-rater variance, as well as to avoid repetition of the same valence and arousal values. We tried to cover the valence-arousal space as thoroughly as possible. As a result, 376 facial expressions were used in the experiment, divided into 47 eight-stimulus sets. The sets were composed so that each had a specific mean level of valence and arousal, while the within-set variances were kept to minimum. In other words, the valence and arousal were kept constant within each set. This enabled us to study valence- and arousal-related responses in block design with good signal-to-noise ratio. While the division into stimulus sets was made based on the ratings obtained in the pilot experiment, the ratings of the pilot experiment were not used in any other phases of the study. Instead, the 17 subjects who participated in the actual experiment gave their valence and arousal ratings after fMRI scanning. These ratings confirmed that the division into fixed valence and arousal sets had been reliable and valid. Fig. 2a shows the average valences and arousals of the stimulus sets as well as the standard errors within each set. From this figure one can see that the variation in valence and arousal within each stimulus set was small. Furthermore, different sets cover the valence-arousal space relatively well. Distinct emotion categories were not the focus of our study and were thus in most cases mixed within a stimulus set. In our study design, all other effects besides valence and arousal of the facial expressions should average out in the parametric data analysis.

Figure 2a

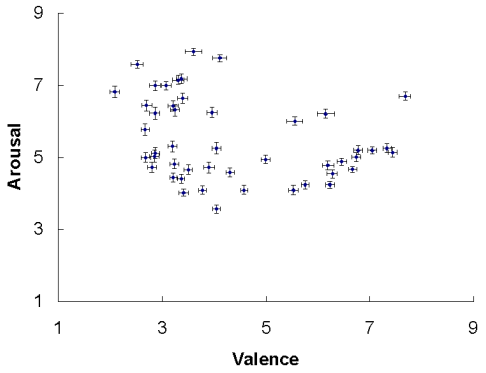

Figure 2b

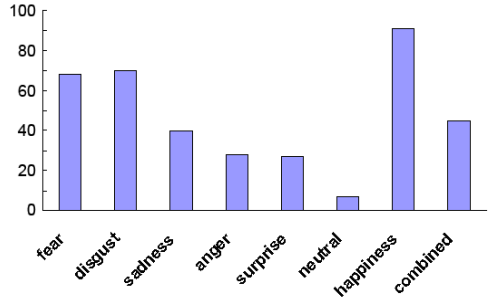

Fig. (2). a) Mean values of perceived valence and arousal for the different stimulus sets (blocks) in the fMRI session, calculated from single stimulus evaluations of the subjects participating in the experiment. Error bars represent standard errors. **b)** Number of different facial expressions used in the experiment. In the "combined" category, there was approximately the same number of combined expressions of happiness and disgust, and happiness and surprise. Approximately 2/3 of the stimuli were of negative valence and 1/3 of positive valence.

The means of valence and arousal evaluations of the 47 stimulus sets, as evaluated by the subjects participating in the fMRI study (see Behavioral measures of perceived valence and arousal), and the distribution of categorized emotions in the stimuli that were shown to the subjects during and after fMRI scanning, are illustrated in Fig. 2. All in all, the stimuli were evaluated in three phases: in the pilot experiment by two volunteers, and both during fMRI scanning and in post-scan behavioral sessions by the subjects who participated in the study. The evaluation times were measured. During fMRI scanning, the subjects provided merely judgments of whether each stimulus block was overall negative or positive. These subject-wise evaluations were used in dividing the fMRI blocks in the analysis phase into negative valence blocks and positive valence blocks. For example, the "neutral" stimulus set (valence ~ 5) was judged as

positive by 14 subjects and as negative by 3 subjects, in a dichotomic forced-choice evaluation.

After the fMRI scanning, the subjects provided evaluations of valence and arousal for each individual stimulus on the scale 1–9 (see Behavioral measures of perceived valence and arousal). These evaluations were averaged per block and used as regressors in the general linear model (GLM). BOLD signal strength was correlated with these evaluations. We wanted to use for each subject his own personal behavioral evaluations, because the perceived valence and arousal of emotional stimuli may differ between individuals. Fig. 2a is based on the averaged valence evaluations from the 17 subjects who underwent the fMRI study. It appeared to be more difficult to cover the positive than the negative side of the valence-arousal space. The figure shows that 15 stimulus sets were positive (valence > 5), 31 sets were negative (valence < 5), and one set was neutral (valence ~ 5). On the basis of the averaged arousal evaluations from the 17 subjects, there were 20 sets with low arousal (3–5), 22 sets with high arousal (5–7), and 5 sets with very high arousal (7–8). All of the sets with very high arousal were of negative valence.

The fMRI paradigm

During fMRI scanning, the facial expressions were presented in sets of 8 pictures. Each expression was presented for 100 ms followed by 1900 ms of white screen. At the end of every picture set the subjects saw a forced-choice evaluation screen that lasted for 6.75 seconds. They were to indicate by a button press whether the emotional states of the persons whose faces were shown were pleasant or unpleasant. The subjects' response times were measured. For all except the most neutral stimulus sets the forced-choice task was easy and its main purpose was to keep the subjects vigilant and attentive during the experiment. Moreover, it has been shown that such an active task enhances activity in the brain areas underpinning perception of facial expressions [38]. The schematic structure of the stimulus presentation is presented in Fig. 3. A single set of pictures could contain several emotion categories (e.g., anger and disgust) and also multiple expressions from the same performer.

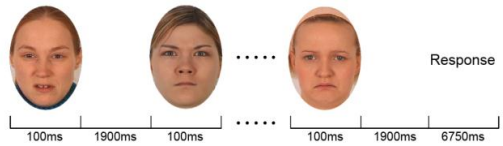

Fig. (3). Example of a single experimental block. Eight facial expressions of 100 ms duration were presented with 1.9-s inter-stimulus intervals, after which a response screen lasting for 6750 ms appeared.

All but one of the negative sets contained expressions from more than one emotion category, typical number of categories being two to three. In the sets with least intensity/arousal, there could be even five different emotion categories (including neutral) within one set. However, with small expression intensities the exact emotion categories are difficult to distinguish and the subjects can typically only perceive whether the expression is negative or positive in nature. For the positive sets there was considerably less heterogeneity, and the number of different emotion categories within each set was typically one or two. There was a significant difference between the number of emotion categories for the negative valence sets and positive valence sets (t-value 5.297, $p < 0.05$). However, within the negative or positive valence sets there was no significant correlation between valence and the number of emotion categories (t-value for negative sets 1.202 and for positive sets 1.450, $p > 0.05$ for both). Arousal did not correlate significantly with the number of emotion categories in a set (t-value 1.000, $p > 0.05$). The number of expressions from the same performer within a set ranged from one to three. This did not show significant variation in relation to valence or arousal, with the correlations being non-significant at 0.139 (t-value 0.934) and 0.173 (t-value 1.160). Each of the 47 picture sets represents a specific mean level of valence and arousal (see Fig. 2a). The order of the sets, and pictures within each set, was randomized for each subject. Before scanning, the subjects were familiarized with the stimulus material with a few random sets, which were also used in the experiment.

Behavioral measures of perceived valence and arousal

Approximately one week after fMRI scanning the subjects were shown again each of

the 376 stimuli in a separate behavioral session, in which they were instructed to evaluate the emotional states (valence and arousal) of the persons whose facial expressions were shown. Each stimulus was evaluated separately. Valence was always evaluated first on a scale ranging from 1 (very unpleasant) through 5 (neutral) to 9 (very pleasant). After this, arousal was evaluated from 1 (soporific) to 9 (highly aroused). Each expression was presented for two seconds on a computer screen, followed by an inquiry. The stimulus presentation time was longer than in the scanner to allow the subjects to clearly observe the facial expressions to enable easy and accurate evaluation. We specifically wanted the experimental procedure to be similar as in our earlier work [25] to make it easier to compare the results. The ratings remained same with different presentation times, as evidenced by non-parametric Spearman correlation between the behavioral ratings of the pictures and the positivity vs. negativity judgments obtained during fMRI scanning ($R = 0.9339$, $p < 0.01$). Subjects were asked to make their evaluations based on the first impression. Evaluation times were measured. The average total evaluation time, excluding picture presentation, was 16 min 4 s for valence (std = 73.3 s) and 10 min 19 s for arousal (std = 27.6s). The mean length of the behavioral experiment was 38 min 55 s. Prior to the behavioral session the subjects saw a nine picture assembly of female face expressing different degrees of valence and arousal in order to familiarize the subjects with the emotion dimensions that were utilized [39] (Emotion Development Lab, Boston, USA, 2006). The subjects also practiced with a few stimuli taken from the actual stimulus material. Due to technical problems one of the subjects did not evaluate sixteen stimuli. These missing values were replaced by mean values of the other subjects.

Scanning procedure and image analyses

The scanning sequence was composed of 47 blocks, where one block consisted of a picture set and a response screen. A 3-Tesla GE Signa MRI device (Milwaukee, WI, USA) was used. Thirteen gradient-echo T2*-weighted EPI volumes were acquired per block using 1750 ms sampling intervals (TE = 32 ms, flip angle = 70 degrees), and twenty-nine 64 x 64 axial images with 3.0 mm slice thickness and 20 cm field of view were collected for each volume. This covered the entire brain apart from the most extreme superior part of the cerebrum and the inferior part of the cerebellum.

Stimuli with a visual angle of 7.9 degrees horizontally and 13.1 degrees vertically were presented to the subjects with a projector via a mirror stationed onto the head coil. The subjects viewed the stimulus sets and evaluated them using a response pad that was placed in their right hand. They were requested to press the right button for positive valence judgment and the left button for negative valence judgment. The subjects were asked to avoid any movements in the scanner apart from the finger movements related to the button presses. Their head was stabilized using unattached, approximately 15x10x2cm, medium hard plates on both sides of the head inside the head coil and tissue paper as padding to soften the support. After the fMRI sequence, subjects' anatomical images were acquired using a T1-weighted gradient echo pulse sequence (FOV = 26 cm, matrix 256 x 256, voxel size = $1mm^3$).

615 EPI volumes were scanned altogether, but the first four ones were discarded to allow for T1-saturation. Total scan time of the fMRI experiment was 17 min 49 s, excluding the four dummy volumes. Data analysis was performed with BrainVoyager software [40]. For every subject the volumes were realigned with respect to the first volume and preprocessed in temporal and spatial domain. Temporally, a linear trend removal was performed and a high-pass filter was used to eliminate frequencies lower than three cycles per time course (cutoff 356.42 s). For spatial filtering we applied a Gaussian kernel with 8 mm full-width-at-half-maximum (FWHM). After these steps the fMRI-volumes were co-registered with the subject's corresponding anatomical image and a Talairach transformation was made.

After subsequent preprocessing steps, the adjusted data comprised of 3D time courses for individual subjects co-registered in Talairach space. After this, we created two first-level models for each individual subject. Multiple parametric models were constructed, because we did not want to restrict the analysis a priori to reveal only one type of correlation. In the first model there was one regressor that measured

mean activation and regressors for the linear and quadratic modulations of perceived valence, as well as for the linear modulation of perceived arousal (predictors in the model were in this order). In the second model there were two conditions, negative valence and positive valence. In the negative valence condition there was one regressor that measured mean activation and regressors for the linear modulations of perceived valence and arousal, but the model was restricted to negative valence stimuli. In the positive valence condition there was one regressor that measured mean activation and regressors for the linear modulations of perceived valence and arousal, but the model was restricted to positive valence stimuli. The predictors were calculated using subjects' individual valence and arousal evaluations that were averaged over a block. In other words, mean values per block of perceived valence and arousal were used. In the case of the second first-level model only negative or positive blocks were used, respectively, to create the predictors. Subjects' personal judgments in the scanner about the negativity/positivity of each block were used to assign each of the blocks to the respective predictors. Hence, the number of negative as well as positive blocks was unequal between subjects. Block length for the first level analysis was calculated from the beginning of the presentation of the first stimulus to the end of the presentation of the response screen. The predictors were normalized and orthogonalized as well as convolved with a standard hemodynamic response function (HRF) to account for the delay of the BOLD response with respect to stimulus presentation.

Second-level analysis was performed using random effects (RFX) analysis with a %-transformation. All subjects (apart from the one male with motion contamination in his data) were grouped, so that condition betas from each subject were entered in the RFX analysis. The first RFX model modeled perceived valence (linear and quadratic) and perceived arousal modulations for the whole valence scale and the second RFX model perceived valence and perceived arousal modulations separately for negative and positive valences, as was the case with the first level models. Voxel-wise t-scores with 16 degrees of freedom were calculated for the group analysis, the statistics were thresholded using $p < 0.001$ uncorrected, and a 135 voxel minimum cluster size constraint (corresponding to 5 original 3x3x3 mm voxels) was applied to control for problems associated with multiple comparisons. The statistical analysis was chosen to be similar as in our earlier study [25] for easier comparison of results across the two studies. There had been variation in the valence and arousal responses across earlier studies [16, 17, 18, 19, 20, 21, 22, 23, 24, 25], which could be due to differences in experimental designs and analysis procedures. Hence, we aimed to have as much in common with our earlier IAPS study as possible in order to make comparison across these two studies straightforward.

For areas showing highest correlations with perceived negative valence and/or positive valence, we created weighted correlation models. These were used to characterize the modulation of BOLD signal as function of perceived negative and positive valence. The data points for the regression plots (see Figures 4 and 5 in the Results) were determined by the BOLD signal (y-axis) and subjects' behavioral valence evaluations (x-axis) as in the earlier RFX analysis. Values averaged over all subjects were used. The percent signal change (PSC) values for every block (depicted on the y-axis) were obtained by using the average signal over the whole experiment as a baseline and by weighting each block with the standard HRF function (scaled so that the sum of block weights equaled one). This way, the hemodynamic response characteristics were taken into account. It should be noted that due to the baseline "activations" and "deactivations" exist only relative to the average signal within each area. In other words, we could have as well used arbitrary units on the y-axis in the graphs. To be able to plot single linear regression curves (linear curve for perceived negative valence and linear curve for perceived positive valence) to this data, we needed to use aggregate values also for the regression weights. The weights for the regression plots were calculated from the subjects' block-wise online evaluations (i.e., whether a block was rated as negative or positive). Most blocks were judged as either negative or positive by all subjects, and these received weights one and zero for the respective weighted correlation models. However, the most neutral block (valence closest to 5), for example, received weights 3/17 and

14/17, since it was evaluated as negative by 3 subjects and positive by the rest. This means that this block contributed more to the linear model of positive valence but also to lesser extent to the linear model of negative valence. Weighting data points in this manner resulted in more accurate model than what would have been obtained simply by dividing blocks into negative and positive by their average perceived valence (threshold at 5).

RESULTS AND OBSERVATIONS

Reaction times measured during fMRI scanning as well as during the behavioral session after the scanning showed no significant (linear or quadratic) correlations with either perceived valence or arousal, indicating that the degree of difficulty in inferring valence and arousal from the various facial expressions was quite similar. However, mean reaction time to one fMRI block was 1.71 s, being 2.82 standard deviations above the the mean reaction time of 1.37 s ($p < 0.01$). Mean valence and arousal ratings of this block were 4.58 and 4.09, respectively. Because longer reaction time suggests that the evaluation task for this block was more difficult than for the other ones, we analysed the fMRI results both with and without this block (see below). Statistically significant correlations with arousal were present for U-shaped (quadratic) valence ($R = 0.418$, $p < 0.05$), negative valence ($R = 0.319$, $p < 0.05$) and positive valence ($R = 0.416$, $p < 0.05$) (see also Fig. 2a). Therefore, it was important that valence and arousal were orthogonalized when used as predictors.

The results from our statistical parametric analyses are shown in Table 1. We did not find any brain areas showing significant linear correlations between BOLD signal and perceived valence calculated across all expressions. However, there were multiple brain areas exhibiting linear correlations between BOLD data and the valence ratings when these were calculated separately for the positive and negative facial expressions. For negative expressions, BOLD activity decreased from neutral to unpleasant in: 1) right intraparietal sulcus/supramarginal gyrus (IPS/SMG), 2) right middle frontal gyrus (MFG) in dorsolateral prefrontal cortex (DLPFC), 3) left anterior insular cortex (AIC), 4) left inferior frontal gyrus (IFG), and 5) left MFG in ventrolateral prefrontal cortex (VLPFC). BOLD activity increased from neutral to unpleasant (perception) in the left retrosplenial cortex (RSC) and left postcentral sulcus. For positive expressions, BOLD activity decreased from neutral to pleasant bilaterally in 1) IPS/SMG, 2) MFG in DLPFC/VLPFC, 3) AIC, 4) superior frontal gyrus (SFG) in DMPFC, and 5) in the left precentral sulcus. Only in one area, left angular gyrus, BOLD activity increased from neutral to pleasant.

Table 1) List of significantly activated brain regions in the valence correlation analyses

		negative valence					positive valence					U-shaped valence				
Brain region	**L/R**	**x**	**y**	**z**	**size (mm^3)**	**t-value**	**x**	**y**	**z**	**size (mm^3)**	**t-value**	**x**	**y**	**z**	**size (mm^3)**	**t-value**
Intraparietal sulcus / Supramarginal gyrus	R	33	-61	49	627	4.548610	36	-61	43	12140 *	-7.332612	39	-49	46	870	-4.564910
DLPFC: middle frontal gyrus	R	49	29	28	199	4.631243	48	32	31	15979 *	-6.021723	48	29	28	510	-4.414343
Retrosplenial cortex	L	-9	-55	10	199	-4.671389						-12	-50	16	304	4.709254
Postcentral sulcus	L	-54	-19	40	157	-4.764987										
Inferior frontal gyrus	L	-54	20	1	1249 *	5.167893										
Anterior insular cortex	L	-30	20	1	1249 *	5.146802	-30	20	10	1108	-5.407875					
VLPFC: middle frontal gyrus	L	-33	56	7	376	4.665880	-30	47	10	175	-4.755190					
Intraparietal sulcus / Supramarginal gyrus	L						-33	-55	37	6001	-6.407336					
DLPFC: middle frontal gyrus	L						-39	26	25	8957 *	-5.879458					
Middle temporal gyrus	R						54	-43	7	187	-4.301093					
Anterior insular cortex	R						39	26	4	15979 *	-5.815949					
VLPFC: middle frontal gyrus	R						33	53	10	1380	-5.217274					
DLPFC/VLPFC: middle/inferior frontal gyrus	L						-45	17	37	8957 *	-7.133788					
DLPFC/VLPFC: middle/inferior frontal gyrus	R						51	20	37	15979 *	-5.614558					
DLPFC/VLPFC: inferior frontal gyrus	L						-51	14	28	8957 *	-5.149184					

DLPFC/VLPFC: inferior frontal gyrus	R						54	20	25	15979 *	-5.314412					
Angular gyrus	L						-39	-79	28	669	5.117026	-42	-79	22	141	4.548374
DMPFC: superior frontal gyrus	L/R						-6	14	49	4604	-6.657107					
Precentral gyrus	L						-33	-1	49	8957 *	-4.996247					
Precentral sulcus	L						-30	5	31	8957 *	-5.840447					
Precuneus	R						9	-73	40	12140 *	-5.727124					
Caudate nucleus	L						-18	5	19	535	-4.446515					
Fornix	R						3	-10	16	268	-4.524065					
Middle temporal gyrus	R						57	-49	-8	461	-4.456576					

Table 1. Significantly activated brain regions in the correlation analyses for perceived valence: hemisphere (L/R), Talairach coordinates (peak activity), cluster size (mm3), and peak t-values. Negative valence indicates linear correlation between BOLD signal and perceived valence with negative faces, positive valence with positive faces, and U-shaped valence shows the second-order correlation across all faces. Results are thresholded using $p < 0.001$, uncorrected, cluster size = 135. Positive t-values denote positive correlation and negative t-values negative correlation. No significant linear correlation across all faces emerged in the analysis. DLPFC = dorsolateral prefrontal cortex, VLPFC = ventrolateral prefrontal cortex, DMPFC = dorsomedial prefrontal cortex. * Multiple regions belong to the same cluster.

Figures 4 and 5 summarize brain areas showing most significant correlations with negative and positive perceived valences. These figures show statistical parametric maps of valence-dependent brain regions (perceived negative valences in Fig. 4 and perceived positive valences in Fig. 5) and regional group-averaged PSC data as a function of perceived valence. In the PSC plots, the y-axis shows the average fMRI-response from a certain stimulus block and the x-axis shows the mean valence rating for that block. Coloring of the data points refers to the percentage of subjects classifying the block as negative (more red) or positive (more green). Red weighted regression lines are fitted to the group averaged BOLD signal changes for expressions evaluated as negative. Green weighted regression lines are fitted to the group averaged BOLD signal changes for expressions evaluated as positive (see Methods: Scanning procedure and image analyses). As can be seen, the linear trend is in all areas opposite for the negative and positive perceived valences. There were no statistically significant ($p > 0.05$) differences between the squared weighted correlations of negative valence model and positive valence model in any of the brain regions.

Figure 4 a) **Figure 4 b)**

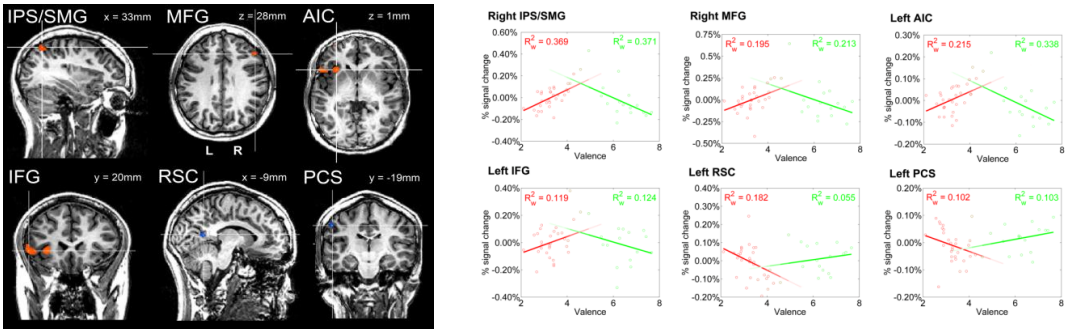

Fig. (4). a) Statistical maps from the correlation analysis between perceived valence and BOLD signal with linear regression restricted to negative faces (portrayed in neurological convention). Positive correlations are in orange and negative correlations in blue; threshold $p < 0.001$, uncorrected, cluster size = 135. **b)** Regional percent signal changes of presented blocks as a function of perceived valence. The circles representing different blocks have been colored green if (the majority of) subjects have classified the block as positive. Respectively, the circles have been colored red if (the majority of) subjects have classified the block as negative. The classifications were obtained online during fMRI scanning and the valence ratings in a post-scan behavioral session. Red straight fading lines represent a weighted regression fit taking into account those blocks evaluated to be negative by one (weight 1/17) or more (weight N/17) subjects. Green straight fading lines represent a weighted regression fit taking into account those blocks evaluated to be positive. R^2-values show the corresponding weighted correlations squared. IPS/SMG = right intraparietal sulcus/supramarginal gyrus; MFG = right middle frontal gyrus; AIC = left anterior insular cortex; IFG = left inferior frontal gyrus, RSC = left retrosplenial cortex, PCS = left postcentral sulcus.

Figure 5 a)

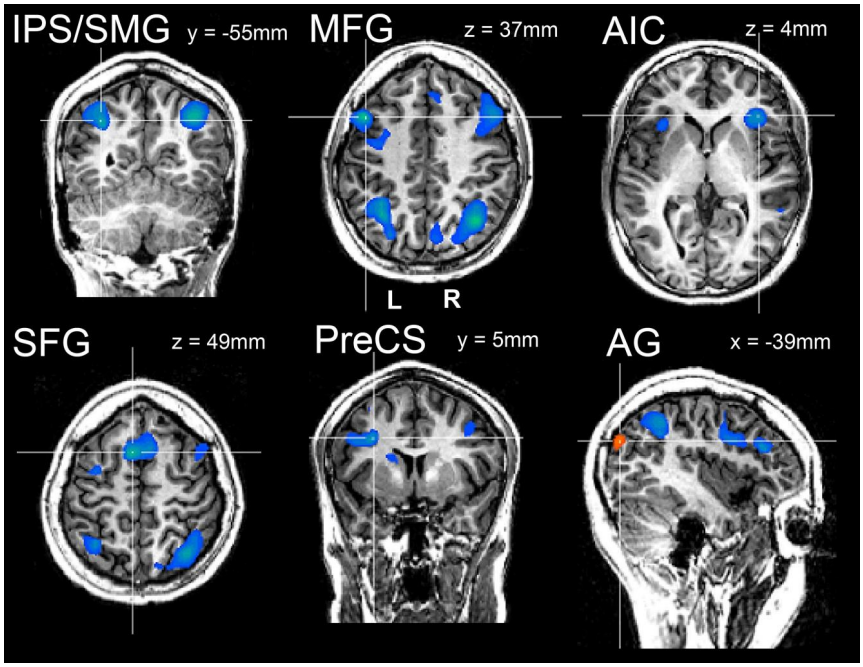

Figure 5 b)

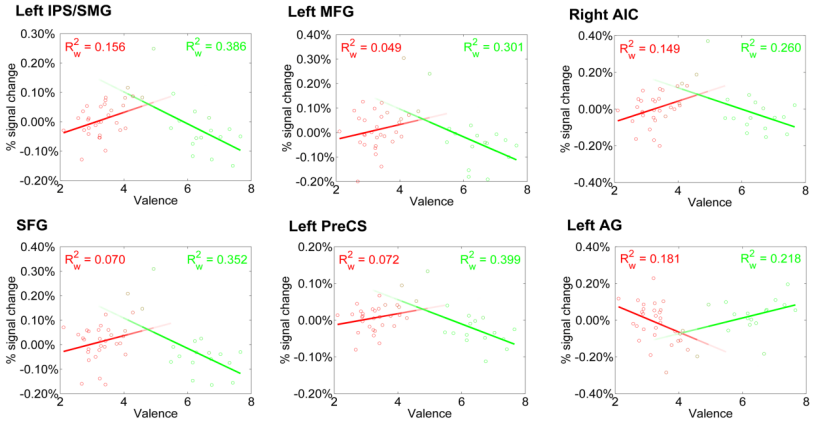

Fig. (5). a) Statistical maps from the correlation analysis between perceived valence and BOLD signal with the linear regression restricted to positive faces (portrayed in neurological convention). Positive correlations are shown in orange and negative correlations in blue; threshold p < 0.001, uncorrected, cluster size = 135. **b)** Regional percent signal changes across the presented blocks as function of perceived valence. The circles representing different blocks have been colored green when the majority of subjects classified the block as positive. Respectively, the circles have been colored red when the majority of subjects classified the block as negative. The classifications have been given online during the fMRI measurement and the valence ratings in a post-scan behavioral session. The red straight fading lines represent a weighted regression fit taking into account those blocks that were evaluated as negative by one (weight 1/17) or more (weight N/17) subjects. The green straight fading lines represent a weighted regression fit taking into account those blocks that were evaluated as positive. R^2-values show the corresponding weighted correlations squared. IPS/SMG = left intraparietal sulcus/supramarginal gyrus MFG = left middle frontal gyrus; AIC = right anterior insular cortex; SFG = left superior frontal gyrus; PreCS = left precentral sulcus; AG = left angular gyrus.

We also found significant non-linear, inverted U-shaped, correlation between BOLD signal and perceived stimulus valence across all expressions in the right IPS/SMG and right MFG in

DLPFC (Fig. 6). In these regions, significant correlations with perceived valence were also found separately for negative and positive stimuli. Thus, significant effects were in fact found with all but the linear modulation of perceived valence across all facial expressions. On the other hand, the left angular gyrus and left RSC showed an upright U-shaped correlation.

Figure 6 a) **Figure 6 b)**

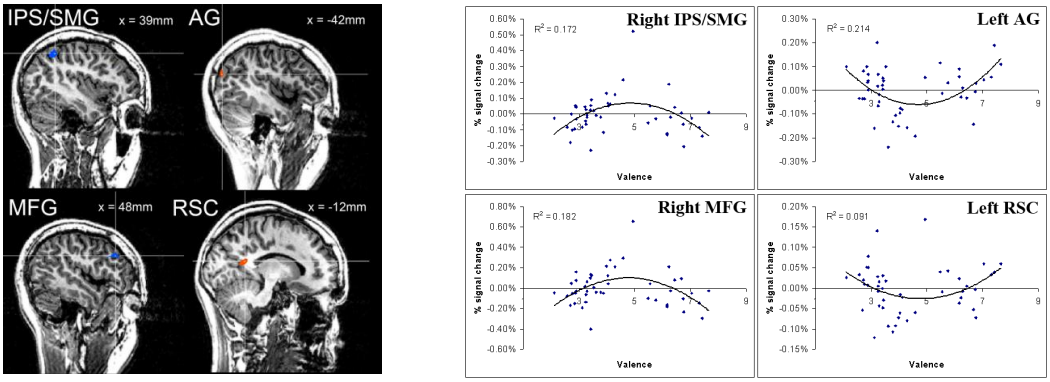

Fig. (6). a) Statistical maps from the correlation analysis of perceived valence and BOLD signal with the second order non-linear (U-shaped) variation. Positive correlations are shown in orange and negative correlations in blue; threshold $p < 0.001$, uncorrected, cluster size = 135. **b)** Shown are scatter-plot diagrams of the second-order correlations of the regions of interest. The diagrams show how each brain region is non-linearly activated as a function of perceived valence around the locus of peak activation. R^2-values show the multiple correlation squared, depicting how well variations in the BOLD signal level can be explained by the variations in perceived valence. IPS/SMG = right intraparietal sulcus/supramarginal gyrus, MFG = right middle frontal gyrus, AG = left angular gyrus, RSC = left retrosplenial cortex.

Linear correlations were also calculated between the BOLD signals and perceived arousal. This was done both with all faces included as well as separately with negative valence and positive valence faces. With $p < 0.001$ and a 135 voxel cluster size threshold, there was a significant correlation between perceived arousal and BOLD signal strength across all faces, independent of valence evaluations, in the left IFG (peak Talairach coordinates -54, 17, 22; cluster size 184 mm3; peak t-value 4.498). There were no significant correlations across only negative or positive faces between BOLD signal and perceived arousal, which of course might have been due to reduced power when dividing the sample.

Two stimulus blocks, the only ones including neutral expressions, were difficult to categorize as either negative or positive. Half of the subjects categorized these blocks as negative and the other half as positive. Furthermore, in one of these two blocks, the response times of the categorizations were significantly above average. For all other blocks the subjects agreed that a given block is negative/positive. Across the seventeen participants, the forced-choice evaluations were distributed as follows: 17/0 (i.e., 17 negative, 0 positive) for 17 blocks, 0/17 (i.e., 0 negative, 17 positive) for 11 blocks, 16/1 for 8 blocks, 1/16 for 3 blocks, 15/2 for 2 blocks, 2/15 for 1 block, 14/3 for 1 block, 3/14 for 1 block, 13/4 for 1 block, 10/7 for 1 block, 9/8 for 1 block. Since the difficulty of the task could have affected the BOLD signals during the two ambiguous blocks, we reanalysed the data with these blocks excluded. This analysis did not affect the significant correlations between the BOLD responses and perceived negative valence or positive valence. Further, the results did not change for the inverted U-shaped dependency findings. However, the upright U-shaped relationship in left angular gyrus fell below statistical threshold in this analysis, and the response in the left RSC did not survive the cluster size constraint. Hence, the two neutral blocks must have elicited lower activity compared to most other blocks in these regions, which can be observed also in the scatter plots in Fig. 6 (note that the block with valence ratings closest to 5 was not amongst the two ambiguous neutral blocks, but these both had valence ratings slightly below 5, i.e. 4.58 and 4.12). Excluding the two ambiguous neutral blocks did not change the results for arousal.

DISCUSSION

Correlation of BOLD activity with perceived valence

Our results revealed several brain areas that showed correlations between BOLD activity and perceived valence of facial expressions, however, BOLD activity in none of these areas correlated linearly with the whole negative-through-neutral-to-positive valence scale. This deviates from observations of many previous studies [16, 17, 18, 19, 20, 21, 22, 23], which could be possibly explained by differences during perception vs. experiencing of emotions. These previous studies did not examine the perceived valence, but rather either experienced valence or valence of the stimuli. In the present study, BOLD activity showed linear parametric relationship with perceived valence from neutral to most negative faces in some areas, and in other areas with perceived valence from neutral to most positive faces. Our results provide additional neurophysiological support for the view that valence is a fundamental factor describing emotions (for a review, see e.g. [10]) and, further, that negative and positive valences have at least partly separate underlying brain mechanisms (see also [17, 24, 25]). These results are also in line with single cell recordings, which have shown separate cells responding to negative valence and positive valence in monkey amygdala [27]. In that study, negative valence was induced in monkeys with an air puff at the face and positive valence with juice administration.

For negative expressions the correlations with BOLD responses were both negative and positive, although in most brain regions the correlations were positive (i.e., the BOLD activity diminished towards more negative expressions). Findings of reciprocal activations have been earlier reported with IAPS pictures in the prefrontal cortex (PFC) [41]. In the current study, the reciprocity was also manifested as U-shaped vs. inverted U-shaped perceived valence dependencies in the left RSC and left angular gyrus vs. in the right IPS/SMG and right MFG in DLPFC. However, only the latter correlations remained significant when two ambiguous neutral blocks were excluded from the analysis. For positive facial expressions the BOLD signals consistently diminished towards more positive stimuli, with the only exception being the left angular gyrus.

Britton and co-workers [42] used both IAPS pictures and facial expressions as stimuli and found that both activated the amygdala, posterior hippocampus, VMPFC, and the visual cortex. The important difference compared to our study was that they investigated emotion categories (happy, sad, anger, fear, and neutral) and classical contrast activations, not correlations with the subjects' ratings. However, in line with these observations, our present results are in many respects similar to those that we previously found using IAPS pictures as stimuli [25]. In both of our studies the general trend has been toward positive correlation between BOLD responses and negative valence stimuli and negative correlation between BOLD responses and positive valence stimuli (see Table 1 and Figs. 4b and 5b; [25]). Furthermore, the correlations with IAPS pictures and facial expressions seem to occur at least partly in the same brain areas (see Table 2). Given that the stimulus presentation times and the experimental paradigms were identical in our studies, this pattern of results suggests a general mechanism for processing negative and positive valences, respectively, for a variety of emotionally evocative visual stimuli including facial expressions. These similarities might be explained by emotional contagion caused by the facial expressions; merely seeing an emotional expression activates the same areas as when one experiences these emotions [32, 35]. For example, it has been shown that a patient with severed insula and putamen is unable to not only experience disgust but also to recognize it in others [43]. Moreover, it has been shown that activation of areas such as the insula during emotion perception also leads to feeling of emotions [44, 45]. Emotional contagion can be important also in evaluation of the stimuli [46].

Table 2 summarizes similarities and differences between the current study and our previous IAPS study. In both studies, the subjects evaluated valence and arousal of the stimuli, but the current study is a special case of emotional perception, since the object is another person's face. Activity in the DLPFC (dorsolateral part of the MFG) showed positive correlation with valence for negative faces in the right hemisphere and negative correlations for positive faces bilaterally.

In other words, the activity was the highest for neutral faces and diminished towards both negative and positive faces. These correlations are similar to what we found using IAPS pictures in our previous study. Activity in the right DLPFC (MFG) showed inverted U-shape correlation with valence across all facial expressions as well as IAPS pictures. In our previous study, we found similar activation pattern also in the left DLPFC, but in the current study this correlation failed to reach significance (although with a more lenient thresholding of $p < 0.005$ uncorrected significant correlation would have been observed).

Table 2) Similarities and differences in the results with facial expressions and IAPS pictures

Brain region	Facial expressions	IAPS pictures
DLPFC	left: –, right: +/–/∩	left: –/∩, right: +/–/∩
DMPFC	left: –, right: –	left: –/∩ , right: –/∩
VLPFC	left: +/–, right: –	left: +, right: +
VMPFC	no activations	left: +, right: +
IPS/SMG	left: –, right: +/–/∩	no activations
AIC	left: +/–, right: –	left: +/–/∩, right: +/–/∩
RSC	left: –/U	no activations
Caudate nucleus	left: –	left: –/∩, right: –
Amygdala	no activations	right: +

Table 2. Similarities and differences of the results of the present study and our recent IAPS study. Left and right denote the hemispheres. Red minus signs denote negative correlation and plus signs positive correlation with negative valence. Green minus signs denote negative correlation and plus signs positive correlation with positive valence. ∩-signs denote negative second-order correlation and U-signs positive second-order correlation with valence for all pictures. Similar trends can be seen in both experiments. DLPFC = dorsolateral prefrontal cortex, DMPFC = dorsomedial prefrontal cortex, VLPFC = ventrolateral prefrontal cortex, VMPFC = ventromedial prefrontal cortex, IPS = intraparietal sulcus, SMG = supramarginal gyrus, AIC = anterior insular cortex, RSC = retrosplenial cortex.

Activity in the DMPFC (SFG) showed significant negative correlations with valence, but only for positive faces. This is again similar to what we found using IAPS pictures. Earlier studies have reported positive linear correlation in the DMPFC across the whole valence range [19] (IAPS stimuli), and both positive [20] (IAPS stimuli) and negative [47] (facial expressions) correlations between BOLD signal and arousal. These contradictory results suggest that while DMPFC seems to be involved in emotional processing, the DMPFC responses may depend on the specific task and experimental setup. Further research is required to elucidate the role of DMPFC in valence and arousal processing.

Integration of emotion and cognition has been suggested to occur in the DLPFC [41, 48], and the DMPFC has been suggested to be important for the cognitive generation of emotional states [49]. Thus, it is likely that the valence perception mechanism in the dorsal prefrontal cortex (dPFC) does not reflect merely bottom-up processing, but that the dPFC participates in cognitive labeling of stimuli as being negative or positive to various degrees.

BOLD activity in the left VLPFC showed positive correlation with negative stimuli, and for positive stimuli there were negative correlations bilaterally. In our earlier study [25], we found positive correlations for negative-valence stimuli bilaterally and no significant correlations for positive stimuli. In a meta-analysis, the VLPFC has been found to be involved in evaluation of emotional stimuli but not evaluation of other's emotions [50]. This suggests that unlike a multitude of other regions (see below), the VLPFC does not participate in mirroring of emotions. Contrary to our previous study, we did not find significant correlations for positive stimuli in the VMPFC. This somewhat contrasts with a previous report of conjoint activations in this region elicited by faces and IAPS pictures [42], but there classical contrast activations were examined instead of parametric correlations. The observed differences in ventral PFC are possibly related to differences in power to elicit emotions, with IAPS pictures possibly being more emotionally evocative than facial expressions.

Activity in the IPS/SMG showed right-hemispheric positive correlation with perceived valence for negative stimuli and bilateral negative correlations for positive stimuli. Across all stimuli, there was an inverted U-shape correlation. These findings provide novel evidence of the role of IPS/SMG in the processing of perceived emotional valence of in particular facial expressions. The IPS has been found to be involved in interpreting goals behind other people's actions, and it has been suggested to be a part of the mirror neuron system [51]. Valence relates to goal-directed, approach vs. withdrawal behavior. Thus, the IPS can be sensitive to valence-based like vs. dislike cues exhibited by facial expressions that help to

interpret goal-directed behavior of others. In other words, facial expressions may convey information about another person's intentions, and the IPS can function through a valence-based mechanism when interpreting these intentions.

The insular cortex, especially its anterior part, has been reported to be involved in emotional processing in several studies (for a review, see [52]). In most studies, its activity has been connected to feeling or perceiving negative emotions, such as fear [53], unpleasant taste [18], and pain [34]. In the present study, the left AIC showed positive correlation with perceived valence of negative faces, but there were also significant negative correlations with perceived valence of positive faces bilaterally. Similar results were observed using IAPS pictures in our previous study. The role of insular cortex in processing positive in addition to negative emotions is supported also by the results of Lewis and co-workers [24]. Thus, evidence is accumulating that the insular cortex is involved in the processing of positive as well as negative valence. Wicker et al. suggest that insular cortex might contain "emotional mirror neurons", which activate both during subjective feelings and third person perception of emotion [35]. Such mechanism might explain the similarity of results obtained in our previous IAPS experiment and the current study.

BOLD responses in the left RSC (i.e., the ventral bank of the posterior cingulate cortex) correlated negatively with perceived valence for negative stimuli but there were no correlations with positive-valence stimuli. In addition, the left RSC demonstrated a U-shaped correlation with perceived valence. These results are consistent with a meta-analysis, which revealed that especially unpleasant and salient stimuli activate RSC [54]. RSC was suggested to be an important but underrated structure in the processing of emotional salience, a property typically corresponding to U-shaped function of valence.

Comparison with earlier parametric facial expression study

Although there has been research on conjoint brain activation patterns elicited by different emotional facial expressions [55], there is only one previous study, that of Gerber and co-workers [47], that has investigated the neural basis of valence and arousal during perception of facial expressions. Their subjects watched individual facial expressions for 18 seconds during fMRI scanning and evaluated online the valence and arousal of the expressions. The valence and arousal ratings were then correlated with the BOLD signal time courses, using both a valence continuum running from negative through neutral to positive, as well as absolute valence (i.e., distance from neutral). Valence ratings correlated negatively with BOLD signal in the left parietal cortex and dorsal anterior cingulate cortex. Arousal ratings correlated negatively with BOLD signal in the amygdala, thalamus, cerebellum, and dorsal pons. Absolute valence ratings correlated positively with BOLD signal in the same regions, with the exception of the thalamus. Since arousal typically increases towards the negative and positive ends of the valence dimension, the observed opposite correlations are not easy to explain. The authors proposed a salience concept to explain their results, admitting the contradiction between their arousal findings and earlier literature.

The results of the present study and those of Gerber and co-workers [47] differ considerably. Contrary to the latter study, we did not find linear correlation with perceived valence in any of the brain regions. Correlations reported in Gerber et al. [47] between BOLD responses and absolute valence should be to some extent comparable with our quadratic (U-shaped) correlations, however, we did not find correlations in similar brain areas. Gerber and co-workers did not analyze negative and positive valence dependencies separately. It is possible that differences in stimulus presentation times explain these differences in findings. In our study, the stimulus presentation was 0.1 s per picture and in Gerber et al. [47] it was 18 s per picture; there are studies indicating differences in how the emotional circuitry of the brain processes stimuli that are presented very briefly vs. for long enough to allow conscious evaluation of stimuli [56]. Additionally, the different results could stem from differences in the evaluation tasks during fMRI imaging. Whereas our subjects had an easy dichotomical (negative vs. positive) judgment task in the scanner for the whole block, in the earlier study the subjects had to evaluate each stimulus online for their valence and arousal. There were

also significant differences in the data analysis of the two studies. Normalization or orthogonalization was not applied to the predictors in Gerber et al. study [47], whereas we applied both of these conventional procedures.

Alternative explanations for the results

Even though we found brain activations, which were gradually dependent on the perceived valence of the stimuli, it is important to consider other possible explanations for the results. In the following, we consider whether emotional intensity or increased processing demands due to stimulus ambiguity could explain them.

In most brain regions, BOLD signals increased towards neutral facial expressions. This phenomenon could be explained by a single variable, emotional intensity, which increases towards both negative and positive ends of the valence spectrum. Decrease in emotional intensity could increase BOLD signal. However, emotional intensity is a more plausible explanation in regions where BOLD signal increased towards negative and positive facial expressions. On the basis of the present experiment, we cannot rule out this explanation. Nevertheless, emotional intensity does not explain valence-dependent BOLD signal in those areas, where the dependence was found only for negative or only for positive valences.

The close-to-neutral stimuli in our experiment were probably more ambiguous than the clearly emotional expressions, since they included neutral expressions as well as expressions with weak intensity. Therefore, the evaluation task in the scanner might have been somewhat more demanding for the neutral stimuli and require more effort. For one neutral block the evaluation times were, indeed, significantly above average. However, our results with negative valence and positive valence models did not change when the neutral blocks were excluded from the analysis. The same applies to the inverted U-shaped model. Hence, task difficulty does not seem to be a probable explanation for significant activations with these models. For the upright U-shaped model, the statistically significant effects vanished, both in the left angular gyrus and RSC. The signal was weaker in these regions during the ambiguous blocks, so the U-shaped effect could be possibly explained by the task difficulty suppressing default mode activation in these regions [57, 58]. However, removal of the two blocks also reduced the statistical power of the analysis, and this caused the U-shaped effects to become non-significant. The combination expressions can also be considered more ambiguous than others, but they were distributed quite evenly amongst the positive valence blocks, and were not included in the blocks where valence was close to neutral. Moreover, they did not influence reaction times during scanning.

Correlation of BOLD activity with perceived arousal

BOLD responses in the left IFG showed significant positive correlation with evaluated arousal across all facial expressions. IFG is an essential part of the mirror-neuron system [36]. In monkeys and humans, it contains neurons that mirror communicative mouth and facial movements [36, 59, 60]. Therefore, it is possible that the observed arousal-related activation is characteristic to the mirror neuron system. BOLD signal strengths in IFG have been found to correlate with the strength of the pain reflected in others' facial expressions [34]. One interpretation of our present findings is that IFG activation is related to the difference in expression intensity between emotional and neutral facial expression. The arousal scale was not covered as well as valence scale; very low arousal ratings did not exist at all, since our expressions did not include expressions of sleepiness, tiredness and fatigue. It cannot be ruled out that the restricted range in arousal led to fewer significant correlations.

Parametric vs. categoric processing of facial expressions

Behavioral experiments have shown that facial expressions are processed simultaneously both as basic emotion categories and in a parametric fashion [61]. Thus, it is important to study both processing mechanisms to understand how facial expressions are processed – also in the brain. The current experiment was designed to reveal the neural underpinnings of the parametric processing, which could be more instantaneous and rudimentary way to process faces. The results show that parametric processing of facial expressions indeed exists also on the neural level. Parametric processing could be utilized, for

example, in instances, where information about the facial expression is scarce, such as in low illumination (cf. [62, 63]) and with faces that are far away (cf. [64]), corresponding to low spatial frequencies [65], or possibly with faces that are moving fast.

Lack of correlations in the amygdala and fusiform gyrus

Activity in both amygdala and fusiform gyrus failed to correlate with valence and arousal, although they are well-known to participate in the processing of emotional faces [42, 66, 67] and emotion dimensions [16, 68, 69]. One possibility is that the amygdala and fusiform gyrus do not process perceptual valence or arousal of faces in parametric fashion but instead as categories. Fusiform activations have been found for classes of negative valence faces and pictures and positive valence faces and pictures [69]. In similar vein, amygdala activations for negative faces and positive faces have been observed [55]. On the other hand, using a quite similar paradigm as ours Gerber and co-workers found correlations with both recognized arousal (negative correlation) and absolute valence (positive correlation) in the amygdala complex [47].

Methodological remarks

Feeling of an emotion after exposure to a stimulus, and evaluating valence of this stimulus are two different aspects of emotion processing. For example, ratings of facial expressions based on emotions induced by seeing the expressions vs. based on perceived emotional properties may differ (see [70]). Here, we specifically studied brain mechanisms underlying perceptual evaluation of facial expressions. However, perception often, if not always, induces emotional feelings, and facial expressions have shown to be emotionally contagious (see for instance [46]). Subjects of the present study evaluated the facial expressions similarly as subjects in our earlier study evaluated IAPS pictures [25], making the results of the two studies directly comparable.

CONCLUSION

Our present results suggest that perception of facial expressions involves brain mechanisms that are tuned to detecting the valence of the facial expressions. These mechanisms may be different for perceived negative and positive valence, supporting the view that there are independent unipolar negative and positive valence dimensions [13, 28, 29]. Significant valence dependencies in the dorsal PFC, VLPFC, and insula partly coincided with those found in our earlier study using IAPS pictures as stimuli [25]. This suggests a common network of areas that is activated when evaluating valence of various types of visual stimuli. Arousal-related activations in the IFG are possibly due to a mechanism evaluating the magnitude of the facial expression, with stronger expressions (relative to neutral expression) resulting in stronger arousal.

ABBREVIATIONS

ACC – anterior cingulate cortex

AIC – anterior insular cortex

BOLD – blood oxygen level dependent

DLPFC – dorsolateral prefrontal cortex

DMPFC – dorsomedial prefrontal cortex

dPFC – dorsal prefrontal cortex

EPI – echo planar imaging

FACS – Facial Action Coding System

fMRI – functional magnetic resonance imaging

FOV – field of view

FWHM – full-width-at-half-maximum

GLM – general linear model

HRF – hemodynamic response function

IAPS – International Affective Picture System

IFG – inferior frontal gyrus

IPS – intraparietal sulcus

MFG – middle frontal gyrus

OFC – orbitofrontal cortex

PFC – prefrontal cortex

PSC – percent signal change

RFX – random effects

RSC – retrosplenial cortex

SFG – superior frontal gyrus

SMG – supramarginal gyrus

std – standard deviation

TE – echo time

VLPFC – ventrolateral prefrontal cortex

VMPFC – ventromedial prefrontal cortex

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interests.

ACKNOWLEDGEMENTS

This study was supported by the Academy of Finland grants 2103470, 130412, 138145, 2103938, the Graduate School in Electronics, Telecommunications and Automation (GETA), and Nokia Research Center.

REFERENCES

[1] Russell JA, Bullock M. Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *J Pers Soc Psychol*. **1985**; 48: 1290-8.

[2] Widen SC, Russell JA. A closer look at preschoolers' freely produced labels for facial expressions. *Dev Psychol*. **2003**; 39: 114-28.

[3] Camras LA, Allison K. Children's understanding of emotional facial expressions and verbal labels. *J Nonverbal Behav*. **1985**; 9: 84-94.

[4] Widen SC, Russell JA. Children acquire emotion categories gradually. *Cogn Dev*. **2008**; 23: 291-312.

[5] Osgood CE, Suci GJ, Tannenbaum PH. *The measurement of meaning*; University of Illinois Press: Urbana, **1957**.

[6] Russell JA. A Circumplex Model of Affect. *J Pers Soc Psychol*. **1980**; 39: 1161-78.

[7] Larsen RJ, Diener E. *Promises and problems with the circumplex model of emotion*; Sage: Newbury Park, CA, **1992**.

[8] Thayer RE. *The biopsychology of mood and activation*; Oxford University Press: New York, **1989**.

[9] Reisenzein R. Pleasure-activation theory and the intensity of emotions. *J Pers Soc Psychol*. **1994**; 67: 525-39.

[10] Barrett LF. Valence is a basic building block of emotional life. *J Res Pers*. **2006**; 40: 35-55.

[11] Miller NE. Reaction formation in rats: an experimental analog for a Freudian phenomenon. *Psychol Bull*. **1937**; 34: 724.

[12] Miller NE. *Liberalization of basic S-R concepts: Extensions to conflict behavior, motivation and social learning*; McGraw-Hill: New York, **1959**.

[13] Cacioppo JT, Gardner WL, Berntson GG. Beyond bipolar conceptualizations and measures: the case of attitudes and evaluative space. *Pers Soc Psychol Rev*. **1997**; 1: 3-25.

[14] Cacioppo JT, Gardner WL, Berntson GG. The Affect System Has Parallel and Integrative Processing Components: Form Follows Function. *J Pers Soc Psychol*. **1999**; 76: 839-55.

[15] Russell JA, Barrett LF. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *J Pers Soc Psychol*. **1999**; 76: 805-19.

[16] Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani DG, Glover G, Gabrieli JD, Sobel N. Dissociated neural representations of intensity and valence in human olfaction. *Nat Neurosci*. **2003**; 6: 196-202.

[17] Grabenhorst F, Rolls ET, Margot C, da Silva MA, Velazco MI. How pleasant and unpleasant stimuli combine in different brain regions: odor mixtures. *J Neurosci*. **2007**; 27: 13532-40.

[18] Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T. Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*. **2003**; 39: 701-11.

[19] Heinzel A, Bermpohl F, Niese R, Pfennig A, Pascual-Leone A, Schlaug G, Northoff G. How do we modulate our emotions? Parametric fMRI reveals cortical midline structures as regions specifically involved in the processing of emotional valences. *Brain Res Cogn Brain Res*. **2005**; 25: 348-58.

[20] Grimm S, Schmidt CF, Bermpohl F, Heinzel A, Dahlem Y, Wyss M, Hell D, Boesiger P, Boeker H, Northoff G. Segregated neural representation of distinct emotion dimensions in the prefrontal cortex-an fMRI study. *Neuroimage*. **2006**; 30: 325-40.

[21] Cunningham WA, Raye CL, Johnson MK. Implicit and explicit evaluation: FMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *J Cogn Neurosci*. **2004**; 16: 1717-29.

[22] Posner J, Russell JA, Gerber A, Gorman D, Colibazzi T, Yu S, Wang Z, Kangarlu A, Zhu H, Peterson BS. The neurophysiological bases of emotion: An fMRI study of the affective circumplex using emotion-denoting words. *Hum Brain Mapp*. **2009**; 30: 883-95.

[23] Colibazzi T, Posner J, Wang Z, Gorman D, Gerber A, Yu S, Zhu H, Kangarlu A, Duan Y, Russell JA, Peterson BS. Neural systems subserving valence and arousal during the experience of induced emotions. *Emotion*. **2010**; 10: 377-89.

[24] Lewis PA, Critchley HD, Rotshtein P, Dolan RJ. Neural correlates of processing valence and arousal in affective words. *Cereb Cortex*. **2007**; 17: 742-8.

[25] Viinikainen M, Jääskeläinen IP, Alexandrov Y, Balk MH, Autti T, Sams M. Nonlinear relationship between emotional valence and brain activity: evidence of separate negative and positive valence dimensions. *Hum Brain Mapp*. **2010**; 31: 1030-40.

[26] Lang PJ, Bradley MM, Cuthbert BN, International affective picture system (IAPS): Affective Ratings of Pictures and Instruction Manual, University of Florida, Gaineswille, FL, 2005.

[27] Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. **2006**; 439: 65-870.

[28] Cacioppo JT, Berntson GG. Relationship Between Attitudes and Evaluative Space: A Critical Review, With Emphasis on the Separability of Positive and Negative Substrates. *Psychol Bull*. **1994**; 115: 401-23.

[29] Watson D, Tellegen A. Toward a consensual structure of mood. *Psychol Bull*. **1985**; 98: 219-35.

[30] Russell JA, Carroll JM. On the bipolarity of positive and negative affect. *Psychol Bull*. **1999**; 125: 3-30.

[31] Bar M, Neta M, Linz H. Very first impressions. *Emotion*. **2006**; 6: 269-78.

[32] Carr L, Iacoboni M, Dubeau MC, Mazziotta JC, Lenzi GL. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc Natl Acad Sci U S A*. **2003**; 100: 5497-502.

[33] Hennenlotter A, Schroeder U, Erhard P, Castrop F, Haslinger B, Stoecker D, Lange KW, Ceballos-Baumann AO. A common neural basis for receptive and expressive communication of pleasant facial affect. *Neuroimage*. **2005**; 26: 581-91.

[34] Saarela MV, Hlushchuk Y, Williams AC, Schurmann M, Kalso E, Hari R. The compassionate brain: humans detect intensity of pain from another's face. *Cereb Cortex*. **2007**; 17: 230-7.

[35] Wicker B, Keysers C, Plailly J, Royet JP, Gallese V, Rizzolatti G. Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*. **2003**; 40: 655-64.

[36] Rizzolatti G, Craighero L. The mirror-neuron system. *Annu Rev Neurosci*. **2004**; 27: 169-92.

[37] Kätsyri J, Human recognition of basic emotions from posed and animated dynamic facial expressions (http://lib.tkk.fi/Diss/2006/isbn951228538X/), Helsinki University of Technology, 2006.

[38] van der Gaag C, Minderaa RB, Keysers C. Facial expressions: what the mirror neuron system can and cannot tell us. *Soc Neurosci*. **2007**; 2: 179-222.

[39] Emotion Development Lab. http://www2.bc.edu/~russeljm/ (accessed February 1, **2006**). Used with permission from Russell, J.A., 2006.

[40] Goebel R, Esposito F, Formisano E. Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp*. **2006**; 27: 392-401.

[41] Northoff G, Heinzel A, Bermpohl F, Niese R, Pfennig A, Pascual-Leone A, Schlaug G. Reciprocal modulation and attenuation in the prefrontal cortex: an fMRI study on emotional-cognitive interaction. *Hum Brain Mapp*. **2004**; 21: 202-12.

[42] Britton JC, Taylor SF, Sudheimer KD, Liberzon I. Facial expressions and complex IAPS pictures: common and differential networks. *Neuroimage*. **2006**; 31: 906-19.

[43] Calder AJ, Keane J, Manes F, Antoun N, Young AW. Impaired recognition and experience of disgust following brain injury. *Nat Neurosci*. **2000**; 3: 1077-8.

[44] Bastiaansen JA, Thioux M, Keysers C. Evidence for mirror systems in emotions. *Philos Trans R Soc Lond B Biol Sci*. **2009**; 364: 2391-404.

[45] Jabbi M, Swart M, Keysers C. Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage*. **2007**; 34: 1744-53.

[46] Wild B, Erb M, Bartels M. Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Res*. **2001**; 102: 109-24.

[47] Gerber AJ, Posner J, Gorman D, Colibazzi T, Yu S, Wang Z, Kangarlu A, Zhu H, Russell J, Peterson BS. An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia*. **2008**; 46: 2129-39.

[48] Gray JR, Braver TS, Raichle ME. Integration of emotion and cognition in the lateral prefrontal cortex. *Proc Natl Acad Sci U S A*. **2002**; 99: 4115-20.

[49] Kober H, Barrett LF, Joseph J, Bliss-Moreau E, Lindquist K, Wager TD. Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage*. **2008**; 42: 998-1031.

[50] Lee KH, Siegle GJ. Common and distinct brain networks underlying explicit emotional evaluation: a meta-analytic study. *Soc Cogn Affect Neurosci*. **2009** Mar 6. [Epub ahead of print].

[51] Hamilton AF, Grafton ST. Goal representation in human anterior intraparietal sulcus. *J Neurosci*. **2006**; 26: 1133-7.

[52] Murphy FC, Nimmo-Smith I, Lawrence AD. Functional neuroanatomy of emotions: a meta-

analysis. *Cogn Affect Behav Neurosci*. **2003**; 3: 207-33.

[53] Morris JS, Friston KJ, Buchel C, Frith CD, Young AW, Calder AJ, Dolan RJ. A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*. **1998**; 121 (Pt 1): 47-57.

[54] Maddock RJ. The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends Neurosci*. **1999**; 22: 310-6.

[55] Yang TT, Menon V, Eliez S, Blasey C, White CD, Reid AJ, Gotlib IH, Reiss AL. Amygdalar activation associated with positive and negative facial expressions. *Neuroreport*. **2002**; 13: 1737-41.

[56] Cunningham WA, Johnson MK, Raye CL, Chris Gatenby J, Gore JC, Banaji MR. Separable neural components in the processing of black and white faces. *Psychol Sci*. **2004**; 15: 806-13.

[57] Fransson P. How default is the default mode of brain function? Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*. **2006**; 44: 2836-45.

[58] Vann SD, Aggleton JP, Maguire EA. What does the retrosplenial cortex do? *Nat Rev Neurosci*. **2009**; 10: 792-802.

[59] Sekiyama K, Kanno I, Miura S, Sugita Y. Auditory-visual speech perception examined by fMRI and PET. *Neurosci Res*. **2003**; 47: 277-87.

[60] Ojanen V, Möttönen R, Pekkola J, Jääskeläinen IP, Joensuu R, Autti T, Sams M. Processing of audiovisual speech in Broca's area. *Neuroimage*. **2005**; 25: 333-8.

[61] Fujimura T, Matsuda YT, Katahira K, Okada M, Okanoya K. Categorical and dimensional perceptions in decoding emotional facial expressions. *Cogn Emot*. **2011;** 26: 587-601.

[62] Feng W, Luo W, Liao Y, Wang N, Gan T, Luo YJ. Human brain responsivity to different intensities of masked fearful eye whites: an ERP study. *Brain Res*. **2009**; 1286: 147-54.

[63] Whalen PJ, Kagan J, Cook RG, Davis FC, Kim H, Polis S, McLaren DG, Somerville LH, McLean AA, Maxwell JS, Johnstone T. Human amygdala responsivity to masked fearful eye whites. *Science*. **2004**; 306: 2061.

[64] Winston JS, Vuilleumier P, Dolan RJ. Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. *Curr Biol*. **2003**; 13: 1824-9.

[65] Loftus GR, Harley EM. Why is it easier to identify someone close than far away? *Psychon Bull Rev*. **2005**; 12: 43-65.

[66] Pegna AJ, Khateb A, Lazeyras F, Seghier ML. Discriminating emotional faces without primary visual cortices involves the right amygdala. *Nat Neurosci*. **2005**; 8: 24-25.

[67] Vuilleumier P, Pourtois G. Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*. **2007**; 45: 174-94.

[68] Anders S, Eippert F, Weiskopf N, Veit R. The human amygdala is sensitive to the valence of pictures and sounds irrespective of arousal: an fMRI study. *Soc Cogn Affect Neurosci*. **2008**; 3: 233-43.

[69] Geday J, Gjedde A, Boldsen AS, Kupers R. Emotional valence modulates activity in the posterior fusiform gyrus and inferior medial prefrontal cortex in social perception. *Neuroimage*. **2003**; 18: 675-84.

[70] Sato W, Yoshikawa S, Kochiyama T, Matsumura M. The amygdala processes the emotional significance of facial expressions: an fMRI investigation using the interaction between expression and face direction. *Neuroimage*. **2004**; 22: 1006-13.