# Neural correlates of emotional valence for faces and words

Daniela Ballotta[1]*, Riccardo Maramotti[1], Eleonora Borelli[2], Fausta Lui[1] and Giuseppe Pagnoni[1]

[1]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy, [2]Department of Medical and Surgical, Maternal-Infantile and Adult Sciences, University of Modena and Reggio Emilia, Modena, Italy

Stimuli with negative emotional valence are especially apt to influence perception and action because of their crucial role in survival, a property that may not be precisely mirrored by positive emotional stimuli of equal intensity. The aim of this study was to identify the neural circuits differentially coding for positive and negative valence in the implicit processing of facial expressions and words, which are among the main ways human beings use to express emotions. Thirty-six healthy subjects took part in an event-related fMRI experiment. We used an implicit emotional processing task with the visual presentation of negative, positive, and neutral faces and words, as primary stimuli. Dynamic Causal Modeling (DCM) of the fMRI data was used to test effective brain connectivity within two different anatomo-functional models, for the processing of words and faces, respectively. In our models, the only areas showing a significant differential response to negative and positive valence across both face and word stimuli were early visual cortices, with faces eliciting stronger activations. For faces, DCM revealed that this effect was mediated by a facilitation of activity in the amygdala by positive faces and in the fusiform face area by negative faces; for words, the effect was mainly imputable to a facilitation of activity in the primary visual cortex by positive words. These findings support a role of early sensory cortices in discriminating the emotional valence of both faces and words, where the effect may be mediated chiefly by the subcortical/limbic visual route for faces, and rely more on the direct thalamic pathway to primary visual cortex for words.

KEYWORDS

fMRI, emotion, facial expression, faces, words, valence, dynamic causal modeling

## 1. Introduction

The ability to recognize other people's emotions is crucial for effective social interactions, where emotional content is predominantly conveyed by facial expressions and words (when words are spoken, rather than read, an essential additional factor is the tonal features of the utterance). For basic emotions, the generation and processing of facial expressions of emotions has been shown to be universal (Nelson et al., 1979; Adolphs, 2002; Cowen et al., 2021; but see different views in, e.g., Hoemann et al., 2019; Berent et al., 2020), with a processing advantage for emotional faces compared to neutral ones (Johansson et al., 2004; Gross and Schwarzer, 2010). More specifically, a prioritization of negative stimuli has been inferred by faster processing times (Fox et al., 2000) and a stronger effect of attentional capture (Alpers and Gerdes, 2007; Bach et al., 2014). An enhancement of perceptual encoding has been observed not only for those stimuli that are recognizable at a very young age (Nelson et al., 1979), but also for stimuli whose

affective valence is learned later in life, such as emotional words (Kanske and Kotz, 2007; Kousta et al., 2009; Yap and Seow, 2014; Goh et al., 2016).

Over the past two decades, the neuronal substrates of emotional processing have represented an intensely researched area in the field of affective neurosciences. Even though functional magnetic resonance imaging (fMRI) studies have demonstrated that the processing of facial and linguistic expressions is supported by two different distributed neural networks (Haxby et al., 2002; Price, 2012), the existence of a common brain substrate underlying general emotional processing has also been hypothesized. In fact, an enhancement of blood oxygen level-dependent (BOLD) signal in visual cortices has been reported for various types of emotional, as opposed to non-emotional, stimuli, including faces (Vuilleumier and Schwartz, 2001; Reisch et al., 2020), pictures of complex scenes (Aldhafeeri et al., 2012) and words (Herbert et al., 2009; Schlochtermeier et al., 2013; Hoffmann et al., 2015).

Nonetheless, while the emotion-related effect of faces in visual processing areas is well supported by the literature, the importance of early sensory cortices for emotion-related word processing is more controversial. An electroencephalographic (EEG) study showed that even when emotion is irrelevant for the task (e.g., in a simple face-word discrimination task), the sensory encoding of emotional content is automatically enhanced for face but not for word stimuli (Rellecke et al., 2011). Similarly, a robust overlapping activation pattern in the extrastriate visual cortex was reported in meta-analytic fMRI studies for pictorial emotional stimuli of different types, but this overlap did not extend to words (Sabatinelli et al., 2011; García-García et al., 2016). A recent fMRI study, directly comparing negative faces, pictures, and words, confirmed the relevance of extrastriate visual areas for faces and pictures, and of left-lateralized frontal and parietal semantic processing areas for lexical stimuli (Reisch et al., 2020).

On the other hand, the importance of early sensory brain cortices for affective experience has been claimed by several studies suggesting the existence of a sensory, modality-specific representation of emotional valence (Satpute et al., 2015; Miskovic and Anderson, 2018). Besides the involvement of early sensory cortices, specific emotion effects for faces were also reported in the "core network" of face processing (Haxby et al., 2000, 2002), including fusiform gyrus, superior temporal sulcus, and inferior occipital cortex (Sabatinelli et al., 2011).

Emotion processing has also been linked to the activity of subcortical brain structures. Contemporary theories ascribe a key role to the amygdala in contextually evaluating and integrating a variety of sensory informations, "tagging" them with appropriate values of emotional dimensions (valence, intensity, and approachability; for a review see Šimić et al., 2021). The amygdala has traditionally been implicated in the modulation of sensory responses to emotional stimuli, particularly fearful ones, and it is known to be anatomically connected to visual cortical regions (LeDoux, 1996; Gschwind et al., 2012). In fact, visual emotional signals appear to be processed by two parallel pathways, both involving the amygdala, albeit at different processing stages: a cortical "high road," and a subcortical "low road." The first route encompasses the thalamic lateral geniculate nucleus, the striate cortex, and the amygdala, and it allows fine-grained, but slow, evaluations of the stimuli. In the second, subcortical route, visual information proceeds to the superior colliculus before being relayed to the amygdala *via* the pulvinar, and it allows for a fast, but coarse,

analysis of the visual input, including potential threat (LeDoux, 1996). A meta-analysis of fMRI studies has demonstrated the involvement of the amygdala for both positive and negative facial expressions (Fusar-Poli et al., 2009), confirming its relevance in the processing of affective content, regardless of valence. The amygdala is also implicated in emotional word processing. It has been shown to rapidly respond to emotional words at an early stage of reading (200 ms after stimulus onset; Naccache et al., 2005), similarly to the occipitotemporal visual word form area (Gaillard et al., 2006), which mediates orthographic processing (Cohen et al., 2000). Finally, a functional coupling between the amygdala and the extrastriate cortex was described in both the left (Herbert et al., 2009) and the right hemisphere (Tabert et al., 2001) during reading of affective words.

Recent studies have applied Dynamic Causal Modeling (DCM) (Friston et al., 2003) — a Bayesian statistical framework to assess effective connectivity in the brain — to examine directional influences within the face processing network (Fairhall and Ishai, 2007; Torrisi et al., 2013; Jamieson et al., 2021). Initially, a hierarchical structure of the "core" and "extended" face networks was proposed, where the "core" network (including mainly visual areas) was hypothesized to exert a feed-forward, bottom-up influence on the "extended" network (prefrontal and limbic regions), considered to be responsible for emotional and social aspects (Fairhall and Ishai, 2007; Da Silva et al., 2011; Torrisi et al., 2013; Sladky et al., 2015). On the other hand, a top-down, feedback effect of the amygdala on visual areas (e.g., occipital face area and fusiform gyrus) was reported by Furl et al. (2013), suggesting that the amygdala mediates an optimization of visual processing depending on the emotional valence of the stimulus. An affective modulation of bidirectional connections between frontal and subcortical structures has also been described. Willinger et al. (2019) showed a modulation of top-down connectivity between the medial prefrontal cortex and the amygdala by positive and negative faces, whereas the bottom-up connectivity between the same regions was modulated by negative and neutral faces only. An increased bidirectional connectivity between frontal areas and the amygdala was also observed during an emotional learning task involving the pairing of emotional words and pictures (Ćurčić-Blake et al., 2012); the modulation was associated with emotional (in particular, negative) stimuli. In summary, as suggested by the review of Underwood et al. (2021), DCM findings suggest a model of emotional processing characterized by dynamic and largely bidirectional modulatory relationships between cortical and subcortical regions.

The aim of the present study was to better understand the functional organization of emotion processing networks, identifying brain circuits coding for valence in the implicit processing of visual stimuli. We chose facial expressions and words as common instances of largely innate and culturally acquired stimuli, respectively (for a recent proposal arguing against this dichotomy, see, however, Feldman Barrett, 2017). The mechanisms by which emotional stimuli may trigger different behavioral reactions (e.g., related to the basic dimensions of approach and avoidance) and attentional movements are particularly relevant for social interactions. Several studies suggest that the neural pathways processing innate and acquired negative valence, as well as the associated avoidance behaviors, may be different (Kim and Jung, 2006; Isosaka et al., 2015; Silva et al., 2016; Kong and Zweifel, 2021), and that a hierarchical organization would prioritize innate over learned negative valence (Isosaka et al., 2015; Kong and Zweifel, 2021). However, central amygdala circuits seem to be critical

in the expressions of both of them, although by different mechanisms (Kong and Zweifel, 2021). On the basis of the known literature, we hypothesized a significant involvement of early visual cortices and amygdala, in addition to face- and word-specific processing areas, and we used DCM together with Bayesian model selection (Parametric Empirical Bayes for DCM, PEB; Zeidman et al., 2019a) to investigate the directional connectivity structure among these regions.

# 2. Materials and methods

## 2.1. Participants

Thirty-six volunteers (21 women; mean age ± SD: 28.0 ± 5.08; range: 18–38) took part in the fMRI study. Inclusion criteria were: being native speakers of the Italian language, right-handedness as assessed by the Edinburgh Inventory (Oldfield, 1971), no history of psychiatric or neurological disease, and no current use of psychoactive medications. The sample size was chosen according to published guidelines for fMRI experiments in healthy volunteers (Friston, 2012). The study was conducted according to the 2013 version of the Declaration of Helsinki and had been approved by the local Ethics Committee (protocol number: CE 134/2014/SPER/AOUMO), with all subjects giving their written informed consent before taking part in the study. Since one participant's data were discarded due to excessive movement during the MR scanning session, the final sample for all the analyses included 35 subjects.

## 2.2. Stimuli

Faces and words were presented on an MRI-compatible display. Faces included photographs of happy (i.e., positive-valenced, $n = 36$), angry (i.e., negative-valenced, $n = 36$), and neutral ($n = 36$) facial expressions performed by 18 female and 18 male models (Radke et al., 2017). In order to reach the required number of trials, which was greater than the number of unique photographs, 12 faces of each category were presented twice (the stimuli to be repeated were randomly selected for each subject). Hair and non-facial contours were masked out in order to minimize their influences, and pictures were balanced for brightness and contrast values. They were presented in black and white on a black background to prevent color from influencing participants' responses. Linguistic stimuli included positive ($n = 18$), negative ($n = 18$), and neutral ($n = 18$) words. Words were selected from the Italian version of the ANEW database (Affective Norms for English Words; Montefinese et al., 2014). As they were carefully balanced for the main distributional, psycholinguistic, and affective features known to affect the time it takes to encode a word, the number of suitable stimuli was limited, therefore 12 items of each category were presented twice, and 6 were presented three times, in order to match the number of face stimuli. Which items were presented two or three times was randomly defined within each subject. More specifically, words were balanced for length in letters, frequency, imageability, and concreteness. Positive and negative words were also balanced for squared valence and arousal (Table 1). Linguistic stimuli were presented in white lowercase Arial font on a black background.

## 2.3. Procedure

An event-related fMRI paradigm was employed. Participants read the task instructions before entering the scanner. Once inside, a two-button response pad was fixed under their right hand. They were asked to keep their gaze fixed on the center of the screen throughout the experiment. Each trial started with a white fixation cross on a black background followed by a stimulus (face or word). Each stimulus was presented within a thin white frame whose left or right side was of a darker (gray) shade and remained on the screen for 2 s (see Figure 1). Participants had to push the left button with their index finger or the right button with their middle finger if the gray sidebar appeared to the left or to the right of the stimulus, respectively. The inter-stimulus interval ranged pseudo-randomly between 2 and 18.8 s. Each participant performed four functional imaging runs; each run consisted of 72 trials, of which 36 faces (12 happy, 12 angry, 12 neutral) and 36 words (12 positive, 12 negative, 12 neutral). The AFNI (Cox, 1996; Cox and Hyde, 1997) *make_random_timing.py* function[1] was used to simulate a series of randomized timing sequences for the trials of each stimulus category, from which the sequence with the best statistical power for the effects of interest was then identified with the *3dDeconvolve* program.[2] Within each category, stimuli were presented in a pseudo-random order, with the constraint that no more than three consecutive stimuli belonging to the same class could occur. Two passive rest blocks were included at the beginning and at the end of each session (range 22.1–24.3 s). Each functional run lasted 8 min and the MRI session included 4 of them.

Participants performed a few practice trials inside the scanner before the experiment started. E-Prime 3.0 software (Psychology Software Tools, Pittsburgh, PA) was used to present the stimuli *via* the ESys functional MRI System[3] remote display, and to collect behavioral responses.

At the end of the experiment, and outside the scanner, participants were asked to rate all the experimental stimuli for their valence and arousal on the Self-Assessment Manikin scale (SAM) (Bradley and Lang, 1994). The two rating questionnaires were delivered and completed on an Excel spreadsheet displayed on a tablet.

## 2.4. Behavioral data analyses

The average reaction times (RT) for the four conditions of interest (positive and negative faces, positive and negative words) were calculated for each volunteer, and the effect of valence within stimulus type (Faces, Words) was assessed with Wilcoxon signed-rank tests because of data non-normality. The average post-scanning ratings of Valence and Arousal for positive and negative faces and words were compared using paired *t*-tests.

---

1  https://afni.nimh.nih.gov/pub/dist/doc/program_help/make_random_timing.py.html

2  https://afni.nimh.nih.gov/pub/dist/doc/htmldoc/programs/3dDeconvolve_sphx.html

3  http://www.invivocorp.com

TABLE 1 Descriptive statistics (mean±SD) of length in letters, frequency of use, imageability, and concreteness, for positive, negative, and neutral words, and of valence and arousal for positive and negative words.

|  | Positive words | Negative words | Neutral words | $F(2,48)$ | $t(34)$ | $p$ value |
|---|---|---|---|---|---|---|
| Length | 7.22 (±2.39) | 6.89 (±2) | 7.06 (±1.89) | 0.11 | – | 0.89 |
| Frequency | 19,305.89 (±25,084.41) | 11,578.56 (±30,717.37) | 34,431.78 (±69,249.36) | 1.15 | – | 0.32 |
| Imageability | 7.49 (±0.95) | 6.93 (±0.94) | 7.46 (±1.40) | 1.46 | – | 0.24 |
| Concreteness | 6.25 (±1.69) | 6.46 (±1.19) | 7.29 (±1.73) | 2.24 | – | 0.12 |
| Valence | 8.28 (±0.25) | 1.82 (±0.23) | – | – | −1.27[§] | 0.21[§] |
| Arousal | 6.32 (±0.74) | 6.63 (±0.75) | – | – | 1.26 | 0.22 |

F- t- and p-value for testing differences of the means are reported in the last three columns. [§]For valence, t- and p-values refer to a test performed on the squared deviations of the valence scores from the center (i.e., 5) of the scale. t-values are reported for valence and arousal, while for the other features F-values are reported.



Push LEFT    Push RIGHT

NEUTRAL    month

POSITIVE    gift

NEGATIVE    war
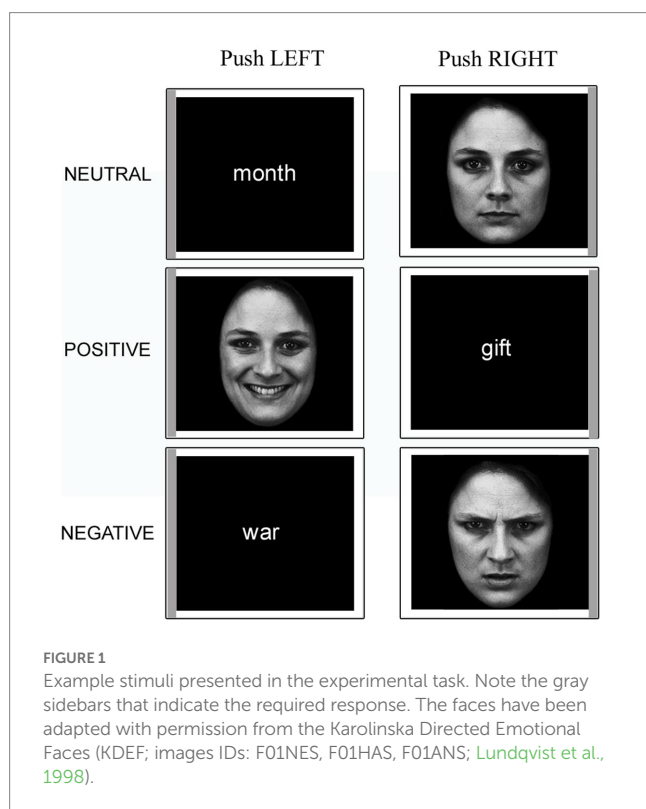
FIGURE 1
Example stimuli presented in the experimental task. Note the gray sidebars that indicate the required response. The faces have been adapted with permission from the Karolinska Directed Emotional Faces (KDEF; images IDs: F01NES, F01HAS, F01ANS; Lundqvist et al., 1998).

## 2.5. fMRI data acquisition and preprocessing

MRI data were acquired with a GE SIGNA Architect 3.0 T MRI scanner and, for each volunteer, included 4 functional runs (gradient-echo echo-planar sequence, 46 axial slices, TR = 1,500 ms, voxel size: 3 × 3 × 2.7 mm with a 0.3 mm gap), and a high-resolution T1-weighted anatomical image (344 sagittal slices, TR = 2184.9 ms, TE = 3.09 ms, voxel size = 1 × 1 × 1 mm).

fMRI data were preprocessed using MATLAB version R2021a (The MathWorks Inc., Natick, Mass) and SPM12 (Wellcome Department of Imaging Neuroscience, London, United Kingdom). Functional volumes were slice-time corrected and realigned to the first volume acquired. The T1-weighted image was coregistered to the mean functional image and segmented using standard SPM tissue probability maps. The estimated deformation fields warp parameters

from subject to MNI (Montreal Neurologic Institute) space were used to normalize the functional volumes to the MNI template implemented in SPM12. Finally, the functional data were smoothed with a 6x6x6 mm FWHM Gaussian kernel.

## 2.6. First- and second-level General Linear Model analyses of fMRI data

First-level (single-subject) analyses were performed as a multiple linear regression with six explanatory variables corresponding to the four stimulus classes of interest: positive faces (*Faces.pos*), negative faces (*Faces.neg*), positive words (*Words.pos*), negative words (*Words. neg*), neutral faces, and neutral words. Each condition was modeled by convolving the stimulus onset vectors for the corresponding stimulus class with a canonical hemodynamic response function (HRF), thus providing regressors of interest for the linear model (one for each class of stimuli) to be fitted to the fMRI data. The six parameters estimated by the motion correction algorithm were used as confound regressors. Linear contrasts were used to compare condition-specific effects and to obtain contrast images to be subsequently entered in a second-level (random-effects) group analysis. The contrasts of interest were: (a) [*Faces vs Rest*]; (b) [*Words vs Rest*]; (c) [*Faces vs Words*]; (d) positive vs. negative [*pos vs neg*]; (e) Faces-positive *vs* Faces-negative [*Faces.pos vs Faces.neg*], (f) Words-positive *vs* Words-negative [*Words.pos vs Words. neg*]. Additionally, an omnibus F-contrast was computed to test against the null hypothesis of the absence of any effect of interest. In the group analysis, gender was entered as a covariate.

SPM cluster-extent-based thresholding (with voxel-wise $p < 0.001$) was chosen to yield a whole-brain family-wise significance level $\alpha < 0.05$ for the group-level statistical maps (Woo et al., 2014). Average beta values for each regressor of interest were extracted from different regions of interest (ROIs) using MarsBar version 0.45[4] and plotted for detailed examination.

In order to identify the brain regions showing a valence effect for both faces and words, the intersection of the thresholded and binarized maps for the contrasts [*Faces.pos > Faces.neg*] and [*Words. pos > Words.neg*] was also computed, following the "Minimum Statistics compared to the Conjunction Null" approach described in Nichols et al. (2005).

---

4   http://marsbar.sourceforge.net/

## 2.7. Dynamic causal modeling

Dynamic causal modeling was performed using SPM12. This is a Bayesian framework to infer directed (effective) connectivity between brain regions (Friston et al., 2003). Since the purpose of the present study was to investigate the neural circuits differentially coding for positive and negative valence in the implicit processing of facial expressions and words, two different DCMs were performed, one for faces (Face-DCM) and one for words (Word-DCM).

### 2.7.1. Volumes of interest selection and time series extraction

The starting point for a DCM analysis is the selection of a set of regions and their putative connections. The specific coordinates for each region were informed by the group-level GLM results and the known neuroimaging literature of emotional processing.

Namely, for the Face-DCM, five volumes of interest (VOIs) were chosen based on the group activation *t*-map peaks as follows. The areas more specifically involved in face processing, i.e., right Fusiform Face Area (FFA) and right Middle Frontal Gyrus (MFG), were identified *via* the [*Faces > Words*] contrast. The regions involved in emotional valence were selected *via* the [*pos > neg*] contrast (bilateral secondary visual area, V2), and the [*Faces > Rest*] contrast (bilateral primary visual area, V1, and right amygdala).

The selection of regions for the Word-DCM—five VOIs as well—was performed with a similar strategy. The left Visual Word Form Area (VWFA) and left Inferior Frontal Gyrus (IFG) were identified *via* the [*Words > Faces*] contrast, the bilateral V2 *via* the [*pos > neg*] contrast, and the bilateral V1 *via* the [*Words > Rest*] contrast. Although the left amygdala was not part of the activated regions for the [*Words > Rest*] contrast, when using a corrected statistical threshold for the whole brain, we included it as a node for the Word-DCM (using symmetrical coordinates to the right amygdala VOI), because of the extant evidence of its involvement in processing of emotional words (Hamann and Mao, 2002; Kensinger and Schacter, 2006; Lewis et al., 2007; Nakamura et al., 2020).

These VOIs were fitted to the individual subjects according to the following automated procedure. For each subject and each VOI, a subject-specific peak for the same contrast was identified within 9 mm from the respective group-level cluster peak. Then, for each of these subject-specific peaks, the first eigenvariate time series was extracted from the preprocessed functional data from a 5 mm-radius sphere centered around it, adjusted for the effects of interest (i.e., regressing out the effects of no interest). All the voxels within this sphere were used in the computation of the first eigenvariate, with no threshold applied. It should be noted that the coordinates initially identified *via* the group contrasts for each VOI were used only as "starting point" for the procedure of subject-wise VOI specification; thus, for every subject and every VOI, the procedure successfully identified a significant local peak within 9 mm from the starting coordinates. The Euclidean distances between the center of the subject-level spheres and the initial group-level peaks were on average 2.55 and 3.28 mm for Face- and Word-DCM, respectively. More details about time series extraction are provided in Table 1, Supplementary material.

### 2.7.2. Model specification

The DCM model space is established by specifying a general connectivity structure to the set of included VOIs. Two one-state bilinear DCMs (Friston et al., 2003) were implemented, one for faces and one for words. Three types of parameters were estimated: (1) endogenous parameters, measuring the average effective connectivity across experimental conditions (specified by the **A**-matrix); (2) modulatory parameters, reflecting changes in effective connectivity due to experimental conditions (specified by the **B** 3D-matrix); (3) driving parameters, showing how the single brain regions respond to experimental stimuli (specified by the **C** matrix). Concerning the specification of extrinsic (i.e., between-region) connectivity, bidirectional connections were enabled in our model between V1 and V2, V2 and FFA, FFA, and MFG for faces, and between V1 and V2, V2 and VWFA, VWFA, and IFG for words; in addition, bidirectional connections were enabled between the amygdala and any other VOI, for both Face-DCM and Word-DCM (Figure 2). Moreover, each brain region was equipped with an inhibitory self-connection, specified by the elements on the leading diagonal of the **A**-matrix, modeling the region's excitatory-inhibitory balance (Bastos et al., 2012).

For Face-DCM as well as for Word-DCM, both faces and words were set as driving input to V1 and the amygdala. Even though the amygdala's activation was not significantly different between positive and negative stimuli in the group analysis, this VOI was included as a target for driving input because it belongs to the subcortical visual pathway and is known to participate in the recognition of emotional signals (Adolphs, 2002).
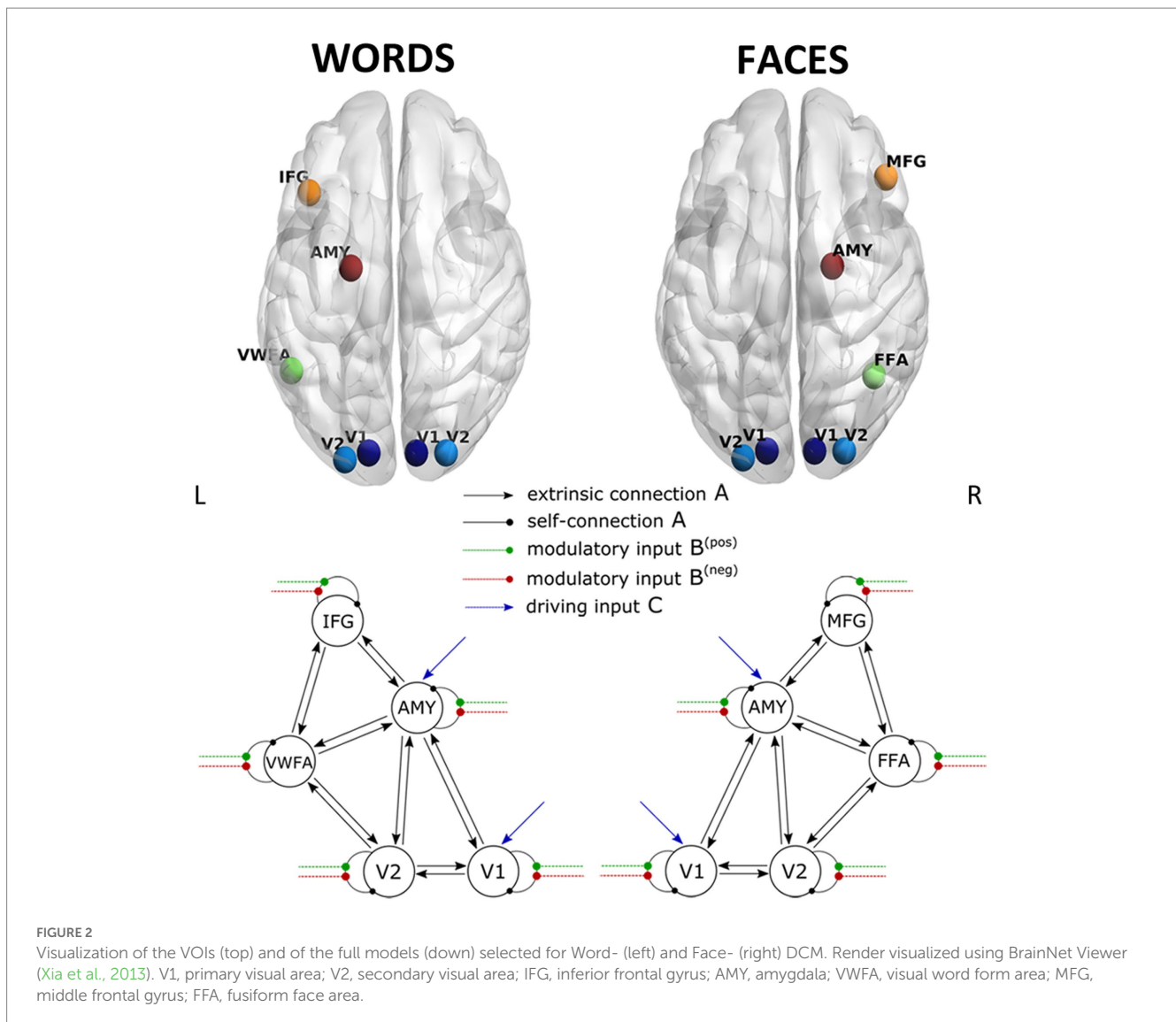
Finally, the modulatory effect of valence for faces or words was enabled on the self-connections of each region of the two models, respectively. As recommended by Zeidman et al. (2019a), restricting the modulatory effects to self-connections improves parameter identifiability. Since self-connections are constrained to be negative (i.e., inhibitory), both endogenous and modulatory self-connection parameters are unitless log-scaling parameters that multiply up or down the default value of −0.5 Hz. As a consequence, the more positive is the self-connection parameter, the more inhibited is the region.

Therefore, four conditions were defined for each DCM: (i) positive faces, negative faces, faces, and words for Face-DCM, (ii) positive words, negative words, faces, and words for Word-DCM. For each condition, "input" vectors representing stimulus timing were mean-centered to improve the model evidence and give the matrix **A** a simpler interpretation (the average connectivity; Zeidman et al., 2019a).

DCM estimation was performed using the function *spm_dcm_fit*, which fits DCM to the data using Variational Laplace. We will refer to the model just specified as the "full" model, since all parameters of interest were therein switched on.

### 2.7.3. Parametric empirical Bayes

In DCM, a "first-level" estimation of each subject's connectivity parameters is typically followed by a "second-level" quantification of the commonalities and differences across subjects. The second-level analysis was performed using the method of Parametric Empirical Bayes (PEB) (Friston et al., 2016). The PEB method rests on having defined and estimated only one "full" DCM at the single-subject level, where all possible connections of interest are present. Other candidate models are

**FIGURE 2**
Visualization of the VOIs (top) and of the full models (down) selected for Word- (left) and Face- (right) DCM. Render visualized using BrainNet Viewer (Xia et al., 2013). V1, primary visual area; V2, secondary visual area; IFG, inferior frontal gyrus; AMY, amygdala; VWFA, visual word form area; MFG, middle frontal gyrus; FFA, fusiform face area.

specified simply by choosing which parameters should be "switched on" and "switched off." More specifically, instead of manually defining a choice set of reduced models, an automatic search was performed using the SPM function *spm_dcm_peb_bmc*. The automatic PEB was implemented separately for matrices **A** and **B** (Zeidman et al., 2019b). For both faces and words, fixed connections were described by 38 s-level parameters: 2 between-subjects effects (group mean and gender) times 19 DCM parameters (5 self-connections, 14 between regions connections). On the other hand, modulatory inputs were described by 20 parameters: 2 between-subjects effects, times 10 DCM parameters (5 for each valence type). Gender and age were used as covariates to account for a potential biasing effect of these variables on the estimation of group mean parameters. The final step of the group analysis is a Bayesian Model Averaging (BMA) procedure, which identifies and reports the parameters with an estimated posterior probability of being nonzero greater than 0.95 (see Supplementary Materials for details).

# 3. Results

## 3.1. Behavioral data and questionnaires

The mean reaction times and post-scanning ratings of all the experimental stimuli are summarized in Figure 3 and Table 2. Since the focus of the present study was the comparison of positive and negative emotional valence of faces and words, the mean RT and post-scanning ratings for neutral stimuli are reported in Table 2 for the sake of completeness, but were not further analyzed.

Differences in RT means related to valence were not significant for Faces or Words, as assessed by Wilcoxon signed-rank tests [*Faces.neg vs Faces.pos*], ($p = 0.10$; [*Words.neg vs Words.pos*], $p = 0.13$), showing that positive and negative stimuli had similar processing demands. Paired $t$-tests comparing the arousal ratings for negative and positive stimuli (faces and words, separately) were also nonsignificant, confirming that positive and negative stimuli were well-matched in this dimension (Faces: $t(34) = 0.045$, $p = 0.96$; Words: $t(34) = -1.68$,
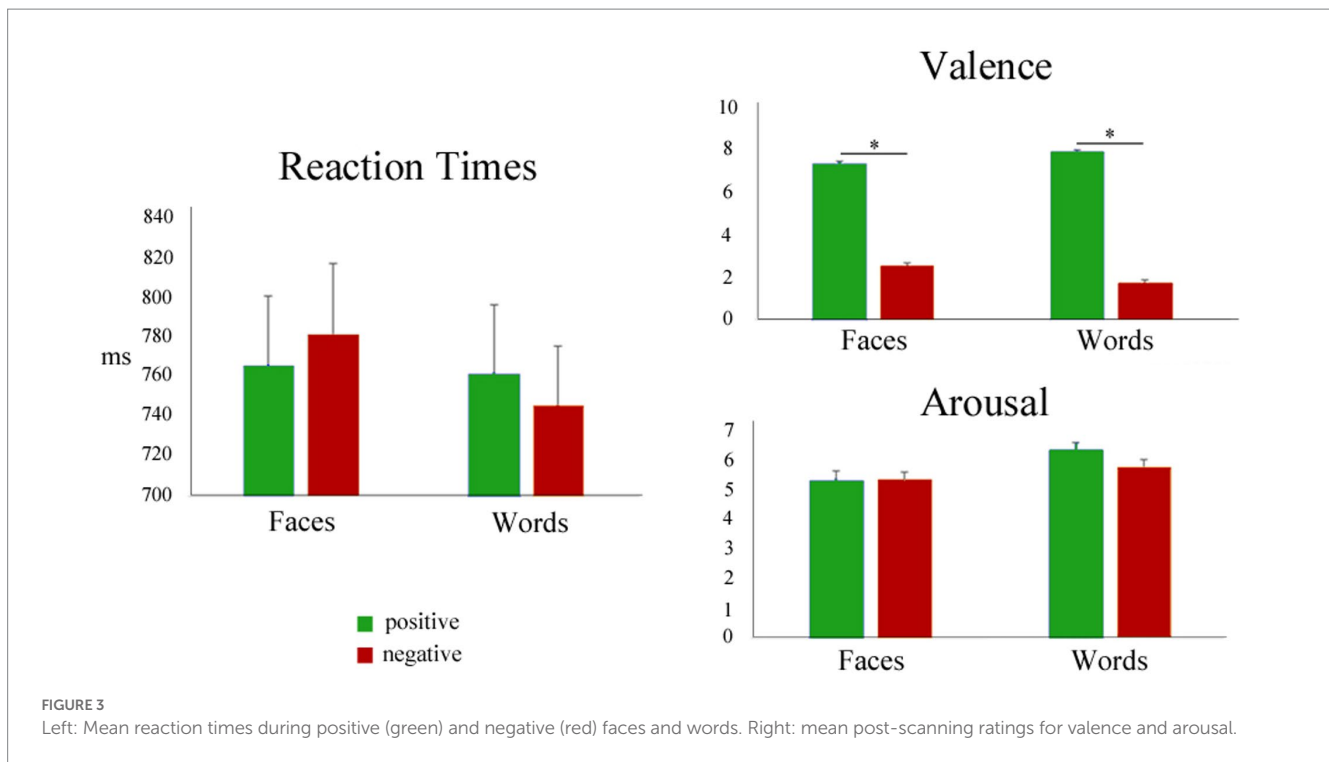
**FIGURE 3**
Left: Mean reaction times during positive (green) and negative (red) faces and words. Right: mean post-scanning ratings for valence and arousal.

TABLE 2 Average post-scanning ratings of valence and arousal (range 1−9) and response times (ms) for positive, negative, and neutral stimuli.

|  | FACES average ratings (SD) | | | WORDS average ratings (SD) | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | Positive | Negative | Neutral | Positive | Negative | Neutral |
| Valence | 7.3 (0.9) | 2.5 (0.7) | 4.6 (0.6) | 7.9 (0.7) | 1.7 (0.5) | 5.0 (0.3) |
| Arousal | 5.3 (1.5) | 5.3 (1.8) | 2.4 (1.3) | 6.3 (1.6) | 5.7 (1.5) | 2.1 (1.4) |
| Reaction times | 765.3 (210.6) | 781.4 (211.9) | 770.7 (202.4) | 761.3 (210.0) | 745.7 (177.7) | 742.7 (205.3) |

SD, standard deviation.

$p = 0.10$). As expected, paired $t$-tests for the valence ratings showed highly significant differences between negative and positive stimuli, for both faces and words (Faces: $t(34) = −20.18$, $p < 0.001$; Words: $t(34) = −34.25$, $p < 0.001$), thus validating the employed category grouping of the experimental stimuli. When analyzing the squared deviations of the valence scores from the center of the scale (i.e., 5), a paired $t$-test showed no significant difference between negative and positive faces ($t(34) = 0.86$, $p = 0.40$), but significantly higher values for negative words compared to positive ones ($t(34) = 3.70$, $p < 0.001$).

## 3.2. GLM analyses results

### 3.2.1. Effect of stimulus type

As expected, the comparison [*Faces > Words*] revealed significant activations of regions belonging to the "core face network," such as bilateral occipital face area (OFA), right fusiform face area (FFA), and superior temporal sulcus. A significant activation of bilateral thalamus and of the extended face network, including right amygdala and right middle frontal gyrus, was also observed (Figure 4 left; Table 3A).

The contrast [*Words > Faces*] identified a network of left-lateralized linguistic regions, including the inferior frontal gyrus (pars opercularis and triangularis), the supramarginal gyrus, and the visual word form area of the fusiform gyrus. The left inferior frontal gyrus and the bilateral angular gyrus were also significantly activated (Figure 4 right; Table 3B).

### 3.2.2. Effect of valence

The contrast [*Faces.pos > Faces.neg*] revealed two clusters in the extrastriate visual cortex (V2, BA 18–19; Figure 5 and Table 4A), with both classes of stimuli showing positive activation with respect to the resting baseline.

The corresponding contrast for words [*Words.pos > Words.neg*] identified a wider occipital network of activated clusters in the bilateral lingual and fusiform gyri, cuneus, and calcarine cortex (Figure 5 and Table 4B). No regions of increased signal for negative stimuli compared to positive ones were detected, either for faces or words.

The areas showing a significant effect of valence for both faces and words were limited to the bilateral secondary visual cortex, with a larger extent on the left side (Figure 5).

## 3.3. Dynamic causal modeling results

The *spm_dcm_fmri_check* function was used to calculate the percentage of explained variance in order to evaluate the performance

**FIGURE 4**
Suprathreshold clusters of activation for the contrasts [*Faces>Words*] (left) and [*Words>Faces*] (right). Bar plots represent the cluster-averaged estimates of GLM coefficients (beta values) for each regressor of interest.
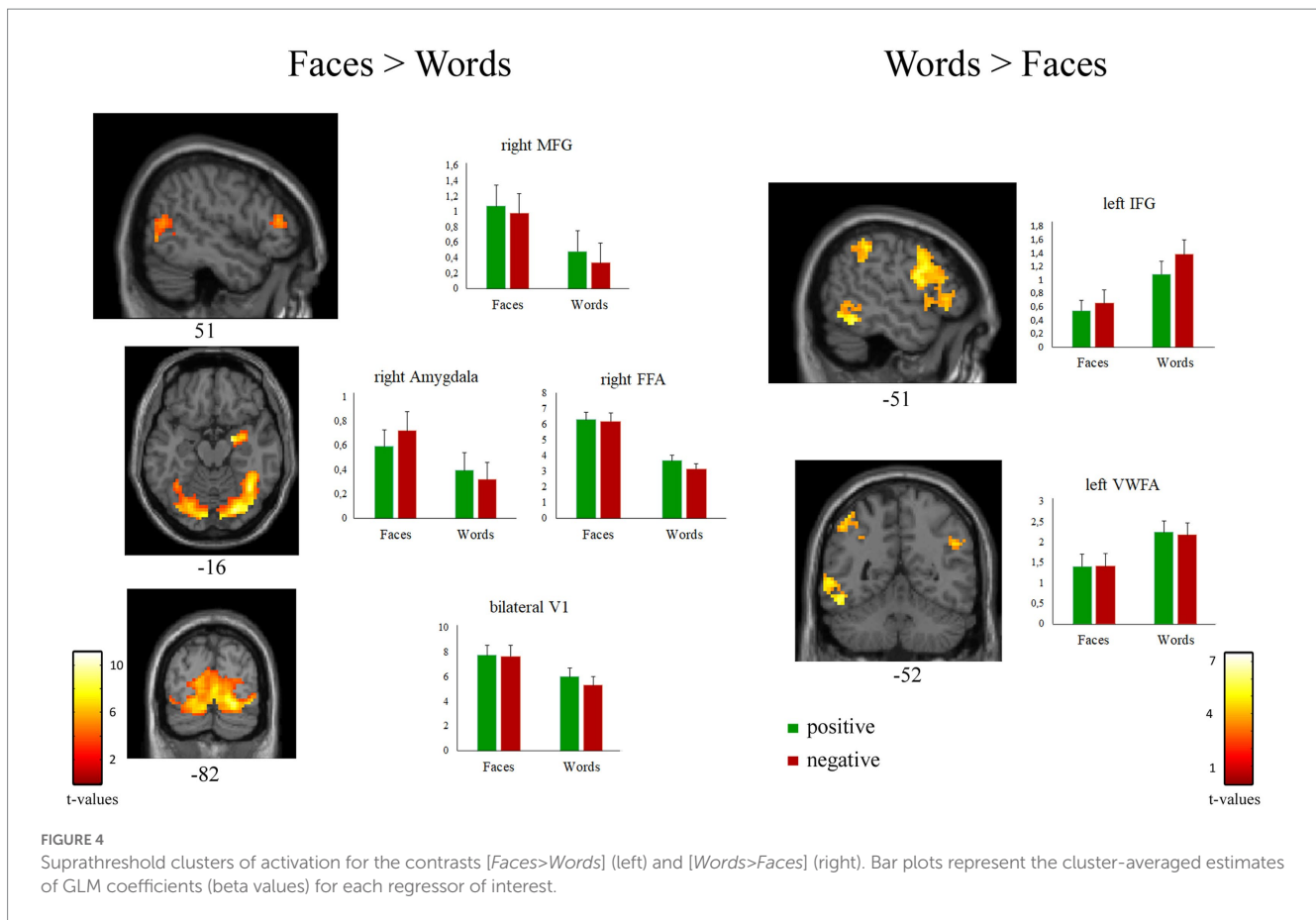
of model inversion. Across all included subjects, an average explained variance of 28.4% (SD 11.4%) for faces and of 21.2% (SD 11.2%) for words was obtained.

### 3.3.1. Face-associated connectivity

The BMA was calculated over 256 models for fixed connections (**A**-matrix) and 109 for modulatory inputs (**B**-matrix). For the group mean parameters, the presence of a strong endogenous self-inhibition was observed in V1, whereas the amygdala's endogenous self-inhibition was weak. The coupling parameters revealed excitatory connections from the amygdala to V1, V2, FFA, and MFG. As can be gleaned by the connectivity scheme in Figure 6, the strength of these connections was further increased for positive faces, as a consequence of the modulation of the amygdala's self-connection. An inhibitory connection was found from FFA to V2; its inhibitory strength was further increased during the processing of negative faces *via* their modulatory input on the FFA's self-connection. Both positive and negative faces modulated the self-connection of the MFG, increasing overall its activation. Further details about self-connections parameters and their modulations are shown in Table 5A.

### 3.3.2. Word-associated connectivity

In the case of the Word-DCM, BMA was calculated over 179 models for the **A**-matrix and 106 for the **B**-matrix. The strong self-inhibition of V1 and the weak self-inhibition of the amygdala were similar to what was found in Face-DCM. In addition, there were excitatory connections from the amygdala to IFG and VWFA. A

significant excitatory connection from V1 to V2 was observed; its strength was further increased during the processing of positive faces *via* their modulatory input on the V1's self-connection. Self-connections of V2, VWFA, and IFG were modulated by both positive and negative words; all of these modulations were negative, indicating a stronger activation of the mentioned regions in presence of both stimuli. Further details are provided in Figure 6 and Table 5B.

## 4. Discussion

The main purpose of the present study was to investigate similarities and differences in neural activations coding for valenced facial expressions and words, using an emotionally-implicit visual processing task. We found evidence for a differential response to negative and positive valence across both faces and words only in early visual cortices (mainly V2). DCM analyses revealed that, for faces, this effect was chiefly mediated by a facilitation of amygdalar activity by positive stimuli and of FFA activity by negative ones; for words, the effect was mainly imputable to a facilitation of V1 activity by positive stimuli.

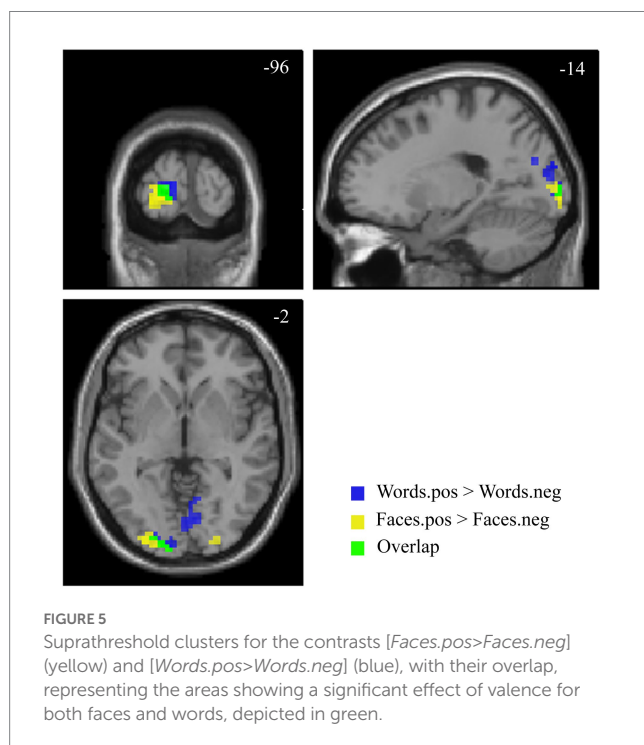Recent perspectives hypothesize the existence of a network of regions coding for valence without a regional specificity for positive and negative stimuli (Chikazoe et al., 2014). The notion is supported by a recent meta-analysis arguing for a flexible affective workspace, processing both positive and negative affect (Lindquist et al., 2016). This valence-general neural space includes anterior insula, rostral and

TABLE 3 Peak coordinates of the contrasts [*Faces>Words*] (cluster size threshold $K \geq 35$ voxels) and [*Words>Faces*] ($K \geq 41$, corrected at $\alpha < 0.05$).

| Anatomical location | BA | Side | Cluster size (voxels) | Peak z-score | MNI coordinates (x, y, z) | | |
|---|---|---|---|---|---|---|---|
| *(A) Faces > Words* | | | | | | | |
| Fusiform gyrus (FFA), occipital fusiform gyrus (OFA), calcarine cortex, lingual gyrus, superior temporal sulcus | 37, 19, 18, 17 | l/r | 2,617 | 7.08 | 42 | −55 | −19 |
| Amygdala | | r | 90 | 6.62 | 21 | −7 | −16 |
| Thalamus | | r | 112 | 6.20 | 21 | −31 | −1 |
| Middle frontal gyrus (dlPFC) | 9 | r | 47 | 4.57 | 51 | 35 | 14 |
| Thalamus | | l | 35 | 4.31 | −27 | −25 | −7 |
| *(B) Words > Faces* | | | | | | | |
| Inferior occipital gyrus | 17 | l | 41 | 5.63 | −24 | −97 | −10 |
| Superior parietal lobule, supramarginal gyrus, angular gyrus | 39, 40 | l | 403 | 4.90 | −30 | −64 | 50 |
| Inferior (VWFA), middle, and superior temporal gyri | 21, 37 | l | 245 | 4.81 | −66 | −43 | 2 |
| Inferior frontal gyrus (pars opercularis, pars triangularis) | 44, 45 | l | 535 | 4.73 | −51 | 5 | 26 |
| Angular gyrus | 39 | r | 54 | 3.92 | 48 | −52 | 32 |

*r*, right; *l*, left; BA, Brodmann area.



**FIGURE 5**
Suprathreshold clusters for the contrasts [*Faces.pos>Faces.neg*] (yellow) and [*Words.pos>Words.neg*] (blue), with their overlap, representing the areas showing a significant effect of valence for both faces and words, depicted in green.
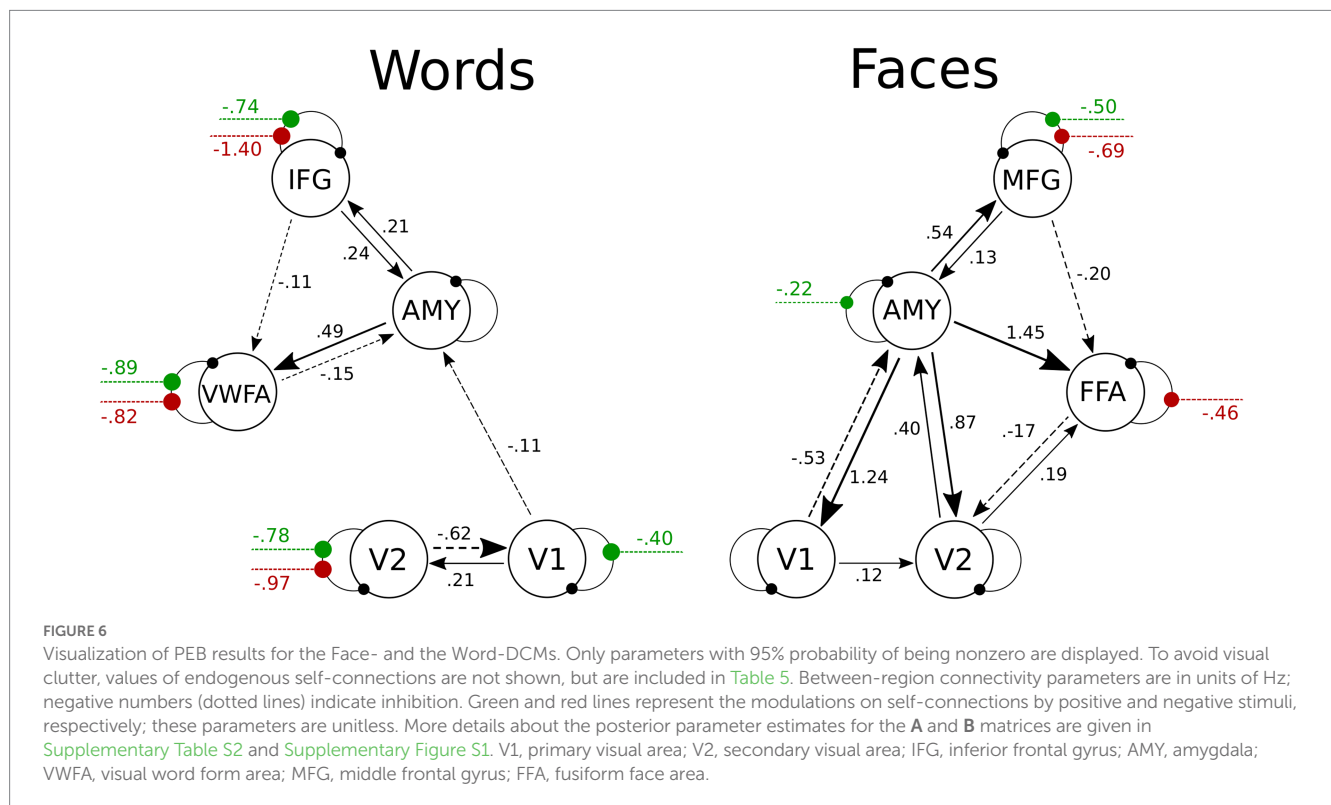
and Anderson (2018) suggested that this general valence coding works together with differentiated modality-specific representations of valence, with a major role of the sensory cortices. Supporting this hypothesis, a recent work using multivariate pattern analysis showed that the specific valence of visual and auditory stimuli can be successfully decoded by the activity of the respective sensory cortices (Shinkareva et al., 2014). A meta-analytic study by Satpute et al. (2015) also demonstrated the involvement of early sensory areas in constructing modality-specific affective experience for all sensory modalities.

Our study showed a differential activation of extrastriate visual cortex for positive and negative stimuli, for both faces and words, confirming previous findings suggesting the importance of early sensory brain cortices for affective experience and valence decoding (Satpute et al., 2015; Miskovic and Anderson, 2018). Even though for emotion-related word processing the extant experimental evidence is a bit more controversial (Kissler et al., 2006; Citron, 2012; Reisch et al., 2020), the present results speak to the importance of the early perceptual stages in valence processing, not only for facial expressions, but also for words. The involvement of early visual cortices during emotional word processing has been in fact reported in a number of fMRI studies. Herbert et al. (2009) identified a robust activation pattern in the left amygdala and in the left extrastriate visual cortex during the reading of pleasant adjectives (compared to unpleasant or neutral adjectives). A single cluster of increased signal in the left extrastriate cortex was reported during a word categorization task, compared to a picture categorization task (with emotional stimuli), by

dorsal ACC, ventromedial prefrontal cortex, amygdala, ventral striatum, thalamus, and occipitotemporal cortex. Recently, Miskovic

TABLE 4 Peak coordinates of the contrast [*Faces.pos>Faces.neg*] (A) and [*Words.pos>Words.neg*] (B).

| Anatomical location | BA | Side | Cluster size (voxels) | Peak z-score | MNI coordinates (x, y, z) | | |
|---|---|---|---|---|---|---|---|
| *(A) Faces* | | | | | | | |
| Inferior occipital gyrus | 19 | l | 100 | 4.70 | −27 | −94 | 5 |
| Inferior and superior occipital gyri | 19 | r | 68 | 4.36 | 27 | −88 | 8 |
| *(B) Words* | | | | | | | |
| Fusiform gyrus, bilateral lingual gyrus, calcarine cortex, cuneus | 17, 18, 19 | r | 746 | 4.97 | 27 | −76 | −13 |

Cluster-extent-based thresholding at $\alpha < 0.05$, with single-voxel $p < 0.001$. r, right; l, left; BA, Brodmann area.



FIGURE 6
Visualization of PEB results for the Face- and the Word-DCMs. Only parameters with 95% probability of being nonzero are displayed. To avoid visual clutter, values of endogenous self-connections are not shown, but are included in Table 5. Between-region connectivity parameters are in units of Hz; negative numbers (dotted lines) indicate inhibition. Green and red lines represent the modulations on self-connections by positive and negative stimuli, respectively; these parameters are unitless. More details about the posterior parameter estimates for the **A** and **B** matrices are given in Supplementary Table S2 and Supplementary Figure S1. V1, primary visual area; V2, secondary visual area; IFG, inferior frontal gyrus; AMY, amygdala; VWFA, visual word form area; MFG, middle frontal gyrus; FFA, fusiform face area.

Flaisch et al. (2015). Finally, a functional coupling between the right extrastriate cortex and the right amygdala was found by Tabert et al. (2001) during unpleasant emotional word processing.

In the present study, we used Dynamic Causal Modeling to elucidate the mechanisms by which the activity of early sensory cortices is modulated by emotional valence of visual stimuli. In accordance with a recent review of functional and effective connectivity studies (Underwood et al., 2021), a dynamic and context-dependent interplay between cortical and subcortical regions was expected, and a key role of the amygdala was plausible for both faces and words. The main novelty of the present study is represented by the DCM results suggesting that different mechanisms may be at play in the extrastriate visual cortices in processing the affective valence of facial expressions and words: (i) V2 activation during emotional face processing seems to depend on excitatory inputs from the amygdala, which is itself modulated by positive stimuli, and on inhibitory inputs from FFA, which is modulated by negative faces; (ii) on the other hand, V2 activity

during emotional word coding appears to rely more on excitatory inputs from V1, whose activity is modulated by positive stimuli.

Therefore, although some previous fMRI studies have reported a statistical association between the activations of the extrastriate cortex and the amygdala in both left and right hemispheres during affective word reading (Tabert et al., 2001; Herbert et al., 2009), our DCM findings do not provide evidence for a significant direct message passing between the amygdala and the secondary visual cortex during emotional words processing. However, a significant bidirectional coupling between the amygdala and the VWFA was observed in the present study, which suggests that the amygdala can detect the affective content of emotional words even before a detailed orthographic decoding, and may in turn influence the reading network (Nakamura et al., 2020).

In summary, this study supports the notion that the emotional valence of both faces and words can be discriminated in early sensory regions *via* two different mechanisms.

TABLE 5  Estimated self-connections parameters obtained using PEB.

|  | Region | Endogenous connectivity | Modulation of positive stimuli | Modulation of negative stimuli |
|---|---|---|---|---|
| *(A) Faces* | V1 | 1.23 (0.05) | – | – |
|  | V2 | – | – | – |
|  | FFA | 0.10 (0.04) | – | −0.46 (0.07) |
|  | MFG | – | −0.50 (0.10) | −0.69 (0.11) |
|  | AMY | −0.83 (0.04) | −0.22 (0.07) | – |
| *(B) Words* | V1 | 0.90 (0.06) | −0.40 (0.08) | – |
|  | V2 | 0.21 (0.03) | −0.78 (0.09) | −0.97 (0.10) |
|  | VWFA | – | −0.89 (0.13) | −0.82 (0.13) |
|  | IFG | – | −0.74 (0.17) | −1.40 (0.16) |
|  | AMY | −0.58 (0.04) | – | – |

The second column contains the average connectivity strength (SD in parenthesis) across experimental conditions (diagonal of **A**-matrix), whereas the last two columns refer to the modulatory influence of positive and negative stimuli. Variables in the Table are the log of scaling parameters that multiply up or down the default self-connection value of-0.5 Hz. For a graphical interpretation of the parameter estimates, see also Supplementary Figure S1. V1, primary visual area; V2, secondary visual area; IFG, inferior frontal gyrus; AMY, amygdala; VWFA, visual word form area; MFG, middle frontal gyrus; FFA, fusiform face area. *Empty cells indicate parameters with probability of being nonzero lower than 0.95. These parameters are not considered meaningful.

Mainly, the amygdala, as a key component of the subcortical visual route, is crucial for valence coding during the processing of faces, whereas for words, the direct thalamic pathway to primary visual cortex may play a stronger role in valence coding. Importantly, this does not necessarily imply that valence is discriminated early *in a temporal sense* in early visual cortices. As suggested by our results, the modulation may occur as a consequence of a top-down signal from hierarchically-higher visual areas (e.g., FFA and VWFA) and frontal regions (e.g., IFG and MFG).

While face and word stimuli differ in various perceptual aspects, the different circuits and patterns of directed connectivity identified in this study may also reflect differences in the innate or acquired capacity to respond to faces and words, respectively. For what concerns perception, the ability to interpret facial expressions of emotions has been shown to be present at a very young age (Nelson et al., 1979), and the production of emotional expressions is also often considered as innate and universal (Izard, 1994; Ekman, 2016; see however, more recent data suggesting a "theory of constructed emotion," Feldman Barrett, 2017). It should also be noted that a generalization of our findings to every exemplar of such categories was beyond the scope of the present study (and likely addressable only with a meta-analysis approach).

## 5. Limitations of the study and further research

To our knowledge, this is the first effective connectivity study using PEB to compare putative neural circuits for faces and words, chosen here as typical instances of largely innate and largely acquired emotional stimuli, respectively; it provides novel insight into the role of early sensory cortices in such processes. However, we acknowledge a number of limitations that should be addressed in further experimental work.

First, our protocol involved only implicit emotional processing. While this was a design choice, it may be argued that a greater (or different) recruitment of neural circuits would be observed when processing of emotional information is explicitly required by the task; on the other hand, implicit task protocols are widely used in psychological and neurophysiological research, and they can be considered more ecological than the explicit tasks in most situations.

Second, emotional visual stimuli were limited to faces and only one type of facial expression was used for each valence class (i.e., happy for positive, angry for negative). How the current findings would change were other sets of visual stimuli employed — e.g., painful expressions, or IAPS stimuli (Lang et al., 2005) — remains to be determined.

Finally, words and faces are very different kinds of objects, therefore our stimuli do not differ only because their recognition is either primarily innate or culturally acquired. Future research could explore innate and acquired valence using stimuli with matched structural characteristics or, alternatively, more sophisticated paradigms where stimuli could be conditioned to modify their innate valence.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving human participants were reviewed and approved by Comitato Etico dell'Area Vasta Emilia Nord, Azienda Ospedaliero-Universitaria di Modena, Università di Modena e Reggio Emilia (Italy). The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any identifiable images or data included in this article.

## Author contributions

DB, EB, FL, and GP contributed to the design of the experiment. DB, EB, and GP contributed to data collection. DB, RM, EB, and GP contributed to data analysis. All authors contributed to the interpretation of the results and manuscript writing.

# Funding

# Conflict of interest

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

# Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

# Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1055054/full#supplementary-material

# References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Curr. Opin. Neurobiol.* 12, 169–177. doi: 10.1016/S0959-4388(02)00301-X

Aldhafeeri, F. M., Mackenzie, I., Kay, T., Alghamdi, J., and Sluming, V. (2012). Regional brain responses to pleasant and unpleasant IAPS pictures: different networks. *Neurosci. Lett.* 512, 94–98. doi: 10.1016/j.neulet.2012.01.064

Alpers, G. W., and Gerdes, A. B. M. (2007). Here is looking at you: emotional faces predominate in binocular rivalry. *Emotion* 7, 495–506. doi: 10.1037/1528-3542.7.3.495

Bach, D. R., Schmidt-Daffy, M., and Dolan, R. J. (2014). Facial expression influences face identity recognition during the attentional blink. *Emotion* 14, 1007–1013. doi: 10.1037/a0037945

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038

Berent, I., Feldman Barrett, L., and Platt, M. (2020). Essentialist biases in reasoning about emotions. *Front. Psychol.* 11:562666. doi: 10.3389/fpsyg.2020.562666

Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9

Chikazoe, J., Lee, D. H., Kriegeskorte, N., and Anderson, A. K. (2014). Population coding of affect across stimuli, modalities and individuals. *Nat. Neurosci.* 17, 1114–1122. doi: 10.1038/nn.3749

Citron, F. M. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain Lang.* 122, 211–226. doi: 10.1016/j.bandl.2011.12.007

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123, 291–307. doi: 10.1093/brain/123.2.291

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., and Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature* 589, 251–257. doi: 10.1038/s41586-020-3037-7

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173. doi: 10.1006/cbmr.1996.0014

Cox, R. W., and Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR Biomed.* 10, 171–178. doi: 10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L

Ćurčić-Blake, B., Swart, M., and Aleman, A. (2012). Bidirectional information flow in frontoamygdalar circuits in humans: a dynamic causal modeling study of emotional associative learning. *Cereb. Cortex* 22, 436–445. doi: 10.1093/cercor/bhr124

Da Silva, E. L., Castellano, G., Sato, J. R., Cardoso, E. F., and Amaro, E. (2011). Emotional face perception in healthy subjects and parkinson's disease: an effective connectivity study. *Commun. Comput. Inf. Sci.* 127 CCIS, 278–289. doi: 10.1007/978-3-642-18472-7_22

Ekman, P. (2016). What scientists who study emotion agree about. *Perspect. Psychol. Sci.* 11, 31–34. doi: 10.1177/1745691615596992

Fairhall, S. L., and Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cereb. Cortex* 17, 2400–2406. doi: 10.1093/cercor/bhl148

Feldman Barrett, L. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Soc. Cogn. Affect. Neurosci.* 12, nsw154–nsw123. doi: 10.1093/scan/nsw154

Flaisch, T., Imhof, M., Schmälzle, R., Wentz, K. U., Ibach, B., and Schupp, H. T. (2015). Implicit and explicit attention to pictures and words: an fMRI-study of concurrent emotional stimulus processing. *Front. Psychol.* 6:1861. doi: 10.3389/fpsyg.2015.01861

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., and Dutton, K. (2000). Facial expressions of emotion: are angry faces detected more efficiently? *Cogn. Emot.* 14, 61–92. doi: 10.1080/026999300378996

Friston, K. J. (2012). Ten ironic rules for non-statistical reviewers. *NeuroImage* 61, 1300–1310. doi: 10.1016/j.neuroimage.2012.04.018

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *NeuroImage* 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7

Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., Van Wijk, B. C. M., et al. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *NeuroImage* 128, 413–431. doi: 10.1016/j.neuroimage.2015.11.015

Furl, N., Henson, R. N., Friston, K. J., and Calder, A. J. (2013). Top-down control of visual responses to fear by the amygdala. *J. Neurosci.* 33, 17435–17443. doi: 10.1523/JNEUROSCI.2992-13.2013

Fusar-Poli, P., Placentino, A., Carletti, F., Landi, P., Allen, P., Surguladze, S., et al. (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *J. Psychiatry Neurosci.* 34, 418–432.

Gaillard, R., Del Cul, A., Naccache, L., Vinckier, F., Cohen, L., and Dehaene, S. (2006). Nonconscious semantic processing of emotional words modulates conscious access. *Proc. Natl. Acad. Sci. U. S. A.* 103, 7524–7529. doi: 10.1073/pnas.0600584103

García-García, I., Kube, J., Gaebler, M., Horstmann, A., Villringer, A., and Neumann, J. (2016). Neural processing of negative emotional stimuli and the influence of age, sex and task-related characteristics. *Neurosci. Biobehav. Rev.* 68, 773–793. doi: 10.1016/j.neubiorev.2016.04.020

Goh, W. D., Yap, M. J., Lau, M. C., Ng, M. M. R., and Tan, L. C. (2016). Semantic richness effects in spoken word recognition: a lexical decision and semantic categorization megastudy. *Front. Psychol.* 7, 1–10. doi: 10.3389/fpsyg.2016.00976

Gross, C., and Schwarzer, G. (2010). Face recognition across varying poses in 7- and 9-month-old infants: the role of facial expression. *Int. J. Behav. Dev.* 34, 417–426. doi: 10.1177/0165025409350364

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., and Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cereb. Cortex* 22, 1564–1576. doi: 10.1093/cercor/bhr226

Hamann, S., and Mao, H. (2002). Positive and negative emotional verbal stimuli elicit activity in the left amygdala. *Neuro Rep.* 13, 15–19. doi: 10.1097/00001756-200201210-00008

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2000). 08_Haxbytics2000. *Trends Cogn. Sci.* 4, 223–233. doi: 10.1016/S1364-6613(00)01482-0

Haxby, J. V., Hoffman, E. A., and Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biol. Psychiatry* 51, 59–67. doi: 10.1016/S0006-3223(01)01330-0

Herbert, C., Ethofer, T., Anders, S., Junghofer, M., Wildgruber, D., Grodd, W., et al. (2009). Amygdala activation during reading of emotional adjectives - an advantage for pleasant content. *Soc. Cogn. Affect. Neurosci.* 4, 35–49. doi: 10.1093/scan/nsn027

Hoemann, K., Crittenden, A. N., Msafiri, S., Liu, Q., Li, C., Roberson, D., et al. (2019). Context facilitates performance on a classic cross-cultural emotion perception task. *Emotion* 19, 1292–1313. doi: 10.1037/emo0000501

Hoffmann, M., Mothes-Lasch, M., Miltner, W. H. R., and Straube, T. (2015). Brain activation to briefly presented emotional words: effects of stimulus awareness. *Hum. Brain Mapp.* 36, 655–665. doi: 10.1002/hbm.22654

Isosaka, T., Matsuo, T., Yamaguchi, T., Funabiki, K., Nakanishi, S., Kobayakawa, R., et al. (2015). Htr2a-expressing cells in the central amygdala control the hierarchy between innate and learned fear. *Cells* 163, 1153–1164. doi: 10.1016/j.cell.2015.10.047

Izard, C. E. (1994). Innate and universal facial expressions: evidence from developmental and cross-cultural research. *Psychol. Bull.* 115, 288–299. doi: 10.1037/0033-2909.115.2.288

Jamieson, A. J., Davey, C. G., and Harrison, B. J. (2021). Differential modulation of effective connectivity in the brain's extended face processing system by fearful and sad facial expressions. *eNeuro* 8:2021. doi: 10.1523/ENEURO.0380-20.2021

Johansson, M., Mecklinger, A., and Treese, A.-C. (2004). Recognition memory for emotional and neutral faces: an event-related potential study. *J. Cogn. Neurosci.* 16, 1840–1853. doi: 10.1162/0898929042947883

Kanske, P., and Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain Res.* 1148, 138–148. doi: 10.1016/j.brainres.2007.02.044

Kensinger, E. A., and Schacter, D. L. (2006). Processing emotional pictures and words: effects of valence and arousal. *Cogn. Affect. Behav. Neurosci.* 6, 110–126. doi: 10.3758/cabn.6.2.110

Kim, J. J., and Jung, M. W. (2006). Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci. Biobehav. Rev.* 30, 188–202. doi: 10.1016/j.neubiorev.2005.06.005

Kissler, J., Assadollahi, R., and Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Prog. Brain Res.* 156, 147–183. doi: 10.1016/S0079-6123(06)56008-X

Kong, M. S., and Zweifel, L. S. (2021). Central amygdala circuits in valence and salience processing. *Behav. Brain Res.* 410:113355. doi: 10.1016/j.bbr.2021.113355

Kousta, S. T., Vinson, D. P., and Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition* 112, 473–481. doi: 10.1016/j.cognition.2009.06.007

Lang, P. J., Bradley, M. M., and Cuthbert, B. N. (2005). *International Affective Picture System (IAPS): Technical Manual and Affective Ratings*. Technical Report A-6. Gainesville, FL.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon and Schuster.

Lewis, P. A., Critchley, H. D., Rotshtein, P., and Dolan, R. J. (2007). Neural correlates of processing valence and arousal in affective words. *Cereb. Cortex* 17, 742–748. doi: 10.1093/cercor/bhk024

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., and Barrett, L. F. (2016). The brain basis of positive and negative affect: evidence from a meta-analysis of the human neuroimaging literature. *Cereb. Cortex* 26, 1910–1922. doi: 10.1093/cercor/bhv001

Lundqvist, D., Flykt, A., and Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Miskovic, V., and Anderson, A. K. (2018). Modality general and modality specific coding of hedonic valence. *Curr. Opin. Behav. Sci.* 19, 91–97. doi: 10.1016/j.cobeha.2017.12.012

Montefinese, M., Ambrosini, E., Fairfield, B., and Mammarella, N. (2014). The adaptation of the affective norms for English words (ANEW) for Italian. *Behav. Res. Methods* 46, 887–903. doi: 10.3758/s13428-013-0405-3

Naccache, L., Gaillard, R., Adam, C., Hasboun, D., Clémenceau, S., Baulac, M., et al. (2005). A direct intracranial record of emotions evoked by subliminal words. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7713–7717. doi: 10.1073/pnas.0500542102

Nakamura, K., Inomata, T., and Uno, A. (2020). Left amygdala regulates the cerebral Reading network during fast emotion word processing. *Front. Psychol.* 11, 1–12. doi: 10.3389/fpsyg.2020.00001

Nelson, C. A., Morse, P. A., and Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child Dev.* 50, 1239–1242. doi: 10.2307/1129358

Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage* 25, 653–660. doi: 10.1016/j.neuroimage.2004.12.005

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4

Price, C. J. (2012). A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage* 62, 816–847. doi: 10.1016/j.neuroimage.2012.04.062

Radke, S., Volman, I., Kokal, I., Roelofs, K., de Bruijn, E. R. A., and Toni, I. (2017). Oxytocin reduces amygdala responses during threat approach. *Psychoneuroendocrinology* 79, 160–166. doi: 10.1016/j.psyneuen.2017.02.028

Reisch, L. M., Wegrzyn, M., Woermann, F. G., Bien, C. G., and Kissler, J. (2020). Negative content enhances stimulus-specific cerebral activity during free viewing of pictures, faces, and words. *Hum. Brain Mapp.* 41, 4332–4354. doi: 10.1002/hbm.25128

Rellecke, J., Palazova, M., Sommer, W., and Schacht, A. (2011). On the automaticity of emotion processing in words and faces: event-related brain potentials evidence from a superficial task. *Brain Cogn.* 77, 23–32. doi: 10.1016/j.bandc.2011.07.001

Sabatinelli, D., Fortune, E. E., Li, Q., Siddiqui, A., Krafft, C., Oliver, W. T., et al. (2011). Emotional perception: meta-analyses of face and natural scene processing. *NeuroImage* 54, 2524–2533. doi: 10.1016/j.neuroimage.2010.10.011

Satpute, A. B., Kang, J., Bickart, K. C., Yardley, H., Wager, T. D., and Barrett, L. F. (2015). Involvement of sensory regions in affective experience: a meta-analysis. *Front. Psychol.* 6:1860. doi: 10.3389/fpsyg.2015.01860

Schlochtermeier, L. H., Kuchinke, L., Pehrs, C., Urton, K., Kappelhoff, H., and Jacobs, A. M. (2013). Emotional picture and word processing: an fMRI study on effects of stimulus complexity. *PLoS One* 8:e55619. doi: 10.1371/journal.pone.0055619

Shinkareva, S. V., Wang, J., Kim, J., Facciani, M. J., Baucom, L. B., and Wedell, D. H. (2014). Representations of modality-specific affective processing for visual and auditory stimuli derived from functional magnetic resonance imaging data. *Hum. Brain Mapp.* 35, 3558–3568. doi: 10.1002/hbm.22421

Silva, B. A., Gross, C. T., and Graff, J. (2016). The neural circuits of innate fear: detection, integration, action, and memorization. *Learn. Mem.* 23, 544–555. doi: 10.1101/lm.042812.116

Šimić, G., Tkalčić, M., Vukić, V., Mulc, D., Španić, E., Šagud, M., et al. (2021). Understanding emotions: origins and roles of the amygdala. *Biomol. Ther.* 11, 1–58. doi: 10.3390/biom11060823

Sladky, R., Höflich, A., Küblböck, M., Kraus, C., Baldinger, P., Moser, E., et al. (2015). Disrupted effective connectivity between the amygdala and orbitofrontal cortex in social anxiety disorder during emotion discrimination revealed by dynamic causal modeling for fMRI. *Cereb. Cortex* 25, 895–903. doi: 10.1093/cercor/bht279

Tabert, M. H., Borod, J. C., Tang, C. Y., Lange, G., Wei, T. C., Johnson, R., et al. (2001). Differential amygdala activation during emotional decision and recognition memory tasks using unpleasant words: an fMRI study. *Neuropsychologia* 39, 556–573. doi: 10.1016/S0028-3932(00)00157-3

Torrisi, S. J., Lieberman, M. D., Bookheimer, S. Y., and Altshuler, L. L. (2013). Advancing understanding of affect labeling with dynamic causal modeling. *NeuroImage* 82, 481–488. doi: 10.1016/j.neuroimage.2013.06.025

Underwood, R., Tolmeijer, E., Wibroe, J., Peters, E., and Mason, L. (2021). Networks underpinning emotion: a systematic review and synthesis of functional and effective connectivity. *NeuroImage* 243:118486. doi: 10.1016/j.neuroimage.2021.118486

Vuilleumier, P., and Schwartz, S. (2001). Emotional facial expressions capture attention. *Neurology* 56, 153–158. doi: 10.1212/WNL.56.2.153

Willinger, D., Karipidis, I. I., Beltrani, S., Di Pietro, S. V., Sladky, R., Walitza, S., et al. (2019). Valence-dependent coupling of prefrontal-amygdala effective connectivity during facial affect processing. *eNeuro* 6, 1–12. doi: 10.1523/ENEURO.0079-19.2019

Woo, C. W., Krishnan, A., and Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *NeuroImage* 91, 412–419. doi: 10.1016/j.neuroimage.2013.12.058

Xia, M., Wang, J., and He, Y. (2013). Brain net viewer: a network visualization tool for human brain Connectomics. *PLoS One* 8:910. doi: 10.1371/journal.pone.0068910

Yap, M. J., and Seow, C. S. (2014). The influence of emotion on lexical processing: insights from RT distributional analysis. *Psychon. Bull. Rev.* 21, 526–533. doi: 10.3758/s13423-013-0525-x

Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A., Price, C. J., et al. (2019a). A guide to group effective connectivity analysis, part 1: first level analysis with DCM for fMRI. *NeuroImage* 200, 174–190. doi: 10.1016/j.neuroimage.2019.06.031

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., et al. (2019b). A guide to group effective connectivity analysis, part 2: second level analysis with PEB. *NeuroImage* 200, 12–25. doi: 10.1016/j.neuroimage.2019.06.032