Research Report

Nonconscious emotional processing involves distinct neural pathways for pictures and videos

Nathan Faivre a,b,*, Sylvain Charron a, Paul Roux a, Stéphane Lehérylic, Sid Kouider a

a Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS/EHESS/DEC-ENS, Paris 75005, France
b Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA
c Centre de Neurométrie de Recherche—CENIR, Centre de Recherche de l’Institut du Cerveau et de la Moelle épinière, Université Pierre et Marie Curie-Paris 6 UMR-S975, Inserm U975, CNRS UMR7225, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

A R T I C L E  I N F O

Article history:
Received 14 July 2012
Received in revised form 22 October 2012
Accepted 29 October 2012
Available online 6 November 2012

Keywords:
Amygdala
Awareness
fMRI
Fusiform face area
Psychophysiological interaction

A B S T R A C T

Facial expressions are known to impact observers’ behavior, even when they are not consciously identifiable. Relying on visual crowding, a perceptual phenomenon whereby peripheral faces become undiscriminable, we show that participants exposed to happy vs. neutral crowded faces rated the pleasantness of subsequent neutral targets accordingly to the facial expression’s valence. Using functional magnetic resonance imaging (fMRI) along with psychophysiological interaction analysis, we investigated the neural determinants of this nonconscious preference bias, either induced by static (i.e., pictures) or dynamic (i.e., videos) facial expressions. We found that while static expressions activated primarily the ventral visual pathway (including task-related functional connectivity between the fusiform face area and the amygdala), dynamic expressions triggered the dorsal visual pathway (i.e., posterior parietal cortex) and the substantia innominata, a structure that is contiguous with the dorsal amygdala. As temporal cues are known to improve the processing of visible facial expressions, the absence of ventral activation we observed with crowded videos questions the capacity to integrate facial features and facial motions without awareness. Nevertheless, both static and dynamic facial expressions activated the hippocampus and the orbitofrontal cortex, suggesting that nonconscious preference judgments may arise from the evaluation of emotional context and the computation of aesthetic evaluation.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Humans constantly react to faces conveying emotional cues, be they visible or not Tamietto and Gelder (2010). Indeed, facial expressions are encoded even when faces are rendered invisible by visual masking or binocular rivalry (Kim & Blake, 2005). Invisible facial expressions trigger emotional reactions that are reflected, peripherally, in the form of increased skin conductance (Esteves, 1994; Williams, Morris, McGlone, Abbott, & Mattingley, 2004) or mimetic facial muscles activity (Killgore & Yurgelun-Todd, 2004) and, centrally, as modulations of electrophysiological components (Liddell, Williams, Rathjen, Shevrin, & Gordon, 2004; Williams et al., 2004) and fMRI response, notably in the amygdala (Morris, Ohman, & Dolan, 1995; Pessoa, Japee, Sturman, & Ungerleider, 2006; Whalen et al., 1998). So far, the studies on nonconscious emotional processing have mainly relied on static images (see Van den Stock et al., 2011, for the perception of dynamic whole-body expressions in blindsight). However, the ecological relevance of static stimuli might be limited, in comparison to dynamic stimuli conveying biological motion cues that are omnipresent in natural vision. Along this line, several studies using visible faces have shown that compared to static expressions, dynamic expressions are better recognized and are associated with enhanced neural activity (Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004; Trautmann, Fehr, & Herrmann, 2009). Yet, it remains unknown whether dynamic cues are processed under nonconscious perceptual conditions. A positive answer would suggest the existence of nonconscious mechanisms of temporal integration, by which the different aspects of the signal changing across time (e.g., successive snapshots of a facial expression) are integrated into a unified percept (e.g., a dynamic facial expression).

Here, we investigated this issue with fMRI by measuring BOLD responses to static and dynamic facial expressions rendered undiscriminable through gaze-contingent crowding (GCC). Crowding is the perceptual phenomenon whereby a peripheral stimulus is perceived as jumbled when it is flanked by similar neighbors (Levi, 2008). Under such conditions, while the stimulus remains...
detectable, it is indistinguishable to the observer. The origins of crowding are commonly attributed to the integration of peripheral signals over inappropriately large distances. Some theories propose that this excess of integration in the visual periphery is due to the presence of large receptive fields in early visual areas (Pelli, 1997), while others consider that it is due to a coarsening of the resolution of spatial attention (Intriligator & Cavanagh, 2001).

When crowding is combined with gaze-contingent control, which consists in substituting the critical peripheral stimulus with an irrelevant content as soon as the observer’s attempted to gaze at it, conscious access to facial expressions is fully prevented (Fairev & Kouider, 2011). Using this approach, we recently reported that both static and dynamic facial expressions can bias evaluative judgments in a nonconscious manner (Kouider, Berthet, & Fairev, 2011), extending previous results with very brief and static stimuli (e.g., Murphy & Zajonc, 1993; Li, Zinbarg, Boehm, & Paller, 2008). Participants were more likely to rate an unknown Chinese character as pleasant when preceded by a face depicting happiness compared to a face depicting anger, while, crucially, they performed at chance-level when discriminating the emotion of the crowded face. Furthermore, complementary controls with a neutral baseline revealed that the nonconscious origin of the preference bias that we obtained stemmed primarily from the processing of facial expressions conveying positive emotions, regardless of their static or dynamic properties (Fairev, Berthet, & Kouider, 2012).

In the present study, we combined this behavioral approach with an fMRI paradigm in order to assess whether nonconscious processing of static and dynamic facial expressions is sustained by common neural pathways. Following our previous behavioral results, we focused more specifically on positive emotions, by measuring the influences of undiscriminable happy compared to neutral faces on pleasantness ratings (see Fig. 1A). This comparison with neutral rather than angry faces prevented from confounds due to the partial overlaps in the brain regions involved in the encoding of negative and positive emotions (Williams et al., 2004; Costafreda, Brammer, David, & Fu, 2008). Combining whole brain analysis, region of interest analysis, and functional connectivity analysis, we could specifically assess the extent to which crowded emotional information is preserved along the subcortical and cortical pathways.

2. Methods

2.1. Participants

Eighteen right-handed volunteers (12 females, 6 males; age 18 to 35) took part in the study, which was approved by the local ethical committee. All participants had a normal vision, with no past history of psychiatric and neurological disease. They gave written informed consent and were paid for their participation. None of them was able to read Chinese.

2.2. Stimuli

Stimuli were recorded from five female actresses recruited from a professional acting academy. Actresses’ face was filmed against a black background, in an equally held illumination room. Each face expressed happiness (HA, with a mouth opening motion revealing the teeth, and a narrowing of the eyes), or was neutral (NE, with a mouth chewing movement and a cyclic opening and narrowing of the eyes). Compared to other baseline conditions relying on mosaics of scrambled faces or morphed faces, this chewing movement provided an ecological, and emotionally neutral control. Emotional paroxysms were defined by two independent observers. Videos were slightly speeded up or down to compensate the differences between the timings of the different actors’ expressions. All videos were matched for average luminance contrast, and image size and were cropped to show the face only such as measuring 3’ × 3’. Snapshots were extracted from each movie, at t = 1100 ms for the emotional paroxysm condition, and at t = 0 ms for the neutral expression condition (as the facial configuration during the chewing movement was not always the same at t = 1100 ms, we preferred to extract the snapshot at t = 0 ms, when all actresses had the mouth closed and the eyes opened). Each video started with a neutral expression (100 ms), followed by emotional progression until paroxysm (1000 ms), and dynamical maintenance at the peak level (100 ms). The flankers were created by blending 2 inverted non emotional faces and 4 inverted objects together, resulting in 2 × 2 non informative patterns. The six flankers surrounding the face were selected randomly on each trial from a set of forty possible flankers. The luminance of the flankers was 40% higher than that of the face. Targets were 3’ × 3’ Chinese pictographs, chosen randomly on each trial from a set of 180 pictographs, without repetition for a single subject.

2.3. Apparatus

Stimuli were displayed against a black background by an Epson video projector (EMP-8300, refresh rate 60 Hz, resolution 1024 × 768, pitch 0.023 [pixel] on a computer running Matlab with the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Eye movements were recorded monocularly with an fMRI compatible eye tracker (Eyenlink 1000 system, SR research, Ontario, Canada) controlled with the Eyelink toolbox (sampling rate 500 Hz; spatial resolution > 1’). Answers were collected from two stickers buttons (fiber optic response device, Current Designs, Philadelphia, USA).

Fig. 1. (A) Experimental procedure: while participants continuously stared at a fixation cross, they were exposed for 1200 ms to a peripheral face surrounded by uninformative flankers. The face could express happiness or be neutral, either statically or dynamically. Importantly, whenever participants stopped gazing at the fixation cross, the peripheral face was substituted by a different neutral face, guaranteeing that facial expressions were never being discriminated. On 75% of trials, participants evaluated the pleasantness of a subsequent Chinese pictograph displayed at the fixation location. On the remaining 25% of trials, randomly assigned, participants had 5200 ms to indicate whether the peripheral face reflected a happy or neutral face. (B) Behavioral results: pleasant response rate as a function of the facial expression, with higher ratings induced by happy compared to neutral faces, both for static and dynamic stimulus types.
2.4. Procedure

The entire experiment consisted in a total of 180 priming trials intermixed randomly with 60 visibility trials and 34 flanckers-only baseline trials, equally divided in 2 blocks of 137 trials. An eye-tracking calibration phase was performed before each block. Each block contained exclusively either dynamic (dyn) or static (stat) condition. Flankers were happy or neutral faces. Two counterbalanced conditions of participants were defined depending on the sequence of blocks they were assigned to (i.e., either [dyn-stat] or [stat-dyn]). Each trial started with a 0.5 × 0.5° fixation cross for 300 ms with a jitter of 200 ms. The face was displayed statically or dynamically for 1200 ms surrounded by six flankers (14° eccentricity between face and fixation cross centered at the quarter top of the screen; 2.6° center to center distance). Importantly, as soon as gaze monitoring was lost or as soon as a participant ceased to gaze at a fixation area consisting in a zone of 5° by 5° surrounding the fixation cross, the face was substituted by the same face with a neutral expression (physically different from the one used as the neutral baseline condition), guaranteeing that the emotional face was never seen foveally. This occurred on 1.5% (SD = 2.42) of trials, corresponding to 0.033% (SD = 0.05) of total looking time. In three out of four trials, the participant’s task was to indicate whether a Chinese pictograph displayed for 150 ms at the fixation location was pleasant (right button press) or unpleasant (left button press). They were instructed to follow a spontaneous intuition, and to provide an answer in less than a second. Randomly, in one out of four trials, a question concerning the facial expression appeared at the fixation location instead of the Chinese pictograph. Participants could read either “Joie?”, or “Neutre?” (Happiness? or Neutral? in French). They had to answer by pressing the right button if they thought the answer was yes, or the left button if they thought it was no, within a delay of 5200 ms. The visibility question concerned uniquely the facial expression and was not related to the surrounding flankers. In the flankers-only baseline trials, only the six flankers and the fixation cross appeared on the screen for 1200 ms, participants were instructed not to press any button. The experiment was divided in 2 blocks of 7 min each.

2.5. Functional localizer

In addition to whole brain analyses, we relied on the use of functional regions of interest. We defined the face responsive regions for each participant in a separate scan session following the main experiment. It consisted in 12 blocks of 20 pictures each, half the blocks containing objects (food, watches, cars, cups) and the other half faces (half men and half women) were shown on the screen against a black background. Faces and objects were presented centrally in order to escape from self-crowding, a situation in which the different features inside an object continue to crowd each other to some extent even in the absence of flankers (Martelli, Majaj, & Pelli, 2005). Each image was displayed for 500 ms and followed by a 500 ms blank screen. A 3 s fixation cross was displayed between each block, the order of blocks (i.e., object or face) was counter-balanced across participants. Participants were requested to attend to the pictures, but not to press any button. The functional localizer lasted 5 min.

2.6. fMRI data acquisition

Whole brain imaging was conducted on a 3T Trio Siemens scanner (Tim 32 channels) at the Centre de Neuroimagerie de Recherche at the Salpetriere Hospital in Paris. Functional T2* weighted images were obtained in an interleaved order using a single-shot echoplanar gradient-echo (EPI) pulse sequence (43 slices oriented on the anterior to posterior commissure axis with no interslice gap, resolution 3 × 3 × 3 mm, repetition time TR = 2100 ms, echo time TE = 25 ms, flip angle 90°). The functional images were processed using the SPM8 software (Wellcome Department of Cognitive Neurology, London, UK). Three initial volumes were discarded to eliminate nonequilibrium effects of magnetization. The functional preprocessing included slice timing (middle reference slice), spatial realignment, spatial normalization (MINI stereostatic space, 2 mm voxel resampling) and spatial smoothing (Gaussian kernel with 8 mm full-width at half maximum). Low-frequency drifts were eliminated using a temporal high-pass filter at 1/128 Hz. Statistical evaluation was performed using the general linear model (GLM). Regressors modeled the eight different trials of interest ([expression preference]dyn/NE [+ expression type] + [expression judgment]) as well as the flankers-only baseline trials and the visibility trials. The convolution used delta functions with a hemodynamic response function (HRF) and its time and dispersion derivatives. In addition, estimated motion parameters were included as covariates of no interest to increase statistical sensitivity. The brain activations of all eight critical conditions were contrasted against the baseline. Linear contrasts lead to the main effect of preference bias in the static condition (i.e., [HAdyn/NEstat > [NEdyn/NEstat]) and dynamic condition ([NEstat > TNEstat] & [NEstat > TNEstat]). For random effect group analyses, the individual contrast images were smoothed (Gaussian kernel with 6 mm full-width at half maximum) and submitted to one sample two-tailed t-tests. All images were thresholded at p < 0.001, uncorrected. A small volume correction (cluster-wise threshold p < 0.05, FWE-corrected for multiple comparisons within the search volume) was conducted relatively to a priori hypothesis regarding activity in the amygdala (dilated 5 ×). The search volume was created from the AAL brain atlas (Tzourio-Mazoyer et al., 2002), and was applied to the SPM dataset using WFU-Pickatlas (Maldjian, Laurienti, Kraft, & Burdette, 2003). For the face-localizer, preprocessing and 1st-level statistics were analogous to the preference bias data. We localized regions in the fusiform gyrus showing more activation for faces compared to objects (at a voxel-wise threshold of p < 0.001, uncorrected). Three out of eighteen participants were discarded due to the absence of activations in the fusiform area (although they were included in the whole brain analysis). Mean FFA localization was x, y, z = −40, −53, −10; x, y, z = −20, 10, −50; −22 for the left and right FFA, respectively. To measure FFA modulation in happy vs. neutral facial expressions, we measured the peak of percent signal change within each participant-specific ROIs.

2.7. Functional connectivity analysis

We conducted a psychophysiological interaction (PPI; Friston et al., 1997) analysis to examine FFA connectivity during the nonconsciously processing of facial expressions in the static and dynamic conditions. For each participant, we extracted the time series of activity from a 5 mm radius sphere centered on the individual peak of activity in the FFA clusters, as previously defined by the face-localizer analysis. The static and the dynamic psychological variables were constructed for each subject as [HA-NE]stat and [HA-NE]dyn, respectively. We constructed two separate GLM with the time series of activity in the FFA as a first regressor, and as second and third regressors, the static or dynamic psychological variable and its interaction with the time series. Separately for the static and the dynamic conditions, we entered the estimated parameters for the interaction regressor in a between-subject, random-effect analysis for obtaining statistical parametric maps. We identified brain activations exhibiting significant contrasts of parameter estimates with a voxel-wise threshold (p < 0.001, uncorrected), and a cluster-wise threshold of p < 0.05, FWE-corrected for multiple comparisons within the search volume. This small volume correction was conducted relatively to the a priori hypothesis derived from the main effect of static facial expressions processing. The search volume was created from the AAL brain atlas (Tzourio-Mazoyer et al., 2002) in order to include the amygdala (dilated 5 ×). It was applied to the SPM dataset using WFU-Pickatlas (Maldjian et al., 2003).

3. Results and discussion

3.1. Behavioral results

Consistently with our previous studies (Kouider et al., 2011; Faivre et al., 2012), we found that the pleasant response rate (mean = 52.4%, SD = 14.9) varied significantly as a function of the facial expression, with higher ratings induced by happy compared to neutral faces [7.4%; t17 = 2.20; p < 0.05], regardless of the static or dynamic nature of the face (F < 1) (Fig. 1B). Participants confessed seeing only a grayish, jumbled collection of pictures at the bottom of the screen, and being unable to discriminate a central face among the flankers. Signal detection analysis revealed chance-level performance for all expression types (mean d′ = 0.19; t17 = 1.15, p = 0.26) regardless of their static or dynamic nature (t17 = 1.37, p = 0.20).

3.2. Whole-brain results for static faces

We first focused on the whole brain activity elicited by static facial expressions, by comparing the influences of happy (HA) vs. neutral (NE) faces on the preference judgment task. We reported all activations tested at a threshold of p < 0.001 (uncorrected), with an additional constraint of contra-lateral activations at a threshold of p < 0.005 (uncorrected). Importantly, all regions showing significant effects reflected increased activity only for the HA in comparison to the NE condition, but not in the converse direction (see Fig. 2 and Table 1). At the subcortical level, we found activations in the right amygdala (cluster-wise threshold p < 0.01 FWE-corrected for the search volume). The left amygdala was activated with a lower threshold of p < 0.002 (cluster-wise threshold p < 0.05 FWE-corrected for the search volume). This structure has been argued to be crucial for the emotional processing of both negative and positive facial expressions (Costafreda et al., 2008). In addition, we observed bilateral activations in the
hippocampus, which has been linked to the contextual processing of emotional faces (Whalen et al., 1998), and in the left and right caudate nuclei, which are involved in response selection during goal-directed action (Chaudhry, Parkinson, Hinton, Owen, & Roberts, 2009). At the cortical level, activations were observed exclusively in temporal and frontal lobes. For frontal regions, we found bilateral activation in the middle orbitofrontal cortex and in the anterior cingulate cortex. These regions are arguably involved in mechanisms of response selection (see Section 4). For temporal regions, activations were found in the bilateral inferior temporal and fusiform cortex, including the fusiform face area (FFA, see below). The FFA is described as preferentially encoding both facial identity (Kanwisher & Yovel, 2006) and facial expressions like happiness (Breiter et al., 1996; Ganel, Valyear, Goshen-Gottstein, & Goodale, 2005) or fear (Winston, Vuilleumier, & Dolan, 2003). This is to our knowledge the first evidence for a modulation of the FFA by nonconscious facial expressions. In addition, the right superior temporal sulcus (STS), previously associated with the nonconscious perception of facial identity (Kouider, Eger, Dolan, & Henson, 2009) and facial expressions (Jiang & He, 2006), was also observed here although with a more lenient threshold of \( p < 0.004 \).

**Table 1**

Cluster extent (CE), statistical values, and MNI stereotaxic brain atlas coordinates for the brain regions activated by viewing happy compared to neutral crowded faces in the static and dynamic conditions.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Anatomy</th>
<th>CE</th>
<th>Z score</th>
<th>( p ) threshold</th>
<th>( x )</th>
<th>( y )</th>
<th>( z )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Static faces</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left hippocampus</td>
<td>52</td>
<td>3.78</td>
<td>0.001</td>
<td>–24</td>
<td>–28</td>
<td>–8</td>
<td></td>
</tr>
<tr>
<td>Right hippocampus</td>
<td>128</td>
<td>3.79</td>
<td>0.001</td>
<td>20</td>
<td>–10</td>
<td>–18</td>
<td></td>
</tr>
<tr>
<td>Right amygdala</td>
<td>128</td>
<td>3.79</td>
<td>0.001</td>
<td>30</td>
<td>0</td>
<td>–20</td>
<td></td>
</tr>
<tr>
<td>Left amygdala</td>
<td>24</td>
<td>3.20</td>
<td>0.002</td>
<td>–32</td>
<td>–6</td>
<td>–16</td>
<td></td>
</tr>
<tr>
<td>Right caudate nucleus</td>
<td>8</td>
<td>3.06</td>
<td>0.004</td>
<td>20</td>
<td>–10</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Left caudate nucleus</td>
<td>7</td>
<td>3.92</td>
<td>0.001</td>
<td>–6</td>
<td>–18</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Right middle orbitofrontal cortex</td>
<td>26</td>
<td>3.87</td>
<td>0.001</td>
<td>14</td>
<td>32</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Left middle orbitofrontal cortex</td>
<td>12</td>
<td>3.49</td>
<td>0.001</td>
<td>–12</td>
<td>32</td>
<td>–10</td>
<td></td>
</tr>
<tr>
<td>Right anterior cingulate cortex</td>
<td>51</td>
<td>3.77</td>
<td>0.001</td>
<td>20</td>
<td>26</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Left middle cingulate cortex</td>
<td>15</td>
<td>2.94</td>
<td>0.003</td>
<td>–8</td>
<td>–2</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>Right fusiform cortex</td>
<td>8</td>
<td>4.04</td>
<td>0.001</td>
<td>28</td>
<td>–44</td>
<td>–4</td>
<td></td>
</tr>
<tr>
<td>Left fusiform cortex</td>
<td>11</td>
<td>2.89</td>
<td>0.004</td>
<td>–26</td>
<td>–44</td>
<td>–12</td>
<td></td>
</tr>
<tr>
<td>Right superior temporal sulcus</td>
<td>26</td>
<td>2.86</td>
<td>0.004</td>
<td>46</td>
<td>–62</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Right inferior temporal cortex</td>
<td>92</td>
<td>2.82</td>
<td>0.002</td>
<td>42</td>
<td>0</td>
<td>–38</td>
<td></td>
</tr>
<tr>
<td>Left inferior temporal cortex</td>
<td>8</td>
<td>3.02</td>
<td>0.001</td>
<td>–52</td>
<td>4</td>
<td>–40</td>
<td></td>
</tr>
<tr>
<td>Right precentral cortex</td>
<td>8</td>
<td>3.41</td>
<td>0.001</td>
<td>42</td>
<td>–2</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td><strong>Dynamic faces</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left hippocampus</td>
<td>41</td>
<td>3.29</td>
<td>0.002</td>
<td>–28</td>
<td>–30</td>
<td>–8</td>
<td></td>
</tr>
<tr>
<td>Right substantia innominata</td>
<td>10</td>
<td>3.30</td>
<td>0.001</td>
<td>14</td>
<td>–4</td>
<td>–8</td>
<td></td>
</tr>
<tr>
<td>Left substantia innominata</td>
<td>8</td>
<td>2.83</td>
<td>0.005</td>
<td>–16</td>
<td>–2</td>
<td>–8</td>
<td></td>
</tr>
<tr>
<td>Right middle orbitofrontal cortex</td>
<td>9</td>
<td>2.84</td>
<td>0.004</td>
<td>34</td>
<td>22</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Left supramedial frontal cortex</td>
<td>21</td>
<td>3.89</td>
<td>0.001</td>
<td>–12</td>
<td>60</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Right posterior parietal cortex</td>
<td>10</td>
<td>3.13</td>
<td>0.001</td>
<td>–2</td>
<td>–56</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Left posterior parietal cortex</td>
<td>19</td>
<td>2.75</td>
<td>0.003</td>
<td>2</td>
<td>–54</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2. Whole-brain functional results. Sagittal, axial and coronal sections showing clusters of activity for HA compared to NE facial expression, in the static condition (top, including parahippocampic regions, amygdala, right superior temporal sulcus and bilateral orbitofrontal cortices), and dynamic condition (bottom, including left hippocampus, bilateral substantia innominata and left posterior parietal cortex). An uncorrected threshold of \( p = 0.005 \) was used to display the contrasts.
3.3. Whole-brain results for dynamic faces

We then turned to the same contrast of HA compared to NE faces but now for dynamic expressions. This contrast revealed activity increases in similar regions as for the static faces, including the left hippocampus ($p<0.002$) and right orbitofrontal cortex ($p<0.004$). Several regions observed in the static condition were not found here, including the caudate nucleus and the amygdala at the subcortical level, and the cingulate and fusiform cortices and FFA at the cortical level. Interestingly, the absence of amygdala activity during the processing of dynamic emotions was compensated here by bilateral activity in the substantia innominate, a subcortical structure often considered as part of the dorsal amygdaloid complex (Sah, Faber, Lopez De Armentia, & Power, 2003). Furthermore, we found additional regions activated specifically by dynamic facial expressions, including the bilateral posterior parietal cortex, and with a more lenient threshold, the left superior medial frontal cortex and the left supra marginal cortex (see Fig. 2 and Table 1).

3.4. Interaction between static and dynamic conditions

At first sight, these results suggest that despite crowding, static and dynamic emotional faces are processed through distinct neural pathways. However, it is of note that following a whole brain analysis strategy, we did not have sufficient statistical power to evaluate the interaction between stimulus type (i.e., static or dynamic face) and stimulus valence (i.e., happy or neutral face). Thus, in order to comfort our observation, we carried out a region of interest analysis for the dynamic condition on the regions obtained from the whole brain analysis for the static condition. In the anatomical cluster corresponding to the amygdala, we ran a $2 \times 2 \times 2$ ANOVA with stimulus valence (happy vs. angry), lateralization (right vs. left hemisphere) and stimulus type (static vs. dynamic faces) as within-subjects factors. We found a significant interaction between stimulus valence and stimulus type ($F(1,17)=23.13$, $p<0.001$), due to an absence of amygdala activation by dynamic expressions ($p>0.6$). For exploratory purposes, we also verified that no effect was found in the fusiform cortex, the inferior temporal cortex, the cingulate cortex, and the right superior temporal sulcus in the dynamic condition (all $p>0.2$). In addition, we carried out a region of interest analysis for the static condition on the regions obtained from the whole brain analysis for the dynamic faces. No effect was found in the substantia innominate, nor in the posterior parietal cortex (both $p>0.6$).

3.5. Functional region of interest in the FFA

The absence of FFA activity in the dynamic condition might have resulted from a whole-brain analysis strategy, which fails to take into account the interindividual anatomic variability of this kind of structures (Saxe, Brett, & Kanwisher, 2006). To address this issue, we relied here on the use of functional regions of interest. Using a functional localizer, we defined the FFA for each participant in a separate scan session following the main experiment (see Section 2), and measured how this functional cluster was modulated by facial expressions. We ran a $2 \times 2 \times 2$ ANOVA with stimulus valence (happy vs. angry), lateralization (right vs. left hemisphere) and stimulus type (static vs. dynamic faces) as within-subjects factors. This analysis revealed a significant interaction of stimulus valence with stimulus type ($F(1,13)=10.50$, $p<0.01$), with a main effect of valence in both left and right FFA for static faces [left hemisphere: $parameter \ estimate=0.82$; $t_{14}=3.46$; $p<0.005$; right hemisphere: $parameter \ estimate=0.85$; $t_{14}=3.29$; $p<0.01$], but not dynamic faces [left hemisphere: $parameter \ estimate=-0.03$; $t_{14}=0.09$; right hemisphere: $parameter \ estimate=-0.28$; $t_{14}=-0.80$; $p<0.43$] (see Fig. 3). This result suggests that the FFA modulation is specific to static faces. Furthermore, we computed a linear regression between the preference bias (as calculated by subtracting the pleasant response rate in the happy vs. neutral condition) and the activity extracted from the FFA. Consistent with the specific role of the FFA in the nonconscious processing of static facial expressions, we found that the increase of FFA activity was correlated with the behavioral bias on preference judgment in the static (adjusted $R^2=0.24$; $p<0.05$), but not dynamic stimulation condition (adjusted $R^2=-0.07$; $p=0.9$). It is of note that our FFA localizer session was performed with static faces (see Section 2) and thus one might argue that the absence of FFA modulations for dynamic facial expressions reflects a mislocalization. However, since brain regions revealed by static and dynamic localizers largely overlap (Fox, Iaria, & Barton, 2009; Schultz & Pilz, 2009), these results rather suggest, as we discuss below, that dynamic stimuli are processed preferentially along dorsal rather than ventral visual pathways.

3.6. Functional connectivity between the FFA and the amygdala

In the static facial expression condition, the PPI analysis revealed a significant functional connectivity between the left FFA and the left amygdala (17 voxels, $x$, $y$, $z=-20$, $-2$, $-20$, voxel-wise threshold $Z=3.25$, $p<0.001$ uncorrected; cluster-wise threshold $p<0.05$ FWE-corrected for the search volume), and

![Fig. 3. Region of interest functional results. Neural activity in the left and right fusiform face area (FFA) induced by happy compared to neutral faces, in the static and dynamic stimulation conditions ($^*p<0.01$).]
between the right FFA and the right amygdala (18 voxels, x, y, Z = 32, 2, -20, voxel-wise threshold Z = 3.25, p < 0.001 uncorrected; cluster-wise threshold p < 0.01 FWE-corrected for the search volume). Importantly, this pattern of results suggests that beyond simple co-activation, the FFA interacts with the amygdala during the nonconscious processing of static facial expressions. It is of note however that the PPI analysis does not provide any clues regarding the direction of these interactions (i.e., from the FFA to the amygdala or from the amygdala to the FFA, see Section 4). As expected with the previous results we obtained, we found that this pattern of functional connectivity interacted with stimulus type (t(4) = 3.35, p < 0.005), as functional connectivity between the FFA and the amygdala was nonexistent in the dynamic condition (functional connectivity between left FFA and left amygdala: Z = 0.09; between right FFA and right amygdala: Z = -1.25). Here, a functional localizer of dynamic faces would be necessary in order to infer which brain regions jointly process crowded emotional videos.

4. Discussion

In this study, we assessed which brain regions are involved during the nonconscious processing of crowded facial expressions. Our results revealed the existence of two patterns of activation depending on the static (i.e., pictures) or dynamic (i.e., videos) nature of facial expressions. For static faces, we found activations in the bilateral amygdalae, in coherence with previous studies using visual masking (Killgore & Yurgelun-Todd, 2004; however, see Whalen et al., 1998 for a decrease in activity compared to fixation) and binocular rivalry (Williams et al., 2004). Interestingly, dynamic facial expressions did not involve the amygdala, but rather the substantia innominata, a sublenticular structure of the basal forebrain that is contiguous with the dorsal amygdala. This region has previously been linked to the conscious perception of facial familiarity (Wright et al., 2003) and facial expression, with a higher sensitivity to emotional arousal rather than emotional valence (Breiter et al., 1996; Kim, Somerville, Johnstone, Alexander, & Whalen, 2003; Whalen et al., 1998).

How are these subcortical structures involved in emotional processing? On the one side, it has been suggested that the amygdaloid complex can be triggered without relying on any visual cortical relay, through a short subcortical route involving superior colliculus and inferior pulvinar (LeDoux, 1998). This hypothesis has been recently comforted by the finding of anatomical connections between the superior colliculus, pulvinar and amygdala in humans (Tamietto, Pulless, de Gelder, Weiskrantz, & Goebel, 2012). Invisible facial expressions are thus considered to bypass striate and extrastriate areas along the ventral visual pathways. In this context, the observation of FFA increase for static facial expressions could reflect feedback from the amygdala, rather than a direct stimulation from visual cortical regions (Dolan, Morris, & de Gelder, 2001). Yet, this classical view, which is primarily based on data obtained from rodents during auditory negative conditioning, has been criticized for being generalized to nonconscious fear perception in the presence of low evidence (Morris et al., 1999; Ohman, Carlsson, Lundqvist, & Ingvart, 2007). On the other side, a recent alternative has challenged the role of a direct subcortical pathway for the processing of emotions (Pessoa & Adolphs, 2010). The authors propose instead that emotional stimuli are processed rapidly along parallel extrastriate cortical routes before triggering the amygdaloid complex. In this framework, the amygdala is triggered by cortical inputs, after elaborated visual processing has taken place in ventral visual regions, and plays a broader role, notably in recognizing the social value of stimuli (Tsuchiya, Moradi, Felsen, Yamazaki, & Adolphs, 2009).

One could argue that since both static and dynamic stimuli were presented peripherally, it is likely that low spatial frequency information was at the origin of the effects we found, suggesting its subcortical origin (Winston et al., 2003). However, it has been shown that magnocellular pathways conveying low spatial frequency information also reach the early visual cortex and then project to parietal and frontal cortices (Bullier, 2004), conveying coarse information about the gist of a visual scene, here probably rough aspects of the facial expressions (Kveraga, Boshyan, & Bar, 2007).

Now regarding the cortical patterns of activation elicited by static facial expressions, our results revealed a main locus of activations in ventral visual pathways, including the FFA. PPI analysis revealed that the FFA was functionally connected with the amygdala. Furthermore, we measured a positive correlation between the modulation amplitude of FFA activity and the amplitude of the behavioral preference bias. From a methodological point of view, this finding contrasts with the absence of effect reported by a previous study using fearful static faces suppressed from awareness by continuous flash suppression (CFS), a variant of binocular rivalry allowing for sustained periods of invisibility (Jiang & He, 2006). Although it remains possible that this discrepancy results from different types of emotional valence between the two studies (i.e., happiness here vs. fear in Jiang & He), it is likely that ventral visual regions, such as the FFA, are less sensitive, or arguably not sensitive at all to suppressed signals in CFS (Almeida, Mahon, Nakayama, & Caramazza, 2008; Lin & He, 2009). In this respect, the Gaze-Contingent Crowding approach seems to overcome this limitation. Indeed, the involvement of ventral regions revealed by the present study corroborates with a previous behavioral finding showing that crowded faces are processed holistically (Kouider et al., 2011), a cognitive operation implying the FFA (Kanwisher & Vowel, 2006). Thus, we suggest that the cortical processing of emotional information resists crowding. To sum up, we propose that despite crowding, static expressions are processed along the ventral visual pathways (i.e., temporal regions, including FFA), trigger parahippocampic regions (including hippocampus, caudate nucleus and amygdala) connected to the prefrontal areas (including cingulate and orbitofrontal cortex) underlying preference judgment (see Fig. 4).

For dynamic facial expressions, the cortical pattern of activations was instead restricted to dorsal regions, including the posterior parietal cortex, which is functionally specialized for the processes of spatially guided behavior and biological motion imagery (Cavanna & Trimble, 2006). Additionally, we found a marginally significant effect in the left supramarginal gyrus, a dorsal structure whose lesions impair judging emotional faces (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000) and body motion (Heberlein, Adolphs, Tranel, & Damasio, 2004). We propose that these regions trigger subcortical structures coding for stimulus arousal (i.e., substantia innominata) which finally orient the preference judgment through orbitofrontal cortex modulation (see Fig. 4).

As described above, unexpectedly, several regions elicited by static faces were not elicited by dynamic faces, including the STS, usually associated with the conscious processing of various biological motions (Puce & Perrett, 2003). According to two influential models of face perception (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000), invariant aspects of faces like identity are encoded in the face-responsive regions of the fusiform gyrus (including the FFA), while changeable aspects like facial expressions primarily activate the face-responsive regions of the temporal sulcus (including the STS). Along this line, a recent study showed that dynamic faces do not produce a greater response than static faces in the STS, but not in the FFA (Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011). Contradictory findings...
show that compared to static faces, dynamic faces expressing happiness are associated with greater activity in the fusiform gyrus (Sato et al., 2004), and in the STS and FFA (Trautmann et al., 2009; Kessler et al., 2011). At first sight, these results dealing with conscious perception might seem at odd with the ones we report, and with the fact that temporal cues play an important role in the nonconscious processing of changeable aspects of faces like emotional expressions. A possible explanation is that in the presence of facial motion information, the dorsal pathway is triggered in the first place following a feed-forward sweep of information processing, while the ventral visual pathway is triggered afterward, through top-down mechanisms taking place only when the stimuli are consciously visible (Lamme & Roelfsema, 2000). Even if distinctions in terms of neuroanatomy and functional dissociation are less pure than originally considered (Farivar, Blanke, & Chaudhuri, 2009), the pattern of results we observed questions the capacity to integrate form information (i.e., ‘snapshots’ of expressive faces by neurons in the ventral pathway) and motion information (i.e., integrated optic flow patterns by neurons in the dorsal pathway) without awareness, as proposed elsewhere to account for the conscious recognition of biological motion (Giese & Poggio, 2003).

As concerns the mechanisms underlying preference judgment, we propose two non-exclusive mechanisms to account for the nonconscious behavioral bias we observed. The first one is the building of implicit associations between the valence of facial expressions and the subsequent preference judgment. Indeed, we found that regardless of their static or dynamic nature, both stimulus types were associated with hippocampic activations, classically described as mediating the contextual evaluation of emotional stimuli, and together with the amygdaloid complex, as mediating implicit learning for visible or invisible stimuli (Morris, Ohman, & Dolan, 1998). The second mechanism is a modulation of the implicit aesthetic preferences computation, which is described as relying on the activity of cortical and subcortical areas implicated in the processing of reward. In agreement, we found that both static and dynamic happy faces preferentially triggered orbitofrontal cortex (bilateral activations with static faces, right activation with dynamic faces), a region which is known among the primates to comprise face-selective neurons activated by temporal cortical inputs (Thorpe, Rolls, & Maddison, 1983) to encode face expression and mediate related social reinforcement (Rolls, Critchley, Browning, & Inoue, 2006), but also to process aesthetic beauty in its medial part (Kawabata & Zeki, 2004) (Fig. 5).
5. Conclusion

This study conjoined two important findings. First, taking advantage of gaze-contingent crowding, a novel approach for studying undiscriminable and long-lasting stimuli, we were able to demonstrate that emotional pictures and emotional videos were processed along different neural pathways. Referring to the dual-stream theoretical framework (Goodale & Milner, 1992), we proposed that static facial expressions were processed through the ventral visual pathway, including the FFA which was functionally connected with the amygdala. We observed that the same emotions expressed dynamically were processed through the dorsal visual pathway, including the posterior parietal cortex, together with dorsal subcortical structures such as the substantia innominata. These results concerning the nonconscious processing of biological motion question the existence of form-motion integration (Giese & Poggio, 2003), and more generally of nonconscious temporal integration, the capacity for the visual system to bind together dynamic stimuli that are invisible. Second, we found that undiscriminable facial expressions were encoded such as to bias a high-level cognitive function such as preference judgment. We proposed that the bias arose from both evaluation of the emotional context and aesthetic preferences computation, respectively supported by the hippocampus and the orbitofrontal cortex, which were activated by static and dynamic facial expressions. This shows that crowded information can modulate a broad set of brain regions supporting high-level cognitive functions, despite being reflected subjectively as a jumbled, uninformative experience.

Acknowledgments

The authors thank two anonymous reviewers for their comments, and Eric Bertasi and Romain Valabregue for their help with data acquisition. This work was supported by a grant from the French Ministry for Higher Education and Research to NF, and grants from the Agence Nationale de la Recherche (ANR) and the European Research Council (ERC) to SK.

Reference


