Dissimilar processing of emotional facial expressions in human and monkey temporal cortex

Qi Zhu [a,1], Koen Nelissen [a,b,1], Jan Van den Stock [c,d,1], François-Laurent De Winter [d], Karl Pauwels [a], Beatrice de Gelder [b,c,d], Wim Vanduffel [a,b,*], Mathieu Vandenbulcke [d]

[a] Laboratory for Neuro- and Psychophysiology, KU Leuven, Leuven, Belgium
[b] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA, USA
[c] Cognitive and Affective Neuroscience Laboratory, Tilburg University, Tilburg, The Netherlands
[d] Brain and Emotion Laboratory Leuven (BELL), Division of Psychiatry, Department of Neuroscience, KU Leuven, Leuven, Belgium

ARTICLE INFO

ABSTRACT

Emotional facial expressions play an important role in social communication across primates. Despite major progress made in our understanding of categorical information processing such as for objects and faces, little is known, however, about how the primate brain evolved to process emotional cues. In this study, we used functional magnetic resonance imaging (fMRI) to compare the processing of emotional facial expressions between monkeys and humans. We used a $2 \times 2 \times 2$ factorial design with species (human and monkey), expression (fear and chewing) and configuration (intact versus scrambled) as factors. At the whole brain level, neural responses to conspecific emotional expressions were anatomically confined to the superior temporal sulcus (STS) in humans. Within the human STS, we found functional subdivisions with a face-selective right posterior STS area that also responded to emotional expressions of other species and a more anterior area in the right middle STS that responded specifically to human emotions. Hence, we argue that the latter region does not show a mere emotion-dependent modulation of activity but is primarily driven by human emotional facial expressions. Conversely, in monkeys, emotional responses appeared in earlier visual cortex and outside face-selective regions in inferior temporal cortex that responded also to multiple visual categories. Within monkey IT, we also found areas that were more responsive to conspecific than to non-conspecific emotional expressions but these responses were not as specific as in human middle STS. Overall, our results indicate that human STS may have developed unique properties to deal with social cues such as emotional expressions.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Research on emotional facial expressions in non-human primates has often attracted scientists because it opens an evolutionary window on emotions and social perception in humans (de Gelder, 2010; de Waal, 2011; Parr and Heintz, 2009; Parr et al., 2005, 2008). Since the advent of functional neuroimaging, facial expressions have been the favorite stimulus class for studying emotion processing in the human brain and insights from animal research have strongly influenced the interpretation of findings in humans. However, in contrast with the large literature of comparative studies on the processing of categorical information (Bell et al., 2009; Pinsk et al., 2009; Rajimehr et al., 2009; Tsao et al., 2003, 2008a), a direct comparison of processing emotional expressions between species has not been reported yet and it remains largely speculative how the primate

brain evolved to deal with emotional cues (Ghazanfar and Santos, 2004). During evolution the repertoire of facial displays evolved in parallel with species-specific social interactions (Burrows et al., 2009; Parr et al., 2005). Hence, although many aspects of processing emotional expressions may be conserved across primate species, the differences between humans and monkeys may primarily be reflected in neural pathways involved in social cognitive processes such as attributing meaning to other's mental states (Brothers, 1989; Joffe and Dunbar, 1997; Parr et al., 2005).

Neural correlates of emotional facial expressions have been reported in humans and monkeys separately. However, the limited number of studies in monkeys hampers a comparison based on the existing neuroimaging literature. Emotion effects in monkeys include activation of face selective ventral prefrontal areas (Tsao et al., 2008b), amygdala (Hoffman et al., 2007), and modulatory effects in non-face-selective inferotemporal cortex (Hadj-Bouziane et al., 2008). In humans, orbitofrontal cortex and amygdala also respond to emotional expressions and are thought to be involved in more basic species-independent emotion operations such as control processes and decoding valence or saliency (Dolan, 2002; Rolls, 2004).

* Corresponding author at: Bldg 149, 13th Street, Charlestown, MA 02129, USA. Fax: +1 6177267422.
   E-mail address: wim@nmr.mgh.harvard.edu (W. Vanduffel).
[1] These authors contributed equally to this work.

Similar to the effects in monkey IT, emotion-dependent activity changes in human ventral temporal occipital face areas are generally interpreted as modulatory effects, as supported by lesion studies of the amygdala (Vuilleumier et al., 2004). In addition, human neuroimaging studies repeatedly documented emotion effects in the superior temporal sulcus (STS). The human STS is not only implicated in processing visual information, including variable facial information such as gaze or expressions (Graham and LaBar, 2012), but also in modality-independent higher order social cognitive functions (Allison et al., 2000; Hein and Knight, 2008; Kujala et al., 2009). Given its proposed role as an interface between perception and more complex social cognitive processes, we considered the STS as a candidate region for human-specific facial emotion effects.

To compare directly the processing of facial emotion cues between species, we used event-related fMRI in monkeys (Vanduffel et al., 2001) and humans with an identical $2 \times 2 \times 2$ factorial design with *dynamic facial expression* (fear and chewing), *species* (human and monkey) and *configuration* (intact versus mosaic scrambled) as factors (Fig. 1). To stay as close as possible to naturalistic conditions, we used dynamic faces. We chose fear as emotional condition because this is the most widely-studied expression in neuroimaging studies of each species separately. Videos of chewing faces served as neutral controls and videos of scrambled faces were used to control for the low-level effects such as motion (Puce et al., 1998). Because the interpretation of emotional expressions is largely species-specific (Hebb, 1946), we took advantage of our factorial design to study which areas responded preferentially to conspecific emotional expressions by contrasting them with heterospecific expressions in both species. Furthermore, to relate our findings anatomically to face-*selective* regions, an independent localizer experiment was also conducted in both species.

Methods

Subjects

Three healthy male rhesus monkeys (M18, M19 and M20; 5–7 kg, 4–5 years old) and twenty-three normal human volunteers (11 male, 24–34 years old, all right-handed and had normal or corrected-to-normal visual acuity) were scanned for the dynamic facial expression experiment. Two of the three monkeys and seven human volunteers (3 male, all right-handed, 23–32 years old) were scanned in the separate localizer experiment. All human participants gave written informed consent in accordance with the Declaration of Helsinki. The ethical committee of the University of Leuven Medical School approved the experiments.

Stimuli

Twenty-four movie clips, acquired from six unfamiliar professional male human actors and six male monkeys, were used for each type of expressions (twelve for each species) in the dynamic facial expression experiment. All dynamic facial expression stimuli were frontal view color movie clips, with the external face contour removed and the mean luminance (9 cd/m^2) equalized (Fig. 1A). The expressions were all gaze-averted but with heads fixed. We chose averted gaze, because unlike similar grimaces in humans, the direct-gaze, teeth-baring expressions of rhesus macaques signal submission towards the observer (de Waal and Luttrell, 1985; Maestripieri and Wallen, 1997). To control for the eye-gaze direction, head orientation and movement asymmetries, the mirror-reversed version of each movie clip was also created. The spatiotemporally scrambled control stimuli were generated from each
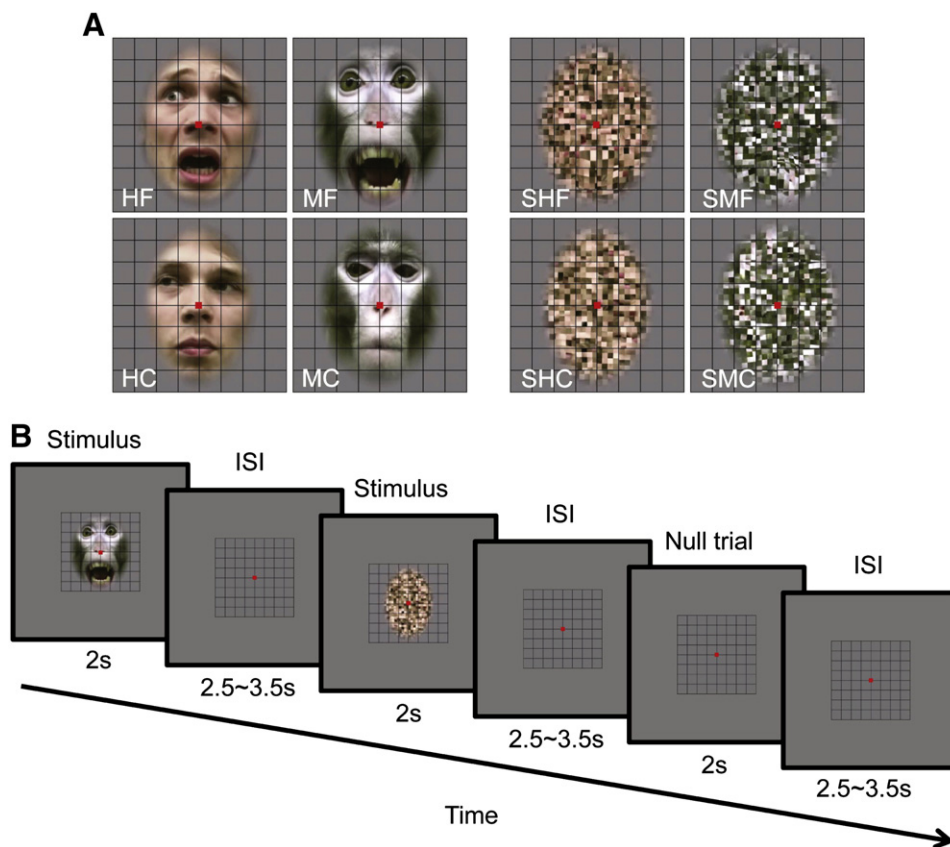
Fig. 1. Stimuli and experimental paradigm. (A) Upper row left panels: intact human *fearful* (HF) and monkey *fearful* (MF) expressions; upper row right: scrambled versions of HF (SHF) and MF (SMF). Lower row left panels: intact human *chewing* (HC) and monkey *chewing* (MC); lower row right panels: scrambled versions of HC (SHC) and MC (SMC). Examples of dynamic displays are provided in supplementary videos 1 to 8. (B) Event-related experimental design. Trials consisted of 2 s stimulus presentation followed by a variable interstimulus interval (ISI) between 2.5 and 3.5 s.

dynamic face video, by applying the temporally scrambled flow field of that movie clip to the mosaic-scrambled start image of the original sequence (Fig. 1A) (for further details on stimulus construction and selection, see supplementary method). The mosaic scrambling was accomplished by dividing the image into a 32×32 grid and shuffling the positions of the grid elements. The flow field of the original movie clips was calculated using an optic flow estimation algorithm developed by Papenberg et al. (2006), then temporally scrambled by spatially dividing the flow field into an 8×8 grid (as shown by the grid lines presented in both the intact and scrambled stimuli, Fig. 1A) and shuffling the frames differently for each grid across temporal blocks with five frames for each block. This way of temporal shuffling completely destroyed the facial expression actions, but kept as much the low-level motion information in the scrambled stimuli as in the original videos, except that the maximum range of motion was restricted to the size of each grid in the scrambled stimuli. It needs to be noted that without temporal shuffling, human subjects clearly recognized the type of expressions in the scrambled videos, hence we chose not to control for the maximum range of motion.

For the localizer experiment, ten object categories, each containing 20 static monochrome images, were presented to both humans and monkeys during the scanning (supplementary Fig. 1). These categories include human and monkey faces, headless human and monkey bodies, inanimate objects with two different aspect ratios (3.09 for objects H and 1.55 for objects M, supplementary Fig. 1), animals, birds, fruits and sculptures. All stimuli were matched in area and mean luminance and were embedded in a random-noise background with the noise being of the same spatial frequency, power spectrum, and mean luminance as the images.

Experimental design and procedure

Dynamic facial expression experiment

We used an event-related design, and stimuli were presented in ten different orders (each run was 550 s long). In each order, every movie clip was presented once for 2 s, followed by a 2.5 s to 3.5 s inter-stimulus interval displaying only the grid (Fig. 1B). Twelve null trials with the grid presented for 4.5 s to 5.5 s were randomly interspersed. All stimuli were presented at a size of 7×7 degrees of visual angle for both species. A central fixation point (8′) was continuously presented and a passive fixation task was performed. Monkeys received liquid rewards for maintaining fixation within a virtual 2×2 degree window. Before scanning, only a few movie clips were shown to human participants for practice, and another set of static object images, unrelated to the present experiment, was used in monkeys for training. After scanning, 19 human subjects participated in three behavioral experiments to assess the emotional significance of the stimuli presented in the fMRI session. In each of the behavioral experiments, a trial consisted of the presentation of a fixation cross of variable duration (1–3 s), followed by a stimulus (2 s) after which a question mark appeared until the response. In the first experiment, participants were instructed to categorize the emotion expressed in the stimulus in a 6-alternative, forced-choice task (anger, disgust, fear, happy, neutral or sad). In the second and third experiments, participants were instructed to indicate separately the arousal and valence of each stimulus, using the Self-Assessment Manikin test (Bradley and Lang, 1994). Monkey eye positions and pupil diameters were monitored during the fMRI scans using a pupil–corneal reflection tracking system (120 HZ, Iscan). Pupil diameter, considered a viable psychophysiological measure of fear (Sturgeon et al., 1989), was used as an index of behavioral significance of the stimuli in monkeys.

Localizer experiment

Stimuli were presented in an event-related fashion, with each stimulus presented for 500 ms, followed by a 2.5 s to 3.5 s interstimulus interval displaying only the noise background. Between successive trials, the noise background was changed to avoid adaptation to the background. A central fixation point (8′) was continuously presented and a passive fixation task was performed by both humans and monkeys. Monkeys received liquid rewards for maintaining fixation, and the reward frequency was increased as the duration of fixation increased. The stimulus sequences were generated using the M-sequences (Buracas and Boynton, 2002), to counterbalance the order of stimulus presentation. Different sequences were randomly selected from 100 pre-generated sequences and used for different runs in both humans and monkeys. Each run lasted 400 s in both species.

fMRI acquisition

Monkeys were scanned on a 3 T Siemens Trio scanner following standard procedures (Ekstrom et al., 2008; Nelissen et al., 2006; Vanduffel et al., 2001), using an 8-channel monkey coil (TR 2 s, TE 17 ms, flip angle 75°, 40 slices, no gap, 1.25 mm isotropic). Before each scanning session, a contrast agent (MION, or Feraheme 8–11 mg/kg) was injected into the monkey's femoral/saphenous vein. The use of the contrast agent improves the contrast–noise ratio approximately threefold at 3 T (Leite et al., 2002; Vanduffel et al., 2001) and enhanced spatial selectivity of the MR signal changes (Zhao et al., 2006), compared with blood oxygenation level-dependent (BOLD) measurements. For the dynamic facial expression experiment, a total of 163, 112 and 102 runs from 5, 5 and 4 sessions were collected, and 144, 103 and 98 runs were analyzed for monkeys M18, M19 and M20 respectively. Only runs in which the monkeys maintained fixation less than 85% of the time, or runs without pupillary records were excluded from the analysis. For the localizer experiment, a total of 106 and 95 runs from four sessions were collected and analyzed for monkey M18 and M19. High-resolution anatomical images were acquired for each monkey during a separate session under Ketamine/Xylazine anesthesia, using a single radial transmit–receive surface coil and a MPRAGE sequence (TR 2200 ms, TE 4.05 ms, flip angle 13°, 208 slices, 0.4 mm isotropic).

Humans were scanned in a 3 T Philips scanner using an 8-channel head coil and a standard EPI-sequence (TR 2 s, TE 30 ms, flip angle 90°, 40 slices, $2.75 \times 2.75 \times 3.5$ mm^3 voxel size). For the dynamic facial expression experiment, a total of 6 runs were obtained in all except three subjects, from whom 1 to 2 runs were omitted due to technical problems. A high-resolution anatomical volume for each subject was acquired in the middle of each scanning session using a MPRAGE sequence (TR 9.6 ms, TE 4.6 ms, flip angle 8°, 182 slices, $0.98 \times 0.98 \times 1.2$ mm voxel size). For the localizer experiment, data from four sessions each containing 8 runs were acquired from each of the human subjects.

Data analysis

Monkey pupillary response

The horizontal and vertical eye position records were first analyzed using ILAB (Gitelman, 2002) and customized Matlab scripts to determine the periods of stable central gaze. Specifically, eye blinks were detected by ILAB and removed from each eye trace prior to analysis. The same methodology as described by Bair and O'Keefe (1998) was then adopted for detecting and extracting the periods of stable central gaze (within a 2.25×2.25 deg window – which differed from the 2×2 deg fixation window during the fMRI experiments). The velocity threshold for detecting saccades was set to 50 deg/s. For each trial, analysis of the pupil size was restricted to a time window from 500 ms before to 4.5 s after the stimulus onset and only the recordings within the aforementioned central gaze periods were considered as valid. Trials having a proportion of valid recordings lower than 75% were excluded from further analysis. Less than 5% of the trials were excluded on average from each run for each subject based on this criterion. For each session, the percent pupil

diameter changes relative to baseline (the average pupil diameter over the 500 ms preceding stimulus onset) were calculated for each condition and then averaged across sessions. To control for different degrees of initial pupillary light reflex after stimulus onset across conditions, the average degree of pupil constriction was calculated from a time window between 375 ms and 425 ms after stimulus onset for each condition, and then subtracted from the pupillary data. This initial pupil constriction time window was determined based on the group data across all the monkey and sessions, centered at the peak of the pupil constriction. The pupillary response to the movie content was calculated within a window from 375 ms to 2 s after picture onset, for each scan session first, and then values from all the sessions were submitted to the second level group analysis across sessions.

fMRI data analysis

Data were analyzed using Freesurfer and FS-FAST (http://surfer.nmr.mgh.harvard.edu/). The human and monkey data were preprocessed in the same way before being submitted to the GLM analysis, except that the slice-time correction was only conducted in humans, and different FWHM values for spatial smoothing were used in humans (5.5 mm) and monkeys (2.4 mm). For GLM analysis, each condition was modeled by convolving a Gamma function (delta = 2.25, tau = 1.25 and exponent = 2 for humans; delta = 0, tau = 8 and exponent = 0.3 for monkeys) at each trial onset over the duration of 2 s reflecting the length of one trial. Trials during which monkeys aborted the fixation were treated as the fixation condition and two extra covariates that were generated from the eye movement traces and the reward schedules were used in monkeys as regressors-of-no-interest. For group analysis, individual human data were resampled to Talairach space using the standard linear Talairach transformation (Fischl et al., 1999), and individual session monkey data were warped to M18's anatomical space using a non-linear transformation in JIP software (http://www.nitrc.org/projects/jip) (Mandeville et al., 2011). A random-effect group analysis (across subjects for humans, and across sessions for monkeys) for the dynamic facial expression experiment and a fixed-effect group analysis for the localizer experiment were conducted in both species (with a cluster-wise correction for multiple comparisons, 10,000 Monte Carlo simulations). The significance maps from the group analysis were projected onto the flattened cortical surface of fsaverage in humans and the M18's surface in monkeys for display.

To plot the profiles of the activated regions, ROIs were selected based on the group activation maps from the dynamic facial expression experiment (p < 0.05, corrected), and then projected back, for each subject in humans, or for each session in monkeys. For the dynamic facial expression experiment, the profiles of these regions are only shown for illustrative purposes to show the amplitude of the fMRI responses in the local maximum. The middle part of the right STS (rSTSm) in humans and area TE and rML in monkeys were defined based on the 3-way interaction between species, expression and configuration. The left anterior inferior temporal cortex (lAIT) ROI in monkeys was defined based on the activation for monkey fearful expressions (fear versus chewing controlled for the activations for scrambled faces). All these ROIs were defined as a cubic volume (3 × 3 × 3 voxels) around the peak activation of each region. The posterior portion of the right superior temporal sulcus (rSTSp) in humans was defined based on the overlap between the emotion effect of human and monkey fearful faces (compared to chewing and controlled for the activations for scrambled faces), therefore to avoid a bias to either of the human or monkey emotion effect, we delineated a same size (27 voxels) cubic volume ROI around the geometric center of the conjoined activation. For the amygdala, ROIs were defined based on the contrast faces versus scrambled faces, and included all the activated voxels at a threshold of p < 0.0005, uncorrected for multiple comparisons. The percent signal change

was calculated relative to fixation and averaged across all voxels within each ROI for each subject in humans or each session in monkeys separately, and then submitted to a second-level random-effect group analysis. For the responses to dynamic facial expressions, within-subject ANOVAs were performed. For the responses to object categories in the localizer experiment, a Wilcoxon signed rank test was performed due to small number of subjects. To facilitate comparisons with BOLD, the sign of the MION percent signal changes was reversed.

Results

Behavioral results

For human subjects, fearful faces of both species were more arousing and their valence was rated more negatively than chewing faces (ps < 0.02, paired t-test). A direct comparison of human and monkey fearful faces revealed that human fearful faces were experienced as more arousing (paired t-test, t(18) = 4.11, p < 0.001) and the valence was perceived more negatively than monkey fearful faces (paired t-test, t(18) = 3.76, p < 0.001). Furthermore, we found a two-way interaction between species and expression: human fearful faces relative to chewing were more arousing and more negative than monkey fearful faces relative to chewing (ANOVA, arousal: F(1, 18) = 36.92, p < 0.001; valence: F(1, 18) = 21.47, p < 0.001) (Fig. 2B). Human subjects categorized human fearful faces accurately, but experienced difficulties in distinguishing between fear and anger when rating monkey fearful faces (Fig. 2A).

In monkeys, after an initial phase of pupil constriction in response to stimulus onset, monkey pupils were significantly more dilated in response to fearful faces relative to chewing faces of both monkeys and humans. There was no significant interaction between the pupillary response to human and monkey fearful expression relative to chewing (Fig. 2D). Also no difference was found in fixation performance between different types of faces and expressions (supplementary Fig. 2).

fMRI results

First, we determined all areas in monkeys and humans that responded[2] to dynamic facial expressions of humans and monkeys irrespective of the expression (single main effect of configuration), then we compared the neural processing of emotional expressions. To relate our findings anatomically to face-selective regions, we used black outlines in Figs. 3 and 4 to label those areas responding more strongly (p < 0.05, uncorrected) to all static faces (human and monkey faces) than control objects (objects H and objects M) in the independent localizer experiment.

Neural processing of dynamic faces

In humans, conspecific and heterospecific dynamic facial expressions (red and green, respectively in Fig. 3A), relative to their spatially and temporally scrambled versions, activated a largely overlapping distributed network (yellow in Fig. 3A). Besides face-selective areas (as defined by the contrast between static faces and control objects), the network also included neighboring occipito-temporal cortex, right inferior frontal gyrus, inferior parietal cortex (Fig. 3A) and bilateral amygdala. Stereotactic coordinates of all reported activations in the human brain are listed in supplementary Table 1. In monkeys, the same contrasts, activated bilateral face-selective areas in the upper

---

[2] 'Selective' and 'responsive' are terms quite intensively used in this paper. We make a clear distinction between them: 'selective' refers to differences between expression conditions (fear and chewing) in the dynamic facial expression experiment, or between faces and objects in the localizer experiment (in this case 'face-selective' is used), while 'responsive' refers to any response compared to scrambled faces or fixation.
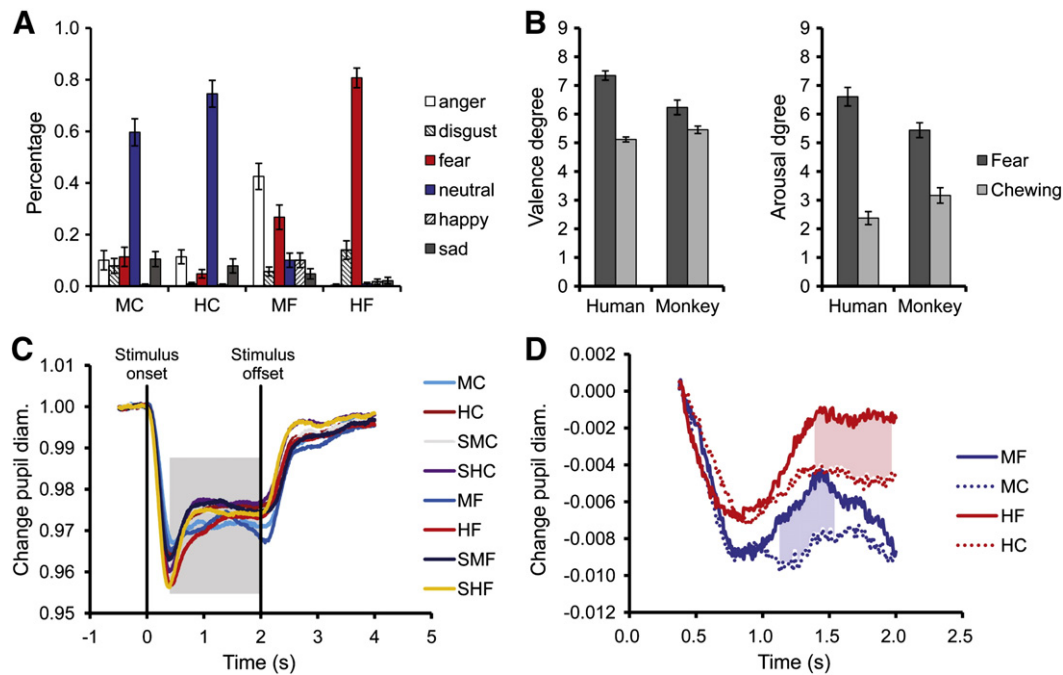
Fig. 2. Behavioral results in humans (A, B) and pupil data in monkeys (C, D). (A) Forced choice categorization of facial expressions (same abbreviations as Fig. 1). Y-axis shows percentage (mean ± s.e.m.) of choices for each category per facial expression. (B) Valence and arousal ratings of facial expressions (mean ± s.e.m.). Scores ranging from 1 to 9 corresponding with highly positive to highly negative valence or lowest to highest arousal. (C) Pupil diameter change (percentage change relative to baseline) per trial for each condition for monkeys. (D) Pupil diameter change between 375 ms and 2000 ms after stimulus-onset (transparent grey time window shown in C) for intact faces relative to scrambled versions. Significant differences between fear and chewing conditions are indicated by blue (monkey) or red (human) shadow (p<0.01). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and lower bank of the STS, with the activity extending posteriorly into TEO and extrastriate areas such as V2, V3 and V4 (Fig. 3B). In addition, prefrontal cortex and left amygdala were face-responsive as well. The right amygdala was responsive to both human and monkey faces at p<0.005, uncorrected for multiple comparisons. Most of this dynamic-face responsive system was conjointly activated by both human and monkey faces (shown in yellow in Fig. 3B).

Neural processing of emotional expressions

We first compared in each species separately the areas that responded selectively to either human or monkey emotional facial expressions, by conducting a two-way interaction between expression and configuration for human faces and monkey faces separately. In humans, human fearful expressions, relative to chewing (controlled for low-level effects such as motion, by using scrambled versions) activated the middle and posterior part of the right STS and upper and lower bank of the posterior part of the left STS (shown in red and yellow in Fig. 4A). Although a similar effect was found in the anterior part of the fusiform gyrus bilaterally, this was relatively weak and only survived when a threshold uncorrected for multiple comparisons was used (supplementary Fig. 3A), or when a direct comparison between fearful and chewing expressions was performed, omitting the scrambled controls (supplementary Fig. 3B). The right posterior STS was the only face-selective region that showed a clear effect of emotional expression (supplementary Fig. 4) and was also the only region that responded to emotional expressions of both species (shown in yellow in Fig. 4A).

In monkeys, bilateral infero-temporal cortex (specifically the convexity of the inferior temporal gyrus) was selective to the fearful expressions of monkeys (shown in green and yellow in Fig. 4A). This effect of emotional expression fell mainly outside the face-selective areas (black outline in Fig. 4A) (activity profiles showing the emotion effect from all face-selective areas can be checked in supplementary Fig. 4). In monkeys, we found that activity in early

visual cortex (mainly restricted to the lunate sulcus) was modulated by emotion for both species (but more extensively for human faces). Given the absence of changes in early visual activity in response to emotional expressions in the human brain, it is unlikely that this effect in monkeys is due to low-level stimulus characteristics. To examine whether specialization for processing conspecific emotional expressions exists in both species, we performed a three-way interaction between species, expression and configuration. In humans, the middle part of the right STS (rSTSm) was specifically activated by human fearful expressions (white outline in rSTS in Fig. 4A) and showed no differential activation between monkey fearful expressions and monkey chewing (paired *t*-test, t(22)=0.71, p=0.48, Fig. 4B profiles). This was the only area in the human brain showing a conspecific–specific response to emotional expressions. In monkeys, we also found conspecific–specific responses bilaterally in posterior TE (white outline in green labeled regions in Fig. 4A) and left lunate sulcus (left V4d). However, in strong contrast with the conspecific responses in human rSTSm, human fear also increased activity in monkey TE in comparison to chewing (paired *t*-test, t(13)=2.31, p<0.05, Fig. 4B profiles).

Although both human STS and monkey IT responded selectively to emotional expressions, we found differences in properties between these regions that make it unlikely that they fulfill the same function in both species. First, monkey IT responded to all dynamic stimuli, including the scrambled displays whereas rSTSp responded only to dynamic facial expressions and rSTSm only to human emotional expressions. Furthermore, our independent localizer experiment with static stimuli showed that monkey IT responded to all non-facial categories tested (ps<0.05, Wilcoxon signed rank test), whereas human rSTSp only responded to faces (p<0.05, Wilcoxon signed rank test) and human rSTSm was not activated at all by any of the visual categories presented (Fig. 5), consistent with the selective response to human emotional expressions. In the monkey, we also
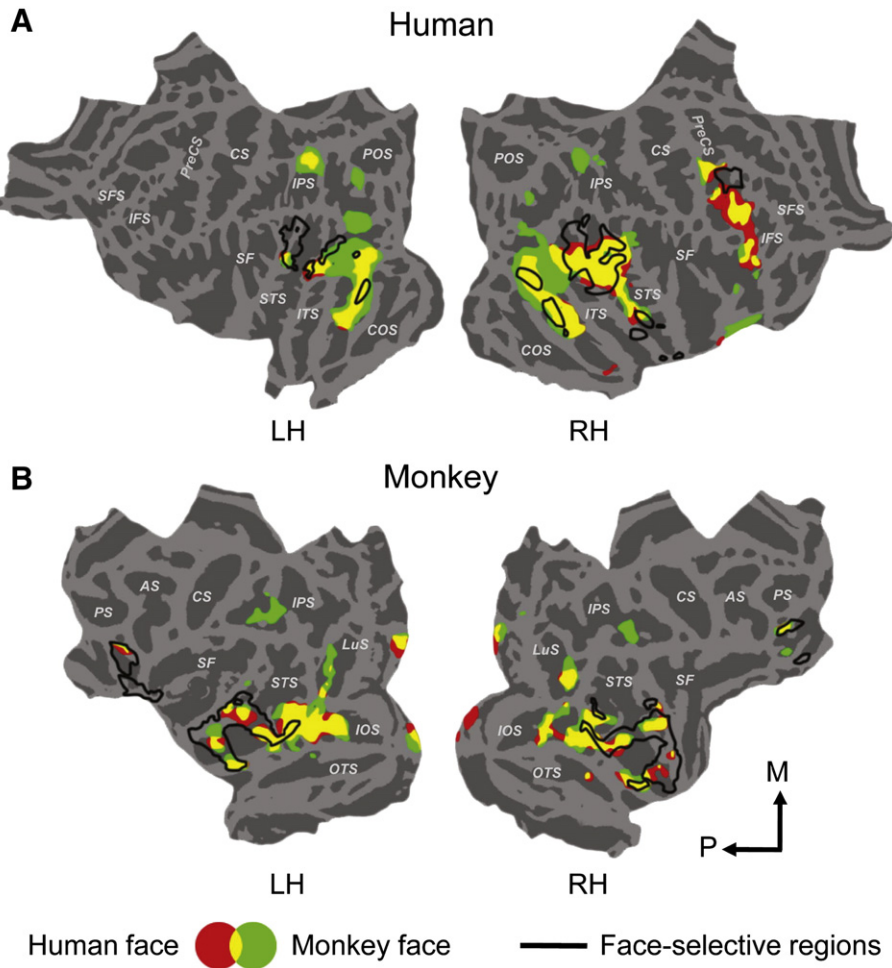
Fig. 3. Neural system responsive to dynamic facial expressions in humans and monkeys. Group-level significance maps of human (red) and monkey (green) dynamic faces compared to their scrambled versions (voxel-level p<0.001, cluster-level corrected p<0.05), irrespective of expression, are shown in humans (A) and monkeys (B). Areas that are activated by both human and monkey faces are shown in yellow. Black outlines represent face-selective areas in both species. LH, left hemisphere; RH, right hemisphere; M, medial; P, posterior. Abbreviations of sulci names: SFS, superior frontal sulcus; IFS, inferior frontal sulcus; PreCS, precentral sulcus; CS, central sulcus; IPS, intraparietal sulcus; POS, parieto-occipital sulcus; SF, sylvian fissure; STS, superior temporal sulcus; COS, collateral sulcus; PS, principal sulcus; AS, arcuate sulcus; CS, central sulcus; LuS, Lunate sulcus; IOS, inferior occipital sulcus; OTS, occipitotemporal sulcus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

investigated whether other areas that responded to conspecific emotional expressions without necessarily a conspecific–specific effect, such as the left anterior inferior temporal cortex (left AIT) (Fig. 4A), would respond solely to conspecific emotional expressions and not to other visual categories, but the answer was negative (Fig. 5).

Although not the primary aim of this study, we found that other species' emotional expressions elicited more distributed and mainly posterior effects in humans, including ventral occipito-temporal and dorsal occipito-parietal cortex, superior parietal lobule, right posterior STS, temporo-parietal junction, but also premotor cortex (shown in green and yellow in Fig. 4A). In monkeys, viewing human fearful faces relative to chewing, was associated with posterior early visual effects and an activation of a face-selective area that has been labeled ML (Moeller et al., 2008). Based on the properties of ML, we can rule out that this activation for human emotional expressions in monkeys corresponds with our conspecific emotion effect in rSTSm in humans: in contrast with rSTSm, area rML was face-selective and responded to all visual categories tested (supplementary Fig. 5).

Finally, the emotion effect in human but not monkey STS was not a matter of differences in responsiveness to dynamic facial stimuli: human and monkey dynamic face stimuli compared to the scrambled ones (single main effect of *configuration*) activated a largely overlapping distributed network including face-selective areas of STS in both species (Fig. 3).

Effects of emotional expressions in the amygdala

Given the evidence favoring amygdala involvement in fear processing across species (Dolan and Morris, 2000; Emery and Amaral, 2000; Phelps and LeDoux, 2005), we specifically looked at the activity profiles in the face-responsive (faces vs. scrambled faces irrespective of the species) parts of the amygdala (Fig. 6). In humans, amygdala responded more strongly to human, but not monkey fearful faces, compared to chewing (paired *t*-test; human fear vs. chewing: t(22)=2.79, p<0.01; monkey fear vs. chewing: t(22)=0.73, p=0.47). There was a significant three-way interaction between species, expression and configuration (ANOVA, F(1, 22)=5.76, p<0.05). However, when studying the two-way species×emotion interaction, leaving out the scrambled versions, the effect was not significant (ANOVA, F(1, 22)=2.38, p=0.14). In monkeys, amygdala responded more strongly to both human and monkey fearful faces than to chewing faces (paired *t*-test; human fear vs. chewing: t(13)=4.82, p<0.001; monkey fear vs. chewing: t(13)=2.33, p<0.05), and there was no significant three-way interaction between species, expression and configuration (ANOVA, F(1, 13)=0.70, p=0.42), or two-way interaction between species and expression when the scrambles were left out (ANOVA, F(1, 13)=1.20, p=0.29).
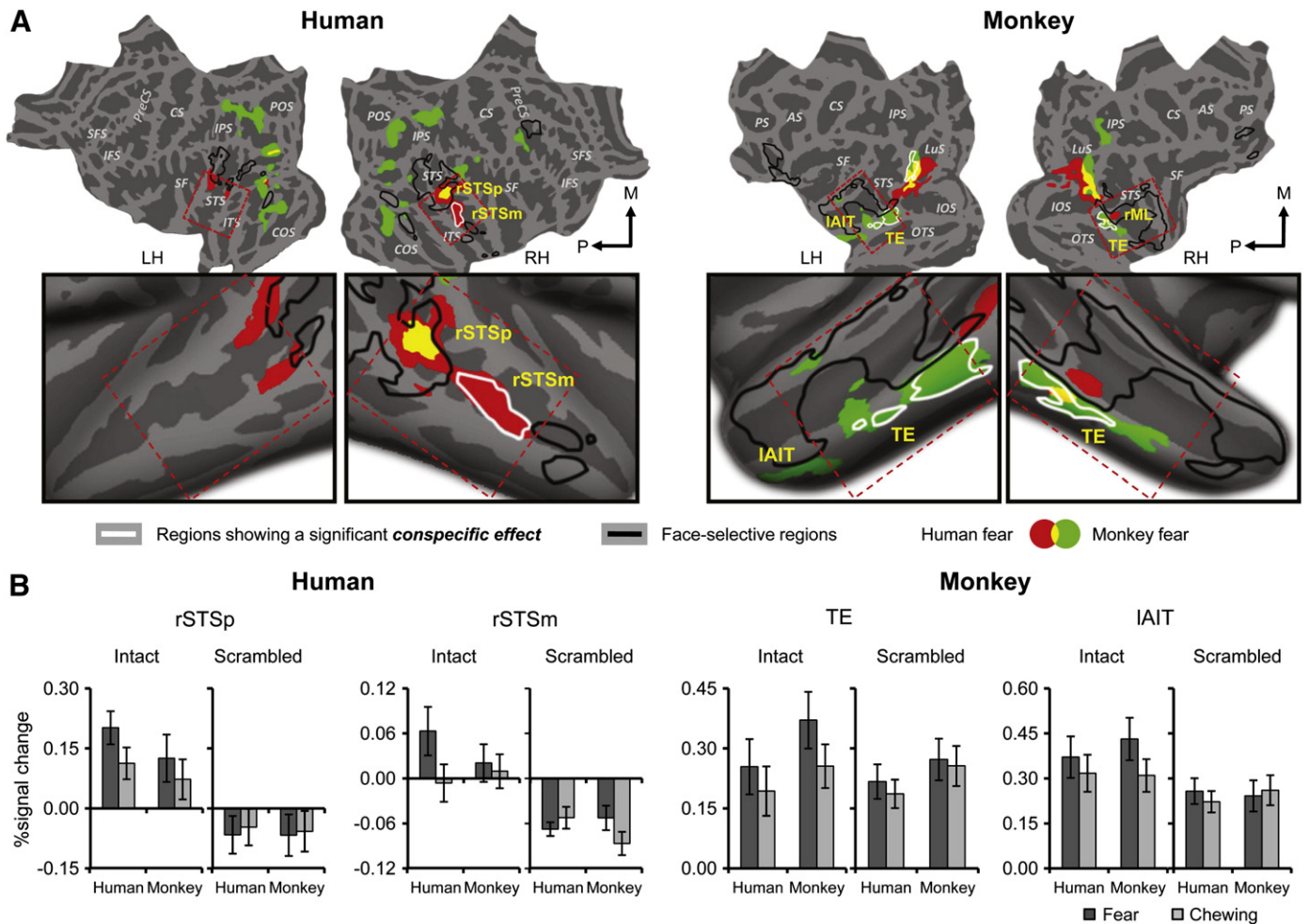
Fig. 4. Areas selective to emotional facial expressions in humans and monkeys. (A) Color-coded surface maps show regions of significant two-way *emotion × configuration* interaction: significance maps (group-level) of human (red) and monkey (green) fearful faces (relative to scrambled fearful faces) compared to chewing faces (relative to scrambled chewing faces) (voxel-level p<0.01, cluster-level corrected p<0.05). Areas that are activated by both human and monkey fearful faces are shown in yellow. Regions of significant conspecific effect (three-way *species × emotion × configuration* interaction, same threshold as two-way interaction) are labeled on the surface maps using white outlines (human as well as monkey specific emotion responsive areas). Black outlines represent face-selective areas in both species. (B) Response to dynamic facial expressions in areas responsive to conspecific emotional expressions in humans and monkeys. Activity profiles (mean ± s.e.m.) show percent signal change relative to fixation (Y-axis) for each of the 8 conditions (X-axis). LH, left hemisphere; RH, right hemisphere; M, medial; P, posterior.

Discussion

Our data reveal differences in neural processing of emotional facial expressions between humans and monkeys, and argue for a more unique role of human STS in facial emotion perception than previously documented. Although human and monkey STS are both responsive to dynamic faces, we found that human but not monkey STS shows significant activity differences between emotional and non-emotional dynamic facial expressions. Second, we provide evidence for further functional specialization within human STS along a posterior to anterior axis. Posterior STS responded selectively to emotional expressions independent of species and the emotion effect in rSTSp fell within a face-selective region. In contrast, the response in rSTSm, anterior to rSTSp, was highly selective for the emotional cue of human faces and appeared outside face-selective areas.

In monkeys, we observed the effects of monkey emotional expressions mainly in the bilateral inferotemporal cortex and also in the early visual cortex. In posterior TE, the activity was significantly higher for conspecific than for human emotional expressions. The emotional effects in monkey IT, appearing outside or at the edge of face-selective areas, confirm earlier findings using static facial expressions (Hadj-Bouziane et al., 2008) and extend those observations to

demonstrate that the posterior part of IT responds particularly to conspecific emotional expressions.

Although tempting to speculate on similarities between human STS and monkey IT in processing emotion cues from dynamic faces, we found important differences in functional properties between these regions with monkey IT being responsive to all visual stimuli presented and human STS being selective for faces (rSTSp) and human emotions (rSTSm). Our interpretation of the data is that human STS developed a high degree of neural specialization for emotional expressions as socially meaningful stimuli (Peelen et al., 2010), whereas emotion effects in monkey IT constitute mainly modulatory responses in the visual processing stream (Hadj-Bouziane et al., 2008). Such modulatory effects in IT have been covered before and are hypothesized to originate from limbic structures, mainly the amygdala (Emery and Amaral, 2000). Support for this hypothesis also comes from the observation that different aspects of facial information are encoded at different latencies during single cell recordings in IT (Sugase et al., 1999). Global information such as species is encoded by an early transient discharge whereas fine information such as emotional expressions is conveyed by a later sustained discharge. The time delay likely reflects feedback from other areas. Furthermore, in agreement with our findings, IT contains neurons that
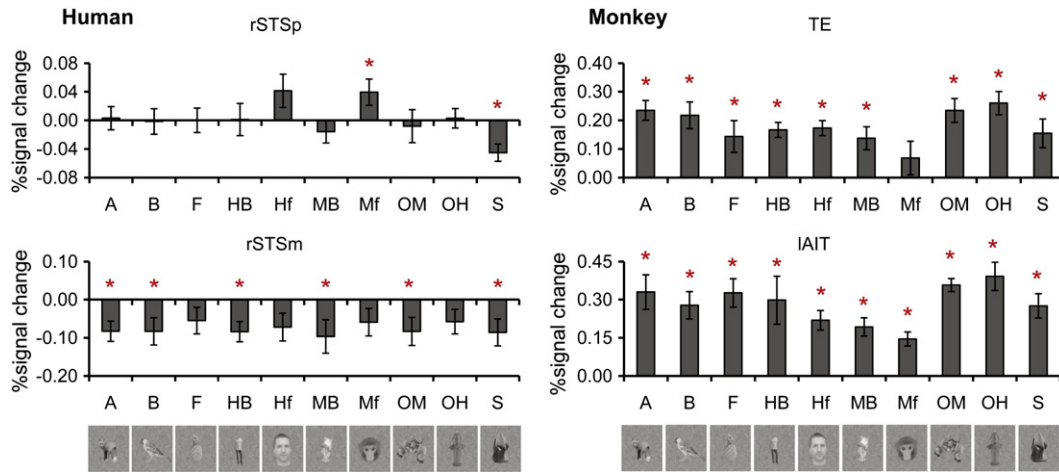
Fig. 5. Response to object categories in areas responsive to conspecific emotional expressions in humans and monkeys. Activity profiles (mean ± s.e.m.) show percent signal change relative to fixation (Y-axis) for each category (X-axis) in the posterior and middle part of the right superior temporal sulcus (rSTSm) in humans, and area TE and left anterior inferior temporal cortex (lAIT) in monkeys. ROIs were defined the same way as in Fig. 4B. Abbreviations for object categories: human faces (Hf), monkey faces (Mf), human bodies (HB), monkey bodies (MB), objects with two different aspect ratios (OH and OB), animals (A), birds (B), fruits (F) and sculptures (S). *: p<0.05 (each category vs. fixation, Wilcoxon signed rank test, uncorrected).

respond at a much higher level to monkey than to human expressions (Sugase et al., 1999). In this study, activity changes to emotional expressions also occurred in the early visual cortex of monkeys. Early visual responses have been reported in human studies showing that attention to stimuli that contain emotional information enhances responses in early visual cortex (Pessoa et al., 2002; West et al., 2011) and is consistent with anatomical studies in monkeys that show feedback projections from the amygdala terminating in TE and in V1 (Freese and Amaral, 2005, 2006). It should be noted that the lack of emotion effects in monkey STS in our study does not mean that monkey STS is not involved in processing emotional expressions. Neurons with preferential responses to emotional expressions in macaque STS have been documented before (Hasselmo et al., 1989; Perrett et al., 1984; Rolls, 2007). However, our findings show that – in contrast with human STS – fMRI response, a measure of averaged regional brain activity, is not significantly higher for emotional compared to non-emotional expressions in monkey STS.

There is growing evidence for an important role of the human STS in the perception of facial emotional expressions (Adolphs, 2002;

Allison et al., 2000; Calder and Young, 2005; Engell and Haxby, 2007; Furl et al., 2007; Haxby et al., 2002; Kret et al., 2011; LaBar et al., 2003; Narumoto et al., 2001; Said et al., 2010; Winston et al., 2004), as well as in other aspects of social perception from faces including gaze perception, lip-reading and other types of meaningful biological motion (Allison et al., 2000; de Gelder, 2006). In line with our hypothesis that STS activation in humans fulfills a social function and is involved in attributing meaning to the expression, there is growing evidence that the posterior STS is implicated in the understanding of others' mental states (Gallagher et al., 2000; Gobbini et al., 2007; Redcay et al., 2010) and encodes supramodal representations of perceived emotions (Peelen et al., 2010). Furthermore, dysfunction of the human STS in clinical populations, such as autistic subjects, leads to complex impairment of social perception (Redcay, 2008; Zilbovicius et al., 2006). The emergence of neural specialization for processing human-specific emotional and social information from faces in middle and anterior parts of the human temporal lobe, especially rSTSm, is not surprising. An important extra-allometric expansion of this part of the brain has occurred in the course of
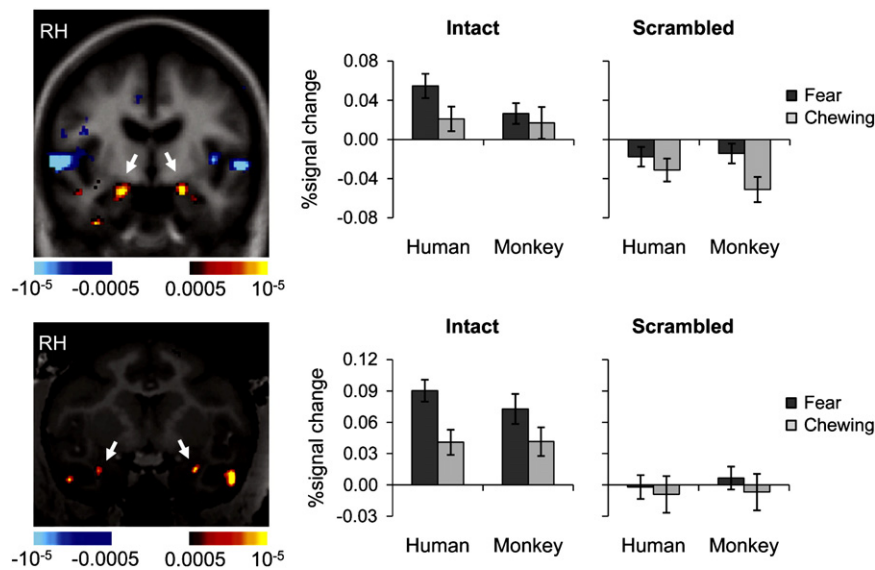
Fig. 6. Response in amygdala to facial expressions. Activity profiles based on all face-responsive voxels (faces vs. scrambled faces irrespective of the species) in human and monkey amygdala (as shown by the white arrows). RH, right hemisphere.

anthropoid evolution (Rilling and Seligman, 2002), which is, at least on the phylogenetic time scale, correlated with increasing social demands (Joffe and Dunbar, 1997). A higher degree of specialization for extracting dynamic information from faces in anterior compared to posterior human STS was recently reported (Pitcher et al., 2011). Another study reported specialization for human facial motion compared to hand motion in right middle STS (Thompson et al., 2007) and fMRI adaptation studies confirm functional specialization within human right STS with sensitivity for human emotional expressions in more anterior parts (Winston et al., 2004). Furthermore, electrical stimulation of human right middle STS disturbs labeling of facial emotions (Fried et al., 1982). Also neurodegeneration of the right anterior temporal cortex leads to severe emotion recognition deficits in patients with frontotemporal dementia (Rosen et al., 2002).

Although the heterospecific faces were primarily meant as controls to study whether emotion effects were specific for the own species, we were surprised to find so little overlap between the effects of conspecific and heterospecific emotional expressions, especially in humans. This contrasts with the important overlap of face-responsive regions in both species (Fig. 3), supporting that face processing in general is largely species-independent whereas processing of emotional cues is much more species-dependent. More posterior, parietal and occipito-temporal, responses to heterospecific expressions have been reported before though in humans (Buccino et al., 2004), but it is not exactly clear what they mean. It is unlikely that these posterior activations were caused by low-level stimulus differences since we control for it by the interaction with the scrambled stimuli. Moreover, activation in the early occipito-temporal cortex was found only in monkeys for human fearful faces (compared to chewing), but not in humans. If it was a low-level effect we should have observed it in both humans and monkeys. Differences in arousal could be another possibility, as dynamic monkey faces (certainly emotional monkey faces) may be more arousing for humans than dynamic human faces are for monkeys. However, the behavioral data presented in Fig. 2B show that this is very unlikely: the degree of arousal for humans is larger for human faces compared to monkey faces. Aspects that are harder to control for are differences in selective spatial attention across stimulus types, which are known to drive portions of parietal cortex and modulate activity in occipital areas. Hence, although speculative, a more parsimonious explanation is that humans paid more attention to the monkey fearful faces than to the human fearful faces. Even so a stronger homospecific (compared to heterospecific) effect was still observed in higher order cortex (rSTSm) in humans, which further strengthens the unique role of STS in dealing with social cues such as emotional expression.

It should also be noted that differences in familiarity may have contributed to the conspecific effect in our results. However, our study design was conceptualized to minimize familiarity effects in monkeys and novelty effects in humans, by contrasting emotional heterospecific with non-emotional heterospecific faces and thereby subtracting the familiarity or novelty effects of heterospecific faces in monkeys and humans respectively.

To conclude, our data suggest that human STS evolved towards an expertise in processing emotional expressions that is not present to a comparable degree in monkeys. More generally, our data underscore the importance of cross-species comparisons (Mantini et al., 2012) to gain insight in the species-typical neural basis of social interactions (Ghazanfar and Santos, 2004). Further comparative studies with species-specific social cues are certainly needed to support our claims and to elucidate what is typically human about our so-called 'social brain'.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.10.083.

References

Adolphs, R., 2002. Neural systems for recognizing emotion. Curr. Opin. Neurobiol. 12, 169–177.

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

Bair, W., O'Keefe, L.P., 1998. The influence of fixational eye movements on the response of neurons in area MT of the macaque. Vis. Neurosci. 15, 779–786.

Bell, A.H., Hadj-Bouziane, F., Frihauf, J.B., Tootell, R.B., Ungerleider, L.G., 2009. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J. Neurophysiol. 101, 688–700.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25, 49–59.

Brothers, L., 1989. A biological perspective on empathy. Am. J. Psychiatry 146, 10–19.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C.A., Rizzolatti, G., 2004. Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. J. Cogn. Neurosci. 16, 114–126.

Buracas, G.T., Boynton, G.M., 2002. Efficient design of event-related fMRI experiments using M-sequences. NeuroImage 16, 801–813.

Burrows, A.M., Waller, B.M., Parr, L.A., 2009. Facial musculature in the rhesus macaque (*Macaca mulatta*): evolutionary and functional contexts with comparisons to chimpanzees and humans. J. Anat. 215, 320–334.

Calder, A.J., Young, A.W., 2005. Understanding the recognition of facial identity and facial expression. Nat. Rev. Neurosci. 6, 641–651.

de Gelder, B., 2006. Towards the neurobiology of emotional body language. Nat. Rev. Neurosci. 7, 242–249.

de Gelder, B., 2010. The grand challenge for frontiers in emotion science. Front. Psychol. 1, 1–4.

de Waal, F.B., 2011. What is an animal emotion? Ann. N. Y. Acad. Sci. 1224, 191–206.

de Waal, F.B.M., Luttrell, L.M., 1985. The formal hierarchy of rhesus macaques — an investigation of the bared-teeth display. Am. J. Primatol. 9, 73–85.

Dolan, R.J., 2002. Emotion, cognition, and behavior. Science 298, 1191–1194.

Dolan, R.J., Morris, J.S., 2000. The functional anatomy of innate and acquired fear: perspectives from neuroimaging. In: Lane, R.D., Nadel, L. (Eds.), Cognitive Neuroscience of Emotion. Oxford University Press, Inc., New York, New York, pp. 225–241.

Ekstrom, L.B., Roelfsema, P.R., Arsenault, J.T., Bonmassar, G., Vanduffel, W., 2008. Bottom-up dependent gating of frontal signals in early visual cortex. Science 321, 414–417.

Emery, N.J., Amaral, D.G., 2000. The role of the amygdala in primate social cognition. In: Lane, R.D., Nadel, L. (Eds.), Cognitive Neuroscience of Emotion. Oxford University Press, Inc., New York, New York, pp. 156–191.

Engell, A.D., Haxby, J.V., 2007. Facial expression and gaze-direction in human superior temporal sulcus. Neuropsychologia 45, 3234–3241.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.

Freese, J.L., Amaral, D.G., 2005. The organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. J. Comp. Neurol. 486, 295–317.

Freese, J.L., Amaral, D.G., 2006. Synaptic organization of projections from the amygdala to visual cortical areas TE and V1 in the macaque monkey. J. Comp. Neurol. 496, 655–667.

Fried, I., Mateer, C., Ojemann, G., Wohns, R., Fedio, P., 1982. Organization of visuospatial functions in human cortex. Evidence from electrical stimulation. Brain 105, 349–371.

Furl, N., van Rijsbergen, N.J., Treves, A., Friston, K.J., Dolan, R.J., 2007. Experience-dependent coding of facial expression in superior temporal sulcus. Proc. Natl. Acad. Sci. U. S. A. 104, 13485–13489.

Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 38, 11–21.

Ghazanfar, A.A., Santos, L.R., 2004. Primate brains in the wild: the sensory bases for social interactions. Nat. Rev. Neurosci. 5, 603–616.

Gitelman, D.R., 2002. ILAB: a program for postexperimental eye movement analysis. Behav. Res. Methods Instrum. Comput. 34, 605–612.

Gobbini, M.I., Koralek, A.C., Bryan, R.E., Montgomery, K.J., Haxby, J.V., 2007. Two takes on the social brain: a comparison of theory of mind tasks. J. Cogn. Neurosci. 19, 1803–1814.

Graham, R., Labar, K.S., 2012. Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention. Neuropsychologia 50, 553–566.

Hadj-Bouziane, F., Bell, A.H., Knusten, T.A., Ungerleider, L.G., Tootell, R.B., 2008. Perception of emotional expressions is independent of face selectivity in monkey inferior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. 105, 5591–5596.

Hasselmo, M.E., Rolls, E.T., Baylis, G.C., 1989. The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. Behav. Brain Res. 32, 203–218.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2002. Human neural systems for face recognition and social communication. Biol. Psychiatry 51, 59–67.

Hebb, D.O., 1946. Emotion in man and animal: an analysis of the intuitive processes of recognition. Psychol. Rev. 53, 88–106.

Hein, G., Knight, R.T., 2008. Superior temporal sulcus—it's my area: or is it? J. Cogn. Neurosci. 20, 2125–2136.

Hoffman, K.L., Gothard, K.M., Schmid, M.C., Logothetis, N.K., 2007. Facial-expression and gaze-selective responses in the monkey amygdala. Curr. Biol. 17, 766–772.

Joffe, T.H., Dunbar, R.I., 1997. Visual and socio-cognitive information processing in primate brain evolution. Proc. Biol. Sci. 264, 1303–1307.

Kret, M.E., Pichon, S., Grezes, J., de Gelder, B., 2011. Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. NeuroImage 54, 1755–1762.

Kujala, M.V., Tanskanen, T., Parkkonen, L., Hari, R., 2009. Facial expressions of pain modulate observer's long-latency responses in superior temporal sulcus. Hum. Brain Mapp. 30, 3910–3923.

LaBar, K.S., Crupain, M.J., Voyvodic, J.T., McCarthy, G., 2003. Dynamic perception of facial affect and identity in the human brain. Cereb. Cortex 13, 1023–1033.

Leite, F.P., Tsao, D., Vanduffel, W., Fize, D., Sasaki, Y., Wald, L.L., Dale, A.M., Kwong, K.K., Orban, G.A., Rosen, B.R., Tootell, R.B.H., Mandeville, J.B., 2002. Repeated fMRI using iron oxide contrast agent in awake, behaving macaques at 3 Tesla. NeuroImage 16, 283–294.

Maestripieri, D., Wallen, K., 1997. Affiliative and submissive communication in rhesus macaques. Primates 38, 127–138.

Mandeville, J.B., Choi, J.K., Jarraya, B., Rosen, B.R., Jenkins, B.G., Vanduffel, W., 2011. fMRI of cocaine self-administration in macaques reveals functional inhibition of basal ganglia. Neuropsychopharmacology 36, 1187–1198.

Mantini, D., Hasson, U., Betti, V., Perrucci, M.G., Romani, G.L., Corbetta, M., Orban, G.A., Vanduffel, W., 2012. Interspecies activity correlations reveal functional correspondence between monkey and human brain areas. Nat. Methods 9, 277–282.

Moeller, S., Freiwald, W.A., Tsao, D.Y., 2008. Patches with links: a unified system for processing faces in the macaque temporal lobe. Science 320, 1355–1359.

Narumoto, J., Okada, T., Sadato, N., Fukui, K., Yonekura, Y., 2001. Attention to emotion modulates fMRI activity in human right superior temporal sulcus. Brain Res. Cogn. Brain Res. 12, 225–231.

Nelissen, K., Vanduffel, W., Orban, G.A., 2006. Charting the lower superior temporal region, a new motion-sensitive region in monkey superior temporal sulcus. J. Neurosci. 26, 5929–5947.

Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J., 2006. Highly accurate optic flow computation with theoretically justified warping. Int. J. Comput. Vis. 67, 141–158.

Parr, L.A., Heintz, M., 2009. Facial expression recognition in rhesus monkeys, *Macaca mulatta*. Anim. Behav. 77, 1507–1513.

Parr, L.A., Waller, B.M., Fugate, J., 2005. Emotional communication in primates: implications for neurobiology. Curr. Opin. Neurobiol. 15, 716–720.

Parr, L.A., Waller, B.M., Heintz, M., 2008. Facial expression categorization by chimpanzees using standardized stimuli. Emotion 8, 216–231.

Peelen, M.V., Atkinson, A.P., Vuilleumier, P., 2010. Supramodal representations of perceived emotions in the human brain. J. Neurosci. 30, 10127–10134.

Perrett, D.I., Smith, P.A., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., Jeeves, M.A., 1984. Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception. Hum. Neurobiol. 3, 197–208.

Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G., 2002. Neural processing of emotional faces requires attention. Proc. Natl. Acad. Sci. U. S. A. 99, 11458–11463.

Phelps, E.A., LeDoux, J.E., 2005. Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48, 175–187.

Pinsk, M.A., Arcaro, M., Weiner, K.S., Kalkus, J.F., Inati, S.J., Gross, C.G., Kastner, S., 2009. Neural representations of faces and body parts in macaque and human cortex: a comparative FMRI study. J. Neurophysiol. 101, 2581–2600.

Pitcher, D., Dilks, D.D., Saxe, R.R., Triantafyllou, C., Kanwisher, N., 2011. Differential selectivity for dynamic versus static information in face-selective cortical regions. NeuroImage 56, 2356–2363.

Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. J. Neurosci. 18, 2188–2199.

Rajimehr, R., Young, J.C., Tootell, R.B., 2009. An anterior temporal face patch in human cortex, predicted by macaque maps. Proc. Natl. Acad. Sci. U. S. A. 106, 1995–2000.

Redcay, E., 2008. The superior temporal sulcus performs a common function for social and speech perception: implications for the emergence of autism. Neurosci. Biobehav. Rev. 32, 123–142.

Redcay, E., Dodell-Feder, D., Pearrow, M.J., Mavros, P.L., Kleiner, M., Gabrieli, J.D., Saxe, R., 2010. Live face-to-face interaction during fMRI: a new tool for social cognitive neuroscience. NeuroImage 50, 1639–1647.

Rilling, J.K., Seligman, R.A., 2002. A quantitative morphometric comparative analysis of the primate temporal lobe. J. Hum. Evol. 42, 505–533.

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

Rolls, E.T., 2007. The representation of information about faces in the temporal and frontal lobes. Neuropsychologia 45, 124–143.

Rosen, H.J., Perry, R.J., Murphy, J., Kramer, J.H., Mychack, P., Schuff, N., Weiner, M., Levenson, R.W., Miller, B.L., 2002. Emotion comprehension in the temporal variant of frontotemporal dementia. Brain 125, 2286–2295.

Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010. Distributed representations of dynamic facial expressions in the superior temporal sulcus. J. Vis. 10, 11.

Sturgeon, R.S., Cooper, L.M., Howell, R.J., 1989. Pupil response: a psychophysiological measure of fear during analogue desensitization. Percept. Mot. Skills 69, 1351–1367.

Sugase, Y., Yamane, S., Ueno, S., Kawano, K., 1999. Global and fine information coded by single neurons in the temporal visual cortex. Nature 400, 869–873.

Thompson, J.C., Hardee, J.E., Panayiotou, A., Crewther, D., Puce, A., 2007. Common and distinct brain activation to viewing dynamic sequences of face and hand movements. NeuroImage 37, 966–973.

Tsao, D.Y., Freiwald, W.A., Knutsen, T.A., Mandeville, J.B., Tootell, R.B., 2003. Faces and objects in macaque cerebral cortex. Nat. Neurosci. 6, 989–995.

Tsao, D.Y., Moeller, S., Freiwald, W.A., 2008a. Comparing face patch systems in macaques and humans. Proc. Natl. Acad. Sci. U. S. A. 105, 19514–19519.

Tsao, D.Y., Schweers, N., Moeller, S., Freiwald, W.A., 2008b. Patches of face-selective cortex in the macaque frontal lobe. Nat. Neurosci. 11, 877–879.

Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van, H.P., Rosen, B.R., Tootell, R.B., Orban, G.A., 2001. Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. Neuron 32, 565–577.

Vuilleumier, P., Richardson, M.P., Armony, J.L., Driver, J., Dolan, R.J., 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat. Neurosci. 7, 1271–1278.

West, G.L., Anderson, A.A., Ferber, S., Pratt, J., 2011. Electrophysiological evidence for biased competition in V1 for fear expressions. J. Cogn. Neurosci. 23, 3410–3418.

Winston, J.S., Henson, R.N., Fine-Goulden, M.R., Dolan, R.J., 2004. fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. J. Neurophysiol. 92, 1830–1839.

Zhao, F.Q., Wang, P., Hendrich, K., Ugurbil, K., Kim, S.G., 2006. Cortical layer-dependent BOLD and CBV responses measured by spin-echo and gradient-echo fMRI: insights into hemodynamic regulation. NeuroImage 30, 1149–1160.

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., Boddaert, N., 2006. Autism, the superior temporal sulcus and social perception. Trends Neurosci. 29, 359–366.