Dynamic gamma frequency feedback coupling between higher and lower order visual cortices underlies perceptual completion in humans

S. Moratti  C. Méndez-Bértolo  F. Del-Pozo  B.A. Strange

ABSTRACT
To perceive a coherent environment, incomplete or overlapping visual forms must be integrated into meaningful coherent percepts, a process referred to as "Gestalt" formation or perceptual completion. Increasing evidence suggests that this process engages oscillatory neuronal activity in a distributed neuronal assembly. A separate line of evidence suggests that Gestalt formation requires top-down feedback from higher order brain regions to early visual cortex. Here we combine magnetoencephalography (MEG) and effective connectivity analysis in the frequency domain to specifically address the effective coupling between sources of oscillatory brain activity during Gestalt formation. We demonstrate that perceptual completion of two-tone "Mooney" faces induces increased gamma frequency band power (55–71 Hz) in human early visual, fusiform and parietal cortices. Within this distributed neuronal assembly fusiform and parietal gamma oscillators are coupled by forward and backward connectivity during Mooney face perception, indicating reciprocal influences of gamma activity between these higher order visual brain regions. Critically, gamma band oscillations in early visual cortex are modulated by top-down feedback connectivity from both fusiform and parietal cortices. Thus, we provide a mechanistic account of Gestalt perception in which gamma oscillations in feature sensitive and spatial attention-relevant brain regions reciprocally drive one another and convey global stimulus aspects to local processing units at low levels of the sensory hierarchy by top-down feedback. Our data therefore support the notion of inverse hierarchical processing within the visual system underlying awareness of coherent percepts.

Introduction
Despite the ease with which we perceive coherent objects in our environment even under poor stimulus conditions, the integration of only partly available visual information into whole percepts is a challenge for the visual system. This integration process has been referred to as perceptual completion or closure. Gestalt psychology considers perceptual completion to arise from processing of a stimulus as a whole, via choosing the simplest interpretation from the interactions of stimulus parts as opposed to the simple summation of single parts themselves (Wertheimer, 1923). Although the Gestalt theoretical framework describes this process at the level of stimulus part interactions, how the brain achieves perceptual completion from a mechanistic point of view is less understood.

Neuroimaging studies highlight a role for ventral visual areas and parietal cortex during perceptual completion of bi-stable (e.g. Rubin vases), and degraded figures (Andrews et al., 2002; Dolan et al., 1997; Kleinschmidt et al., 1998; Sehatpour et al., 2006). Perceptual closure of two-tone Mooney faces (Mooney, 1957) elicits increased hemodynamic responses in face sensitive visual ventral brain regions such as the fusiform face area (Andrews and Schluppeck, 2004; Kanwisher et al., 1998; McKeeff and Tong, 2007). Electroencephalogram (EEG) studies have demonstrated a perceptual closure specific event related potential (ERP) occurring in a time window between 230 ms and 400 ms peaking around 320 ms post-stimulus time (Doniger et al., 2000, 2001; Sehatpour et al., 2006). Source localization of this component also revealed that ventral visual cortex (part of the lateral occipital cortex, LOC) and parietal cortex are active at this latency (Sehatpour et al., 2006). Taken together, these findings accord with the suggestion that not only brain regions for cue invariant object/face recognition such as ventral visual cortex (Haxby et al., 1999; Kanwisher and Yovel, 2006; Malach et al., 1995) but also spatial attention relevant parietal brain regions (Corbetta et al., 1998;
Fernandez-Duque and Posner, 2001) are crucial for perceptual closure and that this process occurs about 230 ms to 400 ms after stimulus onset.

Within the classical view of visual system hierarchy, simple geometric lines and shapes that form complex objects are processed in lower order visual cortex, whereas higher order areas within the ventral visual stream (Mishkin et al., 1983) code invariant object and category information (e.g., Vogels and Orban, 1996) based on feedforward communication from early visual cortex (for a review of these models see Hochstein and Ahissar, 2002). However, recent models of conscious visual perception suggest reverse hierarchical processing (for a review see Hochstein and Ahissar, 2002) whereby higher order visual areas in the ventral and dorsal streams provide top-down feedback to early visual cortex (i.e., predictive coding – Friston, 2003; Rao and Ballard, 1999). In the case of perceptual completion, this top-down feedback is suggested to carry global information to local processing units in early visual cortex (Bullier, 2001; Campana and Tallon-Baudry, 2013; Hochstein and Ahissar, 2002; Lamme and Roelfsema, 2000), which accords with Gestalt theory in that global visual information interacts with local stimulus part processing (Wagemans et al., 2012a, 2012b).

The neuroimaging studies of perceptual completion described above do not report engagement of lower order visual cortex. However, one event-related fMRI study (Altmann et al., 2003) reported both primary and higher order visual cortex activity activation during global shape integration of collinear contours. Although these observations were interpreted as potentially reflecting top-down modulation in global shape perception (Altmann et al., 2003), a measure of interaction between these levels of hierarchy was not provided. Transcranial magnetic stimulation (TMS) studies in humans demonstrate that interrupting recurrent interactions between early and higher visual cortices in the ventral visual stream impairs perception of natural scenes (Keivisto et al., 2011) and perceptual completion of illusionary Kanizsa-type figures (Wolfe et al., 2013). However, perceptual impairment by TMS-evoked disruption of early visual areas (Wolfe et al., 2013) does not directly demonstrate feedback coupling of neuronal activity; an alternative explanation is simply that early visual cortex activates at later latencies independently from any feedback from higher order areas. Thus, although recent evidence suggests that coherent perception relies on feedback from higher to lower order visual cortex, paralleling the global-to-local concept of Gestalt psychology, a characterization of this process in terms of effective connectivity is currently lacking.

To address this, we measured induced neuromagnetic oscillatory brain responses to two-tone Mooney faces that consist of white patches that have to be spatially integrated to perceive a face (Fig. 1; Mooney, 1957). We employ the Mooney face paradigm for two reasons: first, it represents a classical measure of perceptual completion (Mooney, 1957) and brain areas involved in Mooney face perception are well characterized (Andrews and Schluppeck, 2004; Grutzner et al., 2010; Kanwisher et al., 1998; McKeeff and Tong, 2007). Second, Mooney face perception consistently elicits neuronal oscillations in the gamma frequency band (>30 Hz) in the EEG (Rodriguez et al., 1999; Trujillo et al., 2005), intracranial EEG (Lachaux et al., 2005), and MEG (Grutzner et al., 2010). Synchronized oscillatory neuronal gamma band responses

**Fig. 1.** Gamma power changes by Mooney face perception in sensor space: topographies of gamma power changes (50-100 Hz) with respect to baseline in 100 ms intervals are shown for the Mooney face condition (A) and for the scrambled faces (B). Below each, a time frequency plot of mean relative gamma power changes across posterior MEG sensors is depicted. The colorbars indicate percentage of relative gamma power changes compared to baseline.
can be observed during coherent perception of a wide range of visual stimuli (Gruber et al., 2008; Gruber and Muller, 2005; Keil et al., 1999; Martinovic et al., 2008; Muller et al., 1996; Tallon-Baudry et al., 1996, 1997) and are thought to reflect dynamic neuronal interactions between brain areas critical for perceptual synthesis (Engel et al., 2001; Hipp et al., 2011; Martinovic et al., 2008; Muller et al., 1996; Singer and Gray, 1995). Critically, by inventing biophysical neuronal models that estimate the underlying effective connectivity between the identified brain areas based on the cross spectral density of neuroimagnetic gamma band responses (Friston et al., 2012), we determine the directionality of gamma band coupling in terms of feedforward and backward interactions. We predicted that Mooney face perceptual completion would increase gamma band responses in face sensitive fusiform face area (Kanwisher and Yovel, 2006), spatial attention relevant parietal (Corbetta, 1998; Fernandez-Duque and Posner, 2001) and lower order visual cortex (Wokke et al., 2013). Furthermore, the reverse hierarchy-processing hypothesis makes the specific prediction that inversion of effective connectivity models reveals feedback coupling of gamma oscillations from fusiform and parietal to early visual cortex during perceptual completion.

Materials and methods

Participants

Eighteen right-handed volunteers (9 females, 9 males, mean age 31.8 years ± 1.3 s.e.m.) participated in our study after having given written informed consent. All subjects had no history of neurological or psychiatric disease and had normal or corrected to normal vision. The study had full ethical approval.

Experimental design

Stimuli comprised 40 Mooney (Mooney, 1957) and 40 scrambled Mooney faces. Scrambled faces were derived by rotating (45°-90°) the original and randomly rearranging stimulus features. Stimuli were presented to the center of a screen in a magnetically shielded MEG room (visual angle 7° by 10°).

In two experimental runs, 40 Mooney and 40 scrambled faces were presented in random order (total 80 trials per picture category). Stimuli were presented for 200 ms with inter-stimulus interval randomly varying between 2000 and 2500 ms. Participants indicated by button press after stimulus offset when they had perceived a face or not (right and left response buttons counterbalanced across subjects).

Data acquisition and preprocessing

MEG data were recorded continuously (1000 Hz sample rate, 0.1-300 Hz online filter) using a 306-channel system (Elekta® VectorView). The MEG sensors consisted of 102 magnetometers and 102 pairs of orthogonal planar gradiometers. Per-stimulus epochs of 2000 ms (1000 ms baseline) were extracted for each MEG channel and stimulus category. Epochs were discarded from analyses when containing eye artifacts, movement artifacts identified by visual inspection, high amplitudes (3 pT and 1 pT in magnet- or gradiometers, respectively), and pertaining to incorrect trials (a scrambled picture was perceived as a face or vice versa). The number of trials per picture category was equalized by randomly choosing a subset of trials of the picture category that contained more artifact free epochs. This was done in order not to bias experimental conditions with respect to their signal to noise ratios. For statistical comparisons, beamformer power estimations, and connectivity analysis, it has been recommended that experimental conditions should contain the same number of trials (Gross et al., 2013).

Time-frequency spectral changes

Based on previous reports about induced oscillatory power changes elicited by Mooney faces (Grunzner et al., 2010; Lachaux et al., 2005; Rodriguez et al., 1999; Trujillo et al., 2005), we conducted a time-frequency analysis in a time window from —800 ms to 600 ms per-stimulus time and a frequency range between 30 and 100 Hz. For each sensor, epoch, experimental condition, and participant we performed time-frequency decomposition with sliding time windows of 200 ms length in 40 ms steps. We applied 3 tapers (Slepian sequences) obtaining a frequency smoothing of ±10 Hz (for a similar approach see Capilla et al., 2012). Then, time-frequency decompositions of epochs for each experimental condition and subject were averaged and relative power changes with respect to baseline (—800 ms to —200 ms pre-stimulus time) were determined. The multi-taper approach with fixed time windows and frequency smoothing has been recommended for later statistical comparisons, as the number of time-frequency bins is equal across the time-frequency range (Gross et al., 2013). Orthogonal gradiometer data were combined by calculating the modulus of the horizontal and vertical gradients of their relative power changes. Finally, power changes for each MEG channel (magneto- and combined gradiometers) and stimulus category were submitted to statistical analysis.

Statistical analysis at sensor level

We applied a cluster-based nonparametric permutation statistic to determine the time-frequency windows and channel locations of significant relative power differences between Mooney and scrambled faces within a 30 to 100 Hz frequency band. This approach effectively corrects the family wise error rate in the context of multiple comparisons (time-frequency bins and channels) (Maris and Oostenveld, 2007). The permutation test was applied to the magneto- and combined planar gradiometers, separately. Under the null hypothesis of no differences between picture categories, the relative power changes at each sensor from the two experimental conditions can be permuted between conditions. After a permutation step, a paired t-test (initial threshold of p < 0.01) was calculated at each time point, frequency bin, and sensor. Then, significant time-frequency–sensor clusters were formed by temporal, spatial, and spatial adjacency (a cluster threshold contained at least two significant neighbors along the three dimensions). For each cluster the t-values were summed and the greatest sum entered into the permutation distribution. Permutation steps were repeated 1000 times. Empirical cluster sums of t-values that were greater or smaller than the 97.5th percentile (p = 0.025 two tailed test) within the permutation distribution were considered as significant temporal–spectral–spatial clusters of a picture category effect. All preprocessing steps and permutation statistics were done using the FieldTrip toolbox (http://fieldtrip.fcdonders.nl/).

Source reconstruction

The underlying cortical sources of the time–frequency window of interest showing different relative power changes between Mooney and scrambled faces were estimated using a beamformer approach (Van Veen et al., 1997) implemented in FieldTrip. The head and sensor positions of each subject were first co-registered with a MNI canonical template brain (Collins et al., 1998) by realigning it with the individual's fiducials and head shape points. The triangulated skull surface of the template brain served as a single shell volume conduction model (Nolte, 2003). The leadfields for orthogonal dipole pairs tangentially oriented to the scalp surface placed on a three dimensional regular spaced source grid (8 mm distance) were calculated using the method described in Nolte (2003). The source grid was spatially restricted to the gray matter of the template brain. First, based on sensor space data (Fig. 1), cortical responses for the broadband gamma responses (50–100 Hz) were estimated by using a linearly constrained minimum variance (LCMV) beamformer (Van
MEG time series were band-pass filtered between 50 and 100 Hz (see Results and Fig. 1) for each trial. Then, time windows of 100 ms steps beginning at 100 ms and ending at 600 ms post-stimulus time were extracted from the MEG data for each participant and experimental condition. Pre-stimulus segments of the same length were extracted. For every 100 ms time window pre- and post-stimulus segments of both experimental conditions were concatenated and the covariance matrix was calculated to determine the spatial filter coefficients of the LCMV beamformer (Veen et al., 1997). A regularization factor was applied by adding 10% of the mean across the eigenvalues of the covariance matrix to each element of the covariance matrix. Then, each band-pass filtered sensor level MEG epoch was projected into source space through the common spatial filter for each 100 ms time window and the pre-stimulus interval. For each experimental condition, participant, 100 ms time window, and source grid location, power along the optimal dipole orientation as determined by the first eigenvector of the covariance matrix of the two tangentially oriented dipoles was averaged across epochs. For each subject, experimental condition, and 100 ms time window relative power changes with respect to baseline was calculated at each source grid location [(post-stimulus power – pre-stimulus power) / pre-stimulus power].

Second, based on the results of the statistical comparison between the time frequency decompositions of Mooney and scrambled face epochs, MEG time series were band-pass filtered between 55 and 71 Hz (see Results) for each trial. Then, post-stimulus time windows of interest (333 to 538 ms ± 100 ms, see Results) and pre-stimulus segments of the same length were extracted. Pre- and post-stimulus time segments of both experimental conditions were concatenated and the covariance matrix was calculated to determine the spatial filter coefficients of the LCMV beamformer (Veen et al., 1997). As for the broadband gamma response (see above) each band-pass filtered sensor level MEG epoch was projected into source space through the common spatial filter and a regularization factor of 10% was applied. For each experimental condition, participant, and source grid location, power along the optimal dipole orientation as determined by the first eigenvector of the covariance matrix of the two tangentially oriented dipoles was averaged across epochs. Cortical power source grid volumes for the Mooney and scrambled face conditions were then submitted to statistical analysis.

Third, we recalculated the beamformer results with respect of the time and frequency ranges based on the comparison between the time frequency decompositions of Mooney and scrambled face epochs, but dividing the time window in an early (333–433 ms ± 100 ms) and late (433–538 ms ± 100 ms) time segment. At each source grid location the relative power changes for the Mooney with respect to scrambled faces were calculated (Fig. 3C). Finally, based on these relative activity maps relative power changes for the late vs. early time segments were determined (Fig. 3D).

**Statistical analysis of gamma power at source level**

Oscillatory gamma power projected into cortical source space for Mooney and scrambled faces was compared using the same non-parametric cluster based permutation statistics as described for the time frequency sensor level data (using the FieldTrip toolbox). However, as
Table 1
MNI coordinates of peak t values in parietal, fusiform, and occipital cortex and its associated p values (two-sided) and power changes relative to baseline.

<table>
<thead>
<tr>
<th>Region</th>
<th>x, y, z</th>
<th>t (17)</th>
<th>p</th>
<th>Relative power change (± s.e.m.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parietal</td>
<td>56 - 49 40</td>
<td>3.3</td>
<td>0.004</td>
<td>1.5% ± 0.4/-0.3% ± 0.4</td>
</tr>
<tr>
<td>Fusiform</td>
<td>25 - 50 - 10</td>
<td>3.1</td>
<td>0.007</td>
<td>2.5% ± 0.6/0.4% ± 0.5</td>
</tr>
<tr>
<td>Occipital</td>
<td>30 - 89 - 11</td>
<td>2.8</td>
<td>0.01</td>
<td>3.6% ± 0.8/1.4% ± 0.7</td>
</tr>
</tbody>
</table>

...the beamformer solutions (3 dimensional dipole grids in MNI space) already reflect relative power changes within a certain time frequency window, clusters were formed along the spatial dimension only.

Dynamic causal modeling of effective cortical network connectivity

To examine effective coupling between gamma power sources, we applied DCM to the cross-spectral densities (CSD) of these sources in...
the time-frequency range of interest (Friston et al., 2012). The cortical locations of these sources were based on the non-parametric cluster based permutation statistics of oscillatory gamma power differences between Mooney and scrambled faces (Fig. 3). We selected 3 cortical sources (one source in early visual cortex in the right occipital lobe, one in the right fusiform and parietal cortices, respectively; Table 1). The coordinates of these sources correspond to peak t values (t (17) > 2.89, p < 0.01) within the significant source cluster spanning early visual, inferotemporal and parietal cortices. In order to extract the time series of the three sources of interest, the unfiltered sensor level MEG epochs were projected through the spatial beamformer filter. All three source waveforms were baseline corrected using the same baseline as described before. Further, waveforms were normalized by their standard deviations to eschew confounding of DCM results by amplitude differences. Finally, the three cortical source time series were submitted to CSD-based DCM analysis limited to a latency of 233 to 638 ms and a frequency range of 55 to 71 Hz. This time frequency range was chosen based on significant gamma power differences between Mooney and scrambled faces within this temporal spectral window (see Results section and Fig. 3).

Ten different DCMs (Fig. 4) were inverted to estimate the connectivity strength modulations between the three sources by picture category (Mooney vs. scrambled). In light of the reverse hierarchical processing model of the visual system (Hochstein and Ahissar, 2002) connectivity strength modulations between early visual cortex and fusiform or parietal cortex were modeled as purely forward, backward, or both, forwards and backwards. Coupling between fusiform and parietal cortices was modeled in a forward, backward or bi-directional manner. Finally, as gamma power synchronization may occur locally, a tenth model was inverted without modulation of connectivity between any areas. Fig. 4 depicts all DCM models that were inverted. Model parameters were estimated using a variational Bayesian scheme as implemented in SPM12b (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/), with the parameter of interest here being the connectivity modulations between cortical sources by picture category.

All models were tested against each other by employing Bayesian model selection with fixed effects (see Garrido et al., 2007 for details). First, model evidence for each model and subject was estimated by its log evidence, which is a measure indicating the probability of the data given the model. Next, for each DCM model the log evidences were added across subjects (equivalent to multiplying the marginal likelihoods of the models, assuming independence between subjects). The winning model was defined as the model with the log evidence greater than at least a value of 3 relative to all other models (relative likelihood is greater than 20:1). This difference threshold is considered as indicating strong evidence for the winning model reflected in the model posterior probability (Garrido et al., 2007).

Parameter estimation (connectivity modulation by picture category) was done by calculating the conditional density of the parameters at the group level (Bayesian parameter averaging, for details see Garrido et al., 2007). Briefly, the individual parameters and precision matrices are multiplied and summed across participants. The sum of products is then multiplied by the overall precision matrix as defined by the sum across all individual precision matrices. This procedure was applied to the winning model only, yielding mean coupling parameters and their associated conditional probabilities across subjects. Note that the conditional probability expresses how likely the coupling parameter is different from zero change with respect to the baseline condition (scrambled faces). Bayesian model selection and parameter averaging were done using SPM12b.

### Results

Performance, indexed by correct identification of Mooney and scrambled faces, was high (mean percentage correct (± s.e.m.) Mooney 85.8% ± 1.7; scrambled 90.9% ± 1.6; mean decision time Mooney 712 ms ± 22.3; scrambled 772 ms ± 22.2). Both stimulus types induce broadband gamma power (50–100 Hz ± 10 Hz) in posterior MEG sensors with onset latency around 100 ms. Critically, however, Mooney faces elicit prolonged oscillatory gamma band power changes (Fig. 1). Fig. 2 shows the corresponding cortical relative power changes with respect to baseline in this frequency band as estimated by the LCVM beamformer.

Mooney faces significantly increase gamma band activity, relative to scrambled faces, in a frequency range from 55 to 71 Hz (± 10 Hz) and from 333 to 538 ms (± 100 ms) after stimulus onset. Note that although our statistical analysis indicates that gamma power differences started at 333 ms after stimulus onset, due to fixed 200 ms time windows (time resolution ± 100 ms) utilized for time frequency decomposition (Gross et al., 2013), the time interval in which gamma band activity can be said to be increased is therefore between 233 ms and 333 ms after stimulus onset.
and 638 ms post-stimulus onset. This effect is restricted to a right occipito-parietal sensor cluster (cluster-based permutation testing Mooney > scrambled: summed t-value = 3860, p = 0.005; Fig. 3A).

Mean power changes relative to baseline (± s.e.m.) across the significant sensor–time–frequency cluster are 7.3% ± 1.2 for Mooney and 2.2% ± 1.2 for scrambled faces, respectively. No sensor–time–frequency cluster indicated increased gamma activity for scrambled, relative to Mooney, faces (cluster-based permutation testing scrambled > Mooney: summed t-value = −189, p = 0.78).

Source localization on this time–frequency window reveals enhanced gamma power for Mooney faces, relative to scrambled faces, in a cortical source cluster extending from early visual to inferotemporal and parietal cortices (cluster-based permutation testing: summed t-value = 1229, p = 0.04; Fig. 3B). Mean power changes relative to baseline (± s.e.m.) across all voxels within the significant cluster are 2.3% ± 0.5 for Mooney and 0.4% ± 0.4 for scrambled faces. Three peak voxels (t (17) > 2.8; p < 0.01) within this cluster can be localized to early visual, fusiform and parietal cortices of the right cerebral hemisphere (see Table 1 for MNI coordinates and relative power changes for the peak voxels). These source locations served as reference sources for the DCM models. Fig. 3C shows the relative power changes for Mooney with respect to scrambled faces for an early (333–433 ms ± 100 ms) and late (433–538 ms ± 100 ms) time segment of the time window of interest. Fusiform and parietal cortex activities started at earlier latencies, whereas early visual cortex activation occurred at later stages (Figs. 3C and D).

We next employed dynamic causal modeling to test biological plausible, effective connectivity models of coupled gamma oscillators in early visual, fusiform and parietal cortices (coordinates derived from peak voxels; see Table 1) at time frequency bins of significant induced gamma power differences between picture categories (Mooney vs. scrambled faces). Ten DCM models were inverted as depicted in Fig. 4: the observed data is least likely under the model assuming no modulation of connectivity by Mooney face perception between the three cortical brain areas. In contrast, the effective connectivity model of bi-directional coupling between fusiform and parietal cortices and feedback connectivity modulation from these areas to early visual cortex best explains observed gamma power modulations by Mooney faces (Fig. 5A). The best and the second best performing models differed by 5846 with respect to log evidences, suggesting very strong support for the best fitting model (model posterior probability 100%; Fig. 5B). A difference of 3 is considered as evidence favoring one model. The percentage changes of effective connectivity between cortical sources induced by Mooney faces together with their corresponding conditional probabilities are shown in Fig. 5C. Finally, the best model's data fit was reasonable and is provided in Supplementary material.

Discussion

We observe increased gamma power for Mooney vs. scrambled faces in the right early visual, fusiform, and parietal cortices, in keeping with gamma power increases previously reported during perceptual completion in humans (Grunzner et al., 2010; Keil et al., 1999; Rodriguez et al., 1999; Tallon-Baudry and Bertrand, 1999; Trujillo et al., 2005). Increased gamma band activity occurs between 233 and 638 ms post-stimulus time, consistent with the latency of perceptual closure-related event-related potentials (ERPs) (Doniger et al., 2000; Doniger et al., 2001; Sehatpour et al., 2005), and MEG gamma power modulations during perceptual closure of Mooney faces (Grunzner et al., 2010) and perception of complex objects (Gruber et al., 2008). Further, gamma band modulations by Mooney faces were induced and not phase locked (Supplementary material and Fig. S2) as has been reported before for holistic Gestalt perception (Jensen et al., 2007; Tallon-Baudry et al., 1997). In direct support of the anatomical and temporal precision of the effects we report, intracranial data from electrodes in occipital, fusiform and parietal sites during Mooney face perception (Lachaux et al., 2005) overlap with the time interval and anatomical locations observed in the present study.

Critically, we demonstrate that Mooney face perception is associated with feedback coupling of gamma band oscillators from fusiform and parietal to early visual cortex. Importantly, our results indicate that the consistently reported gamma band activity during Mooney face perception (Grunzner et al., 2010; Lachaux et al., 2005; Rodriguez et al., 1999; Trujillo et al., 2005) reflects dynamic neuronal feedback interactions between higher and lower order visual cortices. Thus, our data unite previous findings of gamma power modulations and the concept of reverse hierarchical processing during perceptual completion (Altmann et al., 2003; Bullier, 2001; Campana and Tallon-Baudry,
Feedback coupling downstream to early visual cortex as detected by our DCM analysis is in line with low level visual cortex activation at later latencies during perceptual completion after fusiform cortex engagement (Figs. 3C and D). Similar re-activation patterns in coherent perception for early visual cortex have been reported before (e.g. Jiang et al., 2008).

Gamma band activity is considered a fundamental activity mode for information processing not only for perception but also for higher cognitive functions such as attention and memory (Fries et al., 2007; Jensen et al., 2007). One proposal is that the amplitude of pyramidal cell excitation is re-coded as the time of occurrence of output spikes relative to the inhibitory gamma cycle, stronger inputs leading to earlier responses (Olivier et al., 2003). Within this framework the gamma cycle is viewed as a temporal reference frame for sharing spike-phase coded information in neuronal networks (Fries et al., 2007; Singer and Gray, 1995).

Thus, one interpretation of the top-down feedback coupling from fusiform and parietal to early visual cortex we observe is that neuronal populations in lower sensory cortex are tuned to temporal reference codes of neuronal activity patterns pertaining to face-selective and attention-relevant brain areas. The fusiform face area is a higher order brain region for processing facial stimulus features (Calvert and Seger, 2009; Kanwisher et al., 1998), whereas the parietal cortex is part of a cortical spatial attention network (Corbetta, 1998; Robertson, 2003; Slagter et al., 2007). Gamma frequency feedback coupling from these areas could be viewed as conveying global aspects of a Mooney face, such as facial features and spatial relationships between face parts, to local processing units in early visual cortex by sharing the temporal scheme of neuronal activity throughout a cortical network relevant for perceptual completion. Such a mechanism would indeed parallel the classic Gestalt idea that local stimulus parts are processed as a function of global aspects of the stimulus (Wagemans et al., 2012b).

Further, coherent perception of Mooney faces depends on reciprocal forward and backward connectivity modulations of gamma activity between fusiform face area and parietal cortex. In view of the parietal role in spatial attention (Corbetta, 1998; Robertson, 2003; Slagter et al., 2007), we suggest that gamma frequency coupling between fusiform face area and parietal cortex reflects the detection and integration of facial spatial regularities (Calvert and Seger, 2009) into a whole Gestalt.

Some limitations should be considered in interpreting our results. First, the precision of cortical source localization is limited due to the inverse problem underlying any electromagnetic source modeling (Hauk, 2004). However, our source data are consistent with localizations from fMRI (Anders and Schlappeck, 2004; Andrews et al., 2002; Dolan et al., 1997; Kanwisher et al., 1998; Kleinschmidt et al., 1998; McKeeff and Tong, 2007; Schatpur et al., 2006), EEG source localization (Schatpur et al., 2006), and intracranial recordings (Lachaux et al., 2005) in perceptual completion tasks. Furthermore, that Mooney face perception-induced gamma power modulations are restricted to the right hemisphere is consistent with evidence that face perception in humans is right lateralized (Hillger and Koenig, 1991; Ramon and Rossion, 2012; Rossion et al., 2000), as demonstrated by behavioral visual hemi-field (e.g. Heller and Levy, 1981; Parkin and Williamson, 1987), brain lesion (e.g. Sergent and Signoret, 1992; Wada and Yamamoto, 2001) and functional brain imaging studies (e.g. Kanwisher et al., 1997; Rossion et al., 2012).

A second limitation with respect to the DCM connectivity analysis arises from the beamformer approach for cortical source localization. Conventional beamformers have problems in separating correlated sources. However, simulation studies have shown that the LCMV beamformer utilized in our study can separate sources that are correlated by 50–55% (Belardinelli et al., 2012; Van Veen et al., 1997). Therefore, the LCMV beamformer is a valid method for connectivity analysis. However, highly correlated sources (e.g. 100% correlation) cannot be co-localized by this method.

A third limitation is that Bayesian model selection indicates the probability of the data given a model from a model pool chosen a-priori based on plausible connectivity configurations, i.e. not all possible models are tested. However, we selected our models in line with reverse hierarchical processing in visual perception and its alternatives. Further, we specifically tested a null model that assumed no connectivity modulation between relevant brain areas as it has been argued that gamma band oscillations are better suited for local than for long-distance synchronization (Kopell et al., 2000). Our data are not consistent with this notion, as Bayesian model selection clearly identified the connectivity model without gamma band coupling between early visual, fusiform, and parietal cortices as the least likely model given the data. Thus, our DCM results are in line with long distance gamma band synchronization previously reported in animals (Engel et al., 1991a, 1991b; Konig et al., 1995; von Stein et al., 2000; Saalmann et al., 2007) and humans (Hipp et al., 2011; Rodriguez et al., 1999; Schoffelen et al., 2005, 2011; Siegel et al., 2008).

We note a recent suggestion that feedforward and feedback cortical connectivity may not occur at the same frequencies due to asymmetric intrinsic properties of neurons in superficial and deeper cortical layers (Bastos et al., 2012). Specifically, feedforward connections originate predominately from superficial layers, which typically show neuronal synchronization predominantly in the gamma range, whereas feedback connections arise from deep layers, which prefer lower (alpha or beta) frequencies (Buffalo et al., 2011; Maier et al., 2010; Roopun et al., 2006, 2008). We acknowledge that the class of dynamic causal model employed here (based on cross-spectral density) is agnostic to layer-specific parameters. Although this was the best DCM routine available at the time of analysis, we note that a novel routine within the SPM DCM framework, canonical microcircuits (CMC), has recently been implemented which specifically takes these parameters into account. Thus, although we show top-down coupling in the gamma range during perceptual completion, the CMC framework will enable the relative contributions of different frequency bands to feedforward vs. feedback coupling to be elucidated. However, in non-human primates performing a visual attention task, directed granger causal connectivity in the gamma range has also been observed in the backward direction (V4 to V1), although the forward gamma coupling was stronger (Bosman et al., 2012). Bayesian computational views of perception stress that purely forward architectures of the visual system are not sufficient for perception and that top-down feedback connections from higher order regions are necessary (Friston, 2003). Reverse hierarchical processing in the visual system extends beyond perceptual completion. For example, movement perception is abolished when feedback from higher order area MT to primary visual cortex is disrupted by transcranial magnetic stimulation (TMS) (Koivisto et al., 2010; Pascal-Leone and Walsh, 2001; Silvanto et al., 2005). In fact, neuronal interactions between association and lower order sensory cortices are probably a prerequisite for conscious perception in general (Boly et al., 2011). Thus, our data link current views of distributed cortical gamma power activity in perceptual synthesis with the notion that top-down feedback from higher to lower order visual cortex is required for awareness of coherent visual percepts.

Acknowledgments

This work was supported by a national grant from the Spanish Ministry of Science and Education (PSI-2009-12702), a Ramón y Cajal fellowship to SM (RYC-2009-04974), a PICATA predoctoral fellowship of CEI Moncloa (UCM-UPM) to CM, and a Marie Curie Career Integration Grant to BS. We thank L.T. Trujillo for stimulus material, R Moran for helpful discussion and S Bestmann for comments on an earlier version of this manuscript.

Conflict of interests

The authors do not report any conflict of interests.