Neuronal integration in visual cortex elevates face category tuning to conscious face perception

Johannes J. Fahrenfort[a,b,1], Tineke M. Snijders[c,d], Klaartje Heinen[e], Simon van Gaal[a,b,f,g], H. Steven Scholte[a,b], and Victor A. F. Lamme[a,b]

[a]Brain and Cognition, Department of Psychology, University of Amsterdam, 1018 XA, Amsterdam, The Netherlands; [b]Cognitive Science Center Amsterdam, University of Amsterdam, 1018 WS, Amsterdam, The Netherlands; [c]Helmholtz Institute, Department of Experimental Psychology, Utrecht University, 3508 TC, Utrecht, The Netherlands; [d]Rudolf Magnus Institute of Neuroscience, Department of Child and Adolescent Psychiatry, University Medical Centre Utrecht, 3508 GA, Utrecht, The Netherlands; [e]University College London Institute of Cognitive Neuroscience and Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3AR, United Kingdom; [f]Cognitive Neuroimaging Unit, Institut National de la Santé et de la Recherche Médicale, F91191 Gif-sur-Yvette, France; and [g]NeuroSpin Center, Commissariat à l'Energie Atomique, F91191 Gif-sur-Yvette, France

Edited by Charles Gross, Princeton University, Princeton, NJ, and approved November 12, 2012 (received for review May 10, 2012)

The human brain has the extraordinary capability to transform cluttered sensory input into distinct object representations. For example, it is able to rapidly and seemingly without effort detect object categories in complex natural scenes. Surprisingly, category tuning is not sufficient to achieve conscious recognition of objects. What neural process beyond category extraction might elevate neural representations to the level where objects are consciously perceived? Here we show that visible and invisible faces produce similar category-selective responses in the ventral visual cortex. The pattern of neural activity evoked by visible faces could be used to decode the presence of invisible faces and vice versa. However, only visible faces caused extensive response enhancements and changes in neural oscillatory synchronization, as well as increased functional connectivity between higher and lower visual areas. We conclude that conscious face perception is more tightly linked to neural processes of sustained information integration and binding than to processes accommodating face category tuning.

consciousness | object categorization | figure–ground segregation | perceptual organization | recurrent processing

Image processing and image perception are not the same thing. Faces, for example, are processed even when they are invisible to the observer (1), and face processing has been shown to continue in anesthetized macaques (2). Apparently, even highly complex neuronal tuning responses are not sufficient to achieve perception. So what neural operations turn complex tuning responses into conscious representations? A window into these operations might be obtained by investigating perceptual organization. Perceptual organization is an umbrella term for cortical functions that organize sensory input into coherent and interpretable perceptual structures. It is thought to encompass processes such as figure–ground segregation, object detection, and object categorization (3). Theoretical accounts of human vision going back to Rubin (1915) have suggested that perception requires objects to first be segregated from their background (3–5) and that object recognition follows figure–ground segregation. This notion has been challenged by behavioral studies showing that object recognition influences figure–ground assignment and might even precede it (6). Given their putative role in perception, resolving the relationship between category tuning and figure–ground segregation may tell us how sensory input is transformed into perceptual representations.

To investigate this relationship, we measured signals of category tuning and figure–background processing while subjects viewed visible and invisible objects. Given the speed of object category extraction (7, 8), we hypothesized that basic-level category tuning results from simple-to-complex feedforward computations (9) and is unrelated to object perception (1). We therefore predicted that it should be possible for both visible and invisible objects to elicit category tuning. Figure–ground processing, on the other hand, requires incremental grouping mechanisms in which tuning information from neurons with different receptive field sizes is integrated across visual areas over longer periods of time (10, 11). It has been suggested that mechanisms of neuronal integration are crucial to feature binding (12) and conscious perception (13). These claims are motivated by key properties of everyday perception that can be uniquely explained by mechanisms of information integration: differentiation, the availability of a seemingly infinite number of conscious experiences, and integration, the perceptual unity of each of these experiences (14). Therefore, our second prediction was that only consciously segregated objects show markers of sustained neuronal integration, directly linking neuronal integration to conscious experience.

Using a dichoptic fusion paradigm (1, 11), we presented objects that were either visible or not visible (Fig. 1*A*). Faces, houses, nonsense objects, and homogenous screens (Fig. 1*B*) were constructed using orientation-defined textures of Gabor elements. Monocularly, objects were created using different orientations for object and background (Fig. 1*A*, left and right eye). When properly viewing stimuli with both eyes, textures in the left and right eye fuse together. The fused textures could either be different for object and background (visible, Fig. 1*A*, *Upper*) or the same for object and background (invisible, Fig. 1*A*, *Lower*). Despite having very different perceptual properties when viewed with both eyes, average monocular stimulation was the same for visible and invisible conditions, allowing us to investigate how signals of category tuning and figure–ground processing are impacted when viewing visible and invisible objects. Neural correlates of category tuning were obtained by contrasting objects of different categories among each other. Comparing these objects with homogenous textures made it possible to look at the processes that are active when segregating objects from their background (see Figs. S1 and S2 for examples of these contrasts and *SI Methods* for details). To be sure, we should note that the figure–ground contrast not only highlights figure–ground processes but also category extraction, as individual object stimuli contain a category, whereas homogenous textures do not. Because functional MRI (fMRI) suffers from spatial blood-oxygen level–dependent (BOLD) summation of the underlying neuronal responses (15), the generic figure–ground profile in category-selective regions could also reflect averaged individual category-selective responses, even though we refer to this contrast as the figure–ground contrast for ease of reference.

Results

To minimize attentional and postperceptual differences between visible and invisible conditions (16), subjects performed a distractor task instead of stimulus categorization. On each trial they

Author contributions: J.J.F., K.H., S.v.G., H.S.S., and V.A.F.L. designed research; J.J.F. performed research; J.J.F. and T.M.S. analyzed data; and J.J.F. and T.M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: fahrenfort.work@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1207414110/-/DCSupplemental.

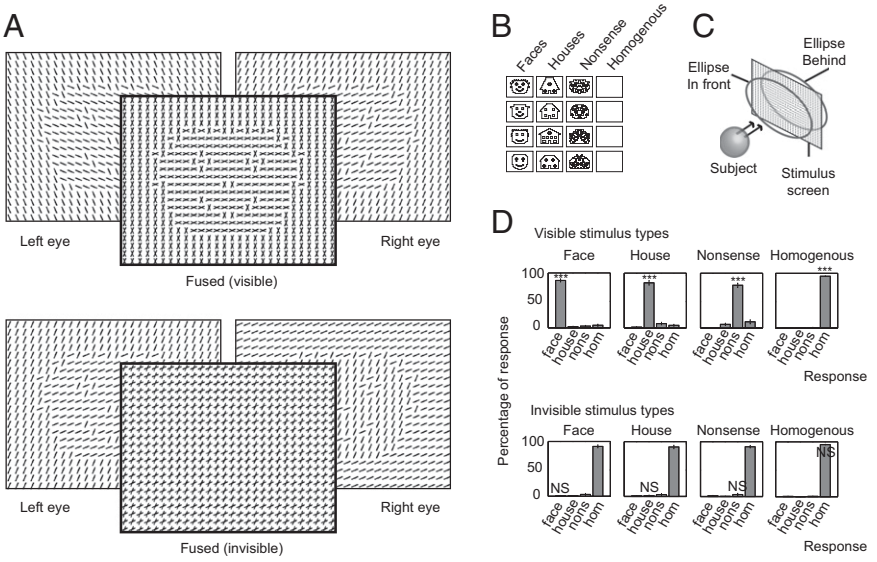

Fig. 1. Experimental setup. (*A*) Illustrations of visible (*Upper*) and invisible (*Lower*) dichoptic fusion using oriented line elements. Objects were defined using a 45° orientation difference between figure and background elements. (*B*) Illustrations of the four stimulus types: faces, houses, nonsense objects, and homogenous screens. Black lines illustrate orientation discontinuities. (*C*) 3D depiction of a subject performing the in front/behind distractor task. Both in front and behind are shown for illustrative purposes. (*D*) Stimulus classification during the post-scanning control task (visible: upper panels; invisible: lower panels). Graphs show the mean response percentage (±SEM) for each response to that stimulus type. Paired *t* test of the hits against false alarms determined whether a category was identified above chance level ($^{N.S.}P > 0.05$; $^{***}P < 10^{-8}$).

had to indicate whether a stereoscopically presented ellipse appeared to be hovering in front or behind the stimulus screen due to a slight offset in left and right eye (Fig. 1*C*; *SI Methods*). The task could only be performed reliably when stereoscopically focusing on the fixation dot, requiring subjects to fuse left and right eye images. In a behavioral control task performed directly after scanning, but still inside the scanner, subjects combined ellipse localization with a stimulus categorization response (face, house, nonsense object, or homogenous; Fig. 1*D*). Task performance on ellipse localization during the control task was the same as during scanning (mean hit rate control = 0.91 vs. scanning = 0.90, $F_{1,15} = 0.445$, $P = 0.515$), confirming that stereoscopic viewing conditions between the experimental and control runs were comparable. Visible categories were classified with high accuracy (mean $d' = 3.47$, $F_{1,15} = 548.65$, $P < 10^{-12}$, all individual $P < 10^{-9}$), whereas none of the invisible categories were identified above chance level (mean $d' = 0.11$, $F_{1,15} = 1.58$, $P = 0.228$, all individual $P > 0.18$).

To identify signals related to category tuning, we used fMRI to isolate classical face- and place-selective areas (17, 18). The face-selective contrast (faces > houses and nonsense objects, visible and invisible combined) activated well-defined group-level clusters in the fusiform cortex (Fig. 2*A*) known as the fusiform face area (FFA). A localizer using photographs of faces and other objects was used to verify that these voxels corresponded to the classical FFA (Fig. S3). The contrast houses > faces and nonsense objects did not reveal place-selective clusters in the parahippocampal place area (PPA). Consequently, the focus of this paper is on face selectivity, although we test for other category-selective responses in all analyses.

We examined BOLD signals in the left and right FFA of functional data that were independent from the data used for FFA selection (*Methods*; Fig. S3). A three-way ANOVA showed that these responses were face selective (faces vs. houses and nonsense objects: $F_{1,15} = 11.26$, $P = 0.004$), larger to visible than to invisible stimuli ($F_{1,15} = 42.68$, $P < 10^{-5}$), and that responses were stronger in the right than in the left hemisphere ($F_{1,15} = 11.70$, $P = 0.004$). Strikingly, face selectivity did not interact with visibility ($F_{1,15} = 1.13$, $P = 0.304$), indicating that category-selective responses to visible and invisible faces were comparable in strength (Fig. 2 *B* and *C*). The absence of an interaction between face selectivity and visibility was confirmed by three ANOVAs that used various combinations of the face category and other categories. Although all ANOVAs showed main effects of visibility and category, none showed an interaction between visibility and category (faces, houses, nonsense objects: $F_{2,14} = 1.55$, $P = 0.247$; faces, houses: $F_{1,15} = 0.404$, $P = 0.534$; faces, nonsense objects: $F_{1,15} = 3.11$, $P = 0.098$). Paired *t* tests further confirmed the presence of face-selective responses in the left and right FFA for both visible and invisible stimuli (invisible left: $t_{15} = 2.06$, $P = 0.029$; visible left: $t_{15} = 2.85$, $P = 0.006$; invisible right: $t_{15} = 2.59$, $P = 0.010$; visible right: $t_{15} = 2.20$, $P = 0.022$; one-tailed), showing that the FFA contains voxels from which face-category information can be extracted, irrespective of whether a face is perceived or not.

To determine the degree to which this is reflected in the time course of the response, we deconvolved the hemodynamic response function (HRF) in the FFA for each of the conditions (19).

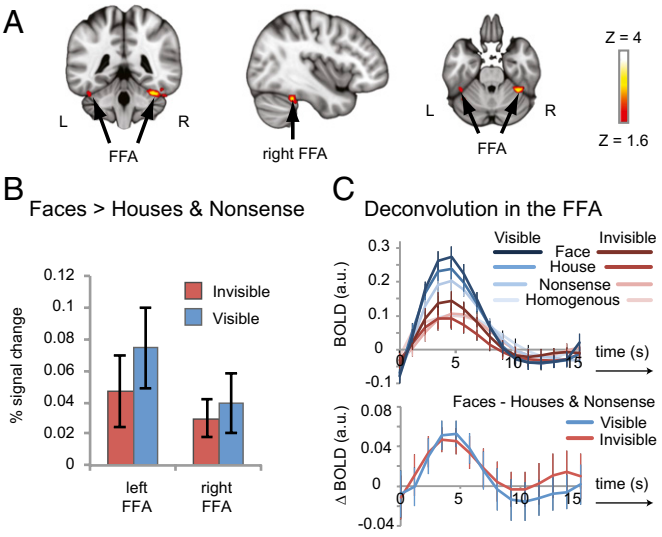

Fig. 2. Univariate category-selective activation to visible and invisible faces. (*A*) The fusiform face area (FFA) was defined as those voxels within ventral fusiform cortex responding more strongly to face stimuli than to houses and other objects. Shown is the averaged response from two datasets, each obtained through a split half procedure (see *Methods* and Fig. S3 for details). MNI coordinates: $x = 40$, $y = -43$, $z = -27$. (*B*) Mean percent signal change obtained from the split half procedure (error bars, ±SEM) for the faces > houses and objects contrast, separately for visible and invisible stimuli in the left and right FFA. (*C*) Deconvolution of the hemodynamic response to each stimulus category in the FFA (*Upper*) and the face-selective part of the response to visible and invisible stimuli in the FFA (*Lower*).

Overall response amplitudes to visible categories were much larger than to invisible categories (Fig. 2*C*, *Upper*), consistent with observations that perceived items lead to stronger responses than items that are not perceived (1, 20). Notably however, FFA responses to visible faces were strongest among all visible conditions, and responses to invisible faces were strongest among all invisible conditions. Isolation of the face-selective part of the time course demonstrates that visible and invisible face-selective HRFs are actually quite similar (Fig. 2*C*, *Lower*), seemingly at odds with many studies that do show a relationship between category representations and stimulus visibility (21). This apparent contradiction is treated in more detail in *SI Discussion*.

Although these findings confirm that visible and invisible faces are able to produce face-selective responses (1), it leaves unresolved to what degree the distributed cortical representation of visible and invisible category information is similar. Multivoxel response patterns are known to be more sensitive to category information than univariate responses and are thought to reflect distributed neural representations (22). We scanned all volumes of each subject using a spherical searchlight kernel (23). Correlations were computed between spatial activation patterns across voxels in the kernel, at each location in the brain, for each possible combination of visible and invisible categories. Next we performed a permutation test at each location to assess whether the pattern correlation around that location carried visibility invariant category information. Specifically, we tested whether within-category visible-invisible correlations were higher than between-category visible-invisible correlations (*SI Methods*). This analysis was performed separately for faces, houses, and nonsense objects. Visible-invisible face correlations (large green dot) were significantly higher than all between-category correlations (small green dots) in a right-lateralized cluster extending across the right occipitotemporal fusiform cortex, confirming the presence of a high-level representation of face-category information that is invariant to stimulus visibility (Fig. 3*A*; familywise error rate controlled at $P < 0.01$). The same cluster emerged when testing whether visible faces were uniquely correlated with invisible faces (large red dot over small red dots, Fig. 3*B*) and when testing whether invisible faces were uniquely correlated with visible faces (large blue dot over small blue dots, Fig. 3*C*). House and nonsense categories did not result in visibility invariant category extraction anywhere. Some may wonder how monocularly presented faces are able to penetrate into high-level visual cortex when their constituent elements seem to be fused at lower levels. This topic is covered in more detail in *SI Discussion*.

To further confirm that visibility invariant face-category extraction was specific to high-level visual cortex and not driven by low level visual information, we attempted to classify the multivoxel patterns of the visible object categories using the invisible category patterns and vice versa in two regions of interest (ROIs): (*i*) left and right ventral occipitotemporal cortex (VOT) (24) and (*ii*) Brodmann area 17 (BA17) (Fig. S4; *SI Methods*). VOT comprises a set of object-, face- (17), and place-selective (18) regions bounded by the fusiform gyrus and the parahippocampal gyrus. BA17 (early visual cortex) is known to be sensitive to low-level image features but insensitive to category information (24). Fig. 3*D* shows that only the right VOT contained multivoxel activation patterns that allowed categorization of invisible faces using visible faces (permutation test: $P = 0.0004$) and vice versa ($P = 0.0012$), confirming that the visible-invisible pattern overlap is specific to high-level (right-lateralized) object-selective cortex and specific to faces.

These results show that visible and invisible faces produce and share univariate and multivariate category-selective responses. Central to our research question, we wanted to determine what is required in addition to category tuning to achieve perception. We speculated that object perception critically depends on sustained spatiotemporal integration during figure–ground processing. To isolate regions putatively involved in figure–ground processing, we contrasted all stimuli (visible and invisible combined) containing a texture-defined surface with textures that do not contain

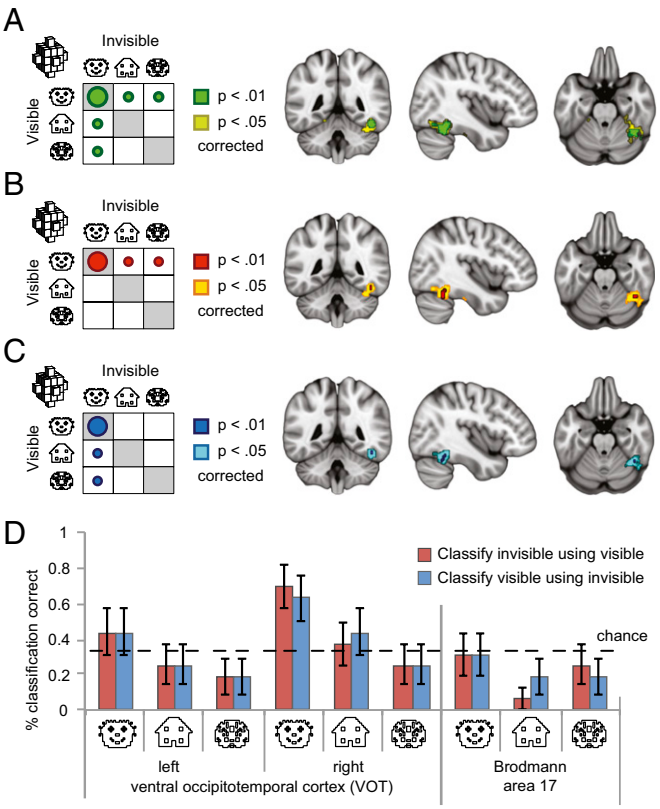

Fig. 3. Analysis of multivoxel pattern overlap between visible and invisible face-selective responses. (*A*) Whole brain analysis of searchlight locations for which the pattern correlations between visible and invisible faces in the kernel were significantly higher than the pattern correlations between nonmatching categories. (*B*) Searchlight locations where pattern responses to visible faces were uniquely correlated with invisible faces over other invisible object categories. (*C*) Searchlight locations where pattern responses to invisible faces were uniquely correlated with visible faces over other visible object categories. *P* values were corrected for multiple comparisons by controlling the familywise error rate. Brains are shown at MNI coordinates: $x = 41$, $y = -46$, $z = -25$. (*D*) Pattern classification of multivoxel responses to visible categories using responses to invisible categories and vice versa (error bars, ±SEM) in VOT and BA17.

a surface (houses, faces, and nonsense objects > homogenous; Figs. S1*C* and S2*A*). This contrast activated a large cluster of areas in the ventral and dorsal visual pathways corresponding to regions that have been implicated in object (24) and figure–ground processing (11) (Fig. 4*A*; cluster creation threshold, $Z > 2.3$; cluster-corrected probability, $P < 0.05$). We divided this cluster into four atlas-defined ROIs, based on the functional significance of these regions during object processing (*Methods*): (*i*) activation in the superior lateral occipital cortex (dorsal foci), (*ii*) activation in the inferior lateral occipital cortex (LOC), (*iii*) VOT, and (*iv*) BA17.

To investigate whether visible and invisible objects contributed equally to the figure–ground signal, we estimated percent BOLD signal changes in functional data that were not used during ROI selection (*Methods*). A three-way ANOVA revealed main effects of figure–ground processing (faces, houses, and nonsense objects vs. homogenous textures: $F_{1,15} = 15.02$, $P = 0.001$), larger responses to visible than to invisible stimuli ($F_{1,15} = 54.78$, $P < 10^{-5}$), and response strength differentiation across all regions of interest (FFA, VOT, LOC, dorsal foci, BA17), with the largest responses in early visual cortex (BA17) weakening along the posterior-anterior dimension ($F_{4,12} = 23.00$, $P < 10^{-4}$, see Fig. S5 for responses to all conditions in all ROIs). Strikingly, although we showed that face-selective responses in our FFA ROI are invariant to stimulus visibility, there was a strong interaction

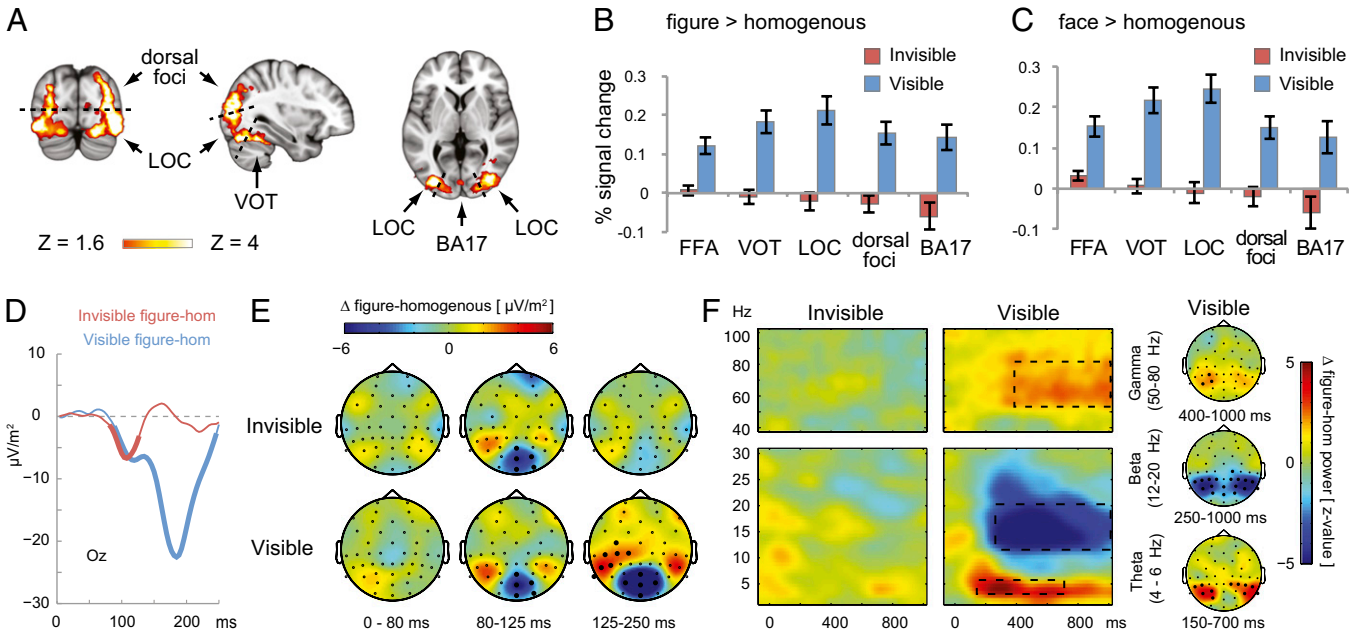

Fig. 4. Figure–ground responses. (*A*) Figure–ground modulation (faces, houses, and nonsense objects > homogenous textures), from which VOT, dorsal foci, LOC, and BA17 were selected. Shown is the averaged response from two datasets, each obtained from a split half procedure (*Methods*). MNI coordinates: $x = 33$, $y = -81$, $z = 3$. (*B*) Figure–ground modulation and (*C*) face-ground modulation (error bars, ±SEM) during the first functional runs. (*D*) The evoked figure–ground EEG response to visible and invisible stimuli at electrode Oz, low-pass filtered at 20 Hz for illustration purposes. Thick lines indicate cluster-corrected significance. (*E*) Topographic distributions of the figure–ground signal evoked by invisible (*Upper*) and visible (*Lower*) figures. Thick electrodes indicate cluster-corrected significance. (*F*) Induced EEG responses to figure–ground modulation. (*Left*) Time frequency representations of the figure–ground signal at electrodes P3, PO3, PO7, P4, PO4, and P08 of bilateral occipitotemporal cortex. (*Right*) Topographic distribution of significant clusters, indicated by the dashed lines in the time-frequency distribution on the left (only present in the visible condition).

effect between stimulus visibility and the figure–ground response ($F_{1,15} = 43.96$, $P < 10^{-5}$), with figure–ground responses occurring only for visible objects. Fig. 4*B* shows figure–ground responses to visible and invisible objects in the four ROIs and in the FFA. Two-tailed paired t tests confirmed that visible objects strongly modulated figure–ground signals in each of the ROIs: VOT ($t_{15} = 5.99$, $P < 10^{-4}$), LOC ($t_{15} = 5.78$, $P < 10^{-4}$), dorsal foci ($t_{15} = 5.23$, $P < 0.001$), and BA17 ($t_{15} = 4.33$, $P = 0.001$), as well as in the FFA ($t_{15} = 5.84$, $P < 10^{-4}$), whereas invisible stimuli produced no significant figure–ground response in any of these areas in either direction (all $P > 0.1$). Note that BA17 seems to show a negative invisible figure–ground and face-ground responses (Fig. 4 *B* and *C*), but these effects are not significant (figure–ground: $t_{15} = -1.75$, $P = 0.1$; face-ground: $t_{15} = -1.51$, $P = 0.153$). A whole-brain analysis using all runs from the dataset showed no figure–ground responses to invisible objects anywhere (cluster creation threshold, $Z > 1.6$; cluster-corrected probability, $P < 0.05$).

The fact that these figure–ground signals occur only when viewing visible objects might be taken as a first indication that face perception is more strongly related to figure–ground processing than to face-category tuning. When inspecting the face-specific figure–ground contrast (faces > homogenous screens; Fig. 4*C*), the result is nearly identical, although invisible faces did activate the FFA more than homogenous textures (FFA: $t_{15} = 2.73$, $P = 0.032$). This finding is unsurprising, given that the face-ground contrast not only captures a figure–ground relationship but also a face-selective response. This invisible face-ground response is not significantly different from the face-selective response in the FFA ($t_{15} = -0.309$, $P = 0.762$; Fig. 2*B*), suggesting that it reflects a face-selective rather than a figure–ground response. Interestingly, the visible face-ground response is about five times larger than the visible face-selective response, showing the large contribution of the object-ground relationship to the overall response. Separate face-ground, object-ground, and house-ground ROIs and their corresponding figure–ground profiles can be found in Fig. S6.

To further investigate the presence of invisible face-ground signals, we performed a whole-brain conjunction analysis of visible and invisible face-ground responses, looking for voxels that were jointly activated by visible face-ground and invisible face-ground contrasts ($P < 0.01$ in both contrasts, uncorrected). This analysis turned up a small bilateral cluster in the FFA/VOT only (peak activations—left: $x = -34$, $y = -60$, $z = -26$; right: $x = 36$, $y = -40$, $z = -30$), but no activations anywhere else. Together, these results show that visible objects cause widespread figure–ground modulation across the visual cortex, whereas invisible objects do not, suggesting that faces that do not elicit a figure–ground response across the visual cortex still evoke category-selective responses in the FFA. See *SI Discussion* for a more detailed account of how this might work.

Crucially, we wanted to know whether consciously segregated objects uniquely display markers of neuronal integration (14). Although information integration is difficult to quantify in practice (25), educated guesses can be made with respect to the types of neural activity that could reflect it. For example, information integration across spatially separated sets of neurons might be coded through neural synchrony (12) or concurrent response enhancement (10). Such mechanisms would manifest as power changes in the time-frequency domain of the EEG or through increased functional connectivity between areas as measured with fMRI (26). To test this hypothesis, we collected EEGs using the same experimental protocol that was used during the fMRI experiment. t tests showed that subjects from both experiments performed the same across all behavioral measures (see Fig. S7 for categorization responses during the EEG experiment). To increase spatial resolution and to filter out deep sources, we generated a scalp current source density (CSD) estimate of the EEG using spherical splines (27). Evoked responses (ERPs) were obtained by averaging the CSD waveforms from stimulus onset onward. We isolated figure–ground responses by subtracting the ERP to homogenous textures from ERPs to stimuli containing a texture-defined surface (faces, houses, and nonsense objects

minus homogenous textures), separately for the visible and the invisible conditions (Fig. 4*D*). Fig. 4*E* shows the topographic distribution of these figure–ground responses over time. Cluster-based permutation testing (*Methods*) revealed a negative cluster over midline occipital electrodes during the 80- to 125-ms interval for both visible ($P = 0.004$) and invisible conditions ($P = 0.002$). Beyond 125 ms, only visible figures evoked a sustained negative response compared with homogenous textures ($P < 0.001$), accompanied by a positive response spreading into the temporo-parietal cortex (left cluster: $P = 0.004$; right cluster $P = 0.22$). Together, these data suggest that visible figure–ground responses are widespread across visual cortex (Fig. 4*B*) and persistent in time, whereas the invisible figure–ground signal is focal in space (Fig. 4*C*) and transient in time (Fig. 4 *D* and *E*). We also tested for category-selective responses in the EEG signal, but neither visible nor invisible faces evoked N170 or other face-selective responses. See *SI Discussion* for further discussion of this finding.

To investigate the presence of neural synchrony in visible and invisible figure–ground responses, we performed a power analysis of the time-frequency spectrum (Fig. 4*F*). Changes in the EEG power spectrum that are not phase-locked to stimulus onset (induced) are more likely to reflect changes in neural synchrony, whereas changes that are phase-locked to stimulus onset (evoked) are more likely to reflect transient afferent activity (28). Therefore, we further isolated the induced response by subtracting the average ERP to each condition from its respective single-trial responses before performing a power spectrum analysis (*SI Methods*). A cluster-based permutation test did not reveal power changes in the time-frequency signal for invisible textures containing a figure compared with homogenous textures, suggesting that the invisible figure–ground response was indeed driven by the evoked portion of the signal. Visible figures, however, exhibited a strong power decrease in the low beta band (12–20 Hz) between 250 and 1,000 ms (cluster-level $P < 0.001$), as well as power increases in the high gamma band (50–80 Hz) at 400–1,000 ms (cluster-level $P < 0.012$) and in the theta band (4–6 Hz) between 150 and 700 ms (cluster-level $P < 0.001$), with a topographic distribution over the bilateral occipitotemporal cortex (Fig. 4*F*). These changes in the time-frequency spectrum show that the processing of visible, but not invisible, figures results in sustained oscillatory (de)synchronization of cell assemblies (28), providing a mechanism by which neurons may label shared surface and boundary elements of the figure–ground display (29). The face-ground response profile was virtually identical, although the face-ground gamma responses were weaker and did not reach statistical significance (Fig. S8). For a more in-depth treatment of functions that may be associated with this spectral fingerprint, see *SI Discussion*.

To probe whether the induced changes in the time-frequency spectrum of the visible figure–background response coincide with increased functional connectivity across the visual cortex, we took the bilateral FFA as a seed region and performed a psychophysiological interaction (PPI) analysis on the fMRI data (30) (*Methods*). Responses to visible faces in the FFA were associated with stronger functional connectivity with early visual areas than responses to invisible faces (Fig. 5), an effect that cannot be attributed to differences in response amplitude between visible and invisible faces (*SI Discussion*). This finding suggests that the processing of visible faces results in stronger intercortical integration than the processing of invisible faces and that this integrative process takes place between higher and lower visual areas on a multisecond timescale. Separate PPI analyses of visible and invisible face-ground processing showed that this effect was not an additive effect for visible over invisible faces, but was a qualitative effect. FFA activity to visible faces had increased functional connectivity with the same cortical areas compared with homogenous textures, but invisible faces showed no increase in functional connectivity with these areas compared with homogenous textures. These results provide topographically specific evidence for the locus of information integration across visual areas during the processing of visible faces.

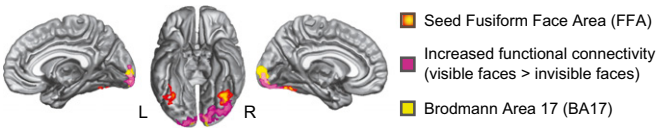

Fig. 5. Increased PPI with FFA activity when viewing visible compared with invisible faces. Ventral and medial renderings of the regions in the cortex that show increased functional connectivity with the FFA when subjects are viewing visible faces compared with when they are viewing invisible faces ($Z > 1.6$, corrected $P < 0.05$). The regions showing increased functional connectivity are indicated in purple. In the right hemisphere, this cluster runs ventrally from the right FFA through large parts of extrastriate cortex to BA17. In the left hemisphere, the cluster is confined to BA17 and BA18.

Discussion

Summarizing, just like visible faces, invisible faces generate category-specific responses but do not generate responses that contain the neural markers of cortico-cortical integration that manifest in the putative figure–ground signals of visible objects. From this, we conclude that traditional theories on human perception may be mistaken in their notion that figure–ground processing is a necessary precursor to object categorization (3–5). We suggest that face category extraction can be achieved through feedforward computations, whereas conscious representations require large-scale neuronal integration through recurrent interactions in visual cortex. Importantly, we combined visible and invisible conditions in all contrasts that were used for ROI selection. This procedure preferentially targets voxels that are selective in both the visible and invisible contrast, if they exist. We do not claim that all voxels that are face selective for visible faces should also be face selective for invisible faces or the other way around. However, because only visible objects generated markers of neuronal integration, these seem to be a better indicator of perceptual organization than category extraction itself. Further note that we only observed category-selective responses to faces. A lack of task relevance to the objects during the task may have resulted in a lack of category-selective responses to other categories (31). Indeed, others have observed category-selective responses to invisible task-relevant categories other than faces (1). Alternatively, invisible category-selective responses do not depend on task relevance but on domain specificity of the regions involved (17, 18). If this is the case, our result might not generalize to other stimulus categories, even under conditions of task relevance. *SI Discussion* contains a more in-depth treatment of this issue.

Finally, it might be argued that the markers of neuronal integration we observed are partially caused by selective attention. We observed markers of neuronal integration despite the fact that our subjects performed a distractor task during scanning in which attentional resources were directed elsewhere. Indeed, evidence suggests that perceptual organization does not depend on attention, but rather provides the structure on which it operates (32). However, even if the visible objects grabbed attention involuntarily, our conclusions remain unchanged. We do not claim that the effects of neuronal integration we observed must be caused by some specific process, but rather that (*i*) none of the effects of neuronal integration that we observed seems to be required for face-category tuning, and (*ii*) correlates of neuronal integration are a better marker for perceptual organization and conscious perception than face-category tuning. We suggest that going from face-category tuning to neuronal integration across the visual cortex marks the transition from unconscious to conscious face representations.

Methods

Subjects. Thirty-eight subjects (18 fMRI, 3 males; 20 EEG, 4 males) viewed texture stimuli in a single-session experiment. Four subjects (two EEG subjects) were excluded because of artifacts or hardware. All subjects were healthy adults with normal vision. All provided written informed consent. The research was approved by the ethical committee of the Psychology Department of the University of Amsterdam.

Equipment. Stimuli were presented at 800 × 600 resolution (16.9° × 12.7° visual angle) at a rate of 60 Hz. Images for the left and the right eye were sent through differently polarized filters while subjects viewed the screen using correspondingly polarized glasses.

Experimental Procedure. All subjects participated in a short training session. Subsequently, subjects performed the experimental task, during which fMRI or EEG imaging data were acquired (see *SI Methods* for detailed acquisition parameters). Directly after fMRI/EEG acquisition, subjects performed the control task to establish stimulus visibility.

fMRI Analysis. Using the Oxford Centre for Functional MRI of the Brain (FMRIB) Software Library (FSL), functional data were motion corrected, slice-time aligned, temporally filtered with a high-pass filter (35 s), and spatially filtered using a Gaussian envelope (univariate analyses: 5 mm; multivoxel analyses: 2 mm). Data were spatially normalized to MNI space using FMRIB's Nonlinear Image Registration Tool (FNIRT). All ROI analyses were performed using a split half procedure. Each trial was pseudorandomly assigned to one of two datasets, each containing exactly half of the data. Subsequently, one dataset was used to draw ROIs, exporting responses in those ROIs from the other dataset, and the same procedure was repeated after switching the datasets used for ROI selection and export. Finally, the exported data were averaged and used for statistical testing (Fig. S3). Whole-brain searchlight analyses were performed on the entire dataset using custom code in Matlab (Mathworks). For each subject and each run, a general linear model was created. A predictor convolved with a standard HRF modeled each condition. Single subject parameter estimates were generated by combining runs using a fixed-effects higher-level analysis. Individual subject statistics were combined using FMRIB's Local Analysis of Mixed Effects (FLAME1) for univariate and PPI analyses, correcting for multiple comparisons using cluster thresholding. Multivoxel analyses were combined across subjects using FMRIB's randomize function for nonparametric permutation testing (33), correcting for multiple comparisons using threshold-free cluster enhancement (TCFE) (34).

Functional Connectivity Analysis. A nine-column design matrix was constructed: (*i*) a regressor for the contrast of interest, e.g., visible vs. invisible faces, convolved with the HRF; (*ii*) a time series regressor containing the time course of FFA activation; (*iii*) an interaction regressor containing the interaction term between 1 and 2, and (*iv–ix*) HRF-convolved regressors for the other six stimulus types: visible and invisible houses, nonsense objects, and homogenous screens. The regression coefficient of the interaction term gives a measure of increased functional connectivity with the FFA for visible compared with invisible faces (30).

EEG Analysis. The data were preprocessed using the Brain Vision Analyzer with a high-pass filter (0.5 Hz), a notch filter (50 Hz), and segmented into (−300, 1,300) millisecond periods. Segments containing transients exceeding ±200 μV were removed before ocular artifact correction. Finally, segments containing transients exceeding ±50 μV were removed, and a CSD transformation was applied to obtain reference-free data. All subsequent analyses were performed in Matlab with the FieldTrip toolbox (35) (Radboud University). Differences between conditions were assessed using cluster-based permutation testing (36).

ERPs. We visualized three time bins based on previously established properties of the ERP during the processing of figure–ground displays (37): 0–80 ms (does not differentiate between figure and no-figure textures), 80–125 ms (first figure–ground response but does not correlate with perception), and 125–250 ms (first figure–ground response to correlate with perception). We performed a permutation test over each averaged bin to establish clusters of electrodes exhibiting an evoked response in that time window.

ACKNOWLEDGMENTS. We thank Tomas Knapen and two anonymous reviewers for helpful comments on earlier drafts of this manuscript. T.M.S. is supported by Vici Grant 453-07-004 from the Netherlands Organisation for Scientific Research and V.A.F.L. is supported by Advanced Investigator Grant 230355 from the European Research Council.

1. Moutoussis K, Zeki S (2002) The relationship between cortical activation and perception investigated with invisible stimuli. *Proc Natl Acad Sci USA* 99(14):9527–9532.
2. Ku SP, Tolias AS, Logothetis NK, Goense J (2011) fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. *Neuron* 70(2): 352–362.
3. Marr D (1982) *Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information* (W. H. Freeman, San Francisco).
4. Rubin E (1958) *Figure and Ground* (Van Nostrand, New York), pp 194–203.
5. Nakayama K, Zijiang JH, Shinsuke S (1995) Visual surface representation: A critical link between lower-level and higher-level vision. *An Invitation to Cognitive Science*, eds Osherson DN, Gleitman LR, Kosslyn SM (MIT Press, Cambridge), pp 1–70.
6. Peterson MA, Gibson BS (1994) Must figure-ground organization precede object recognition? An assumption in peril. *Psychol Sci* 5(5):253–259.
7. Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62(2):281–290.
8. Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381(6582):520–522.
9. Serre T, Oliva A, Poggio T (2007) A feedforward architecture accounts for rapid categorization. *Proc Natl Acad Sci USA* 104(15):6424–6429.
10. Roelfsema PR (2006) Cortical algorithms for perceptual grouping. *Annu Rev Neurosci* 29:203–227.
11. Zipser K, Lamme VAF, Schiller PH (1996) Contextual modulation in primary visual cortex. *J Neurosci* 16(22):7376–7389.
12. Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci* 18:555–586.
13. Tononi G, Edelman GM (1998) Consciousness and complexity. *Science* 282(5395): 1846–1851.
14. Tononi G (2004) An information integration theory of consciousness. *BMC Neurosci* 5:42.
15. Avidan G, Hasson U, Hendler T, Zohary E, Malach R (2002) Analysis of the neuronal selectivity underlying low fMRI signals. *Curr Biol* 12(12):964–972.
16. Koch C, Tsuchiya N (2007) Attention and consciousness: Two distinct brain processes. *Trends Cogn Sci* 11(1):16–22.
17. Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17(11): 4302–4311.
18. Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. *Nature* 392(6676):598–601.
19. Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *Neuroimage* 9(4):416–429.
20. Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 3(8): 837–843.
21. Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21(4):753–759.
22. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539):2425–2430.
23. Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci USA* 103(10):3863–3868.
24. Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27: 649–677.
25. Tononi G, Sporns O (2003) Measuring information integration. *BMC Neurosci* 4:31.
26. Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001) The brainweb: Phase synchronization and large-scale integration. *Nat Rev Neurosci* 2(4):229–239.
27. Perrin F, Pernier J, Bertrand O, Echallier JF (1989) Spherical splines for scalp potential and current density mapping. *Electroencephalogr Clin Neurophysiol* 72(2): 184–187.
28. Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin Neurophysiol* 110(11):1842–1857.
29. Castelo-Branco M, Goebel R, Neuenschwander S, Singer W (2000) Neural synchrony correlates with surface segregation rules. *Nature* 405(6787):685–689.
30. Friston KJ, et al. (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6(3):218–229.
31. Lavie N, Ro T, Russell C (2003) The role of perceptual load in processing distractor faces. *Psychol Sci* 14(5):510–515.
32. Qiu FTT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. *Nat Neurosci* 10(11):1492–1499.
33. Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum Brain Mapp* 15(1):1–25.
34. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1):83–98.
35. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
36. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
37. Fahrenfort JJ, Scholte HS, Lamme VAF (2007) Masking disrupts reentrant processing in human visual cortex. *J Cogn Neurosci* 19(9):1488–1497.

Supporting Information

Fahrenfort et al. 10.1073/pnas.1207414110

SI Discussion

If Category Tuning Is Invariant to Stimulus Visibility, Why Do Many Studies Report a Relationship Between Category Tuning and Perception? Many rivalry studies show that neural object representations follow the perceived stimulus (1, 2), suggesting that category tuning and visibility are intertwined. However, the apparent correlation between object representation and visibility during rivalry may be a side effect of neural competition. During rivalry, objects are presented in the same location. Because they occupy the same retinotopic space, it is often suggested that ambiguity is resolved by suppressing neural representations originating from the nondominant object and/or eye (3). Indeed, evidence suggests that information about a suppressed stimulus is able to survive beyond area V1 (4) and that category information is sometimes retained in ventral (5) and parietal (6) cortex. These findings indicate that information is not completely absent when a stimulus is invisible but rather is suppressed to a degree that detecting it has become difficult.

This interpretation may also explain the apparent contradiction between binocular rivalry and dichoptic fusion studies with respect to the presence of category-selective responses to invisible stimuli. As in binocular rivalry, objects are perceptually invisible during dichoptic fusion. However, rather than causing competition at the level of the object representation and/or eye, image elements are perceptually fused instead of engaging in competition. There is no competition at the object level, as there is only one object in the scene. Indeed, in dichoptic fusion studies, object representations are invariably identified quite easily (7, 8), despite stimulus invisibility.

Importantly, not all objects, visible or invisible, evoke category tuning. There are many reasons why an object does or does not elicit a category specific response. Some objects, such as faces, may evoke so much bottom-up activation that they have a strong tendency to result in category tuning, even when invisible and unattended (9) or task irrelevant (10). In other instances, selective attention may work to overcome competition between objects by biasing some objects to be processed over others (11, 12). Attention has been shown to bias the processing of invisible stimuli (13), even when these belong to high-level categories (14). Finally, task relevance of the object category that is being viewed is known to boost category tuning, even when the object itself is not spatially attended (15). The resolution of competition between object representations will often result in the co-occurrence of category tuning and conscious vision, which may lead to the misguided belief that consciousness and category tuning are critically related. However, evidence suggests that the (in)visibility of an object is not the factor that determines whether a category-specific response occurs. More likely, bottom-up sensory-driven mechanisms and top–down attentional influences bias the processing of some objects over others, leading some objects to evoke category tuning and others not. However, category specific responses are not impervious to the phenomenal status of an object. A visible object may be more likely to fall prey to the biases and selection mechanisms of the attentional system than an invisible object and may even result in more extensive category selective responses. The point is, however, that visibility is not critically involved in generating category selectivity or vice versa.

How Does Monocularly Presented Category Information Reach High-Level Visual Cortex When Its Constituent Elements Are Fused at Lower Levels? A large number of cells in early visual cortex are monocular, meaning that they retain information about the eye of origin. In macaques, ocular dominance columns are well segregated, and virtually all cells in layer 4C of area V1 respond exclusively to either the right eye or the left eye (16). In higher visual areas, such as object-selective cortex, cells are thought to be mostly binocular, meaning that these neurons respond to input from both eyes. Although it has been established previously that perceptually fused monocular information can penetrate into the object selective cortex (7), some may wonder how this might work. The first thing to note is that even in monkey inferotemporal cortex, about half the neurons show ocular dominance, meaning that they are preferentially tuned to either the left or the right eye (17). Apparently, substantial information about eye of origin is retained even in high-level visual cortex.

Moreover, even though the Gabor elements in this study are perceptually fused, and thus do not engage in interocular competition at a perceptual level, one should be cautious in assuming that all of the underlying information must therefore also be fused at the neural level. Hybrid models of binocular rivalry, such as that by Tong et al. (3) (see figure 2 in their paper) outline the progression from eye-based to pattern-based representations in visual cortex. There is a parallel progression from neurons representing low-level stimulus characteristics to abstract object properties, such as object category, in the high-level visual cortex (18). Evidence suggests that such abstract (e.g., viewpoint invariant) information is already computed during the first feedforward sweep (19), but little is known about the way in which this transformation occurs. However, it would likely require intermediate stages where the tuning properties become unstuck from the low-level veridical stimulus representation.

For monocular category information to penetrate into the object-selective cortex, the only requirement would be that monocular cells project onto binocular cells that carry such intermediate stage information. These binocular cells could receive input from monocular cells and respond to abstract feature conjunctions at the neural level, even though these feature conjunctions are not visible perceptually. Although monocular information may be lost at that point, this does not mean that there are no neural responses that represent information that is present in both eyes respectively. The hypothesis would be that during dichoptic fusion, neurons from different ocular columns ultimately project onto a single higher-order pattern. If pathways from both columns project onto the same higher-order pattern, there is no reason to assume that the intermediate and higher-order representations are lost at the moment that monocular information is lost, even if these higher-order representations are not visible. Bear in mind that none of the higher-order representations are in competition with each other during dichoptic fusion, as the same object is presented to the left and right eye.

In support of this idea, the dissociation of perceptually visible information and neural tuning responses is a property that is abundantly present throughout the visual cortex and beyond. For example, the orientation of invisible line elements can be decoded from V1 responses (20), and the meaning of perceptually invisible words is processed up to the level of semantic understanding (21).

Could Our Results Reflect Low-Level Feature Selectivity Instead of Category Tuning? Our stimulus categories were constructed in such a way that they contain both contour and feature similarities. For example, although lacking a clear semantic interpretation, nonsense objects contained similarities with faces and houses; they contain local Gabor configurations that resemble mouths and contours that resembled houses or faces. Houses also contained features that resemble eyes or mouths. Nevertheless, we

only find visibility invariant category responses for faces and not for nonsense objects and houses, indicating that these activations are caused by the configuration as a whole and not by the features that make up the configuration. Moreover, we looked at the visible-invisible overlap between multivoxel activations in Brodmann area 17 (BA17) (Fig. 3*D*), which is known to be involved in the processing of low-level stimulus characteristics. If the similarities were caused by the propagation of low-level stimulus features, one would expect this to result in above chance classification in BA17, but no above chance classification was found in early visual cortex. Finally, the face-selective activations we found were specific to the fusiform face area (FFA) and ventral occipitotemporal cortex (VOT). Several studies have shown that these regions are not selective for local contours or features (22–24).

How Does Object Categorization Work When There Is No Figure–Ground Segregation? How Are the Grouping Mechanisms During Category Extraction Different from Those Needed for Perceptual Organization? At first glance, it may seem puzzling how an object can be categorized without being segregated from the rest of the image before categorization (25, 26). A theoretical account of how this might work is outlined in the incremental grouping theory by Roelfsema and colleagues (27, 28). According to this theory, image features are initially grouped and classified in feedforward base groupings, causing complex tuning properties in high visual areas. These base groupings are computed rapidly because they reflect the selectivity of feedforward connections, but they do not reflect perceptual organization, as evidenced by the fact that putative correlates of these groupings also occur for objects and shapes that cannot be perceived (29, 30). Moreover, base groupings are hard-wired and therefore limited in the number of feature combinations that can be represented (although there is evidence that viewpoint invariant input can be categorized during the first feedforward sweep; ref. 19). Therefore, a flexible grouping mechanism called incremental grouping is required. During incremental grouping, base groupings are used to guide a selection process, which depends on horizontal and recurrent interactions. In this selection process, the responses of neurons coding features that are bound in perception are enhanced, effecting figure–ground segregation and conscious perception. Incremental grouping is flexible and therefore provides the dynamic mechanism for perceptual organization of the wide range of visual input with which we are confronted every day. We hypothesize that in this study, both invisible and visible faces are detected and categorized in feedforward base groupings, whereas only visible faces are grouped incrementally during a prolonged integration phase.

Why Are There No Face-Selective Responses in the EEG Signal? The EEG signal contained no classical N170 (N200) or other face-selective components (31), either in the visible or in the invisible face-selective contrasts. This result may be surprising, given that it has often been suggested that the N170 reflects a mandatory component of face processing that is not attenuated by task-related factors such as attention (32) and is largely unaffected by the mode of image presentation, such as schematic line drawings of faces (33).

However, these notions have recently been challenged in studies that show large influences of attention and perceptual load on the N170 (34–36). A 2009 report by Sreenivasan and colleagues (36) showed that the modulatory role of attention on face processing interacts with the discriminability of the face stimulus, with N170 responses to suboptimal face presentations benefiting from attention, whereas easily discriminable faces are already at peak. Moreover, a 2009 study from Mohamed and colleagues (34) showed that conditions of high perceptual load can lead to a concomitant face-N170 reduction and house-N170 increase, to the point where face selectivity is all but abolished. During our experiment, subjects had to judge whether an ellipse virtually hovered in front or behind the stimulus screen, resulting from a seven-pixel offset of the ellipse presented in the left compared with the right eye (*SI Methods*). This was a perceptually challenging task because of the small retinal disparity used to create the illusion. In addition, our stimuli were not contrast-defined as schematic face stimuli usually are, but were second-order defined by the orientation of Gabor patches. Together, these factors may have created one of the most unfavorable modes of presentation imaginable for eliciting a face-selective N170 response. The fact that we still find face-selective responses in the FFA suggests that the N170 EEG response is not merely a reflection of face-selective responses in the FFA. Indeed, source reconstructions of the N170 effect have implicated many cortical regions, including the occipital face area (OFA), the FFA, and the posterior temporal sulcus (pSTS), but also distributed sources across the anterior fusiform gyrus together with activations in a parieto-temporal-occipital network (see ref. 37 for a review).

What Does the Spectral Fingerprint of Conscious Figure–Ground Segregation Mean? Figure–ground segregation is thought to result from the neuronal labeling of image elements as belonging to a surface (surface segregation) or to the border of a surface (border-ownership). Single-unit (38–40) and functional MRI (fMRI) (41, 42) reports indicate the involvement of V1 and V2 cells, of which the receptive field size provides the acuity to perform this labeling operation with sufficient spatial detail. Others have stressed the involvement of higher-order areas of which the receptive field size and tuning properties are large enough to capture entire objects (43, 44). Physiologically constrained models of figure–ground segregation solve this apparent paradox by suggesting that feedforward computations achieve initial tuning to proto-object or base groupings through lateral inhibition, whereas feedback connections select or enhance neural responses carrying spatially precise surface (45) and/or border-ownership information (46). Feedback provides the means for integrating and binding information carried by distant neurons with different tuning properties and different receptive field sizes. Most recently, the idea that surface selection is guided by higher visual areas through feedback has been confirmed in studies using frequency tagging for figure and background (47). The binding of cells carrying same-surface and/or same-border-owner information can either be coded through covarying response enhancement (38, 39, 48), neural synchrony (49–51), or both. Indeed, we observed stimulus-induced (nonevoked) power changes in the time-frequency domain of the EEG for visible objects only (Fig. 4). Specifically, visible objects induced topologically specific increases in the high gamma (50–80 Hz) and theta bands (4–6 Hz) and a strong power decrease in the low-beta band (12–20 Hz) over the bilateral occipitotemporal cortex.

Although it is difficult to attribute specific functional properties to modulations in certain power bands, some speculation is in order. There is a widespread notion that different frequency bands reflect different spatial scales in cortical synchronization (52). Low-frequency modulations are often linked to long-distance oscillatory interactions, whereas high-frequency modulations are thought to reflect local interactions (53, 54). A recent hypothesis is that the spectral properties within certain brain regions are not only informative with respect to the distance over which communication between brain areas takes place but that these can be mapped onto different cognitive functions (55). Specifically, these authors propose that gamma band oscillations emerge during the encoding of information during local interactions, whereas beta band synchronization is involved in long-distance integrative cognitive functions such as top–down control and decision-making.

Consistent with this idea, induced gamma band oscillations are often linked to the binding of feature elements (56, 57). For example, induced gamma has been associated with feature binding of Kanisza elements to illusory triangles (58) and with other examples of gestalt perception (59). Moreover, evidence indicates that perceptual grouping is correlated with activity in the high-gamma range (60) and that gamma is linked to the spatial location where figure–ground segregation takes place (61). These findings all provide clues that gamma band activity in the visual cortex plays a role during feature integration (50) and figure–ground organization (27). Such findings are somewhat clouded by the fact that gamma band oscillations in the visual cortex have also been related to top–down attention (62–64). However, attentional selection is known to be mediated by long-range control of the prefrontal cortex over the visual cortex (65), which is typically marked by low-frequency oscillations (55, 65). A more parsimonious account of attention-related gamma band modulation in visual cortex is therefore that it is a reflection of top–down interactions with local integrative processes in the sensory cortex (66), rather than a reflection of attentional selection itself.

The role of beta oscillations is not as well delineated. Although low-frequency oscillations, particularly in the beta band and in the prefrontal and parietal cortex, have been linked to long-range cognitive integrative functions, their role is less specific than that of gamma during local integration and encoding (55). Interestingly, an increase of gamma band oscillations as a result of visual stimulation often occurs concurrently with low-frequency suppression. For example, Kinsey and colleagues (61) found beta suppression during object processing and figure–ground segregation, concurrently with an increase in the high gamma range. However, unlike the increase in gamma, beta band suppression was largely independent of the spatial location of the target, leading them to speculate that it is a general effect of attention. Event-related (non–phase-locked) desynchronization is caused by small patches of neurons or neuronal assemblies working in a relative independent or desynchronized manner as a result of stimulus presentation. Pfurtscheller and Lopes da Silva (57) point out that occipital event-related desynchronization in the low beta band is an electrophysiological marker of increased cellular excitability during which processing of sensory information is enhanced. As such, it may be related to cortical reevaluation of visual input. A purely speculative hypothesis is that the desynchronization of beta oscillations we observed is functionally related to the breaking up of homogenously organized visual input as a result of stimulus presentation.

Last, theta band oscillations have been linked to information encoding in the hippocampus (67). Intracortical recordings in rodents suggest that theta is involved in mediating hippocampocortical feedback loops and that it may be engaged in shaping the phase of cortical gamma band activity (68). Possibly, this constitutes the mechanism by which information in spatially widespread cortical assemblies is transferred to the associative networks of the hippocampus. In EEG, memory-related cross-frequency coupling between theta and gamma has been shown mostly at frontal sites (69), but at least one study reported that gamma-theta cross-frequency coupling at posterior recording sites is correlated with short-term memory performance (70). Although little is known about the relationship between consciousness and memory encoding, an intriguing possibility may be that only consciously perceived objects result in memory encoding, whereas invisible objects do not.

Could Differences in the Functional Connectivity Results Be Secondary to Differences in Response Amplitude Between the Conditions in the Experiment? When assessing functional connectivity in resting state analyses, differences in functional connectivity can be driven either by changes in connectivity between the nodes of the network or simply by changes in activity within the network, without changes in connectivity. The latter is caused by the fact that changes in the signal-to-noise ratio (SNR) can bring about differences in the correlation between two signals. As a result, simple changes in activity can bring about artificial changes in functional connectivity (71), which could be considered worrying, given that there were large amplitude differences between visible and invisible faces. However, the psychophysiological interaction (PPI) analysis we applied does not have this problem, because it separately models these changes in activation. In our analysis, the functional connectivity effect is defined as a visibility-specific change in the relationship between the FFA and other brain areas, over and above what can be explained by the main effect of the visible-invisible face contrast. As a result, the SNR issue does not apply in our PPI analysis, and we can be confident that the functional connectivity we observed is not caused by differences in signal strength between visible and invisible faces (72).

Additionally, it might be considered surprising that we found large amplitude differences between conditions in BA17 (the anatomical equivalent of functionally defined V1). Neurons in this area have small receptive fields that have traditionally been found to respond only to simple stimulus characteristics such as line orientation (16). Estimates of V1 receptive field size in macaques are ∼0.25° in the fovea and less than 1° even at sizeable eccentricities (73). Estimates of early visual receptive field size in humans using subdural electrodes also point to (much) smaller than 1° sizes (74). As the Gabor elements in this study subtended a visual angle comparable to the receptive field size of V1 neurons (0.58°) and conditions were balanced out at this level (both monocularly and binocularly and both within the visible conditions and within the invisible conditions), such differences are not to be expected in BA17 when only considering feedforward tuning responses. It is therefore unlikely that the large amplitude differences for the different conditions in early visual areas are caused by bottom up factors. More likely, the differences in response amplitude in early visual areas emerged as a result of recurrent interactions with neurons that have receptive field sizes of sufficient size to signal the global organization of the stimulus (27, 38).

Do These Results Generalize to Category Selectivity in General or Are They Specific to Faces? The results of our study only show category-selective responses for visible and invisible faces but not for categories other than faces. Hence, one may wonder whether this result can be generalized to other stimulus categories. Faces are intrinsically relevant to the human visual system, giving them priority over other stimuli even when not spatially attended (9) and even when task irrelevant (10). Other object categories have shown category selectivity when made task relevant (75), even when presented in cluttered scenes that are not spatially attended (15). Therefore, it may be that objects from other categories will show visibility-invariant category selectivity once they are made task relevant, as in Moutoussis and Zeki (7). However, this is not an open and shut case that follows from our data. Indeed, the mechanisms behind face processing have been the subject of heated debate. On one end of the spectrum is the position that face processing is always different from processing other object categories, also referred to as domain specificity, a hypothesis championed by Kanwisher and colleagues (76). On the other end of the spectrum is the hypothesis that visual processing of faces only seems special because people have greater expertise in recognizing faces than in performing within-class discrimination of other object classes, also known as the expertise hypothesis by Gauthier and colleagues (77). Naturally, whether one or the other is true might impact the interpretation of our results. If faces are special (domain specificity), it may be that they are unconsciously processed, even though invisible objects for which no domain specific module exists never elicit

similar category-selective responses. On the other hand, it may be that sufficient expertise with an object category and/or task relevance while processing an object category is sufficient to generate category-selective responses, even when the object is invisible. Whether our results can be extrapolated from faces to other categories and viewing conditions, for example, by making the object category being viewed task relevant, is an issue for future research.

SI Methods

Stimuli. Texture stimuli were generated on a 28 × 22 matrix of oriented 2D Gabor elements. Objects were constructed using a 45° orientation difference between elements representing the figure and elements representing the background (Fig. 1*A*). Each stimulus instance (Fig. 1*B*) was created using all combinations of orientations (22.5°, 67.5°, 112.5°, and 157.5°) containing a 45° difference. Consequently, each oriented Gabor element would occur equally often in each location of the visual field for each condition, both on a monocular and binocular level, balancing out local physical stimulation across conditions (Figs. S1 and S2). Mask screens consisted of Gabor elements of which half were randomly rotated by 0° and the other half were rotated by 90°. The element in the one eye was always matched by an orthogonally oriented element in the other eye, resulting in a dichoptically fused matrix of plus-like Gabor elements. Illusory in-front-behind depth perception was created using a small retinal disparity (0.14°) between ellipses (15.45° × 11.76°) in the left and the right eye. The disparity was small enough to make depth detection a difficult task when viewing stimuli with both eyes and impossible to do monocularly.

Stimulation Protocol. A red fixation dot was present throughout the experiment. Each 1,800-ms trial (Fig. S1*A*) started with a 167-ms gray screen. Subsequently, a 250-ms stimulus sequence was presented four times in succession to maximize signal strength (50-ms mask screen, 67-ms gray screen, 83-ms object screen, 50-ms gray screen). Target ellipses of the in-front-behind task were visible during the first 1,533 ms of a trial.

fMRI sessions contained three runs (three subjects had a four-run session); EEG sessions contained six blocks. A run/block contained 32 visible and invisible instances of the four stimulus types (256 trials) in random order. fMRI runs started and ended with a 16-s fixation-only period and contained a pseudorandom intertrial interval (ITI) of 0, 1.8, or 3.6 s, which was optimized for hemodynamic response function (HRF) deconvolution. fMRI image acquisition was not locked to stimulus onset. EEG blocks contained a random ITI between 300 and 800 ms to preclude coincidental phase-stimulus relationships. After imaging, subjects were presented with an identical 256 trial run/block during which they performed the control task.

fMRI Acquisition Parameters. Data were acquired on a Philips 3T Intera scanner using the following parameters for anatomical [T1 turbo field echo sequence; 182 coronal slices; flip angle (FA), 8°; echo time (TE), 4.6 ms; repetition time (TR), 9.7 s; slice thickness, 1.2 mm; field of view (FOV), 256 × 256 mm; in-plane voxel resolution, 1 × 1 mm] and functional data [T2*-weighted sequence; 35 coronal slices; FA, 90°; TE, 28 ms; TR, 2.3 s; slice thickness, 3 mm; FOV, 220 × 220 mm; in-plane voxel resolution, 2.8 × 2.8 mm].

Regions of Interest. Regions of interest (ROIs) were selected using functional data from a split half procedure and anatomical criteria obtained from their known involvement in face processing (FFA) (76), category selectivity (VOT) (18), object selectivity, and figure–ground processing [inferior lateral occipital cortex (LOC) and dorsal foci in the lateral occipital cortex] (18), as well as figure–ground processing and low-level vision (BA17) (16, 38, 39). FFA voxels responded more to faces than to houses and objects, at an uncorrected threshold of $Z > 2.3$, forming contiguous clusters located within the atlas-defined occipitotemporal fusiform cortex. To verify whether these clusters correspond to the functionally defined FFA, their location was verified using a classical FFA localizer containing photographs of faces and other objects. All other areas were defined as clusters of voxels responding more to stimuli containing a figure–ground relationship (faces, houses, and nonsense objects) than to homogenous textures (thresholded at $Z > 2.3$ at a corrected cluster significance threshold of $P < 0.05$) within their corresponding atlas-defined regions. BA17 voxels were further constrained to have >50% probability of being part of BA17, as defined by the Juelich Histological Atlas.

FFA Localizer. The blocked-design FFA localizer contained photographs of isolated faces and other objects (houses, bottles, cats, chairs, scissors), as well as phase-scrambled images. Each block had a duration of 16 s, containing eight images that were shown for 1 s with 1-s intervals. Subjects had to indicate whether the object was a repeat of an object they had just seen or not. The localizer contained 16 stimulus blocks (4 face blocks, 4 house blocks, 4 blocks containing other objects, and 4 blocks containing phase-scrambled images), with 8-s rest periods in between. It started and ended with a 16-s baseline period. The FFA was localized by contrasting blocks containing images of faces with blocks containing images of houses and other objects.

Univariate Analyses. Individual subject parameter estimates were exported using a split half procedure in which data export and ROI selection were carried out on independent selections (halfs) of the entire dataset. Parameters estimates were exported for each subject, condition, and ROI and transformed into percent signal change. During export, the average percent signal change across all voxels in each ROI was calculated by weighting individual voxel values according to the relative contribution of those voxels to the ROI based on the ROI's Z values. The same procedure was carried out twice, switching around datasets used for ROI selection and for data export, thus making sure that neither of the two datasets suffered from the nonindependence error (78). The two resulting datasets were averaged before statistical testing.

Multivoxel Pattern Analyses. Single-subject voxelwise t values for each condition were used as input images, combining all runs from each subject. Subject's images were normalized by subtracting the average response across visible and invisible conditions from the response to the individual categories in each voxel (75). Searchlight kernel response patterns (kernel radius: 10 mm) to each of the visible categories were correlated with response patterns to each of the invisible categories at all voxel locations in the brain. This procedure resulted in nine correlations (Fig. 3 *A–C*, *Left*) for each subject at each voxel location. Before statistical testing, all correlations were Fisher-transformed using $0.5 \times \ln[(1 + r)/(1 - r)]$. Within-category correlations had matching categories (e.g., visible face with invisible face), whereas between-category correlations were nonmatching (e.g., visible face with invisible house). For each category, a permutation test assessed whether the within-category correlation was (*i*) higher than all between-category correlations (Fig. 3*A*); (*ii*) higher than invisible between-category correlations (Fig. 3*B*); and (*iii*) higher than visible between-category correlations (Fig. 3*C*). This test was performed for all voxels, correcting for multiple comparisons by controlling the familywise error rate on threshold-free cluster–enhanced (TFCE) images (79). Pattern classification within ROIs was done by determining for each subject whether the within-category correlation was higher than the between-category correlations (classification success) or not (classification miss).

EEG Acquisition. EEGs were acquired on a 48-channel BioSemi Active Two system at a sample rate of 256 Hz.

Spectral Analysis. Induced responses were isolated by subtracting the condition-specific average evoked response (ERP) waveform from each trial segment. Time-frequency representations (TFRs) were calculated for each trial. For the low frequencies (2–30 Hz), a sliding time window of $\Delta T = 0.5$ s was used, and the data in each time window were multiplied with a Hanning taper. For the high-frequency band (20–100 Hz), we applied a multitaper approach (80) using a sliding time window of $\Delta T = 0.4$ s and seven orthogonal Slepian tapers, resulting in a frequency smoothing filter of 20 Hz. The difference in power between figure–ground versus homogenous textures across trials was quantified using a Z value for each subject, reducing the contribution of subjects with large variance. Z values were tested against zero at a group level using cluster-based permutation testing across all dimensions (frequency, electrode, time), separately for the visible and the invisible conditions. This analysis yielded significant results only in the visible condition in three frequency bands: theta (4–6 Hz), low-beta (12–20 Hz), and gamma (50–80 Hz). Cluster-based permutation tests were subsequently performed on the mean TFR Z values in these frequency bands, now clustering over time points and electrode sites only. TFRs were visualized over instructively pooled electrodes and in topographical maps of the resulting time windows.

1. Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21(4):753–759.
2. Sheinberg DL, Logothetis NK (1997) The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci USA* 94(7):3408–3413.
3. Tong F, Meng M, Blake R (2006) Neural bases of binocular rivalry. *Trends Cogn Sci* 10(11):502–511.
4. Leopold DA, Logothetis NK (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379(6565):549–553.
5. Sterzer P, Haynes JD, Rees G (2008) Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *J Vis* 8(15):1–12.
6. Fang F, He S (2005) Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nat Neurosci* 8(10):1380–1385.
7. Moutoussis K, Zeki S (2002) The relationship between cortical activation and perception investigated with invisible stimuli. *Proc Natl Acad Sci USA* 99(14):9527–9532.
8. Schurger A, Pereira F, Treisman A, Cohen JD (2010) Reproducibility distinguishes conscious from nonconscious neural representations. *Science* 327(5961):97–99.
9. Finkbeiner M, Palermo R (2009) The role of spatial attention in nonconscious processing: A comparison of face and nonface stimuli. *Psychol Sci* 20(1):42–51.
10. Lavie N, Ro T, Russell C (2003) The role of perceptual load in processing distractor faces. *Psychol Sci* 14(5):510–515.
11. Kastner S, Ungerleider LG (2000) Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci* 23:315–341.
12. Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
13. Bahrami B, Lavie N, Rees G (2007) Attentional load modulates responses of human primary visual cortex to invisible stimuli. *Curr Biol* 17(6):509–513.
14. Shin K, Stolte M, Chong SC (2009) The effect of spatial attention on invisible stimuli. *Atten Percept Psychophys* 71(7):1507–1513.
15. Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460(7251):94–97.
16. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195(1):215–243.
17. Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotemporal cortex of the Macaque. *J Neurophysiol* 35(1):96–111.
18. Grill-Spector K, Malach R (2004) The human visual cortex. *Annu Rev Neurosci* 27:649–677.
19. Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* 62(2):281–290.
20. Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat Neurosci* 8(5):686–691.
21. Dehaene S, et al. (1998) Imaging unconscious semantic priming. *Nature* 395(6702):597–600.
22. Hasson U, Hendler T, Ben Bashat D, Malach R (2001) Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *J Cogn Neurosci* 13(6):744–753.
23. Andrews TJ, Schluppeck D, Homfray D, Matthews P, Blakemore C (2002) Activity in the fusiform gyrus predicts conscious perception of Rubin's vase-face illusion. *Neuroimage* 17(2):890–901.
24. Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293(5534):1506–1509.
25. Peterson MA, et al. (1993) Shape-recognition inputs to figure ground organization in 3-dimensional displays. *Cognit Psychol* 25(3):383–429.
26. Peterson MA, et al. (1994) Must figure-ground organization precede object recognition? An assumption in peril. *Psychol Sci* 5(5):253–259.
27. Roelfsema PR (2006) Cortical algorithms for perceptual grouping. *Annu Rev Neurosci* 29:203–227.
28. Roelfsema PR, Lamme VAF, Spekreijse H (2000) The implementation of visual routines. *Vision Res* 40(10-12):1385–1411.
29. Kovács G, Vogels R, Orban GA (1995) Cortical correlate of pattern backward masking. *Proc Natl Acad Sci USA* 92(12):5587–5591.
30. Fahrenfort JJ, Scholte HS, Lamme VAF (2007) Masking disrupts reentrant processing in human visual cortex. *J Cogn Neurosci* 19(9):1488–1497.
31. Bentin S, Allison T, Puce A, Perez E, McCarthy G (1996) Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 8(6):551–565.
32. Cauquil AS, Edmonds GE, Taylor MJ (2000) Is the face-sensitive N170 the only ERP not affected by selective attention? *Neuroreport* 11(10):2167–2171.
33. Harris A, Nakayama K (2007) Rapid face-selective adaptation of an early extrastriate component in MEG. *Cereb Cortex* 17(1):63–70.
34. Mohamed TN, Neumann MF, Schweinberger SR (2009) Perceptual load manipulation reveals sensitivity of the face-selective N170 to attention. *Neuroreport* 20(8):782–787.
35. Jacques C, Rossion B (2007) Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cereb Cortex* 17(5):1055–1065.
36. Sreenivasan KK, Goldstein JM, Lustig AG, Rivas LR, Jha AP (2009) Attention to faces modulates early face processing during low but not high face discriminability. *Atten Percept Psychophys* 71(4):837–846.
37. Rossion B, et al. (2011) The N170: understanding the time-course of face perception in the human brain. *The Oxford Handbook of ERP Components*, eds Luck SJ, Kappenman E (Oxford Univ, Oxford, UK).
38. Lamme VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15(2):1605–1615.
39. Zipser K, Lamme VAF, Schiller PH (1996) Contextual modulation in primary visual cortex. *J Neurosci* 16(22):7376–7389.
40. Qiu FTT, Sugihara T, von der Heydt R (2007) Figure-ground mechanisms provide structure for selective attention. *Nat Neurosci* 10(11):1492–1499.
41. Skiera G, Petersen D, Skalej M, Fahle M (2000) Correlates of figure-ground segregation in fMRI. *Vision Res* 40(15):2047–2056.
42. Scholte HS, Jolij J, Fahrenfort JJ, Lamme VAF (2008) Feedforward and recurrent processing in scene segmentation: Electroencephalography and functional magnetic resonance imaging. *J Cogn Neurosci* 20(11):2097–2109.
43. Kastner S, De Weerd P, Ungerleider LG (2000) Texture segregation in the human visual cortex: A functional MRI study. *J Neurophysiol* 83(4):2453–2457.
44. Grill-Spector K, Kushnir T, Edelman S, Itachak Y, Malach R (1998) Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21(1):191–202.
45. Roelfsema PR, Lamme VAF, Spekreijse H, Bosch H (2002) Figure-ground segregation in a recurrent network architecture. *J Cogn Neurosci* 14(4):525–537.
46. Craft E, Schütze H, Niebur E, von der Heydt R (2007) A neural model of figure-ground organization. *J Neurophysiol* 97(6):4310–4326.
47. Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM (2006) Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26(45):11695–11708.
48. Qiu FT, von der Heydt R (2005) Figure and ground in the visual cortex: v2 combines stereoscopic cues with gestalt rules. *Neuron* 47(1):155–166.
49. Milner PM (1974) A model for visual shape recognition. *Psychol Rev* 81(6):521–535.
50. Singer W, Gray CM (1995) Visual feature integration and the temporal correlation hypothesis. *Annu Rev Neurosci* 18:555–586.
51. Castelo-Branco M, Goebel R, Neuenschwander S, Singer W (2000) Neural synchrony correlates with surface segregation rules. *Nature* 405(6787):685–689.
52. Engel AK, Fries P, Singer W (2001) Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2(10):704–716.
53. von Stein A, Sarnthein J (2000) Different frequencies for different scales of cortical integration: From local gamma to long range alpha/theta synchronization. *Int J Psychophysiol* 38(3):301–313.
54. Kopell N, Ermentrout GB, Whittington MA, Traub RD (2000) Gamma rhythms and beta rhythms have different synchronization properties. *Proc Natl Acad Sci USA* 97(4): 1867–1872.
55. Donner TH, Siegel M (2011) A framework for local cortical oscillation patterns. *Trends Cogn Sci* 15(5):191–199.
56. Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3(4):151–162.
57. Pfurtscheller G, Lopes da Silva FH (1999) Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin Neurophysiol* 110(11):1842–1857.
58. Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci* 16(13):4240–4249.
59. Keil A, Müller MM, Ray WJ, Gruber T, Elbert T (1999) Human gamma band activity and perception of a gestalt. *J Neurosci* 19(16):7152–7161.
60. Vidal JR, Chaumon M, O'Regan JK, Tallon-Baudry C (2006) Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals. *J Cogn Neurosci* 18(11):1850–1862.
61. Kinsey K, Anderson SJ, Hadjipapas A, Holliday IE (2011) The role of oscillatory brain activity in object processing and figure-ground segmentation in human vision. *Int J Psychophysiol* 79(3):392–400.
62. Wyart V, Tallon-Baudry C (2008) Neural dissociation between visual awareness and spatial attention. *J Neurosci* 28(10):2667–2679.
63. Gruber T, Müller MM, Keil A, Elbert T (1999) Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clin Neurophysiol* 110(12): 2074–2085.

64. Fries P, Womelsdorf T, Oostenveld R, Desimone R (2008) The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area V4. *J Neurosci* 28(18):4823–4835.
65. Zanto TP, Rubens MT, Thangavel A, Gazzaley A (2011) Causal role of the prefrontal cortex in top-down modulation of visual processing and working memory. *Nat Neurosci* 14(5):656–661.
66. Reynolds JH, Chelazzi L (2004) Attentional modulation of visual processing. *Annu Rev Neurosci* 27:611–647.
67. Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Res Brain Res Rev* 29(2-3):169–195.
68. Sirota A, et al. (2008) Entrainment of neocortical neurons and gamma oscillations by the hippocampal theta rhythm. *Neuron* 60(4):683–697.
69. Schack B, Vath N, Petsche H, Geissler HG, Möller E (2002) Phase-coupling of theta-gamma EEG rhythms during short-term memory processing. *Int J Psychophysiol* 44(2):143–163.
70. Sauseng P, et al. (2009) Brain oscillatory substrates of visual short-term memory capacity. *Curr Biol* 19(21):1846–1852.
71. Friston KJ (2011) Functional and effective connectivity: A review. *Brain Connect* 1(1): 13–36.
72. O'Reilly JX, Woolrich MW, Behrens TE, Smith SM, Johansen-Berg H (2012) Tools of the trade: Psychophysiological interactions and functional connectivity. *Soc Cogn Affect Neurosci* 7(5):604–609.
73. Smith AT, Singh KD, Williams AL, Greenlee MW (2001) Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cereb Cortex* 11(12): 1182–1190.
74. Yoshor D, Bosking WH, Ghose GM, Maunsell JH (2007) Receptive fields in human visual cortex mapped with surface electrodes. *Cereb Cortex* 17(10):2293–2302.
75. Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293(5539):2425–2430.
76. Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17(11): 4302–4311.
77. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC (1999) Activation of the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nat Neurosci* 2(6):568–573.
78. Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: The dangers of double dipping. *Nat Neurosci* 12(5):535–540.
79. Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1):83–98.
80. Percival DB, et al. (1993) *Spectral Analysis for Physical Applications: Multitaper and Conventional Univeriate Techniques* (Cambridge Univ Press, Cambridge, NY), p xxvii.

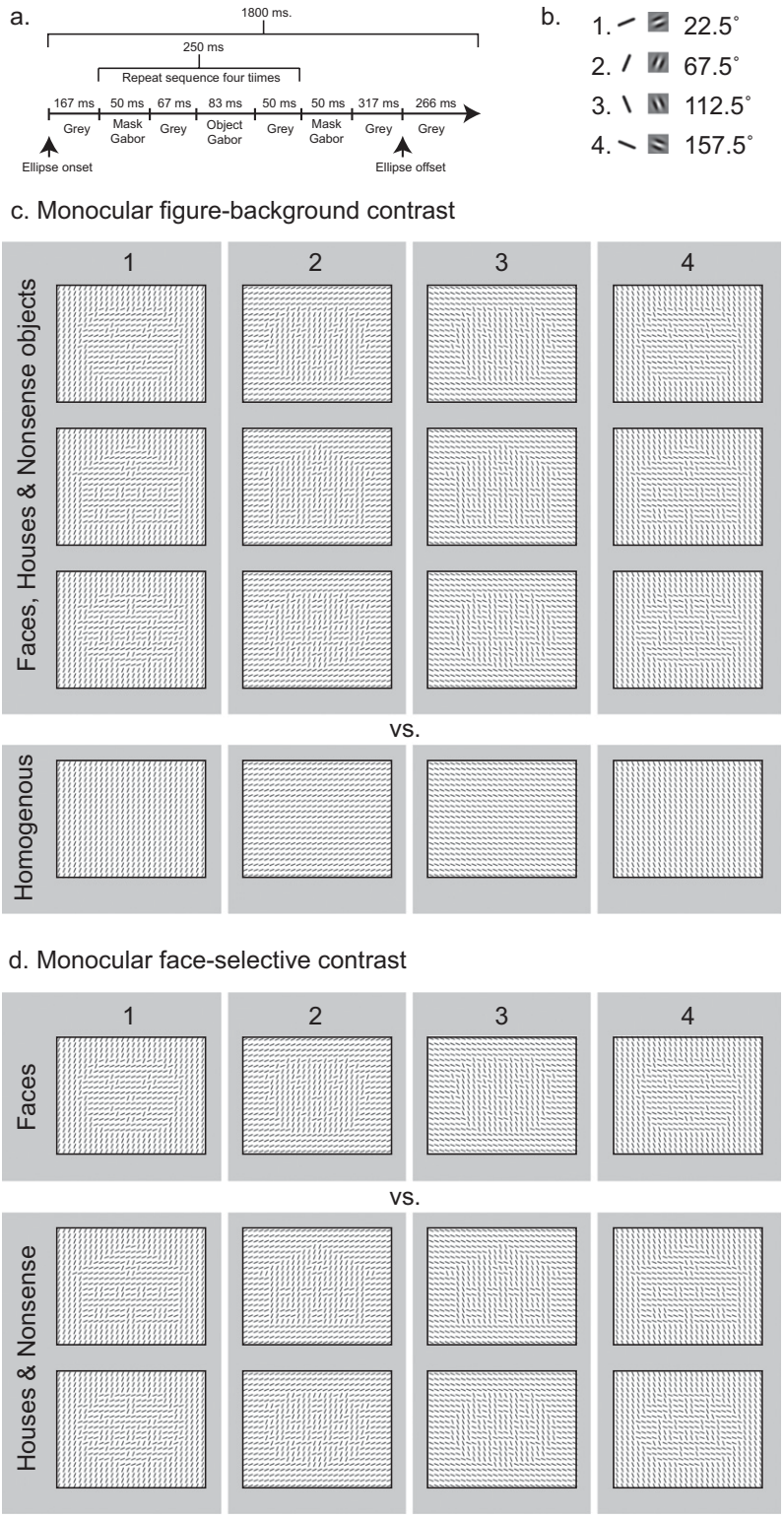

Fig. S1. Stimulus timing and schematics of stimulus and contrast construction. (*A*) Schematic depiction of the timeline of a trial. A trial lasted 1,800 ms. Four 250-ms repetitions of a stimulus sequence contained gray screens (Gray), homogenous array of fused plus-like elements (Mask Gabor; *SI Methods*), and textures of Gabor patches containing an object or a homogenous texture (Object Gabor). (*B*) The four oriented Gabor patches used for the construction of objects, schematically represented by lines in the left and right eye of Fig. 1*A*. Gabor elements had luminance features that roughly correspond to the receptive field structure of V1 simple cells and a size comparable to that of the receptive field size of V1 neurons (0.58°). (*C*) Simplified example of a monocular figure–background contrast in the experiment. For presentation purposes, the Gabor patches used in the experiment (as in *B*) were replaced with line elements. In all conditions, all locations on the screen were stimulated equally often by each oriented Gabor element. Hence, local physical stimulation at the level of a Gabor patch was on average the same for all stimulus categories (faces, houses, objects, and homogenous screens). (*D*) Simplified example of a monocular face-selective contrast. As for the figure–background contrast, local physical stimulation was balanced out at the level of a Gabor patch in all category-selective contrasts.

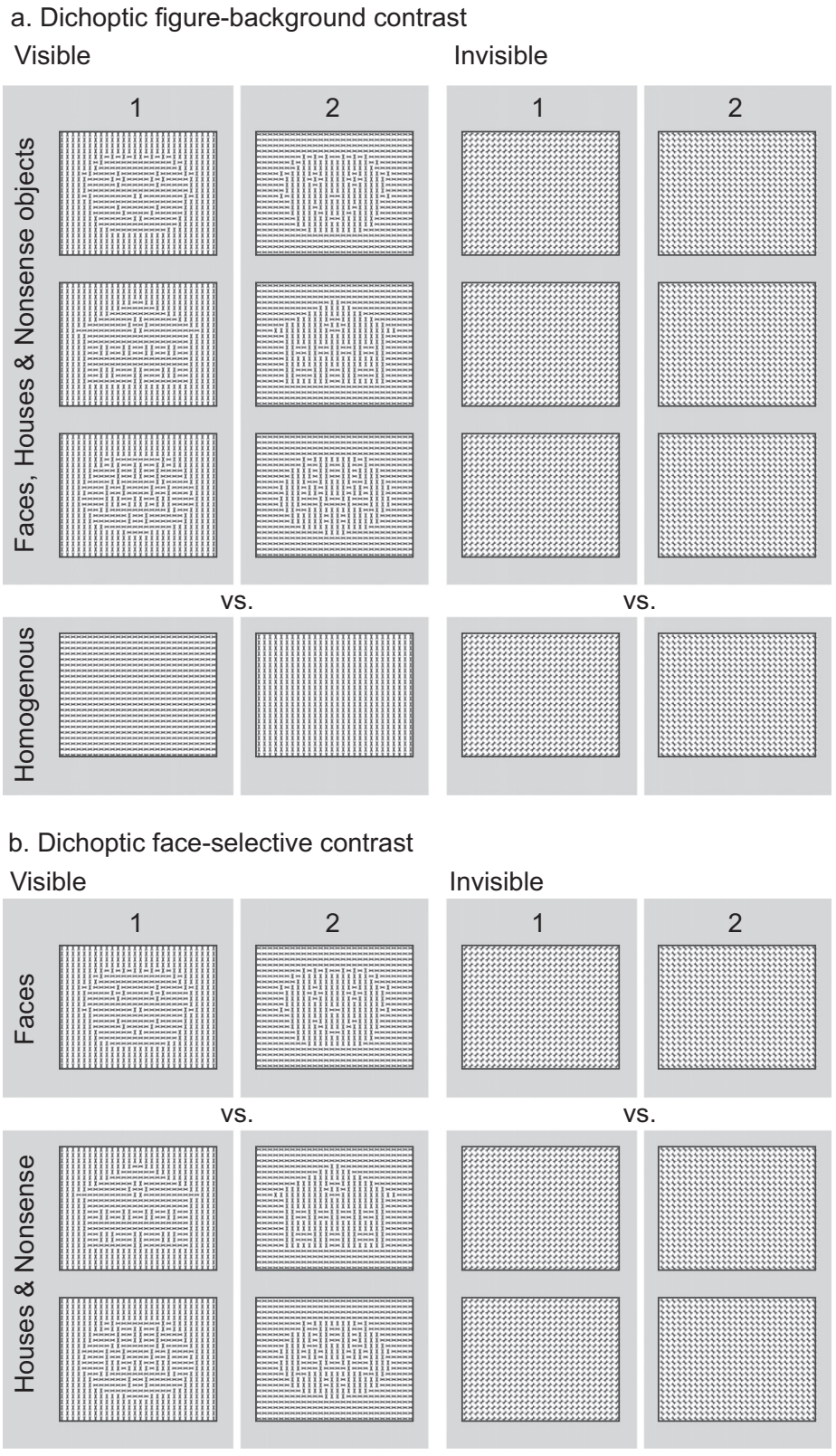

Fig. S2. Simplified examples of dichoptic figure–background and face-selective contrasts (visible and invisible). (*A*) Simplified example of a figure–background contrast in the experiment for visible (*Left*) and invisible (*Right*) stimuli. Shown are the fused images that subjects would perceive when input from the left and right eye was combined (Fig.1*A*). Visible images were created by combining panels 1 and 4 from Fig. S1 *C* and *D* (panel 1 in the left and panel 4 in the right eye and vice versa) and panels 2 and 3 (idem). Invisible images were created by combining panels 1 and 2 (left eye/right eye and vice versa) and panels 3 and 4 (idem). Although invisible images would not contain segregable objects when fused, they still contained objects at the level of the individual eyes. As was done at a monocular level, dichoptic local physical stimulation was balanced out at the level of the fused individual Gabor patches in the visible and invisible figure–background contrasts. (*B*) Simplified example of a face-selective contrast. As for the figure–background contrast, local physical stimulation was balanced out at the level of fused Gabor patches in all category-selective contrasts.

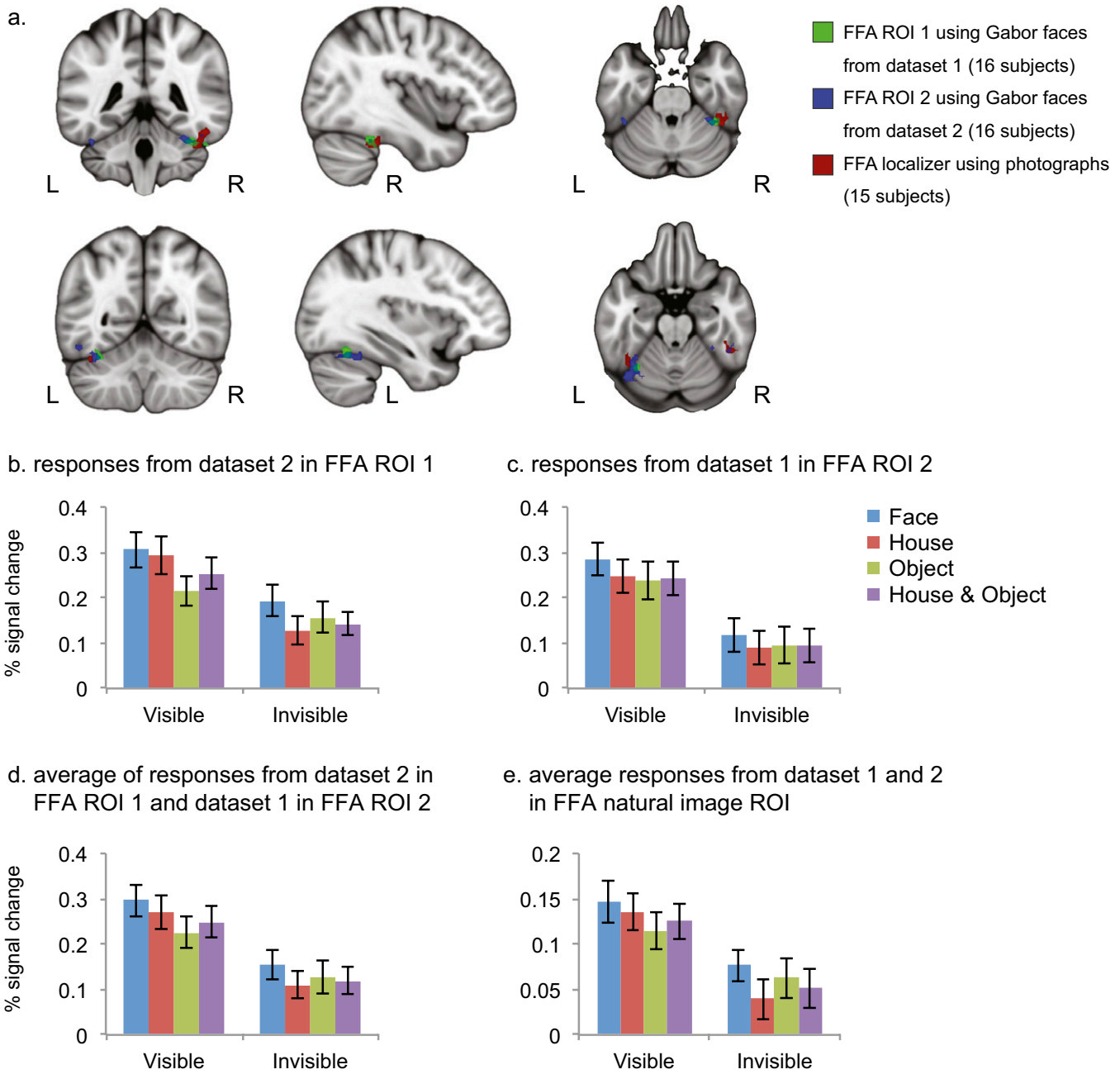

Fig. S3. Split half procedure for ROI selection. ROI selection was performed using a split half procedure ensuring independence of data used for ROI selection and data used for statistical testing. Each trial was pseudorandomly assigned to one of two datasets, each dataset containing half of the data. Subsequently, one dataset was used to draw ROIs, exporting responses in those ROIs from the other dataset, repeating the same procedure after switching the datasets used for ROI selection and export. Finally, the exported data from both datasets were averaged and used for statistical testing. An independent FFA mapper containing isolated natural images of faces, houses, and objects was used to verify that FFA ROIs were located in the same regions that are involved in generic face-selective processing. (*A*) FFA ROIs from datasets 1 and 2 (thresholded at $Z > 1.6$ for visualization purposes) and the FFA mapper containing natural images (thresholded at $Z > 2.3$). ROIs are in close proximity and partially overlapping. (*B*) Exported responses on each of the stimulus categories in the visible and invisible conditions from dataset 2 in the ROI that was obtained using dataset 1. The difference between the blue bar (faces) and the purple bar (houses and objects) signifies the face-selective contrast. (*C*) Exported responses from dataset 1 in the ROI that was obtained using dataset 2. (*D*) Averaged responses from dataset 1 (*C*) and dataset 2 (*B*). (*E*) Averaged responses in the natural image ROI from datasets 1 and 2. Although the main effect of face selectivity in invisible and visible conditions is still visibly present in the natural image FFA, the statistical significance of the effects reported in the main text depend on ROI selection using the split half procedure rather than using a generic FFA localizer that uses photographs.

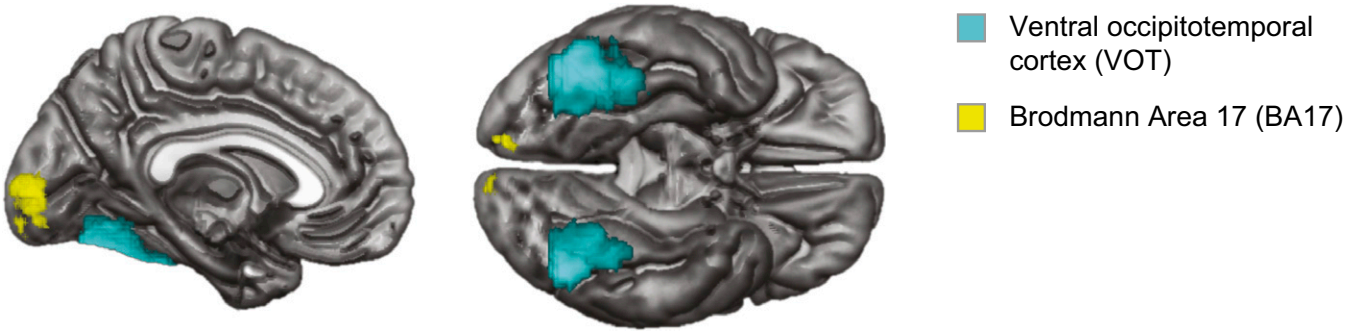

Fig. S4. Medial and ventral renderings of the VOT and BA17. VOT is known to be sensitive to object category information and was defined here as those voxels in the anatomically constrained VOT responding more strongly to textures containing objects than to homogenous textures. BA17 is insensitive to category information but exhibits tuning to low-level image features. It was defined as voxels having an atlas-defined probability of more than 50% of belonging to BA17, as well as responding more strongly to textures containing objects than to homogenous textures.

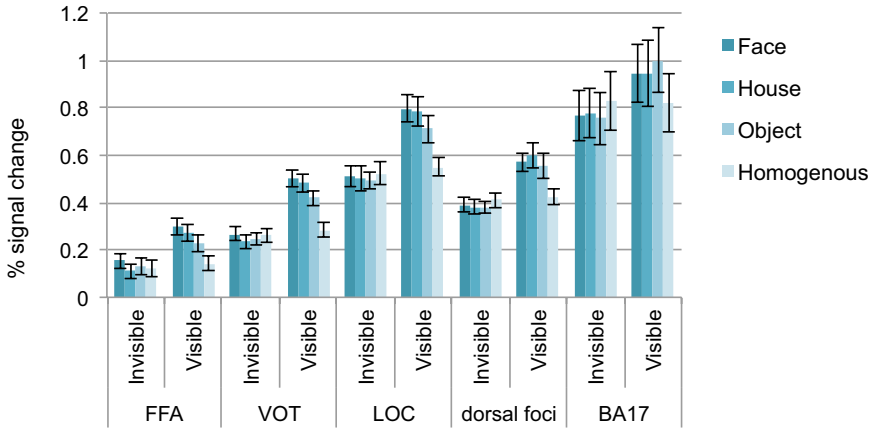

Fig. S5. Estimates of percent signal change . Responses to the four stimulus types (faces, houses, nonsense objects, and homogenous) in the two visibility conditions (visible, invisible) across the five regions of interest (FFA, VOT, LOC, dorsal foci, and BA17).

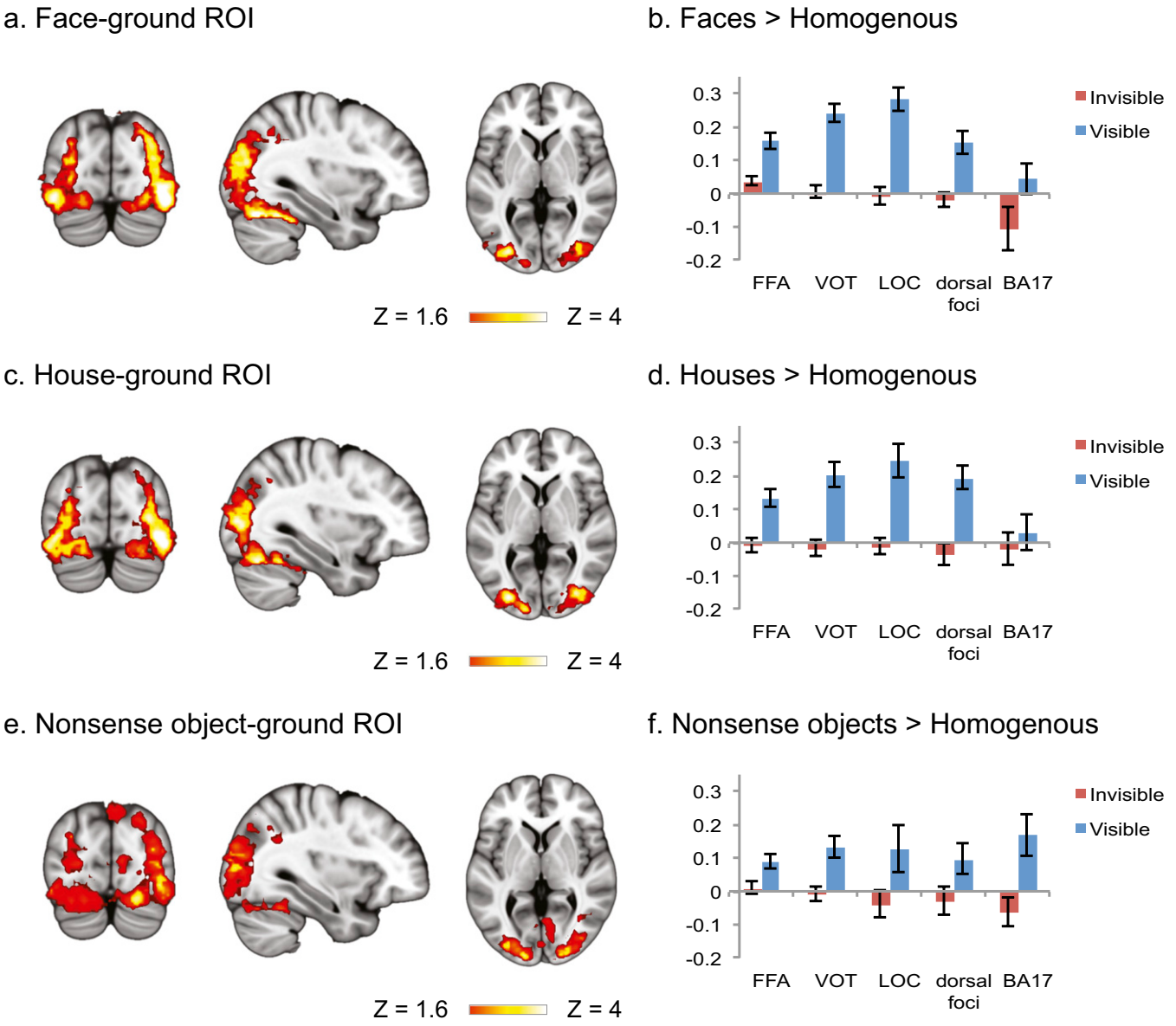

Fig. S6. Object-ground ROIs and the corresponding visible and invisible object-ground responses for each of the three stimulus categories. (*A*) Face-ground ROI. (*B*) Face-ground response profile for visible and invisible faces. (*C*) House-ground ROI. (*D*) House-ground response-profile for visible and invisible houses. (*E*) Nonsense-object ROI. (*F*) Nonsense object-ground profile for visible and invisible nonsense objects. As in the other figures, each of the ROIs is an average of two ROIs (thresholded at $Z > 2.3$) obtained from the split half procedure, whereas the object-ground response profile data were the averaged exported data resulting from that same procedure. Note that there are small differences between *B* and Fig. 4*C*, especially in the amount of face-ground modulation in BA17. This finding may be surprising, as both are face-ground profiles so they should show the same responses. However, Fig. 4*C* shows data from the figure–ground ROIs rather than the face-ground ROIs. As the face-ground ROIs are based on less data (faces > homogenous) than the figure ground ROIs (faces, houses, and objects > homogenous), the latter has a higher SNR, which might have resulted in a more stable selection of face-ground selective voxels across the split half procedure. This procedure may have disproportionally impacted BA17 given the relatively few voxels that respond to the figure–ground contrast in that region compared with the other regions.

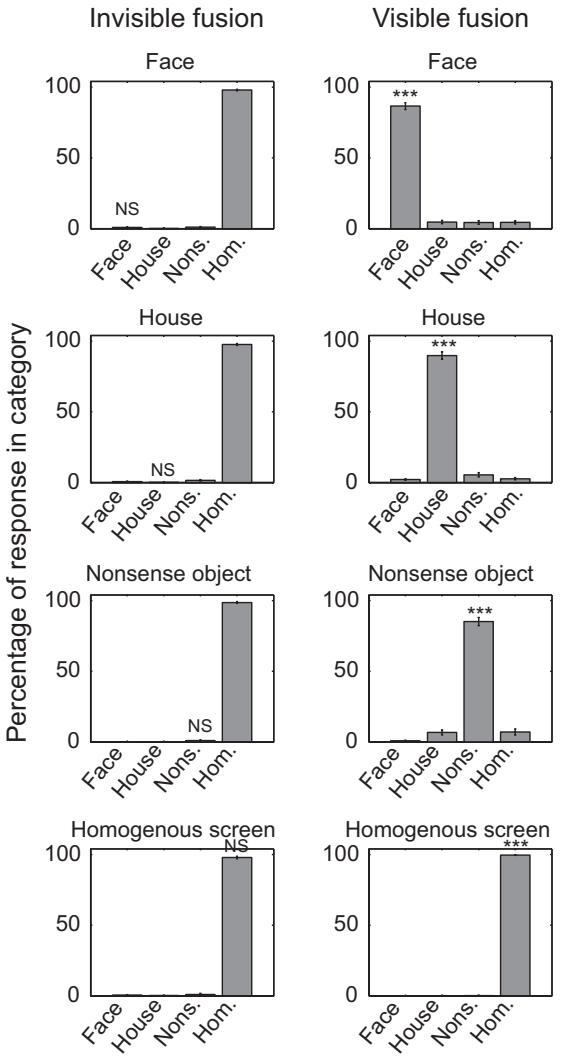

Fig. S7. Behavioral data of stimulus visibility in the EEG experiment. Stimulus classification during the postscanning control task for the invisible (*Left*) and visible conditions (*Right*). Each graph represents a stimulus type, showing the mean response percentage (±SEM) to that stimulus type for each response. For each stimulus type, a paired *t* test of the hits against false alarms determined whether that category was identified above chance level in that condition ($^{N.S.}P > 0.05$; $^{***}P < 10^{-13}$), confirming that performance on each of the invisible categories was at chance. Task performance on ellipse localization during the control task was the same as during EEG acquisition (mean hit rate control = 0.90 vs. EEG = 0.93, $F_{1,15} = 1.561$, $P = 0.228$), confirming that stereoscopic viewing conditions between the experimental and control runs were comparable.

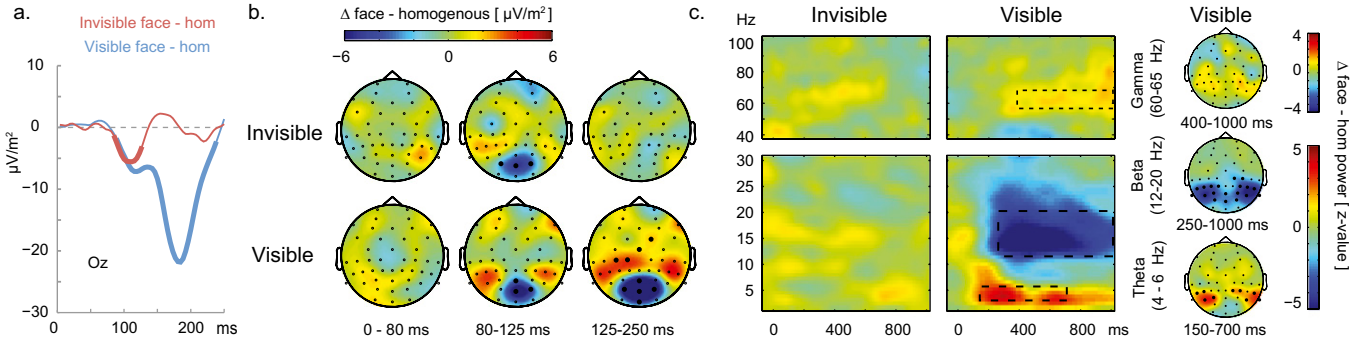

Fig. S8. Face-ground responses. (*A*) The evoked face-ground EEG response to visible and invisible stimuli at electrode Oz. Thick lines indicate cluster corrected significance at $P < 0.05$. (*B*) Topographic distributions of the face-ground signal evoked by invisible (*Upper*) and visible (*Lower*) stimuli. Thick electrodes indicate significant time-electrode clusters. (*C*) Induced EEG responses to face-ground modulation. (*Left*) Time frequency representations of the face-ground signal at electrodes P3, PO3, PO7, P4, PO4, and P08 of the bilateral occipitotemporal cortex. (*Right*) Topographic distribution of the frequency bands that were outlined in the time frequency panels.