NeuroImage 146 (2017) 71–80

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Scene segmentation in early visual cortex during suppression of ventral stream regions

Pablo R. Grassi[a,b,c,*], Natalia Zaretskaya[a,b,c], Andreas Bartels[a,b,c,*]

[a] Centre for Integrative Neuroscience, University of Tübingen, 72076, Tübingen, Germany
[b] Department of Psychology, University of Tübingen, 72076, Tübingen, Germany
[c] Max Planck Institute for Biological Cybernetics, 72076, Tübingen, Germany

ARTICLE INFO

Keywords:
fMRI
V1
Feedback
Gestalt
Ventral stream
Bistable perception

ABSTRACT

A growing body of literature suggests that feedback modulation of early visual processing is ubiquitous and central to cortical computation. In particular stimuli with high-level content that invariably activate ventral object responsive regions have been shown to suppress early visual cortex. This suppression was typically interpreted in the framework of predictive coding and feedback from ventral regions. Here we examined early visual modulation during perception of a bistable Gestalt illusion that has previously been shown to be mediated by dorsal parietal cortex rather than by ventral regions that were not activated. The bistable dynamic stimulus consisted of moving dots that could either be perceived as corners of a large moving cube (global Gestalt) or as distributed sets of locally moving elements. We found that perceptual binding of local moving elements into an illusory Gestalt led to spatially segregated differential modulations in both, V1 and V2: representations of illusory lines and foreground were enhanced, while inducers and background were suppressed. Furthermore, correlation analyses suggest that distinct mechanisms govern fore- and background modulation. Our results demonstrate that motion-induced Gestalt perception differentially modulates early visual cortex in the absence of ventral stream activation.

1. Introduction

The organization of a visual scene into a set of coherent objects is an extraordinary feat of the visual system. It involves the binding of local visual features into unified structures, segregation of fore- and background, and filling in of missing surface- and contour information. Monkey electrophysiology studies on perceptual grouping have shown that neurons in early visual cortex respond selectively to border-ownership and illusory contours (Peterhans and von der Heydt, 1989) and that foreground surfaces get filled in (Poort et al., 2012; Roelfsema et al., 2007). In accord with this, human fMRI studies have consistently shown modulation of early visual cortex during object or Gestalt processing, suggesting Gestalt-related feedback from higher level regions (Kok and de Lange, 2014; Murray et al., 2002). These findings are compatible with the framework of predictive coding, in that high-level predictions interact with sensory signals in early cortex (Friston, 2005; Rao and Ballard, 1999).

However, potential sources for this feedback have not been conclusively established. All shape or Gestalt stimuli previously used to report early visual modulation also involved activation of ventral shape or object processing regions in the ventral lateral occipital cortex (LOC) compared to control stimuli (Fang et al., 2008; Hirsch et al., 1995; Mendola et al., 1999; Murray et al., 2002; Stanley et al., 2003; Strother et al., 2012). Simple “Kanizsa” shapes (Mendola et al., 1999; Stanley and Rubin, 2003) as well as moving diamonds, 3D shapes, or structure-from-motion stimuli have been shown to activate LOC compared to their control stimuli (De-Wit et al., 2012; Fang et al., 2008; Halgren et al., 2003; Murray et al., 2003, 2002), and electrophysiology has identified grouping-related responses in V4 that may account for so-called filling-in through feedback (Cox et al., 2013). To our knowledge no prior study examined early visual cortex modulations during Gestalt percepts that are not accompanied by ventral stream activation.

In a prior study we had examined a bistable motion illusion that induced mutually exclusive percepts of either a dynamic illusory Gestalt or unbound local moving elements (Grassi et al., 2016; Zaretskaya et al., 2013). This stimulus was accompanied by suppression of early visual cortex and of ventral regions, with parietal cortex being the sole activated region and also causally involved in Gestalt percept generation (Grassi et al., 2016; Zaretskaya et al., 2013).

In the present study we examined the precise nature of early visual cortex modulation, as, in contrast to all preceding studies, LOC and

* Correspondence to: Vision and Cognition Lab, Centre for Integrative Neuroscience, University of Tübingen, Otrfried-Müller-Str. 25, 72076, Tübingen, Germany.
E-mail addresses: Pablo.grassi@cin.uni-tuebingen.de (P.R. Grassi), andreas.bartels@tuebingen.mpg.de (A. Bartels).

http://dx.doi.org/10.1016/j.neuroimage.2016.11.024
Received 20 July 2016; Accepted 10 November 2016
Available online 12 November 2016
1053-8119/

other ventral stream regions were suppressed during Gestalt perception. We examined regions of interest corresponding to the physical inducers, the foreground, the illusory borders, and the background, in areas V1 and V2. Despite the ventral stream suppression, we found a specific modulation pattern in the early visual cortex: illusory contour regions and foreground were enhanced, while the inducers and background were suppressed. In addition, the degree of modulation in different sub-regions correlated across V1 and V2, but not with each other, suggesting that enhancement of foreground and suppression of background is governed by distinct processes.

2. Material and methods

2.1. Participants

Seventeen volunteers (21–29 years old, ten females, three left-handed, one author) participated in this study after signing an informed consent form. All had normal or corrected-to-normal vision and no history of neurological impairments. The study was conducted according to the Declaration of Helsinki and was approved by the ethics committee of the University Clinic Tübingen.

2.2. Imaging parameters

Data were acquired on a 3 T Siemens Prisma system with a 64-channel head coil (Siemens, Erlangen, Germany). Functional data were acquired using a slice-accelerated multiband gradient-echo echo planar imaging (EPI) sequence using T2* weighted blood oxygenation level dependent (BOLD) contrast (multiband acceleration factor: 4, GRAPPA acceleration factor: 2), with the following parameters: repetition time (TR)=870 ms, echo time (TE)=30 ms, flip angle=56°, FOV=192×224 mm and an isotropic voxel size of 2×2×2 mm. This high-resolution sequence allowed us to measure the spatial distribution of BOLD responses in early visual areas with a higher accuracy compared to standard EPI sequences. Whole-brain functional images consisted of 56 slices. The initial ten images of each experimental fMRI run were discarded for equilibration of T1 signal. A high-resolution T1-weighted anatomical scan was performed for each participant (ADNI, 192 slices, voxel size 1 mm^3, TR=2 s, TE=3.06 ms, FOV=232×256 mm). Data for the main experiment and for the structural scan were acquired in the same session. Data for retinotopic mapping were acquired on a separate session.

2.3. Retinotopic mapping

Using standard retinotopic protocols we delineated areas V1–V4 using phase-encoding (Engel et al., 1994; Sereno et al., 1995) in seven subjects. For the remaining ten subjects, V1 and V2 were defined using anatomical labels generated by FreeSurfer (Hinds et al., 2009) and standard anatomical criteria (see ROI definition below) (Wandell et al., 2007). The stimulus used for retinotopic mapping consisted of a wedge shaped checkerboard (100% contrast, 6 Hz contrast inversion flicker, check sizes increased logarithmically with eccentricity, wedge covered 90°) on a gray background rotating around the fixation dot at the center of the screen with a cycle duration of 55.7 s. Image acquisition and preprocessing was identical to the main experiment. Surface reconstruction (Dale et al., 1999) as well as visual area definition was done in FreeSurfer (Fischl et al. 1999; http://surfer.nmr.mgh.harvard.edu/).

2.4. Stimulus, task, and experimental procedure

The bistable motion stimulus was generated using MATLAB 2010a (MathWorks, Natick, MA) with the Psychophysics Toolbox 3 extensions (Brainard, 1997; Pelli, 1997) on a Windows computer and presented using a linearized projector (1024×768 resolution, 60 Hz). The visual

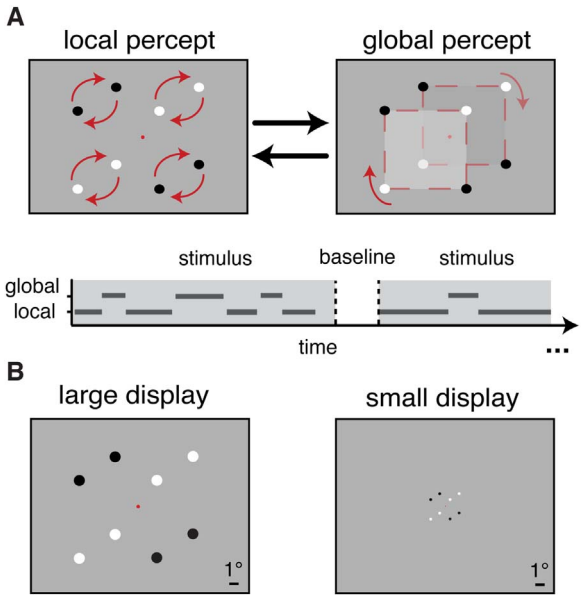

Fig. 1. Illustration of stimuli and the two possible perceptual interpretations. (A), The bistable motion stimulus led to alternations between the perception of local motion of dot pairs and of two illusory squares sliding over each other in planar motion (global Gestalt perception), respectively. During the stimulus presentation, subjects indicated periods of global or local motion perception by button presses. (B), Two stimulus conditions: large display and small display (ratio 1:4). Scale bar equals one visual degree.

display covered 22×16.3° visual degrees and was viewed at 82 cm distance. The stimulus consisted of four pairs of dots moving in-phase on circular paths (Anstis and Kim, 2011) (see Fig. 1A). This stimulus is perceptually bistable, and can be perceived as local motion of individual dots or as two illusory squares sliding over each other (i.e. global Gestalt motion). These two percepts alternate during continuous viewing.

During the experiment, subjects were asked to fixate the red dot and report their current percept by pressing and holding down one of two buttons on an fMRI-compatible keyboard with their right hand (one button was for the local, the other one for the global percept). Participants were instructed not to report anything if unsure of their percept. Simultaneous key presses were not used for further analysis.

The stimulus was presented in two different sizes in separate conditions (Fig. 1B). This allowed us to examine effects of eccentricity on BOLD responses. In the “large display” condition, individual dots had a size of one visual degree, the radius between dots and center of the dot pair was 2°, and the distance between each dot pair and center was 5°. The edge-length of the illusory square was 7°. Values in the “small display” were scaled to a fourth of those from the “large display” (i.e. 0.25° dot size, 0.5° dot-pair radius, 1.25° pair-center distance, edge-length 1.75°). Both stimuli had a red fixation dot (0.125°) in the center of the screen and were presented on a gray background (368.7 cd/m^2) at 100% contrast.

The dot rotation speed was adjusted for each participant individually prior to the experiment in order to achieve balanced durations of local and global perception. This led to a range of dot rotation speeds of 2–2.5 rotations per second (mean ± S.D.: 2.27 ± 0.13) across subjects. Dots from the same and opposite doublets had the same color (black or white) and alternated between each stimulus presentation.

Each fMRI run consisted of four stimulus presentations (90 s each) intermitted by 15 s of baseline consisting of fixation only. Every participant underwent 4–5 experimental runs. The experimental conditions (i.e. large and small stimulus) were presented in a pseudorandomized counterbalanced sequence.

2.5. Preprocessing and GLM analysis

For preprocessing, functional volumes were slice-time corrected, motion-corrected and coregistered with a high-resolution anatomical scan. Slice-time correction was done using SPM8 (Wellcome Department of Imaging Neuroscience, London, UK), all other analysis steps including subsequent GLM analysis used the FreeSurfer software (Fischl et al., 2008). After preprocessing, volume data were projected onto individual inflated cortical surfaces. All further analysis was conducted on the cortical surfaces. We performed a cortical surface based analysis for each subject separately using a standard GLM approach. For this, functional data was smoothed on the surface using a Gaussian filter with 3 mm full-width at half-maximum. Seven regressors of interest were included. They modeled the full durations of the baseline, of the stimulus periods for large and small displays, and onsets of global and local percepts, separately for large and small displays. We also included six nuisance regressors with the motion realignment parameters and a regressor consisting of a column of ones to model offset effects of each run (speed or other unspecific effects).

2.6. ROI definition

We restricted our surface-based ROI analysis to V1 and V2. Area V3 and higher areas have substantially larger receptive field sizes that make it difficult or impossible to delineate sub-regions representing the different stimulus aspects. The boundaries of visual areas V1 and V2 were determined manually in seven subjects based on the responses to the retinotopic mapping runs. In the remaining ten subjects they were determined based on the anatomical V1 and V2 labels generated by FreeSurfer (Hinds et al., 2009) and standard anatomical criteria (Wandell et al., 2007). Subsequent analyses were performed on data pooled across both approaches, but also separately for retinotopically and anatomically defined visual regions to verify their consistency (see Section 3.5.2).

ROIs within V1 and V2 were defined based on responses to the physical presentation of stimuli versus blank screen, i.e. the contrasts "large stimulus > baseline", and "small stimulus > baseline", respectively. This provided neural responses to the moving dot locations independent of the perceptual state. ROIs were defined in V1 and V2 on reconstructed inflated hemispheres (see Fig. 2). In a separate control analysis we defined ROIs independent of functional responses of the main experiment, based solely on their inferred retinotopic location only, yielding a consistent outcome (see Section 3.5.2).

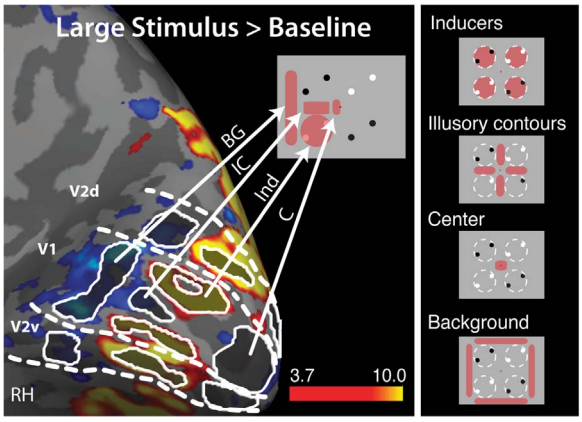

Fig. 2. Representative early visual cortex responses for the contrast "large stimulus > baseline" used for ROI definition in V1 and V2 (thresholded at t=3.7). This spatial localizer contrast reflects the retinotopic location of the physical stimulus. Stimulus-specific ROIs are superimposed on the BOLD activity (red–yellow). Dashed lines indicate area boundaries of V1 and V2. The right panel illustrates the stimulus space represented by the different ROIs. Ind: inducer region; IC: illusory contour region; BG: background region; C: center of illusory figure; V2d, dorsal V2; V2v, ventral V2; RH, right hemisphere.

Fig. 2 illustrates all four types of ROIs that were defined: inducer regions, illusory contour, center, and background (for an example of the ROI definition for the small display see Supplemental Fig. S1). The 'inducer regions' corresponded to the retinotopic locations of the dot pairs and were identified for each hemisphere dorsally and ventrally from the calcarine sulcus (min threshold for detection: t=3.7; $p < 0.0001$, uncorrected). Inducer regions were defined in V1 and V2, for small and large stimuli. The definition of both inducer regions during the small stimulus presentation was successful in 52 cases in V1 and 66 cases in V2 out of an expected total of 68 inducer regions (17 subjects, two hemispheres, two inducers per hemisphere). Based on the retinotopic locations of the inducers we defined further stimulus-specific regions. For the large stimulus in V1 it was possible to define the gap between the representations of the inducers as 'illusory contours', the 'center' of the large square formed by the inducers (defined in N_{hemi}=34), and the 'background' (N_{hemi}=34), i.e. a patch of cortex eccentric to the 'inducer' and 'illusory contour' regions. For V2, and for the small stimulus in V1 and V2, we defined only 'background', and 'center', in addition to the inducers (N_{hemi}=34 in all). For the small stimulus in both areas and for both stimuli in V2 it was not possible to reliably identify 'illusory contours'.

We separated the 'illusory contour' regions (i.e. gap between the representation of the inducers corresponding to the areas where the illusory borders of the transparent square move) from the 'center' regions in V1 during the large stimulus presentation to examine to what extend these regions elicit similar responses. Previous work using an apparent motion stimulus revealed an increase of activity in V1 sites representing an illusory motion path (Muckli et al., 2005). We hypothesize that similar responses could be elicited by our illusory stimulus. We hence defined 'illusory contour' ROIs, along the horizontal meridian (N_{hemi}=32) and, whenever possible, also along the vertical meridians (N_{hemi}=12). Please note, that since the representation 'illusory contour' regions along the vertical meridian lay exactly on the V1–V2 border, for simplicity we assigned those regions to V1.

2.7. ROI data analysis

To determine whether the specific ROIs in areas V1 and V2 showed perceptual modulations, we extracted the beta estimates within each individual ROI, averaged all values and normalized them to BOLD percent signal change using the respective ROI mean signal as normalization reference. Next, average beta estimates for global and local perceptual onsets were compared using paired-sample t-tests. To control for multiple comparisons, we applied a Bonferroni-Holmes-correction for the number of ROIs (thirteen tests). All t-tests survived correction.

To verify that the different visual field definition approaches (retinotopic-mapping and anatomic-based definition) were equivalent, we additionally performed ROI analyses in both subgroups of subjects.

We also wanted to determine to what extent the percept-driven modulation of ROI responses were correlated between ROIs. Similar modulation strength of inducers, figure and background would imply a common driving force, whereas uncorrelated modulation would suggest different underlying processes. We quantified the preference of each specific region to global perception in each experimental condition using the contrast value for global versus local (one value per subject per hemisphere). In order to relate the responses in the different regions and measure their similarity, we computed the Pearson correlation coefficients between all global percept preference indices with n=34 hemispheres. An alternative approach could be to perform this analysis within each subject, using trial-wise fluctuations. However, since many of the stable perceptual periods lasted only one or a few TRs, trial-wise activity would be extremely noisy and confounded by processes not related to the perceptual fluctuations.

We hence conducted this analysis across hemispheres using more reliable mean contrast values.

2.8. Whole-brain analysis

Most previous studies working on perceptual grouping showed an increase of activity during global illusory shape perception in shape-responsive ventral lateral occipital cortex (LOC), for both unambiguous (Mendola et al., 1999; Stanley et al., 2003; Strother et al., 2012) and bistable stimuli (De-Wit et al., 2012; Fang et al., 2008; Murray et al., 2002). Compatible with this, monkey electrophysiology studies revealed neural selectivity for illusory contours in the infero temporal (IT) cortex (Sáry et al., 2008) and showed that neurons in V4 responded selectively when the illusory foreground surface fitted their receptive fields (Cox et al., 2013). In contrast, the bistable, purely perceptually driven motion illusion used in this study is likely mediated by dorsal parietal areas. Furthermore, it has been shown to occur without the involvement of shape-responsive ventral areas like V4 and both sub-divisions of LOC, the lateral occipital gyrus (LO) and the posterior fusiform gyrus (pF) (Grassi et al., 2016; Zaretskaya et al., 2013).

We conducted a whole-brain group analysis on the surface to provide an illustration of ventral suppression during global Gestalt perception and also to illustrate how early visual cortex modulations might have averaged into negative values in our previous studies (Grassi et al., 2016; Zaretskaya et al., 2013). For this analysis, we projected single subject preprocessed volume data onto the average template of the FreeSurfer software. Thereafter, we spatially smoothed the data along the surface with a Gaussian kernel of 8 mm full-width at half-maximum and performed a standard fixed-effects GLM analysis for every subject (as in the ROI analysis). Subsequently, we conducted a random-effects group analysis and tested the contrast "global > local" to illustrate clusters being deactivated during global Gestalt perception. Note that a detailed description of whole-brain results using this bistable stimulus can be found in our previous work (Grassi et al., 2016; Zaretskaya et al., 2013).

2.9. Eye tracking

We monitored the fixation accuracy of 16 subjects throughout the fMRI experiment using a video-based infrared eye-tracker with long-range optics (ASL EyeTrac 6 Eye Tracking System, Applied Science Laboratories) recorded using ViewPoint Eyetracking Software (Arrington Research, Scottsdale, USA) at a sampling rate of 60 Hz. The preprocessing of raw data included blink detection and linear interpolation, temporal smoothing with a 200 ms running average window and high-pass filtering. Thereafter, we computed average Euclidean distances between gaze position and fixation dot as well as blink and saccade rates. We tested potential differences between conditions by means of two-way repeated measurements ANOVA (2×2 rANOVA) with factor 'stimulus size' (large/small) and 'percept' (global/local).

3. Results

3.1. Behavioral results

Over all subjects the average median duration of the global percept was 7.13 ± 3.23 s (mean ± SD) during the large display and 5.38 ± 1.53 s during the small display. The average duration of local percepts was shorter: 4.73 ± 1.58 s and 3.98 ± 1.81 s, respectively. Main effects 'stimulus size' and 'percept type' were significant (2×2 rANOVA, 'stimulus size': $F_{1,16}$=10.84, p=0.0046; 'percept type': $F_{1,16}$=17.51, p=0.0007), with no significant interaction ($F_{1,16}$=2.59, p=0.1268). The distributions of dominance durations were well fitted with a gamma function, as tested by the Kolmogorov-Smirnov goodness-of-fit test (all p > 0.1).

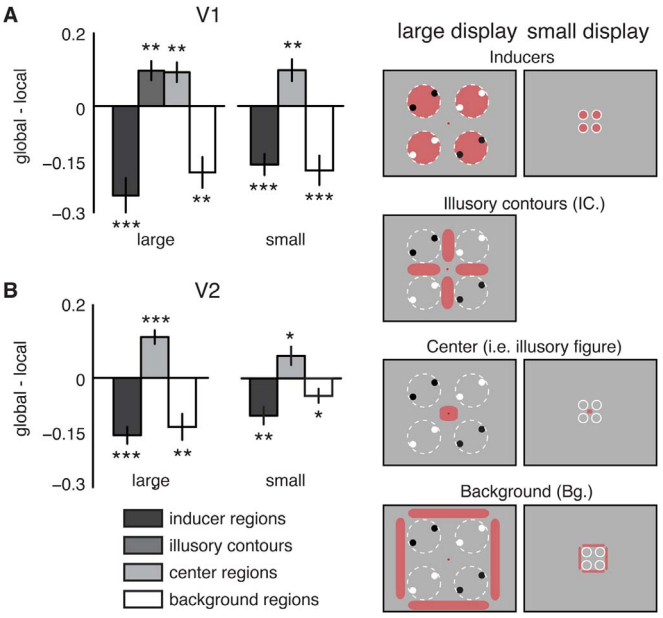

Fig. 3. Percept-driven modulations in V1 and V2 during viewing of the bistable stimulus, shown for regions representing inducer, illusory contours, center, and background. All examined regions of interest are shown in the right panel. For detailed information of the ROI definition, see the Material and Methods section. (A) BOLD signal modulations shown as beta contrast (global – local) in V1 during the large (left) and small (right) stimulus display. (B) BOLD signal in the secondary visual cortex (V2) during the large (left) and small (right) display, respectively. *** p < 0.001, corrected; ** p < 0.01, corrected; * p < 0.05, corrected. Plots show mean ± SEM (n=34).

3.2. Percept specific ROI analysis

We investigated activity modulations in early visual cortex during purely *perceptual changes* associated with perceiving the bistable stimulus as global Gestalt or as locally moving elements. These activity modulations were quantified for spatially specific subregions within V1 and V2, corresponding to the locations of the physical dots stimuli ('inducer regions'), the illusory Gestalt ('illusory figure': broken down to 'illusory contours' and 'center' in V1 for the large stimulus), and the background of the presented stimuli ('background regions'). These subregions were defined separately for large and small stimuli, which allowed us to rule out that any modulations were confounded by eccentricity, as e.g. inducer regions for small stimuli fell within the center regions of large stimuli, etc. (see Fig. 3). Analyses were carried out separately for both stimulus sizes, and for ROIs defined in V1 and V2.

Analyses for both stimulus sizes revealed highly consistent response patterns, both in V1 and in V2 (see Figs. 3 and 4). Fig. 3 depicts responses for the contrast (*global percepts – local percepts*) for all ROIs. In Fig. 4 we show for completeness the responses for each perceptual state separately (global and local), with baseline being fixation only (i.e. no stimulus except fixation dot).

Fig. 3 A reveals that the inducer regions in V1 showed a relative suppression of BOLD signal during Gestalt perception compared to local perception (t_{33}=5.11, p < 0.0001). This differential suppression was present also during viewing of the small stimulus (t_{33}=5.56, p < 0.0001). In secondary visual cortex (V2), this result was replicated for both stimulus sizes (t_{33}=6.71, p < 0.0001; t_{33}=4.23, p=0.0002, respectively) (Fig. 3B). Note that this spatially specific decrease of neural activity cannot be explained by differences in physical stimulus properties between global and local conditions, since physical stimulus characteristics remained constant.

In contrast, regions of the early visual cortex that corresponded to the illusory Gestalt showed an overall increase of neural activity during Gestalt perception, although no physical stimulus was present in these

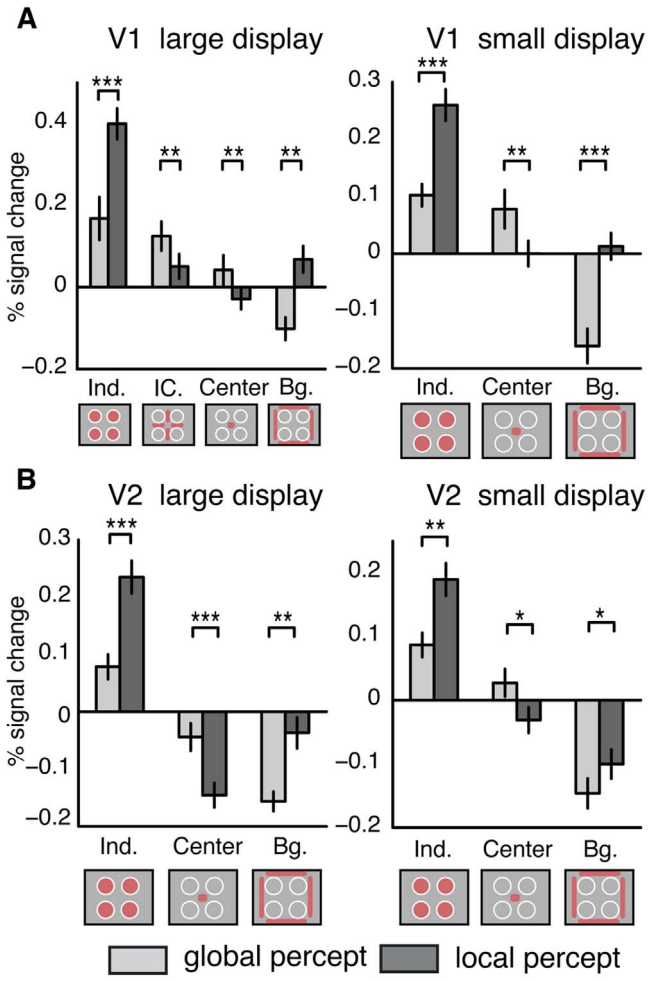

Fig. 4. Neural responses for each of the perceptual states (global and local) relative to the baseline (fixation only) periods. (A) Differential modulations in the primary visual cortex (V1) during the large (left) and small (right) stimulus display for the different regions of interest. (B) Neural responses in the secondary visual cortex (V2) during the large (left) and small (right) display. Fixation-only trials served as baseline in all plots. *** $p < 0.001$, corrected; ** $p < 0.01$, corrected; * $p < 0.05$, corrected. Plots show mean ± SEM (n=34).

regions. In V1, the 'illusory contour' regions (t_{31}=3.72, p=0.0008), and the 'center' regions of the illusory Gestalt (t_{33}=3.46, p=0.0015) were up-modulated during global perception in the large stimulus. This selective enhancement of 'center' activity was also observed during the small stimulus display in V1 (t_{33}=3.27, p=0.0025). In V2, this effect was also present during the large (t_{33}=5.95, $p < 0.0001$) and small (t_{33}=2.45, p=0.0198) stimuli, even though somewhat weaker in the latter. Thus, both V1 and V2 regions corresponding to the illusory figure showed an increased response during global perception. Crucially, eccentricity effects cannot account for this selective enhancement, since both, large and small displays showed a similar modulation pattern.

Proximate background regions of the illusory Gestalt showed a decrease of activity during global perception. This was the case in visual area V1 (t_{33}=4.3, p=0.0001) and V2 (t_{33}=3.75, p=0.0007) during the large display, and also during the small display condition (V1: t_{33}=4.39, p=0.0001; V2: t_{33}=2.48, p=0.0185, respectively).

Although the differential modulation was consistent across V1 and V2 (Fig. 3), the two regions appeared to differ in their absolute response signs relative to the baseline in the 'center' and 'background' regions (see Fig. 4). Responses to local perception compared to baseline were positive in V1, but negative in V2. Also responses to the global perception in the 'center' ROI were negative compared to baseline in V2 during the large stimulus presentation, but not in all other cases.

3.3. Relation between inducers, center, and background

The modulations in early visual cortex raise an important question. In particular, since distinct sub-regions in early visual cortex are modulated in opposite directions and represent entirely distinct aspects of visual perception, one wonders to what extent they are related. Similar modulation strengths of inducers, figure and background would imply a common driving force, whereas uncorrelated modulation would suggest different underlying processes. To examine the relation between distinct ROIs we computed the Pearson correlation coefficients between the magnitudes of their percept-driven modulations (contrast global versus local) across hemispheres. A high correlation (positive or negative) between two ROIs would indicate a common driving source for both, whereas no correlation would imply independent processes driving their modulation.

The results revealed a high correlation between corresponding ROIs (i.e. inducer correlated with inducer, center with center and background with background) even between distinct visual areas and display sizes (see Fig. 5). In contrast, non-corresponding ROIs exhibited no systematic correlation, even within the same visual area and display size.

Moreover, we specifically tested whether indices of corresponding ROIs obtained from different display sizes (e.g. inducer small display vs. inducer large display) had higher correlations than non-corresponding ROIs, also from different display sizes (e.g. inducer small display vs. background large display). Results confirm that indices from corresponding ROIs had higher correlations than those from non-corresponding ROIs across both display sizes (two-sample t-test, t_{64}=5.34, $p < 0.0001$) (Fig. 5B, upper bar plot) and, most importantly, between distinct display sizes (t_{34}=4.07, p=0.0003) (Fig. 5B, lower bar plot).

3.4. Whole-brain results

In order to examine whether a whole-brain group analysis would also reveal positive and negative responses in early visual regions (as revealed by the ROI analysis), we conducted a whole-brain analysis with the contrast "global > local". As in our previously published analyses (Grassi et al., 2016; Zaretskaya et al., 2013), we found a large deactivation in the early visual cortex, extending ventrally deep into the fusiform gyrus, for both experimental conditions (see Fig. 6). In fact, in contrast to our ROI results, we could not find any active voxel favoring global motion perception in the early visual cortex (even with a liberal threshold of t=2). This directly illustrates how early visual cortex modulation might have averaged into negative values in our as well as in other previous studies through group averaging (Grassi et al., 2016; Zaretskaya et al., 2013). Moreover, the whole-brain analysis replicated the deactivation of shape-responsive ventral areas (LO, pF and V4), together with activation of the posterior parietal cortex during global Gestalt perception (Grassi et al., 2016; Zaretskaya et al., 2013) (see Fig. 6 and Supplemental Fig. 2). In prior studies using illusory or real shapes, LOC has been consistently activated, including stimuli using sparse inducers such as "Kanizsa" shapes (De-Wit et al., 2012; Fang et al., 2008; Hirsch et al., 1995; Mendola et al., 1999; Murray et al., 2002; Stanley et al., 2003; Strother et al., 2012).

3.5. Control analyses and eye-tracking results

3.5.1. Invariance of results with respect to percept durations

Differences in percept durations between global and local percept types could potentially confound our fMRI results. To test whether this difference affected our results, we performed two control analyses. First, we conducted a separate ROI analysis with a subset of subjects with the opposite duration bias, i.e. longer local than global motion

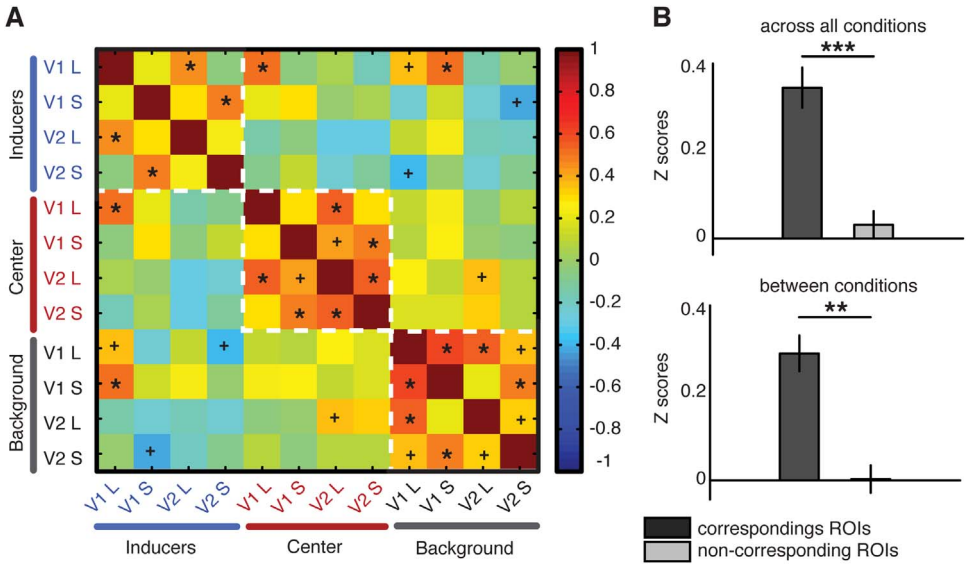

Fig. 5. Correlations between percept-driven modulations of corresponding and non-corresponding ROIs across display sizes and visual regions. (A) Correlation matrix relating global-vs.-local preference of each ROI (i.e. inducer, center and background regions), of each area (V1 and V2), for each display size (S: small, L: large) across hemispheres. * $p < 0.05$, FDR-adjusted corrected; + $p < 0.05$, uncorrected. (B) Bar plots show the Fisher-Z transformed correlation coefficients (mean ± SEM) for correlations between corresponding ROIs (e.g. inducer vs. inducer) and non-corresponding ROIs (e.g. inducer vs. background) averaged across both display sizes (upper bar plot) and between different display sizes (lower bar plot). *** $p < 0.0001$, uncorrected; ** $p < 0.001$, uncorrected.

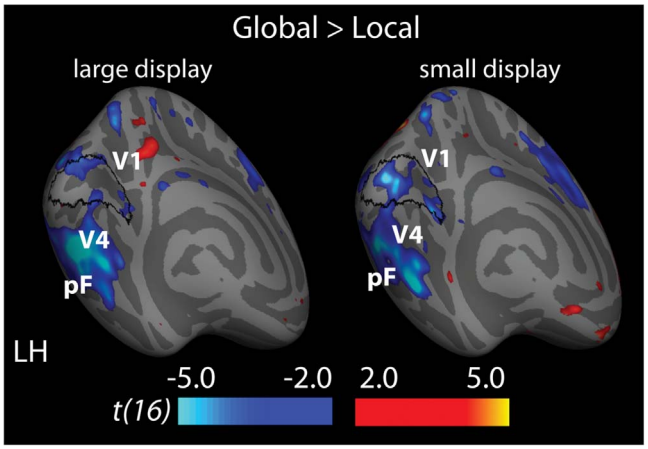

Fig. 6. Whole-brain results showing negative differential responses (contrast "global > local") during both stimulus displays (large and small) for the left hemisphere (medial rear view; thresholded at the liberal value of t=2). Responses in the right hemisphere were extremely similar. The black label depicts the prediction of V1 area borders generated by the FreeSurfer software (Hinds et al., 2009). Deactivations extend from the early visual cortex into shape-responsive areas V4, posterior fusiform gyrus (pF) and lateral occipital gyrus (LO, see Supplemental Fig. 2). LH, left hemisphere.

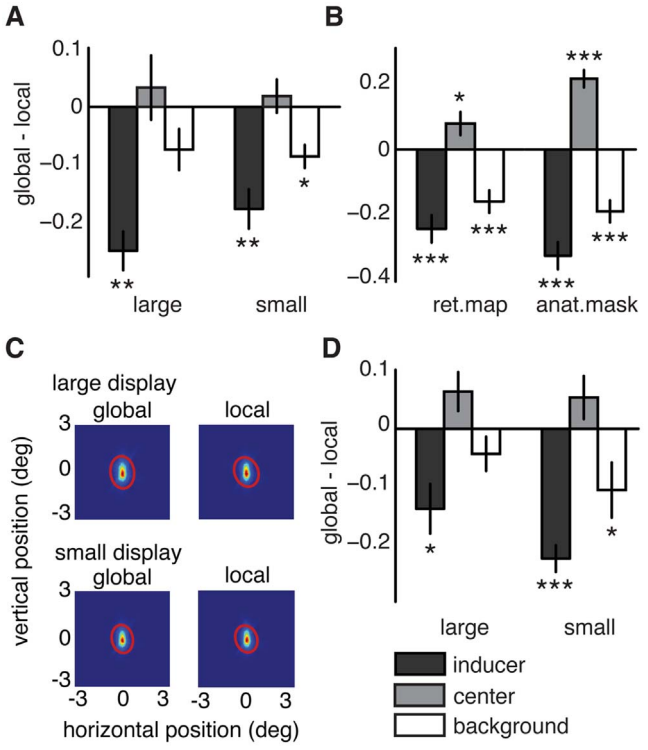

Fig. 7. Control ROI analyses and eye-tracking results. (A) Control analysis for percept duration. Same analysis as in Fig. 3, i.e. contrast values (global – local) for the three examined regions (inducer, center, background), but from a subset of subjects with longer local than global motion perception. Data are pooled across visual areas (V1 and V2). (B) Control analysis for definition of visual areas V1 and V2. Same analysis as in Fig. 3, but shown separately for participants in whom V1 and V2 were defined using retinotopy, or using an anatomy-based approach, respectively (see methods). Data are pooled across stimulus size and visual areas. (C) 2D-histograms on the eye positions (in visual degrees) of all measured subjects (n=16) to visualize fixation accuracy. Red lines depict the 68% confidence ellipse. (D) Control analysis for fixation accuracy. Contrast values from subjects with better fixation accuracy during global compared to local perception. Fixation-only trials served as baseline in all bar plots. *** $p < 0.001$, uncorrected; ** $p < 0.01$, uncorrected; * $p < 0.05$, uncorrected.

perception for both display sizes separately (large display: 2 subjects, small display: 3 subjects), to test if percept durations had an effect on the activation pattern described above. The results show the same tendency as in the original analysis (see Fig. 7A), even reaching significance in some cases. If percept durations rather than percept type had driven the results, the opposite pattern should have emerged. The present results hence strongly suggest that percept type, not durations, drove the results. Second, we correlated the mean contrast values (*betas from global motion – betas from local motion perception*) from all individual regions with the difference in dominance durations (*mean global duration – mean local duration*) to examine whether differential dominance durations correlated with differential fMRI signal effects. Among 12 correlations we examined, there was only one marginal negative correlation, namely of the background ROIs in V1 during the large stimulus presentation with the difference in dominance duration ($r=-0.54$, $p=0.024$). However, this correlation did

not survive correction for multiple comparisons. Hence, this analysis shows that percept durations had no systematic relationship with differential BOLD signal amplitudes in our paradigm, whereas percept type did.

Apart from this direct, empirical evidence that percept type, not duration, drove our results, also theoretical considerations would lead one to expect the same: in a GLM-based analysis as performed here, trial durations should only affect noise-estimates and hence significance levels, but not estimated magnitudes. Sign-switches (as observed between e.g. center and inducer ROIs) and the correlation results would hence be difficult to account for by potential biases introduced by perceptual durations. Finally, in two prior studies examining parietal percept-related modulations, our analyses showed that results remained stable when we purposefully analyzed sessions with systematically longer or shorter global or local percept durations, respectively (Grassi et al., 2016; Zaretskaya et al., 2013). On these grounds, we deem it unlikely that the difference in overall dominance durations confounded our fMRI results.

3.5.2. Invariance of results with respect to ROI definition

We performed two control analyses to test robustness of results with respect to alternative ROI definitions. First, we tested whether results were equivalent in our two subgroups of participants in whom visual areas V1 and V2 were defined using retinotopic mapping or the anatomy-based definition using FreeSurfer, respectively. We hence conducted separate ROI analyses for each subgroups of subjects. In accord with the main ROI analysis, the results revealed the same activation pattern, regardless of the approach used to define visual fields (see Fig. 7B). Second, we performed an alternative ROI definition for the 'inducers', 'center' and 'background' regions in V1 during the large stimulus presentation based on an anatomical template of the human striate retinotopy (Benson et al., 2014, 2012) to test if similar retinotopically specific modulations could be found with ROIs defined independently of main experiment (see Supplemental Fig. S3 A). The analysis of these independently defined ROIs shows the same activation pattern as in the original analysis (see Supplemental Fig. S3 B), revealing significant deactivations in the 'inducer' and 'background' regions and a positive tendency in the 'center' regions.

3.5.3. Invariance of results with respect to eye movements

Analysis of eye-tracking data revealed a high fixation accuracy in all conditions across all subjects (mean deviation over all subjects: 0.4638 ± 0.2033 SD, see Fig. 7C). There were no differences between experimental conditions (2×2 rANOVA, main effect 'stimulus size': $F_{1,15}$=2.65, p=0.1244). There was a marginal increase of gaze position error during global compared to local perception (main effect 'percept type': $F_{1,15}$=5.45, p=0.0339). To test if this difference affected the fMRI results we performed a control correlation analysis for all ROIs (pooled over corresponding ROIs in V1 and V2). Mean contrast values (*betas from global motion – betas from local motion perception*) from all regions (inducers, center and background) and differences in distance to fixation during both perceptual states were uncorrelated in both experimental conditions (all *r*-values < 0.3). Moreover, we conducted a separate ROI analysis with subjects with inverted fixation accuracy values (large display: 3 subjects, small display: 7 subjects). This analysis also showed the same tendency as the original ROI analysis (see Fig. 7D). Thus, the small difference in fixation accuracy between the two perceptual states (0.014° ± 0.0368) is unlikely to be associated to the early visual cortex modulations reported above.

Finally, statistical analysis of blink and saccades rates showed no significant difference regarding the main effect 'stimulus size' (blinks: $F_{1,15}$=1.51, p=0.2384; saccades: $F_{1,15}$=2.3, p=0.1591) nor for perceptual differences (blinks: $F_{1,15}$=1.88, p=0.1901; saccades: $F_{1,15}$=3.46, p=0.0824). All interactions were non-significant (p > 0.1).

4. Discussion

Using a bistable motion stimulus we examined activity modulations in early visual cortex that were driven purely by perceptual changes associated with Gestalt perception while the physical stimulation remained constant. Importantly, previous studies indicated that the Gestalt percept in this case was mediated by dorsal parietal rather than ventral regions (Grassi et al., 2016; Zaretskaya et al., 2013). Gestalt perception suppressed BOLD activity in regions of V1 and V2 corresponding to the local inducers and to the background. Regions representing the illusory Gestalt, i.e. its center and illusory contours, showed a significant increase. Our results were not only replicated in V1 and V2, but also using two distinct stimuli, where the second stimulus was reduced four times in size. Hence, the observed effects generalize across stimulus sizes and are independent of eccentricity effects. We found that the modulation strengths of corresponding sub-regions of the visual scene were correlated across visual areas (V1/V2) and stimulus sizes (large, small), whereas distinct sub-regions within the same area were not correlated. This suggests that different grouping-related processes may underlie the modulation of background, foreground, and inducers. These results provide first evidence of retinotopically resolved differential modulation in early visual cortex during bistable perception and scene segmentation in the absence of ventral activation.

These results are in line with a recent fMRI study using static Kanizsa shapes that showed a similar bi-directional response pattern: deactivation of the inducers together with an enhancement of the illusory figure (Kok and de Lange, 2014) that appear to occur in lower cortical layers (Kok et al., 2016). It is important to note though that the Kanizsa stimulus was compared to a physically distinct control, which does not fully preclude potential modulations driven by processes arising within early regions as proposed by prior models (Heitger and von der Heydt, 1993; Madarasmi et al., 1994; Tani et al., 2014; Zhaoping, 2005). Second, Kanizsa-type cues are static shape cues that have previously been shown to be processed in ventral stream object-related regions (Cox et al., 2013; Mendola et al., 1999; Sáry et al., 2008; Stanley et al., 2003). In contrast, in our paradigm Gestalt perception relied entirely on motion cues and led to suppression of ventral stream regions, with the only up-modulated regions being in parietal cortex (Grassi et al., 2016; Zaretskaya et al., 2013). The potential feedback mechanisms in the two paradigms may hence involve entirely distinct pathways.

4.1. Proposed explanations for the observed modulations

There are several, mutually non-exclusive, processes related to perceptual grouping that may – in isolation or combination – account for the segregated modulation patterns observed in the early visual cortex (Bartels, 2014).

The first set of processes is related to figure-ground segmentation. Neurons in V1 and V2 have been shown to distinguish the side of the contour "owned" by the foreground object (Zhou et al., 2000). Changes in border-ownership alone have been shown to up-modulate fMRI signal predominantly in V2 (Fang et al., 2009). This is consistent with enhanced firing of neurons whose receptive fields are inside the foreground (Lamme, 1995; Zipser et al., 1996). The relatively delayed response, and its predominant appearance in upper and lower layers suggests feedback as its driving force (Lamme and Roelfsema, 2000; Poort et al., 2012). The foreground enhancement observed here and in previous studies on illusory figure perception (Kok and de Lange, 2014; Seghier et al., 2000; Strother et al., 2012) is likely accounted for by such feedback (Kok et al., 2016). However, the suppression of neural representation of the inducers cannot be easily accounted for by any of the above described mechanisms of figure-ground segmentation.

An alternative explanation for the selective responses reported here involves attentional processes, known to modulate activity by facilitat-

ing neural representations in the attended locations (Somers et al., 1999; Tootell et al., 1998) and suppressing activity in nonattended locations (Gouws et al., 2014; Slotnick et al., 2003; Smith et al., 2000). During Gestalt perception, attentional resources may have been focused on the foreground, with comparably less resources on the inducers. Attention could also account for the inducer suppression: during local perception, each of the four sets of dot-pairs draws attentional resources, and hence was up-modulated. Such re-distribution of attention is also known as biased competition (Desimone and Duncan, 1995; McMains and Kastner, 2011).

The third set of processes that can account for the differential modulation is predictive coding (Friston, 2005; Muckli and Petro, 2013; Mumford, 1992; Rao and Ballard, 1999). Following this hierarchical model, sensory information in the early visual cortex is compared to feedback predictions from higher-level regions that attempt to account for sensory input. This theory predicts that the direction of early visual cortex modulations depends on how congruent top-down predictions and bottom-up inputs are. If predictions match the incoming data, both signals cancel, reducing responses in sensory cortex. High responses occur when there are no high-level predictions that match the bottom-up data, such as during local perception.

Hence, both, attentional and predictive coding accounts fit well with our data and with previous research (Kok and de Lange, 2014), showing a significant decrease of activity in regions where top-down interpretations match sensory input (i.e. inducer regions) and an increase of BOLD responses in areas where this is not the case (i.e. illusory figure areas) during Gestalt perception. Crucially, attentional and predictive coding accounts are not mutually exclusive, since attentional modulation can be understood in terms of predictive coding (Brown and Friston, 2013; Spratling, 2008). In accord with this, an extensive body of evidence supports the functional relationship of the mechanisms of the high-level processes of binding, scene segmentation and visual attention (Bartels, 2009; Fang et al., 2008; McMains and Kastner, 2011; Poort et al., 2012; Qiu and Sugihara, 2007; Wannig et al., 2011; Yokoi and Komatsu, 2009). Our findings do not only provide empirical evidence for these hierarchical models, but also show by means of a bistable stimulus that these early cortex modulations are purely dependent on the observer's conscious perceptual state.

4.2. Background suppression

Intriguingly, most studies on perceptual grouping and figure-ground segmentation failed to report background suppression. To our knowledge only one recent fMRI study of figure-ground segmentation showed background suppression (Strother et al., 2012), with similarly sparse evidence for background suppression in monkey electrophysiology (Gilad et al., 2013; Hupé et al., 1998). Our results are reminiscent of a concurrent fMRI and electrophysiology study on monkeys that revealed suppression of both BOLD signal and neural activity beyond stimulated regions of V1 (Shmuel et al., 2006). Our findings suggest that similar inhibitory mechanisms may also be involved during the perception of an illusory Gestalt. Clearly, more studies are needed to explain under which conditions background suppression does occur.

4.3. Sub-processes of perceptual organization

Our correlation analysis revealed a strong correlation of modulation strengths between corresponding sub-regions of the visual layout (e.g. background-with-background) but almost no correlation between distinct sub-regions (e.g. inducer-with-background). This implies independent sub-processes involved in mediating activity of inducer, center, and background, rather than a unique push-pull mechanism. This pattern of relationships appears to be common to both visual areas, and independent of display size or eccentricity. These results could be fMRI-analogues to effects revealed in monkey electrophysiology showing distinct processes involved in initial scene analysis, figure-ground segregation, and selective attention, each of which affect distinct scene-subregions at different time-windows (Poort et al., 2012; Roelfsema et al., 2007).

Although the pattern of differential perceptual modulation was consistent across all conditions and regions, we did find some differences in the sign of the modulation relative to fixation-only baseline in the 'center' and 'background' regions between V1 and V2. Especially during the 'local perception' in the background regions, V1 and V2 appeared to differ in the direction of the signal change compared to baseline. Even though it is unclear what drives these differences, they are consistent with previous work on high-order feedback processing in early visual cortex that showed differences between V1 and V2 in the processing of illusory contours (Ramsden et al., 2001), color (Bannert and Bartels, 2013) and size constancy (Sperandio et al., 2012). Future studies are needed to specifically investigate differences in feedback-related responses to V1 and V2.

4.4. Dorsal source of modulation

Most previous studies on perceptual grouping used explicit shape-cues and focused their research on shape-responsive regions along the ventral stream, known to be involved in shape and object representation (Altmann et al., 2003; Cox et al., 2013; Mendola et al., 1999; Murray et al., 2002; Stanley and Rubin, 2003). For example, Kanizsa-style stimuli have been shown to activate V4 and object-responsive regions (Cox et al., 2013; Mendola et al., 1999; Sáry et al., 2008; Stanley et al., 2003). Correspondingly, ventral shape-processing regions have been proposed as most likely source of potential feedback signals (Cox et al., 2013; Halgren et al., 2003; Murray et al., 2002; Stanley and Rubin, 2003).

Despite this, it is still unclear which high-level sources mediate early visual cortex modulation during scene segmentation. There could be multiple sources, each modulating distinct aspects in early cortex. For simpler motion cues, such as illusory motion, early modulation has been observed and attributed to motion regions that were not activated here (Alink et al., 2010; Vetter et al., 2015).

However, the bistable illusion used here differs from previously used simple motion but also from static shape stimuli: it relies entirely on sparse motion cues, and has been shown to deactivate ventral stream regions as well as mid-level motion regions, while instead activating the posterior parietal cortex (PPC) (Grassi et al., 2016; Zaretskaya et al., 2013). We have previously demonstrated a causal role of PPC in perceptual grouping for this stimulus, making it a reasonable candidate for sending grouping-related feedback signals to early visual cortex (Zaretskaya et al. 2013).

Indeed, a few prior studies demonstrated that the PPC is also involved in grouping and object processing. For example, it has been shown that the PPC contains invariant object representations (Konen and Kastner, 2008), and that it responds to perceptual grouping in humans (Xu and Chun, 2007; Zeki and Stutters, 2013) and monkeys (Yokoi and Komatsu, 2009). It is thus likely that grouping-related processes can be mediated by both the ventral as well as the dorsal hierarchically organized visual pathways and that areas within both streams are capable of modulating early visual cortex responses via feedback connections.

5. Conclusion

Our findings provide first evidence that scene segmentation based on purely perceptual processes during viewing of bistable stimuli elicits spatially segregated modulations in the early visual cortex in the absence of ventral stream involvement. Representations of illusory foreground and contours were enhanced, whereas inducers and background were suppressed. The results are in accord with both, biased competition models of attention as well as with predictive models of

visual perception. We hypothesize that the PPC is the cortical region responsible for the described perceptual modulations through recurrent connections.

Acknowledgments

We would like to thank Matthias Nau and Andreas Schindler for their assistance with the collection of the retinotopy data. This work was funded by the Centre for Integrative Neuroscience Tübingen (the German Excellence Initiative of the German Research Foundation (DFG) Grant no. EXC307), by DFG Grant BA4914/1-1 and by the Max Planck Society, Germany. The authors declare no competing financial interests.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuroimage.2016.11.024.

References

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. 30, 2960–2966. http://dx.doi.org/10.1523/JNEUROSCI.3730-10.2010.

Altmann, C.F., Bülthoff, H.H., Kourtzi, Z., 2003. Perceptual organization of local elements into global shapes in the human visual cortex. Curr. Biol. 13, 342–349.

Anstis, S., Kim, J., 2011. Local versus global perception of ambiguous motion displays. J. Vis. 11, 1–12. http://dx.doi.org/10.1167/11.3.13.Introduction.

Bannert, M.M., Bartels, A., 2013. Decoding the yellow of a gray banana. Curr. Biol. 23, 2268–2272. http://dx.doi.org/10.1016/j.cub.2013.09.016.

Bartels, A., 2009. Visual perception: converging mechanisms of attention, binding, and segmentation? Curr. Biol. 19, R300–R302. http://dx.doi.org/10.1016/j.cub.2009.02.014.

Bartels, A., 2014. Visual perception: early visual cortex fills in the gaps. Curr. Biol. 24, R600–R602. http://dx.doi.org/10.1016/j.cub.2014.05.055.

Benson, N.C., Butt, O.H., Brainard, D.H., Aguirre, G.K., 2014. Correction of distortion in flattened representations of the cortical surface allows prediction of V1–V3 functional organization from anatomy. PLoS Comput. Biol., 10. http://dx.doi.org/10.1371/journal.pcbi.1003538.

Benson, N.C., Butt, O.H., Datta, R., Radoeva, P.D., Brainard, D.H., Aguirre, G.K., 2012. The retinotopic organization of striate cortex is well predicted by surface topology. Curr. Biol. 22, 2081–2085. http://dx.doi.org/10.1016/j.cub.2012.09.014.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Brown, H.R., Friston, K.J., 2013. The functional anatomy of attention: a DCM study. Front. Hum. Neurosci. 7, 1–10. http://dx.doi.org/10.3389/fnhum.2013.00784.

Cox, M.A., Schmid, M.C., Peters, A.J., Saunders, R.C., Leopold, D.A., Maier, A., 2013. Receptive field focus of visual area V4 neurons determines responses to illusory surfaces. Proc. Natl. Acad. Sci. USA 110, 17095–17100. http://dx.doi.org/10.1073/pnas.1310806110.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

Desimone, R., Duncan, J., 1995. Neural mechanisms of selective visual attention. Annu. Rev. Neurosci. 18, 193–222. http://dx.doi.org/10.1146/annurev.ne.18.030195.001205.

De-Wit, L.H., Kubilius, J., Wagemans, J., de Beeck, H.P.O., 2012. Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. J. Vis. 12, 1–14. http://dx.doi.org/10.1167/12.11.12.Introduction.

Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.-J., Shadlen, M.N., 1994. fMRI of human visual cortex. Nature 369, 525. http://dx.doi.org/10.1038/369525a0.

Fang, F., Kersten, D., Murray, S.O., 2008. Perceptual grouping and inverse fMRI activity patterns in human visual cortex. J. Vis. 8, 2–9. http://dx.doi.org/10.1167/8.7.2.Introduction.

Fang, F., Boyaci, H., Kersten, D., 2009. Border ownership selectivity in human early visual cortex and its modulation by attention. J. Neurosci. 29, 460–465. http://dx.doi.org/10.1523/JNEUROSCI.4628-08.2009.

Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9, 195–207.

Friston, K., 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 815–836. http://dx.doi.org/10.1098/rstb.2005.1622.

Gilad, A., Meirovithz, E., Slovin, H., 2013. Population responses to contour integration: early encoding of discrete elements and late perceptual grouping. Neuron 78, 389–402. http://dx.doi.org/10.1016/j.neuron.2013.02.013.

Gouws, A.D., Alvarez, I., Watson, D.M., Uesaki, M., Rogers, J., Morland, A.B., 2014. On the role of suppression in spatial attention: evidence from negative BOLD in human subcortical and cortical structures. J. Neurosci. 34, 10347–10360. http://dx.doi.org/10.1523/JNEUROSCI.0164-14.2014.

Grassi, P.R., Zaretskaya, N., Bartels, A., 2016. Parietal cortex mediates perceptual gestalt grouping independent of stimulus size. Neuroimage 133, 367–377. http://dx.doi.org/10.1016/j.neuroimage.2016.03.008.

Halgren, E., Mendola, J., Chong, C.D.R., Dale, A.M., 2003. Cortical activation to illusory shapes as measured with magnetoencephalography. Neuroimage 18, 1001–1009. http://dx.doi.org/10.1016/S1053-8119(03)00045-4.

Heitger, F., von der Heydt, R., 1993. A computation model of neural contour processing: figure-ground segregation and illusory contours. In: Proceedings of the 4th International Conference Comput. Vision. pp. 32–40.

Hinds, O., Polimeni, J.R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., Schwartz, E.L., Fischl, B., Triantafyllou, C., 2009. Locating the functional and anatomical boundaries of human primary visual cortex. Neuroimage 46, 915–922. http://dx.doi.org/10.1016/j.neuroimage.2009.03.036.

Hirsch, J., Delapaz, R.L., Relkin, N.R., Victor, J., Kim, K., Lit, T.A.O., Borden, P., Rubini, N., Shapley, R., 1995. Illusory contours activate specific regions in human visual cortex: evidence from functional magnetic resonance imaging. Proc. Natl. Acad. Sci. USA 92, 6469–6473.

Hupé, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., Bullier, J., 1998. Cortical feedback improves discrimination between figure and background by V1–V3 neurons. Nature 394, 784–787. http://dx.doi.org/10.1038/29537.

Kok, P., de Lange, F.P., 2014. Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. Curr. Biol. 24, 1531–1535. http://dx.doi.org/10.1016/j.cub.2014.05.042.

Kok, P., Bains, L.J., Van Mourik, T., Norris, D.G., De Lange, F.P., 2016. Selective activation of the deep layers of the human primary visual cortex by top-down feedback. Curr. Biol. 26, 371–376. http://dx.doi.org/10.1016/j.cub.2015.12.038.

Konen, C.S., Kastner, S., 2008. Two hierarchically organized neural systems for object information in human visual cortex. Nat. Neurosci. 11, 224–231. http://dx.doi.org/10.1038/nn2036.

Lamme, V.A.F., 1995. The neurophysiology visual cortex figure-ground segregation in primary visual cortex. J. Neurosci. 15, 1605–1615.

Lamme, V.A.F., Roelfsema, P.R., 2000. The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. 23, 571–579.

Madarasmi, S., Ting-Chuen, P., Kersten, D., 1994. Illusory Contour Detection Using MRF Models 7. vol. 7, pp. 4343–4348.

McMains, S., Kastner, S., 2011. Interactions of top-down and bottom-up mechanisms in human visual cortex. J. Neurosci. 31, 587–597. http://dx.doi.org/10.1523/JNEUROSCI.3766-10.2011.

Mendola, J.D., Dale, A.M., Fischl, B., Liu, A.K., Tootell, R.B., 1999. The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. J. Neurosci. 19, 8560–8572.

Muckli, L., Petro, L.S., 2013. Network interactions: non-geniculate input to V1. Curr. Opin. Neurobiol. 23, 195–201. http://dx.doi.org/10.1016/j.conb.2013.01.020.

Muckli, L., Kohler, A., Kriegeskorte, N., Singer, W., 2005. Primary visual cortex activity along the apparent-motion trace reflects illusory perception. PLoS Biol. 3, e265. http://dx.doi.org/10.1371/journal.pbio.0030265.

Mumford, D., 1992. On the computational architecture of the neocortex – II. The role of the corticocortical loops. Biol. Cybern. 66, 241–251. http://dx.doi.org/10.1007/BF00202389.

Murray, S.O., Kersten, D., Olshausen, B.A., Schrater, P., Woods, D.L., 2002. Shape perception reduces activity in human primary visual cortex. Proc. Natl. Acad. Sci. USA 99, 15164–15169. http://dx.doi.org/10.1073/pnas.192579399.

Murray, S.O., Olshausen, B.A., Woods, D.L., 2003. Processing Shape , Motion and Three-dimensional Shape-from-motion in the Human Cortex. pp. 508–516.

Pelli, D.G., 1997. The videotoolbox software for visual psychophysics transforming numbers into movies. Spat. Vis. 10, 437–442.

Peterhans, E., von der Heydt, R., 1989. Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. J. Neurosci. 9, 1749–1763.

Poort, J., Raudies, F., Wannig, A., Lamme, V.A.F., Neumann, H., Roelfsema, P.R., 2012. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. Neuron 75, 143–156. http://dx.doi.org/10.1016/j.neuron.2012.04.032.

Qiu, F.T., Sugihara, T., 2007. Figure-ground Mechanisms Provide Structure for Selective Attention. pp. 1492–1499. doi: http://dx.doi.org/10.1038/nn1989.

Ramsden, B.M., Hung, C.P., Roe, a W., 2001. Real and illusory contour processing in area V1 of the primate: a cortical balancing act. Cereb. Cortex 11, 648–665. http://dx.doi.org/10.1093/cercor/11.7.648.

Rao, R.P., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87. http://dx.doi.org/10.1038/4580.

Roelfsema, P.R., Tolboom, M., Khayat, P.S., 2007. Different processing phases for features, figures, and selective attention in the primary visual cortex. Neuron 56, 785–792. http://dx.doi.org/10.1016/j.neuron.2007.10.006.

Sáry, G., Köteles, K., Kaposvári, P., Lenti, L., Csifcsák, G., Frankó, E., Benedek, G., Tompa, T., 2008. The representation of Kanizsa illusory contours in the monkey inferior temporal cortex. Eur. J. Neurosci. 28, 2137–2146. http://dx.doi.org/10.1111/j.1460-9568.2008.06499.x.

Seghier, M., Dojat, M., Delon-Martin, C., Rubin, C., Warnking, J., Segebarth, C., Bullier, J., 2000. Moving illusory contours activate primary visual cortex: an fMRI study. Cereb. Cortex 10, 663–670.

Sereno, M.I., Dale, a M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., Tootell, R.B., 1995. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268, 889–893. http://dx.doi.org/10.1126/science.7754376.

Shmuel, A., Augath, M., Oeltermann, A., Logothetis, N.K., 2006. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. Nat. Neurosci. 9, 569–577. http://dx.doi.org/10.1038/nn1675.

Slotnick, S.D., Schwarzbach, J., Yantis, S., 2003. Attentional inhibition of visual

processing in human striate and extrastriate cortex. Neuroimage 19, 1602–1611. http://dx.doi.org/10.1016/S1053-8119(03)00187-3.
Smith, A.T., Singh, K.D., Greenlee, M.W., 2000. Attentional suppression of activity in the human visual cortex. Neuroreport 11, 271–277. http://dx.doi.org/10.1097/00001756–200002070-00010.
Somers, D.C., Dale, A.M., Seiffert, A.E., Tootell, R.B., 1999. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. Proc. Natl. Acad. Sci. USA 96, 1663–1668. http://dx.doi.org/10.1073/pnas.96.4.1663.
Sperandio, I., Chouinard, P.A., Goodale, M.A., 2012. Retinotopic activity in V1 reflects the perceived and not the retinal size of an afterimage. Nat. Publ. Gr. 15, 540–542. http://dx.doi.org/10.1038/nn.3069.
Spratling, M.W., 2008. Predictive coding as a model of biased competition in visual attention. Vision. Res. 48, 1391–1408. http://dx.doi.org/10.1016/j.visres.2008.03.009.
Stanley, D.A., Rubin, N., 2003. fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. Neuron 37, 323–331. http://dx.doi.org/10.1016/S0896-6273(02)01148-0.
Stanley, D.A., Rubin, N., Place, W., York, N., 2003. fMRI Activation in Response to Illusory Contours and Salient Regions in the Human Lateral Occipital Complex 37. pp. 323–331.
Strother, L., Lavell, C., Vilis, T., 2012. Figure–ground representation and its decay in primary visual cortex. J. Cogn. Neurosci. 24, 905–914. http://dx.doi.org/10.1162/jocn_a_00190.
Tani, I., Yamachiyo, M., Shirakawa, T., Gunji, Y.-P., 2014. Kanizsa illusory contours appearing in the plasmodium pattern of Physarum polycephalum. Front. Cell. Infect. Microbiol. 4, 10. http://dx.doi.org/10.3389/fcimb.2014.00010.
Tootell, R.B., Hadjikhani, N., Hall, E.K., Marrett, S., Vanduffel, W., Vaughan, J.T., Dale, a M., 1998. The retinotopy of visual spatial attention. Neuron 21, 1409–1422. http://dx.doi.org/10.1016/S0896-6273(00)80659-5.
Vetter, P., Grosbras, M.H., Muckli, L., 2015. TMS over V5 disrupts motion prediction. Cereb. Cortex 25, 1052–1059. http://dx.doi.org/10.1093/cercor/bht297.
Wandell, B.A., Dumoulin, S.O., Brewer, A.A., 2007. Visual field maps in human cortex. Neuron 56, 366–383. http://dx.doi.org/10.1016/j.neuron.2007.10.012.
Wannig, A., Stanisor, L., Roelfsema, P.R., 2011. Automatic spread of attentional response modulation along Gestalt criteria in primary visual cortex. Nat. Neurosci. 14, 1243–1244. http://dx.doi.org/10.1038/nn.2910.
Xu, Y., Chun, M.M., 2007. Visual grouping in human parietal cortex. Proc. Natl. Acad. Sci. USA 104, 18766–18771. http://dx.doi.org/10.1073/pnas.0705618104.
Yokoi, I., Komatsu, H., 2009. Relationship between neural responses and visual grouping in the monkey parietal cortex. J. Neurosci. 29, 13210–13221. http://dx.doi.org/10.1523/JNEUROSCI.1995-09.2009.
Zaretskaya, N., Anstis, S., Bartels, A., 2013. Parietal cortex mediates conscious perception of illusory gestalt. J. Neurosci. 33, 523–531. http://dx.doi.org/10.1523/JNEUROSCI.2905-12.2013.
Zeki, S., Stutters, J., 2013. Functional specialization and generalization for grouping of stimuli based on colour and motion. Neuroimage 73, 156–166. http://dx.doi.org/10.1016/j.neuroimage.2013.02.001.
Zhaoping, L., 2005. Border ownership from intracortical interactions in visual area v2. Neuron 47, 143–153. http://dx.doi.org/10.1016/j.neuron.2005.04.005.
Zhou, H., Friedman, H.S., von der Heydt, R., 2000. Coding of border ownership in monkey visual cortex. J. Neurosci. 20, 6594–6611.
Zipser, K., Lamme, V.A.F., Schiller, P.H., 1996. Contextual modulation in primary visual cortex. J. Neurosci. 16, 7376–7389.