

Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts

Lee H. de-Wit

Laboratory of Experimental Psychology,
University of Leuven, Belgium



Jonas Kubilius

Laboratory of Experimental Psychology and Laboratory
of Biological Psychology, University of Leuven, Belgium



Johan Wagemans

Laboratory of Experimental Psychology,
University of Leuven, Belgium



Hans P. Op de Beeck

Laboratory of Biological Psychology,
University of Leuven, Belgium



Activity in the primary visual cortex reduces when certain stimuli can be perceptually organized as a unified Gestalt. This reduction could offer important insights into the nature of feedback computations within the human visual system; however, the properties of this response reduction have not yet been investigated in detail. Here we replicate this reduced V1 response, but find that the modulation in V1 (and V2) to the perceived organization of the input is not specific to the retinotopic location at which the sensory input from that stimulus is represented. Instead, we find a response modulation that is equally evident across the primary visual cortex. Thus in contradiction to some models of hierarchical predictive coding, the perception of an organized Gestalt causes a broad feedback effect that does not act specifically on the part of the retinotopic map representing the sensory input.

Keywords: predictive coding, fMRI, Gestalts, grouping, perceptual organization

Citation: de-Wit, L. H., Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision*, 12(11):12, 1–14, <http://www.journalofvision.org/content/12/11/12>, doi:10.1167/12.11.12.

Introduction

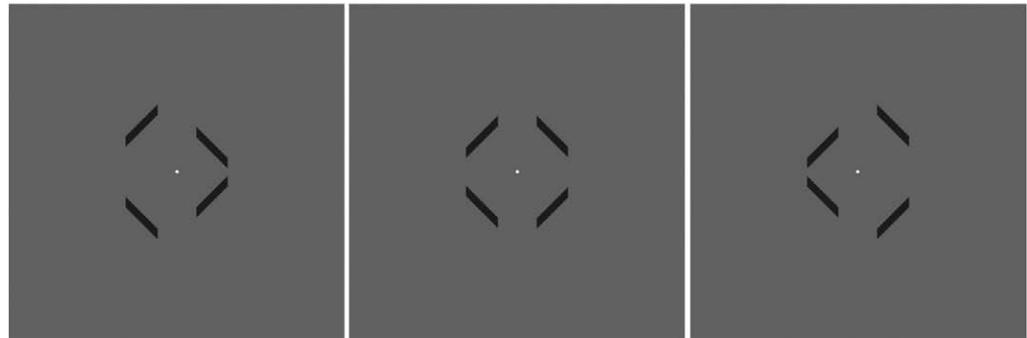
There are powerful and explicitly implementable models of the manner in which feedforward transformations along the ventral visual pathway could convert information from simple local feature detectors into progressively more complex, global and useful interpretations (e.g., Riesenhuber & Poggio, 1999). However, it is currently unclear how feedback processes within the visual hierarchy should be modeled or understood (Muckli, 2010). This lack of understanding regarding the role of feedback comes despite their potential importance. Feedback connections are not just anatomically abundant but also have a distinct pattern of connectivity and wiring that is clearly not a simple mirror reversal of the pattern of feed-forward connections (Sillito, Cudeiro & Jones, 2006). Furthermore, there is increasing evidence that early visual areas are affected by computations that are thought to be computed at later stages of the system (e.g., Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010; Fang,

Kersten, & Murray, 2008; Gilbert & Sigman, 2007; Kosslyn et al., 1999; Murray, Boyaci, & Kersten, 2006; Murray, Kersten, Olshausen, Schrater, & Woods, 2002; Roelfsema, Lamme, & Spekreijse, 1998; Sterzer, Haynes, & Rees, 2006; Williams et al., 2008). In addition, there is evidence that the feedback of neural activation from higher to lower visual areas has a direct impact upon on the contents of conscious awareness (Pascual-Leone & Walsh, 2001) consistent with certain theoretical proposals that feedback loops are critically required for conscious perception (Lamme, 2006).

Our understanding of feedback is undoubtedly hindered by the lack of precision with which its effects can be singled out. For example, while directly recorded neural differences in V1 responses in the first peak response following the physical presentation of a stimulus can probably be assumed to reflect the feedforward input from the retina via the lateral geniculate nucleus (LGN), the interpretation of neural activation after this peak response is much more ambiguous. The inability to single out the consequences of feedback is even more evident when studied via the

Local Interpretation.

In this interpretation of the stimulus the 4 lines are perceived as moving separately. From left to right the bitmaps illustrate the motion path of the four lines.

Global Interpretation.

In this interpretation of the stimulus the 4 lines are perceived as moving together as one rigid body object. The dashed lines in this figure are used to illustrate the 'Diamond' shape that results from this amodal-completion.

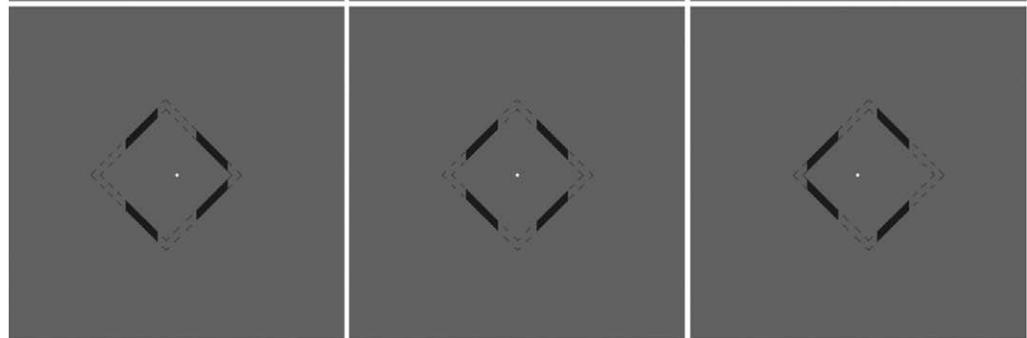


Figure 1. The stimuli from left to right illustrate the full motion profile of the Diamond Figure. The dashed lines in the bottom three Figures are used to highlight the perceptual interpretation of the amodally completed Diamond shape.

blunt time course of functional magnetic resonance imaging (fMRI). Numerous studies have attempted to circumvent these limitations by investigating the modulation of early visual processing elicited via perceptual interpretations that could not plausibly be computed at early stages of processing. Murray and colleagues provide an elegant example of this in their investigation of the primary visual systems response to a bistable stimulus (Fang et al., 2008; Murray, Kersten, Olshausen, Schrater, & Woods, 2002, Experiment 3). The bistable stimulus they used consists of four bars that are arranged in such a way that they can be interpreted as one whole (diamond) shape moving behind a set of modally completed occluders, or simply as four separate lines (Lorenceanu & Shiffrar, 1992). With the right size, luminance and motion parameters these two 'local' and 'global' interpretations of the same stimulus can be rendered equally plausible such that neither interpretation is stable, and one's perception switches from one interpretation to the other (see Figure 1). The use of these perceptual switches in deriving conclusions about feedback relies on the assumption that grouping of distinct elements into a unified Gestalt is performed in stages higher in the visual system than V1, where one finds larger receptive fields and more complex response properties. Given the profound disruption to the perception of even very basic perceptual groupings or Gestalts in patients with preserved primary visual function and higher ventral stream lesions (de-Wit, Kentridge, & Milner, 2009;

Goodale et al., 1995; James, Culham, Humphrey, Milner, & Goodale, 2003), this assumption will be assumed to be valid in this article.

Despite the evidence that the perceptual integration of sensory input into an organized Gestalt has to be computed at stages higher than V1, Murray and colleagues found a striking reduction in primary visual activity when participants switched from a local (individual lines) to a global (whole shape) interpretation of this stimulus. It should be noted that this effect may be somewhat specific to the parameters and procedure used in this study (see Caclin et al., 2012, and our Discussion). Nevertheless, if the perception of a Gestalt sometimes reduces the BOLD signal in V1, one can question what this reduction reveals about the computational role of feedback. The reduction in V1 activation is consistent with a 'hierarchical predictive coding' framework in which activations in lower visual areas partly reflect the representation of 'error' (or unexplained input) that can be 'explained away' via the feedback of predictions from higher to lower areas of the system (Friston, 2009; Rao & Ballard, 1999). Indeed, the result by Murray and colleagues is potentially related to other demonstrations of reductions in activation levels at earlier stages of processing when presented with predictable stimulus dynamics (Alink, Schwiedrzik, Kohler, Singer, & Muckli, 2010) or predictable stimulus contingencies (den Ouden, Daunizeay, Roiser, Friston, & Stephane, 2010). In this case the predictability comes from the assumption that

integrating sensory input into an organized Gestalt increases the predictability of that input and hence reduces the error represented at early stages of processing, which leads to an overall reduction in early activity.

The idea that the brain actively seeks to predict upcoming sensory input can be seen as consistent with a long tradition that regards perception as a ‘hypothesis-generating’ process (Gregory, 1997; Helmholtz, 1925). Predictive coding (Rao & Ballard, 1999), however, is not simply a computational level (Marr, 1982) explanation of what the brain is trying to achieve, but rather an algorithmic level theory about the way in which the different levels in the visual hierarchy interact to achieve this goal. We have, therefore, adopted the term ‘hierarchical predictive coding’ in order to clarify that we are talking about the way in which perception can be regarded as a process of making predictions specifically through feedback in a hierarchical system. More specifically, in this framework, feedback conveys predictions that seek to explain away potential error evoked by unexplained sensory input.

This hierarchical predictive coding model of feedback stands in contrast with another proposed role of feedback, that does not seek to reduce error, but simply to enhance sensory input that is consistent with interpretations or hypotheses that are generated at a higher level of processing (Murray, Schrater, & Kersten, 2004). Indeed, given the inherent ambiguity of sensory input, feedforward processing might lead to the generation of potential hypotheses regarding the gist of a scene and the objects that are contained in it, which generate feedback signals that simply enhance sensory input consistent with these hypotheses (Bar et al., 2006). In this framework, the visual system would rather use predictions to form a kind of positive feedback loop. A positive feedback loop would intuitively lead one to predict an increased response to predictable stimulus properties. However, it is possible that positive feedback could lead to an overall decrease in fMRI activation via an increase in the efficiency of lower level representations by suppressing noise, and thus representing the same information with less overall activity (see Kok, Jehee, & de Lange, 2012). Critical to the current study, both hierarchical predictive coding and efficient coding explanations assume that any reduction would be specific to the sensory or bottom-up representation of that input.

To summarize, given the clear anatomical potential for feedback and the lack of consensus about the role of feedback, we decided to explore in more detail the reduction in primary visual activation associated with a bistable shape configuration. More specifically, we set out to test the assumption (common to both predictive coding and efficiency explanations of feedback) that the effect should be specific to the retinotopic location at

which the stimulus was presented. Indeed this assumption was made explicit in Fang et al. (2008), who stated that “*we are confident that the modulations in the fMRI signal that we observed occurred in the retinotopic representation of the stimulus and not in immediately adjacent retinotopic regions*” (p. 6). In contrast, when we investigated this directly it became evident that, although there was a clear reduction in activation at the retinotopic location of the stimuli, this reduction did not appear to be specific to, or even selectively stronger in, the retinotopic location of the stimulus. As will be discussed further, this result is consistent with other demonstrations that feedback effects are observable at retinotopic locations other than those where the stimuli are classically represented (Ester, Serences, & Awh, 2009; Williams et al., 2008), but would not have been directly expected if the effect of feedback on V1 would be to explain away error signals or enhance early sensory representations.

Methods

Participants

Fifteen participants completed a pilot behavioral experiment to assess the average duration of stable perceptual periods while observing the bistable diamond stimulus. None of these participants had experience with this stimulus prior to testing in this experiment. On the basis of this pilot, two participants were excluded from completing the fMRI experiment, because the average duration of their stable percepts was less than 4 seconds. Functional and high-resolution anatomical MRI images were therefore obtained from 13 participants (ages: 21–37). For the first fMRI participant we did not include a baseline measure of fixation activation to which the data analysis for all other participants is normalized, therefore this participant’s data is not included in the results described below. All participants reported normal or corrected-to-normal vision. Participants were provided with financial compensation and written informed consent was obtained. The study was approved by the ethical committee of the Faculty of Psychology and Educational Sciences and the committee for medical ethics at the University of Leuven.

Stimuli and procedure

All 13 participants completed three types of fMRI runs: Diamond, Stimulus-Specific-Checkerboard, and Large-Square-Wave-Grating, which will be described below. Moreover, all the participants also completed

Meridian-Mapping and Object-Versus-Scrambled-Object runs and a T1 weighted structural run (these data were available from a previous experiment for six participants, see Kubilius et al., 2011).

In the Diamond runs participants were presented with 16 seconds of fixation at the start of the run followed by 403 seconds of the bistable moving diamond / lines stimulus followed by a further 13 seconds of fixation. Participants had to fixate on a small central dot during the entire run. The Diamond stimulus consisted of four oriented lines that moved such that they could be interpreted either as four lines moving separately (local condition) or as one solid ('diamond') shape moving horizontally behind three invisible occluding bars (global condition; Figure 1). The visible line segments were 2.6° in length and 0.45° in width, they moved horizontally at $1.3^\circ/\text{s}$ and switch direction every second. When symmetric around fixation the center of each line was at located at 2.8° eccentricity. While viewing the stimulus, participants used two buttons to indicate whether they perceived the stimulus as separate lines or as one shape. Participants were only required to press one of the buttons when their percept switched from one interpretation to another. However, they were also allowed to press again at any point if they had, for example, lost track and wanted to make sure that the correct percept was registered. Repeat presses that did not involve a switch in percept were then ignored. Participants completed either four or five repetitions of this Diamond run.

The Meridian-Mapping runs adopted a standard procedure from Tootell et al. (1995) in which horizontal and vertical checker board wedges (of 15°) were presented to define the boundaries between primary visual areas (as used in Kubilius et al., 2011). Each run contained 8 seconds of fixation at the start and end with 10 blocks of vertical or horizontal wedges, presented for 16 seconds, in a counterbalanced order across runs. During their presentation, three parameters of the vertical or horizontal wedges changed: (a) the number and size of squares across the width of the wedge (from 20 to 40 cycles), (b) the number and size of squares across the length of the wedge (between 4 and 18 cycles), and (c) the color (see Figure 2).

In the Stimulus-Specific-Checkerboard runs participants were presented with 16 second blocks with different pairs of stationary flickering checkerboard in which the phase randomly shifted every 100 ms. There were three conditions (Figure 2b): a fixation condition, where only a fixation dot was present, pairs of checkerboards corresponding to the upper left and lower right lines of the Diamond stimulus, and pairs of checkerboard corresponding to the lower left and upper right lines of the Diamond stimulus. Because the Diamond stimulus was moving, the checkerboards

covered the entire region where the lines of the Diamond appeared. Participants had to fixate on a small central dot throughout the run. Blocks of conditions were presented in a counterbalanced order (Fix-A-B-A-B-Fix-B-A-B-A-...) and alternate runs would begin either A-B or B-A. These runs were 400 seconds long.

Participants were also presented with six Background ROI (Large-Square-Wave-Grating) runs of 160 seconds. These runs started and ended with 16 seconds of fixation and alternated between left and right oriented square wave gratings with an 8.5° radius but with a 1.7° gap around fixation. The wedges of this grating were 0.45° in width. The phase of the grating changed randomly every 100 ms (Figure 2).

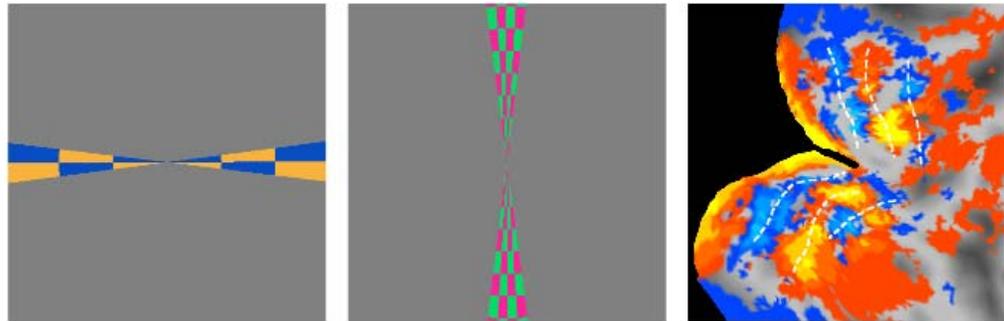
The Object-Versus-Scrambled-Object runs adopted a standard Lateral occipital complex (LOC) localizing procedure (Grill-Spector et al., 1998; as used in Kubilius et al., 2011). There were two conditions: intact objects and scrambled objects. The set of 20 intact objects was retrieved from <http://www.imageafter.com> and <http://www.morguefile.com> and consisted of images of human-made objects, foods, and plants. All images were full color and 256×256 pixels in size. Each image featured a prominent object presented on a simple background that was not uniform to ensure that images of both intact objects and scrambled objects subtended the same area. The set of 20 scrambled objects was created from the 20 images of intact objects by dividing each image into 256 tiles (16×16 pixels) and shuffling them randomly within the image. Intact and scrambled images were presented at four locations around fixation and participants performed an 'odd one out' detection task. Stimuli were presented approximately 7° away from a central fixation dot and subtended 5° of visual angle. Each trial consisted of a 300-ms stimulus presentation (with a fixation dot present) followed by a 500-ms interstimulus interval with only a fixation dot. There were eight blocks of 20 intact images each, eight blocks of 20 scrambled images each, and five fixation blocks. Blocks were counterbalanced and lasted for 16 seconds each. Participants were asked to press a key when they spotted an immediately repeated image (a one-back task). Images were repeated between two and four times per block.

Scanning parameters

The experiment was carried out using a 3T Phillips Intera scanner (Phillips, Best, The Netherlands) with an eight-channel SENSE head coil with an echo-planar imaging sequence. We recorded from 18 slices during the Diamond, Stimulus-Specific-Checkerboard and Large-Square-Wave-Grating runs oriented downwards

Meridian Mapping.

Contrasting the horizontal and vertical meridians reveals the borders between V1, V2 and V3. This procedure was derived from Tootell et al. (1995)



Sensory-Specific ROI.

Two pairs of flickering checker boards were contrasted to identify the retinotopic regions in which the lines of the bistable diamond stimulus (see figure 1) would be represented.



Background ROI.

A large flickering square wave grating was used to define a larger ROI, from which voxels active in the Sensory-Specific ROI were removed in order to construct a Background ROI.

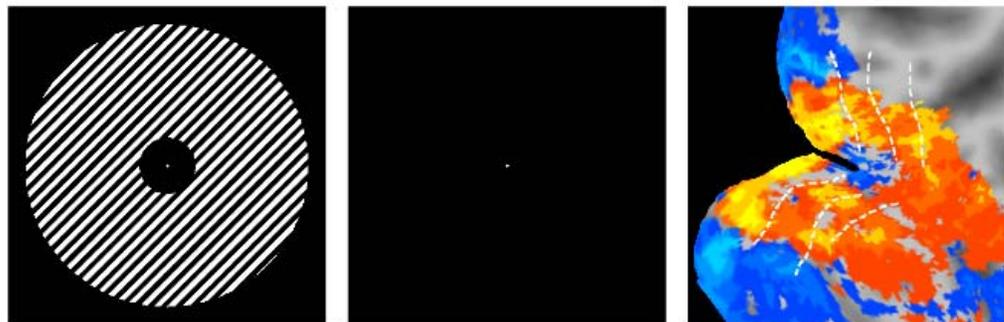


Figure 2. Illustrating the stimuli used to select the regions of interest. The first row shows the two retinotopic mapping meridians, and the way in which the contrast is used to define the boundaries between the primary visual areas (Tootell et al. 1995). The second row shows the two Stimulus-Specific-Checker Board stimuli, and the resulting activation in the contrast between them (threshold at $t = 3.7$), overlaid on the retinotopic boundaries. The third row presents the Large Square Wave Grating, and the resulting activation in the contrast to fixation ($t = 3.7$).

for a full occipital and inferotemporal cortex coverage with a voxel size $2.75 \times 2.75 \times 2.75$ mm and interslice distance of 0.2 mm (acquisition matrix 80×80). The Diamond, Checkerboard and Large Square Wave Grating Runs consisted of 432, 400 and 160 measurements respectively with a TR of 1000 ms, flip angle 60° , and an echo time of 30 ms. This short TR was used to maximize the timing precision of participants perceptual state when viewing the Diamond stimulus (which switched on average every 8 seconds). The Meridian-Mapping and Object-Versus-Scrambled-Object scans were acquired with 37 slices and consisted of 168 measurements with a TR of 2000 ms, flip angle 90° , and an echo time of 30 ms.

The T1-weighted anatomical scan had $.85 \times .98$ mm in plane resolution and 1.37 mm between the slices

(acquisition matrix 256×256), 9.6 ms TR, 4.6 ms TE, 182 coronal slices, and lasted for 383 s.

fMRI analysis

The data were preprocessed using the Statistical Parametric Map package (Version 8; Wellcome Trust Centre for Neuroimaging, London, England). A standard preprocessing procedure was followed to align, coregister, normalize and smooth the data. Regions of interest were identified on the basis of smoothed data (with a 5.5 mm full-width-half-maximum Gaussian kernel). The unsmoothed, normalized data from the Diamond runs were extracted directly from each run after a linear and quadratic trend correction. The ‘local’ and ‘global’ event-related time

courses were calculated on the basis of the button presses of the participant. TRs that contained a switch were assigned to the ‘new’ percept only when the change occurred in the first half (500 ms) of that TR, otherwise the new percept was taken to start from the start of the next TR. Resulting perceptual durations of only 1 TR (around 3% of all percepts, see Figure 6), were excluded from the final analysis. For each ‘event’ (nonrepeat button press) we extracted a time course of which the length was equal in time to the duration of the percept. As a result, the early time points in the shown time courses are based upon more trials than the later time points. All time courses from the diamond runs are expressed in percent signal change versus fixation.

Identifying regions of interest

Regions of interest (ROIs) were defined in CARET 5 (Van Essen et al., 2001) on a flat surface of the brain (Figure 2, right column). First, V1-V2, V2-V3, and V3-higher regions’ borders were defined using a contrast horizontal > vertical wedge from Meridian-Mapping runs. Using this border information, Stimulus-Specific ROIs for the Diamond in V1, V2, and V3 were selected based on activation (min threshold, $t| = 3.7$, uncorrected $p < 0.0001$) elicited in the contrast between the two checker board locations. ROIs were identified in the left and right hemisphere and dorsally and ventrally from the calcarine sulcus corresponding to the retinotopic locations of the four lines composing the Diamond, leading to four retinotopic ROIs for each visual area V1, V2, and V3.

A larger Background ROI was selected on the basis of the activation elicited by the Large-Square-Wave-Grating runs (with a threshold of $t = 3.7$ in contrast to fixation). Any voxels in this Background ROI that were also present in the Stimulus-Specific ROI were removed. In this way one could ensure that none of the voxels in the Background ROI were taken from areas of retinotopic cortex in which the lines of the Diamond stimulus were presented. In order to be absolutely certain that there was no ‘spillover’ of signal from the Stimulus-Specific ROI to the Background, a more Conservative Background was also constructed by artificially inflating the size of the Stimulus-Specific ROI by one voxel in all three dimensions around the border of this ROI, and then subtracting this enlarged Stimulus-Specific ROI from the Background ROI to give a more Conservative Background.

However, because the stimulus-specific ROI was already relatively large the resulting Background ROI was sometimes very small (Figure 2 displays a subject for which this problem was not present). This was

especially true for the higher areas (V2/V3), which is not surprising given the increase in receptive field size. If, after removing voxels also present in the Stimulus-Specific ROI, the resulting Background ROI was both half the size of the mean ROI size for that area across participants and was below one standard deviation of the mean for that visual area, then the ROIs for that visual area for that participant were excluded. This procedure resulted in the exclusion of one participant for V1, two participants for V2 and three participants for V3. This procedure excluded one participant’s data in all three areas, so this participant’s data was also removed from the lateral occipital complex (LOC) analysis (although the same result for LOC was found when this participant’s data was included).

For the shape-selective lateral occipital (LO) and posterior fusiform (pFs) cortex, we used the objects > scrambled objects contrast (Grill-Spector et al., 1998). LO was defined as a lateral shape-selective region, while posterior Fusiform sulcus (pFs) was defined on the basis of shape-selective activation on the ventral surface. These two ROIs were also analyzed together as the LOC.

Results

We aimed to replicate the effect of the perceptual interpretation of the bistable stimulus in the Stimulus-Specific ROI in V1, and then to investigate to what extent this same effect occurs in the Background ROIs. To this end the fMRI signal while observing this stimulus was computed as a function of the participant’s perceptual switches (recorded via a button press). The change in BOLD signal, just before and after a button, is illustrated in Figure 3. The change is illustrated for the Stimulus-Specific, Background and the Conservative-Background ROIs in V1, V2, and V3. The solid line (reflecting a switch to the global stimulus interpretation) clearly goes down after a switch in V1 and V2, whereas the dashed line (reflecting a switch to the local interpretation of the bistable stimulus) shows an increase in activity after a switch.

In order to test for the sensitivity to the change in percept in each area a three-factor repeated-measures ANOVA (with 16 Time points, two Grouping levels: Local vs. Global, and three ROIs: Specific vs. Non-Specific) was used. First focusing on V1, there was an interaction between Time and Grouping, $F(15, 150) = 4.21$, $p < 0.011$, replicating the previous finding that the BOLD signal over time is influenced by the subjective perception of the participant. The same three-factor repeated-measures ANOVA also revealed an effect of Grouping over Time, $F(30, 270) = 3.6$, $p < 0.015$, in V2 but not in V3, $F(15, 120) = 1.04$, $p = 0.4$.

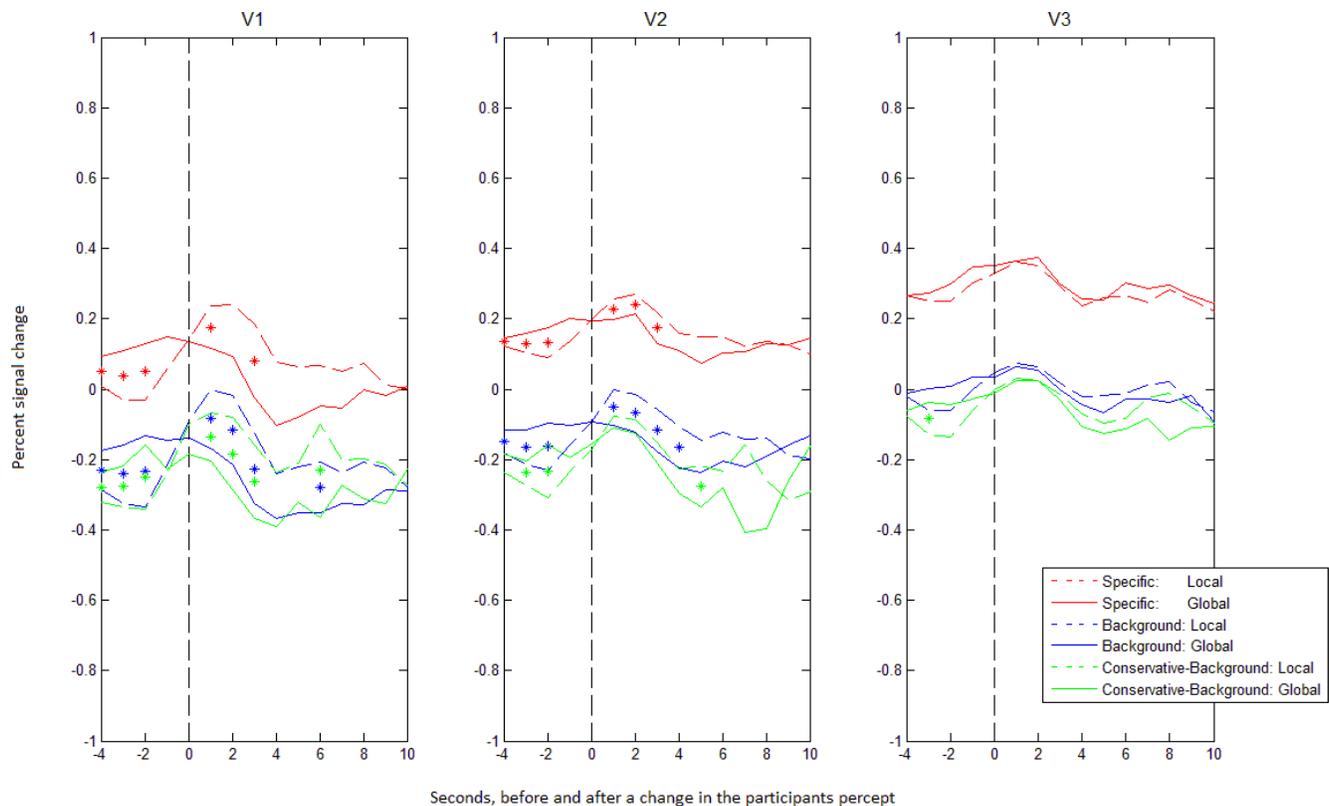


Figure 3. The change in BOLD signal is plotted relative to the participants button press (indicated by a vertical dashed line). For V1, V2, and V3 the response is plotted for the Stimulus-Specific, Background and the Conservative-Background ROIs. The solid line plots a switch from a local to a global stimulus interpretation, and the dashed line plots a switch from the global to the local. Significant differences ($p < 0.05$, two tailed, uncorrected paired sample t test) between the responses to the local and global percepts in each ROI are illustrated with a star.

Surprisingly, however, none of these regions show a difference (interaction) between the change in percept over time and the distinct ROIs in each region, V1: $F = 0.99$; V2: $F(30, 270) = 1.4$, $p = .26$; V3: $F(30, 240) = 0.375$, $p = 0.37$. This suggests that the influence of the perceptual interpretation of the stimulus in V1 and V2 is not restricted to the retinotopic location of the stimulus. Although the interaction term in the three-factor ANOVA is the appropriate means with which to test for a differential sensitivity to the subjective interpretation of the percept, we also completed individual t tests (uncorrected for multiple comparisons) comparing the size of the difference between the two perceptual interpretations for each ROI in V1 at every time point. This post-hoc analysis did not reveal a single time point at which the size of the effect was larger in the Stimulus-Specific ROI. The only significant difference was for a stronger difference in the Background ROI at time point 1, but this was only marginal ($p = 0.046$) and would not hold up when correcting for multiple comparisons. The same paired t tests in V2 and V3 also revealed no evidence of a differential sensitivity to the strength of the perceptual modulation of the BOLD signal.

Given the clear sensitivity to the grouping of the percept in the Background ROIs, it was important to confirm that the selection of Stimulus-Specific, Background and the Conservative-Background ROIs had been defined correctly as distinct areas of the retinotopic map. The reduced overall activation level relative to fixation for the Background ROI is already a potential indication that distinct ROIs had been identified (significant in V1, $F(2, 20) = 26.7$, $p < 0.001$, and V2, $F(2, 18) = 4.6$, $p = 0.042$, but not in V3, $F(2, 16) = 4$, $p = 0.36$, because this is consistent with previous demonstrations of suppressed BOLD responses in retinotopic areas adjacent to the presented stimuli in V1) (Bressler, Spotswood, & Whitney, 2007). However, in order to demonstrate more explicitly that the ROIs were chosen correctly, we contrasted the first 16 seconds of spatially smoothed (with a 5.5 mm full-width-half-maximum Gaussian kernel) BOLD signal change following the onset of the Diamond stimulus in each run with the preceding 16 seconds of fixation (plotted in Figure 4).

This contrast enables us to test whether each ROI responds to the sensory onset of the stimulus. In V1 this contrast leads to a clear response to the Diamond onset

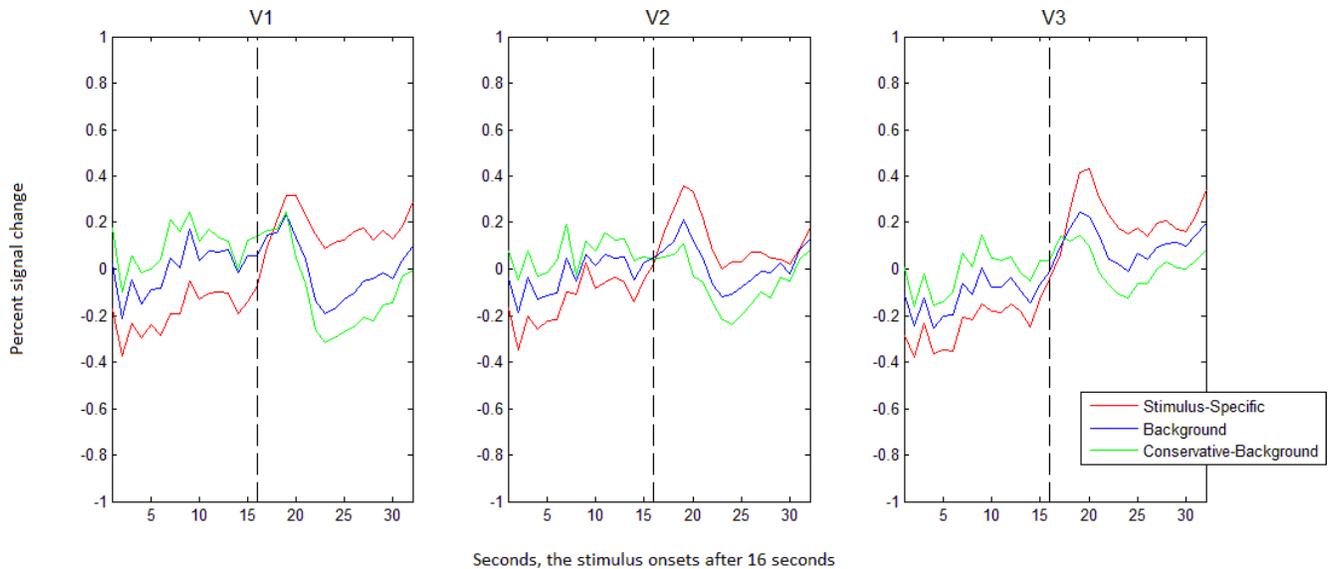


Figure 4. Plotting the time course of activation at the start of each run. Each run begins with 16 seconds of fixation before the stimulus appears, indicated by the dashed vertical line.

in the Stimulus-Specific ROI, $t(10) = 3.6$, $p = 0.0048$, and no response to the onset of the Diamond per se in the Background ROI ($p = 0.96$) and if anything a reverse trend in the more Conservative-Background ROI, $t(10) = 1.9$, $p = 0.089$. For V2 we found a clear stimulus onset response in the Specific ROI, $t(9) = 2.4$, $p = 0.039$, and no response in the Background ROI ($p = 0.57$) or the more Conservative-Background ($p = 0.83$). In V3, there was a response to the onset in the Sensory-Specific ROI, $t(8) = 3.6$, $p = 0.0073$, there was, however, a response in the Background ROI, $t(8) = 2.6$, $p = 0.03$, but not the more Conservative-Background ROI ($p = 0.83$). The fact that the difference between the specific and Non-specific ROIs was clearer in V1/V2 than V3 might be related to smaller receptive field sizes in V1/V2 and a clearer retinotopic organization in V1/V2 in general (Dumoulin & Wandell, 2008). The fact that the Background ROI in V3 responds to the onset, but that the more Conservative Background ROI does not, offers, in itself, something of a sanity check for our definitions of these ROIs, because the more Conservative Background should by definition be less susceptible to spill over from the Sensory ROI. Critically, this analysis clarifies that while the Background ROI shows the same sensitivity to the perceptual interpretation of the stimulus, it does not respond to the onset of the stimulus per se in V1, V2, and the Conservative-Background responds in neither V1, V2, nor V3.

Fang et al. (2008) extended Murray et al.'s (2002) results by demonstrating that in addition to the clear reduction of activation in V1, there was also a reversal of this pattern in a region they labeled as LOC. We did not find a clear increase in the whole of LOC (LO and pFs, see Methods, $F(15, 150) = 1.23$, $p = 0.236$, but

upon discussing this discrepancy with the original authors (F. Fang, personal communication, August 17, 2012) it appears plausible that the region they labeled LOC in fact only contained what we have defined as the more posterior shape selective region LO. The increase in LO is not as large as that reported by Fang et al. but it is still quite robust, $F(15, 150) = 3.96$, $p < 0.001$). It is also likely that the more anterior shape selective region on the ventral surface that we have labeled pFs corresponds to what Fang et al. labeled temporal object areas (TOA). Although there is a similar pattern of results in pFs to that found in LO (see Figure 5) these are not significant, $F(15, 150) = 0.74$, $p = 0.7$, a pattern broadly consistent with the results of Fang et al.

Looking just at the participants' behavioral data, the frequency of each perceptual interpretation for the bistable figure reveals a slight difference in the distributions for each perceptual interpretation (plotted in Figure 6). In order to avoid any potential confounding influence this might have, the analysis of the fMRI data was also repeated using perceptual durations from one perceptual interpretation only when they could be matched with a duration of equal length of the other interpretation. The resulting distribution of 'Equalized Frequencies' is also plotted in Figure 6. This procedure tended to reduce the number of perceptual durations at the extreme of the distribution; this is clearest in the absence of any durations lasting only 1 second in the Equalized Frequencies (although as noted in the Methods, durations of only 1 second were also not included in the main analysis reported above). The analysis of the fMRI data set using only this set of balanced durations

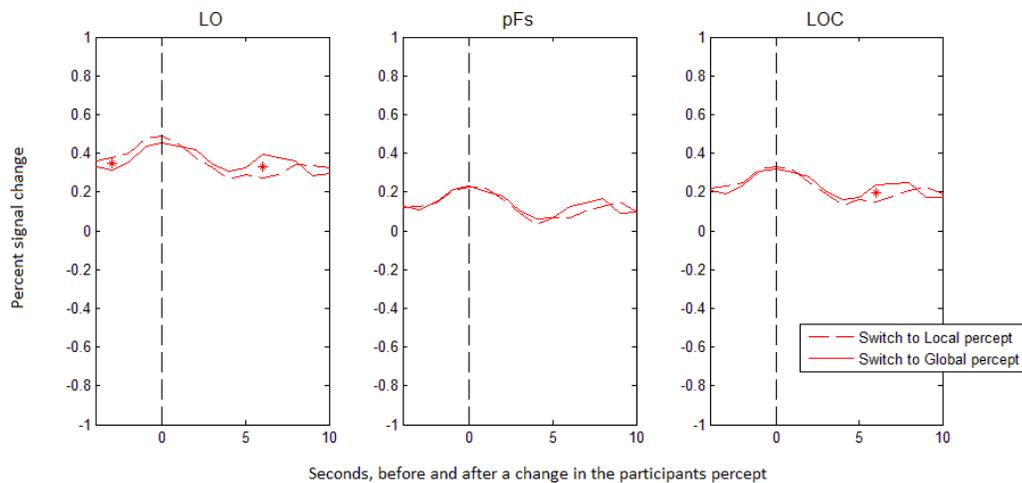


Figure 5. The change in BOLD signal is plotted relative to the participants button press separately for LO and pFs, and in the combined region 'LOC.' The stars indicate significant differences between the local and global percepts using an uncorrected paired samples *t* test.

revealed the same findings as those reported above, highlighting that the slight difference in the durations of each percept do not account for the BOLD signal differences between these interpretations.

Discussion

One of the biggest open questions in visual neuroscience centers on the functional role of the abundant feedback connections to early visual areas. Most models of feedback assume that it acts especially on the bottom-up sensory input. The hierarchical predictive coding model in particular assumes that feedback is directed to the sensory representation of incoming input to provide a comparison between top down predictions and bottom up signals. This model has previously been motivated based on data showing that organized (and thus predictable) percepts lead to reduced activation levels on the primary visual cortex (Fang et al., 2008; Murray et al., 2002). The results of the current study replicate this reduction, but find that this reduction is not specific to the sensory representation of this stimulus; rather, it also occurs in retinotopic regions that do not respond to the sensory onset of the stimulus. Importantly however this reduction is not found in all visual areas (which might suggest a general arousal effect); rather, our results also replicate an increase in activation for the organized percept in the higher-level shape selective region LO (Fang et al., 2008).

The studies by Murray and colleagues (2002, Experiment 3; Fang et al., 2008) upon which the current investigation was based, were particularly significant in guiding models of feedback dynamics because they documented a modulation in primary

visual activation that clearly could not be explained by changes in the bottom-up stimulus to the system. The reduced activation in V1 following the perceptual organization of the bistable input has been cited as potential evidence that when sensory input can be interpreted at a higher stage in the visual hierarchy, this interpretation can be fed back as a prediction that leads to a reduction in activation either by canceling out already predicted input or by reducing the hypothesized error induced by the representation of this input. Just based on average fMRI data alone, however, reductions in signal do not provide conclusive evidence for hierarchical predictive coding, rather this coarse reduction in activation at the level of voxels could be consistent with an efficient coding model in which feedback sharpens the underlying representation and inhibits noise (Kok et al., 2012; Murray et al., 2004). The current results, however, are hard to reconcile with either model of feedback effects, as both would predict an effect that would be specific to the incoming sensory stimulation that would be the target of such feedback.

This nonsensory specific reduction in V1 activation appears more consistent with the notion of a flexible reverse hierarchy (Hochstein & Ahissar, 2002) in which the system flexibly allocates resources so as to preferentially read out information from the most appropriate level of the system for representing the stimulus. Indeed this interpretation is consistent with behavioral data, both for this diamond illusion (He, Kersten, & Fang, 2012) and for other stimuli (Poljac, de-Wit, & Wagemans, 2012; Sayim, Westheimer, & Herzog, 2012) that the extent to which visual input can be grouped into a higher level Gestalt leads to a reduction in the lower level sensory representation of the 'parts' of that Gestalt.

This idea that the visual system might preferentially allocate processing resources to the most appropriate

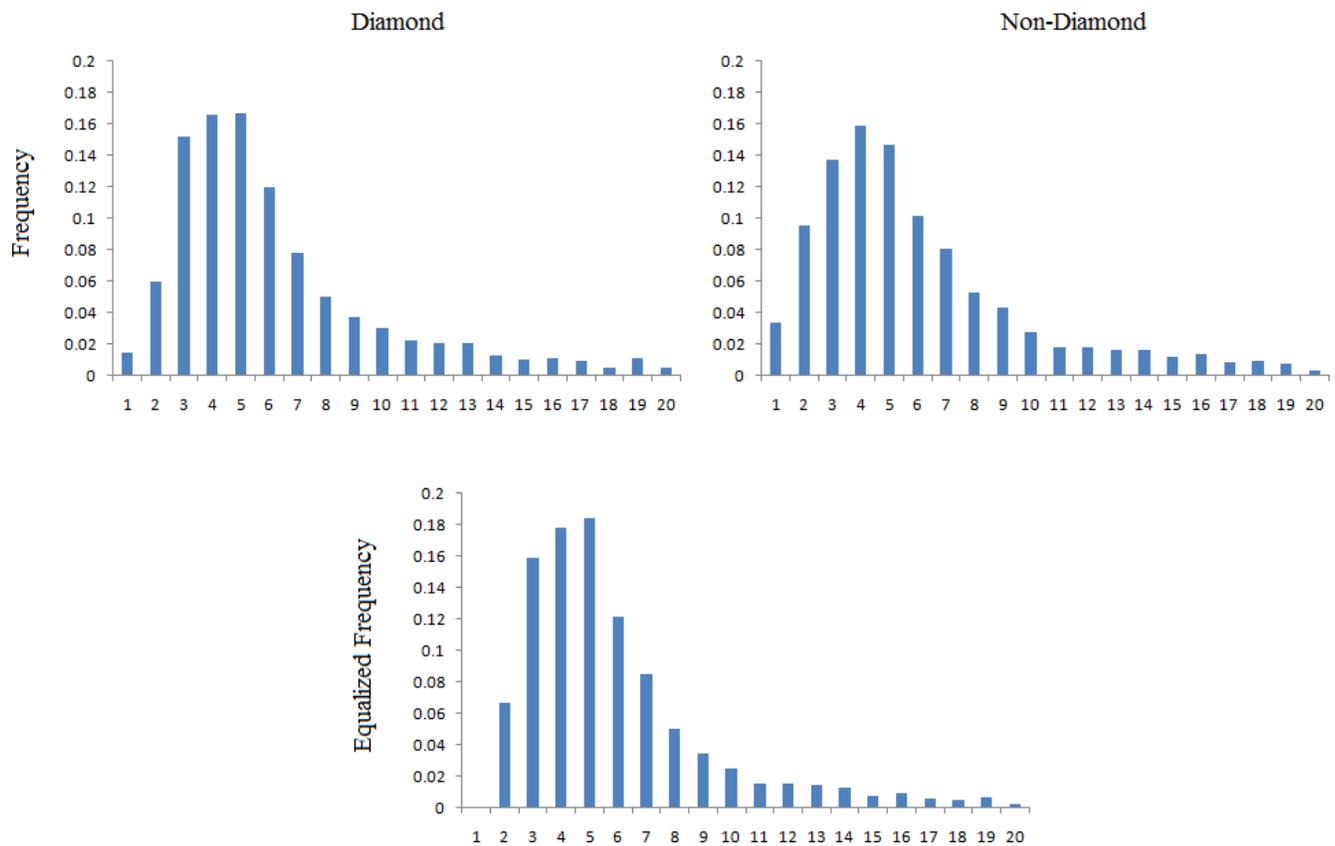


Figure 6. The top two graphs illustrate the frequency of the different perceptual durations for the two percepts; the graph below shows the distribution resulting when only durations of equal length were selected from each condition. This analysis tended to reduce the number of perceptual durations at the extremes of the distribution.

stage of processing for representing this stimulus could be regarded as a form of attentional selection to different stages in the visual hierarchy. Attention is, however, a potentially problematic term that can be used in the context of any modulation in activation levels without necessarily providing any insight into the underlying mechanism. It is therefore pertinent to note that the change in BOLD signal reported here is different from classical observations of selective spatial attention in two important ways. First attention to a given stimulus appears to be generally as spatially specific as the retinotopic representation of that stimulus (Brefczynski & DeYoe, 1999; Saygin & Sereno, 2008; Tootell et al., 1998). Indeed, attending to a stimulus is an established means of improving the procedure for defining retinotopic maps (Bressler & Silver, 2010), which could obviously not be the case if the change in activation was as homogeneous as that observed in the current work. The second distinction is that attentional effects are typically found to be larger at higher stages of processing (Moran & Desimone, 1985; Saygin & Sereno, 2008; Tootell et al., 1998), in contrast to the progressive reduction in the size of the modulation seen here (and in Fang et al., 2008; Murray et al., 2002) such that there was a reduced effect in V2

and no effect in V3. Thus while the current results could be described as a form of attentional selection to a given stage in the processing hierarchy, it should be clear the nature of this modulation is very different from that observed in more classical explains of selective spatial attention, and is unlikely to involve a common mechanism.

The idea that the nonstimulus specific reduction in BOLD signal reflects some form of flexible resource allocation to different levels of the visual hierarchy is, however, just a speculative possibility for explaining these data. The more solid conclusion to draw from these results is that they reject the hypothesis drawn from hierarchical models of predictive coding. This result therefore provides an important challenge for an increasingly popular approach for modeling cortical dynamics in the hierarchy of the visual system (Friston, 2012; Kersten, Mamassian, & Yuille, 2004; Mumford, 1992; Rao & Ballard, 1999). Part of the appeal of the predictive coding framework, however, is not just the role it plays in the visual system, but how it serves as a more general framework for thinking about how the human mind works, or what it is trying to achieve (Clark, 2012). Our results do not necessarily question this broader notion that the brain attempts to predict

upcoming sensory input, rather it challenges a specific implementation of hierarchical predictive coding in which specific predictions about sensory input are fed back to earlier stages of processing. Even within the domain of perception, there are other potential implementations of predictive coding, for example neurons at one level of the system may attempt to predict the next stimulus to appear in a sequence over time (Perrett, Xiao, Barraclough, Keysers, & Oram, 2008). Thus the current results do not speak to the idea that the brain might be attempting to predict upcoming input in general, rather they question one source of evidence that these predictions are implemented via stimulus specific feedback signals within the visual hierarchy.

The demonstration of ‘nonsensory’ modulation of V1 in the current study is at one level consistent with numerous observations of feedback influencing V1, either in the absence of sensory input (Ester, Serences, & Awh, 2009; Harrison & Tong, 2009; Kosslyn et al., 1999) or at retinotopic locations other than the actual presentation of the stimulus (Williams et al., 2008). What seems to set all of these observations apart from the results here, however, is that these previous studies all seem to involve some active use of the activity elicited in V1, either in terms of imagining stimuli or holding stimuli in working memory. Indeed Williams et al. only found patterns of activation in nonstimulated parts of the retinotopic map when participants had to actively perform a demanding task with the shapes presented more peripherally.

In the study of Ester, Serences, and Awh (2009) the authors explore more carefully the nature of the more widespread feedback effects to the primary visual cortex, by testing whether stimuli presented to one hemifield are also maintained in patterns of activation in the other hemifield during a working memory task. Such a manipulation could be informative with regards to the current study in order to more fully understand the nature of the feedback process at work. Future work could also provide further insight into the nature of this V1 suppression by looking at GABA concentrations, which are both important in inhibiting signals (Sumner, Edden, Bompas, Evans, & Singh, 2010) and appear to play an important role in other examples of bistable perceptual interactions (van Loon et al., 2012).

Another open question in our understanding of the modulation in V1 (and V2) regards its source. Fang et al. argued that the opposite sensitivity of LO to this bistable stimulus made LO a plausible source of this feedback, consistent with this areas role in perceptual organization more generally. The opposite signal change in LO does not however prove that this region is responsible for the modulation in V1. To test that more causal methods (such as transcranial magnetic stimulation [TMS]) would be needed to rule out a

potential role of other areas, such as areas important for grouping moving stimuli (Shipp & Zeki, 1989; Sterzer, Haynes, & Rees, 2006), or dorsal areas in the parietal lobe that also represent simple shapes (Lehky & Sereno, 2007; Xu & Chun, 2007), and indeed complex objects (Konen & Kastner, 2008).

The role of visual areas involved in motion processing is also highlighted in a recent study using a variation of the same bistable Diamond stimulus (Caclin et al., 2012), which was published after we had carried out our study. Importantly this study fails to find a decrease in primary visual activation when switching to an organized (bound) stimulus, in fact they find an increase in activation. As highlighted by Caclin et al., however, there are a number of differences between their study and that of Fang et al. (and therefore our replication). First there are differences in the stimulus used, not only in terms of its color but also its motion path. In addition there are differences in the procedure used, in that in the present study (and that of Fang et al.) only spontaneous transitions in the interpretation of the sensory input were presented, whereas Caclin et al. also studied cases in which the change in percept was induced by changes to the stimulus. Although they did not find an effect of induction type (spontaneous or induced) per se, it may be that prior exposure to induced changes could change the manner in which participants interpret the stimulus. It is perhaps not surprising that our study replicates the reduction found in Murray and colleagues’ previous work, given that we explicitly set out to replicate the parameters used in the study of Fang et al. Clearly, future work will be required to pin down exactly what conditions are required for cases of perceptual grouping to lead to reductions in primary visual activation.

In summary, this work tests a key assumption regarding the retinotopic specificity of feedback to V1. This retinotopic specificity was an important assumption, not only in models of hierarchical predictive coding, but also rival models that explained reductions in activation as a fine tuning of information via an enhancement of representations consistent with top-down signals. The nonstimulus specific nature of the effect reported here provides an important challenge to these models. The result is still consistent with behavioral observations that the representation of ‘early’ stimulus features seem to be reduced when they can be organized into a Gestalt, which could, in turn, be indicative of some form of selective resource allocation to different levels in the visual hierarchy. The current results, however, do not enable us to fully characterize the nature of the more global modulation of V1. Furthermore while they do find an opposite relationship to the perceptual grouping in LO, more causal work would be needed to test the exact role of this area. It is important to note that while these results

call into question a particular implementation of hierarchical predictive coding, they do not necessarily undermine the more general idea that the brain could be optimally understood as a predictive coding system. Finally, future work will also be required to pin down why certain stimulus parameters lead to reductions in primary visual activation when perceptual input is organized into a Gestalt and others do not (Caclin et al., 2012).

Acknowledgments

This work was supported by a Methusalem Grant (METH/08/02) awarded to Johan Wagemans from the Flemish Government. L. D. is a postdoctoral fellow, and J. K. a research assistant of the Research Foundation – Flanders (FWO). This research was carried out using a number of free and open access software tools, including SPM, Caret, PsychoPy (Python) and Scribus. We also thank Ronald Peeters for technical assistance with the fMRI scanning, and Bart Cooreman and Eveline Smolders for their help in reconstructing a number of the cortical flat maps in Caret.

Commercial relationships: none.

Corresponding author: L. H. de-Wit.

Email: lee.dewit@ppw.kuleuven.be.

Address: Laboratory of Experimental Psychology, University of Leuven, Belgium.

References

- Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *Journal of Neuroscience*, *30*(8), 2960–2966.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., & Coauthors, (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(2), 449–454.
- Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the ‘spotlight’ of visual attention. *Nature neuroscience*, *2*(4), 370–374, doi:10.1038/7280.
- Bressler, D. W., & Silver, M. A. (2010). Spatial attention improves reliability of fMRI retinotopic mapping signals in occipital and parietal cortex. *NeuroImage*, *53*(2), 526–533, doi:10.1016/j.neuroimage.2010.06.063.
- Bressler, D., Spotswood, N., & Whitney, D. (2007). Negative BOLD fMRI response in the visual cortex carries precise stimulus-specific information. *PLoS ONE*, *2*(5), e410.
- Caclin, A., Paradis, A.-L., Lamirel, C., Thirion, B., Artiges, E., Poline, J.-B., & Coauthors, (2012). Perceptual alternations between unbound moving contours and bound shape motion engage a ventral/dorsal interplay. *Journal of Vision*, *12*(7):11, 1–24, <http://www.journalofvision.org/content/12/7/11.abstract>, doi:10.1167/12.7.11. [PubMed] [Article]
- Clark, A. (in press). Whatever next? Predictive brains, situated agents and the future of cognitive science. *Behavioral and Brain Sciences*.
- den Ouden, H. E. M., Daunizeau, J., Roiser, J., Friston, K. J., & Stephan, K. E. (2010). Striatal prediction error modulates cortical coupling. *Journal of Neuroscience*, *30*(9), 3210–3219.
- de-Wit, L. H., Kentridge, R. W., & Milner, A. D. (2009). Shape processing area LO and illusory contours. *Perception*, *38*(8), 1260–1263.
- Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage*, *39*(2), 647–660.
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *The Journal of Neuroscience*, *29*(48), 15258–15265.
- Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7):2, 1–9, <http://www.journalofvision.org/content/8/7/2>, doi:10.1167/8.7.2. [PubMed] [Article]
- Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301.
- Friston, K. J. (2012). What does functional MRI measure? Two complementary perspectives. *Trends in Cognitive Sciences*, *16*(10), 491–492, doi:10.1016/j.tics.2012.08.005.
- Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, *54*(5), 677–696.
- Goodale, M. A., Jakobson, L. S., Milner, A. D., Perrett, D. I., Benson, P. J., & Hietanen, J. K. (1994). The nature and limits of orientation and pattern processing supporting visuomotor control in a visual form agnostic. *Journal of Cognitive Neuroscience*, *6*(1), 46–56.
- Gregory, R. L. (1997). *Eye and Brain: The Psychology*

- of Seeing (1966). London: Weidenfeld and Nicolson.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzhak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, 6(4), 316–328, doi:10.1002/(SICI)1097-0193(1998)6:4.
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458(7238), 632–635.
- He, D., Kersten, D., & Fang, F. (2012). Opposite modulation of high- and low-level visual after-effects by perceptual grouping. *Current Biology*, 22(11), 1040–1045.
- Helmholtz, H., von. (1925). *Treatise on physiological optics* (J.P.C. Southall, Ed.). Washington, DC: Optical Society of America. [Original work published 1910].
- Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36(5), 791–804.
- James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: An fMRI study. *Brain*, 126(11), 2463–2475.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, 55, 271–304, doi:10.1146/annurev.psych.55.090902.142005.
- Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less is more: Expectation sharpens representations in the primary visual cortex. *Neuron*, 75(2), 265–270.
- Konen, C. S., & Kastner, S. (2008). Two hierarchically organized neural systems for object information in human visual cortex. *Nature Neuroscience*, 11(2), 224–231.
- Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J. P., & Thompson, L. W., et al. (1999). The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science*, 284(5411), 167–170.
- Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2011). Emergence of perceptual gestalts in the human visual cortex: the case of the configural-superiority effect. *Psychological Science*, 22(10), 1296–1303, doi:10.1177/0956797611417000.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences*, 10(11), 494–501.
- Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology*, 97(1), 307–319, doi:10.1152/jn.00168.2006.
- Lorenceau, J., & Shiffrar, M. (1992). The influence of terminators on motion integration across space. *Vision Research*, 32(2), 263–273.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715), 782–784.
- Muckli, L. (2010). What are we missing here? Brain imaging evidence for higher cognitive functions in primary visual cortex V1. *International Journal of Imaging Systems and Technology*, 20(2), 131–139.
- Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, 66(3), 241–251, doi:10.1007/BF00198477.
- Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience*, 9(3), 429–434.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), 15164–15169.
- Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks: The Official Journal of the International Neural Network Society*, 17(5–6), 695–705, doi:10.1016/j.neunet.2004.03.010.
- Pascual-Leone, A., & Walsh, V. (2001). Fast back-projections from the motion to the primary visual area necessary for visual awareness. *Science*, 292(5516), 510–512.
- Perrett, D. I., Xiao, D., Barraclough, N. E., Keysers, C., & Oram, M. W. (2009). Seeing the future: Natural image sequences produce ‘anticipatory’ neuronal activity and bias perceptual report. *Quarterly Journal of Experimental Psychology (2006)*, 62(11), 2081–2104, doi:10.1080/17470210902959279.
- Poljac, E., de-Wit, L., & Wagemans, J. (2012). Perceptual wholes can reduce the conscious accessibility of their parts. *Cognition*, 123(2), 308–312, doi:10.1016/j.cognition.2012.01.001.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87, doi:10.1038/4580.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical

- models of object recognition in cortex. *Nature Neuroscience*, 2(11), 1019–1025.
- Roelfsema, P. R., Lamme, V. A., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, 395(6700), 376–381.
- Saygin, A. P., & Sereno, M. I. (2008). Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. *Cerebral Cortex*, 18(9), 2158–2168, doi:10.1093/cercor/bhm242.
- Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science*, 21(5), 641–644.
- Shipp, S., & Zeki, S. (1989). The organization of connections between areas V5 and V1 in Macaque monkey visual cortex. *The European Journal of Neuroscience*, 1(4), 309–332.
- Sillito, A., Cudeiro, J., & Jones, H. (2006). Always returning: Feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences*, 29, 307–316.
- Sterzer, P., Haynes, J.-D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *NeuroImage*, 32(3), 1308–1316.
- Sumner, P., Edden, R. A. E., Bompas, A., Evans, C. J., & Singh, K. D. (2010). More GABA, less distraction: A neurochemical predictor of motor decision speed. *Nature Neuroscience*, 13(7), 825–827.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., et al. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409–1422.
- Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, 15(4), 3215–3230.
- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association: JAMIA*, 8(5), 443–459.
- Van Loon, A., Scholte, H. S., Johan-Saaltink, E. S., Knapen, T., Donner, T., & Lamme, V. G. *GABA concentrations predict individual differences in bistable perception*. Association for the Scientific Study of Consciousness, Brighton, UK, May 2012.
- Williams, M., Baker, C., Op de Beeck, H., Shim, W., Dang, S., Triantafyllou, C., and Coauthors, (2008). Feedback of visual object information to foveal retinotopic cortex. *Nature Neuroscience*, 11(12), 1439–1445.
- Xu, Y., & Chun, M. M. (2007). Visual grouping in human parietal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), 18766–18771.