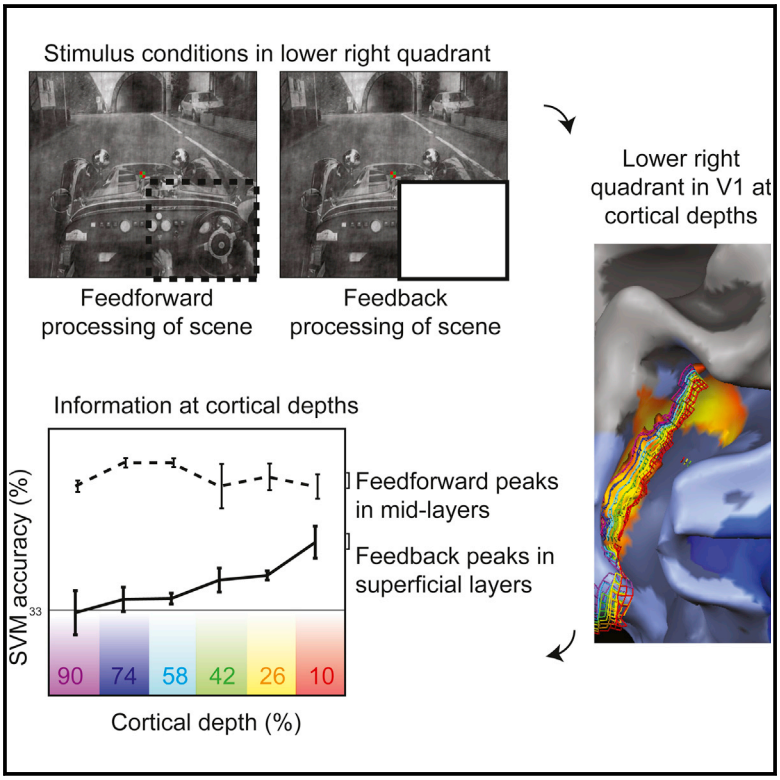
Report

Current Biology

Contextual Feedback to Superficial Layers of V1

Graphical Abstract



Authors

Lars Muckli, Federico De Martino, Luca Vizioli, ..., Kamil Ugurbil, Rainer Goebel, Essa Yacoub

Correspondence

lars.muckli@glasgow.ac.uk

In Brief

Muckli et al. have discovered that the superficial layers of visual cortex V1 receive information when not directly stimulated. This information contains contextual feedback from higher visual areas. The data provide empirical evidence for layer-specific cortical feedback relevant for the neurobiology of predictive coding.

Highlights

- High-resolution MRI shows functional information patterns in non-stimulated V1
- Non-stimulated V1 receives cortical feedback information to superficial layers
- Feedback to non-stimulated V1 superficial layers is predictive of visual context

Muckli et al., 2015, Current Biology *25*, 2690–2695
October 19, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.cub.2015.08.057

CellPress

Current Biology
Report

Contextual Feedback to Superficial Layers of V1

Lars Muckli,[1,*] Federico De Martino,[2,4] Luca Vizioli,[1] Lucy S. Petro,[1] Fraser W. Smith,[3] Kamil Ugurbil,[4] Rainer Goebel,[2,5] and Essa Yacoub[4]

[1]Centre for Cognitive Neuroimaging (CCNi), Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, 58 Hillhead Street, G12 8QB Scotland, UK

[2]Faculty of Psychology and Neuroscience, Department of Cognitive Neurosciences, Maastricht University, Oxfordlaan 55, 6229 EV Maastricht, the Netherlands

[3]School of Psychology, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK

[4]Center for Magnetic Resonance Research (CMRR), Department of Radiology, University of Minnesota, 2021 Sixth Street SE, Minneapolis, MN 55455, USA

[5]Department of Neuroimaging and Neuromodeling, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA Amsterdam, the Netherlands

*Correspondence: lars.muckli@glasgow.ac.uk

http://dx.doi.org/10.1016/j.cub.2015.08.057

SUMMARY

Neuronal cortical circuitry comprises feedforward, lateral, and feedback projections, each of which terminates in distinct cortical layers [1–3]. In sensory systems, feedforward processing transmits signals from the external world into the cortex, whereas feedback pathways signal the brain's inference of the world [4–11]. However, the integration of feedforward, lateral, and feedback inputs within each cortical area impedes the investigation of feedback, and to date, no technique has isolated the feedback of visual scene information in distinct layers of healthy human cortex. We masked feedforward input to a region of V1 cortex and studied the remaining internal processing. Using high-resolution functional brain imaging (0.8 mm^3) and multivoxel pattern information techniques, we demonstrate that during normal visual stimulation scene information peaks in mid-layers. Conversely, we found that contextual feedback information peaks in outer, superficial layers. Further, we found that shifting the position of the visual scene surrounding the mask parametrically modulates feedback in superficial layers of V1. Our results reveal the layered cortical organization of external versus internal visual processing streams during perception in healthy human subjects. We provide empirical support for theoretical feedback models such as predictive coding [10, 12] and coherent infomax [13] and reveal the potential of high-resolution fMRI to access internal processing in sub-millimeter human cortex.

RESULTS

To isolate feedback processing, we exploited the retinotopic organization of the visual cortex and blocked informative feedforward activation by occluding visual scene input in a contiguous subregion of the visual field. Using retinotopy, we isolated voxels that responded only to the occluded portion of the scene; in these voxels, we recorded brain activity by high-resolution and high-field fMRI and separated it into six different cortical depth layers. Scene-specific information did not directly stimulate the classical receptive fields of neurons in these voxels (i.e., through visual input), thus any scene responses are due to contextual feedback stimulating non-classical receptive fields. We studied multivariate activation patterns restricted to individual cortical depths to test for the presence of contextual scene information.

In the first experiment, we presented three visual scenes controlled for global luminance, contrast, and energy ("car on street," "people at market," and "ship in harbor," as in [14]). The visual scenes were either presented in full (as "feedforward" stimulation) or with the right lower visual field quadrant occluded by a mask ("feedback" condition; Figure 1A). Voxels responding in the feedforward condition contained a mixture of feedforward, lateral, and feedback signals; however, for simplicity, we refer to feedforward and feedback conditions. The second experiment consisted of two one-quarter-occluded images ("crowd of people" and "vintage car"; luminance, contrast, and energy controlled), which were presented in the original space and in two spatially shifted versions (2° and 8° visual angle). This shifting of the surrounding context allowed us to test how a parametric change of contextual information modulates feedback. We presented "target" and "surround" checkerboard mapping stimuli in both experiments (Figure 1B) to isolate feedback-receiving voxels and eliminate spill-over from feedforward-stimulated regions (e.g., mediated by horizontal interactions within V1) [14].

We recorded functional brain imaging data with gradient echo (GE) and spin- and GE-based 3D-GRASE fMRI sequences at 0.8 mm^3 resolution. We segmented the cortex using anatomic MRI scans (bias-field-corrected T1 over PD normalized) [15] and adjusted deep, inner, and superficial outer gray matter boundaries along the local GE-image intensity values to eliminate pial blood vessels and to correct for GE-EPI distortions. We used relative cortical depth values to create Laplace-based equipotential grid lines at six depths (from deep, inner [at the gray-white matter boundary] to superficial, outer [next to pia]

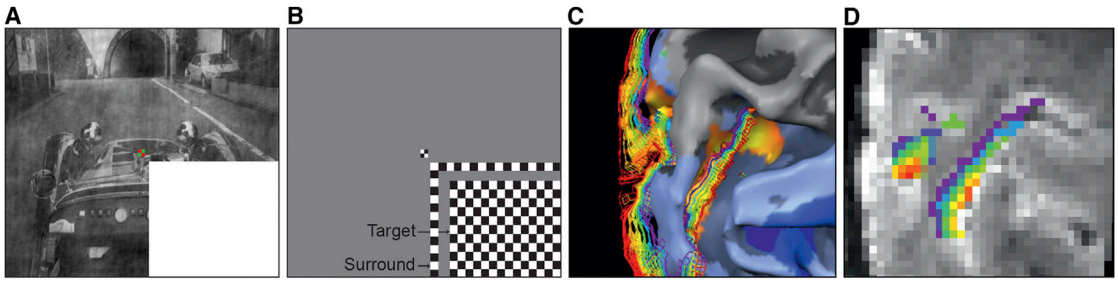


Figure 1. Experimental Procedure

(A) Example stimulus for the "feedback" condition, in which the lower right quadrant was occluded by a white mask (see [14]). The "feedforward" condition comprised the full image (not shown).

(B) "Target" and "surround" checkerboards (presented individually during scanning) to locate voxels responding to the lower right visual field.

(C) Left hemisphere cortical reconstruction of subject 2 in experiment 1, overlaid with a contrast of target response greater than surround response (light blue V1 and V3; dark blue V2). The cortical grid mesh depicts reconstructed depth layers from deep/inner (purple; close to white matter) to superficial/outer (red; close to the pial surface) cortical boundary.

(D) Corresponding regions of interest to (C) overlaid onto GE-EPI images.

90%, 74%, 58%, 42%, 26%, and 10% depths; Figures 1C and 1D). The depth grid lines covered the cortical representation of the occluded image section in the lower right visual field quadrant of retinotopic areas V1d, V2d, and V3d (Figure 2), mapped independently but also as part of each run.

In these "non-stimulated" patches of cortex, the GE-EPI data showed stronger responses in the superficial depths (close to the pial surface), whereas the 3D-GRASE data showed comparable responses across all cortical depths [16]. The bias of GE fMRI imaging to stronger responses in superficial cortical depths is due to larger blood vessels that lie on the cortex pial surface (Figure S1). Larger blood vessels could washout BOLD signal from a wider cortical area, compromising its retinotopic specificity (but see [17]). However, by including only voxels that respond retinotopically to the target and not the surround, we were able to exclude voxels exhibiting non-specific signals from draining veins. As a result, the filtered GE-fMRI responses in the superficial depths were still present but less pronounced and displayed constant retinotopic specificity across cortical depths (Figure S1).

To map the information at different cortical depths, we performed a multivoxel pattern analysis using two approaches: (1) decoding based on support vector machine (SVM) classification and (2) GLM-based encoding (Figure S2).

In decoding analyses, single-block SVM classification was significant at each depth for each subject (permutation tested at 5%; no corrections) during feedforward stimulation of V1. The highest performance was at 80% correct classification (chance 33%) at a cortical depth of 58% (third deepest depth; Figure 3A). For the feedback condition (i.e., the occluded images), only the superficial, outermost depth (10%) was significant in all four subjects, second-most outer depth (26%) was significant in three of four subjects, mid-depth (42%) was significant in two of four, and no subjects showed significance at the 58% mid-level, where feedforward information peaked. Deepest, inner depths (90% and 74%) were only significant in one subject (subject 3). We tested for the main effects of signal (feedforward and feedback) and depth using a 2 (signal) × 6 (depth) linear mixed model (see the Experimental Procedures). We found main effects of signal and depth and an interaction. For the feedforward signal, SVM accuracy was significantly larger for depth 58% compared to depth 90% ($t(15) = 4.6$; $p = 0.0003$), whereas for the feedback signal, SVM accuracy at depth 10% (i.e., the superficial layer) was significantly larger than depth 58% ($t(15) = 4.9$; $p = 0.0001$), depth 74% ($t(15) = 3.0$; $p = 0.002$), and depth 90% ($t(15) = 5.5$; $p = 0.00005$).

Within a classical hierarchical framework, it is reasonable to assume that the white occluders elicit identical patterns of brain activity. However, if lateral and feedback connections transmit contextual information to the non-stimulated region of cortex as shown by the decoding analysis, then we can also identify the univariate effect of single voxels and investigate the spatial extent in this informative structure. We addressed this in our encoding analysis. In brief, this involved a re-randomization analysis following the permutation of the explanatory variables in the general linear model of induced responses. This enabled us to evaluate the null distribution of the percent variance explained and associate the observed responses with a p value. We found that informative voxels explained significantly more variance for the original GLM (one predictor per scene) compared to a permutation GLM (randomized condition assignment). Subjects showed a J-function with some informative voxels in the deepest, innermost depth (90%), the least-informative voxels in the second innermost depth (74%), and the strongest peak in the superficial, outer depth (10%). We consistently found the highest percentage of informative voxels in the outer depths of V1 in all subjects (Figure S2).

Subsequent to the GE measurement, we scanned the same four subjects again for four runs with identical trials but used the 3D-GRASE sequence to acquire fMRI data that we expected to be less sensitive to large draining veins [16]. The 3D-GRASE sequence has reduced contrast to noise ratio, and we could only cover V1 because of the small acquisition field of view. For two subjects (1 and 3), feedforward stimulation did not lead to consistent informative readout in any of the six depths. For the other two subjects (2 and 4), classification based on visual stimulation was clear and significant. In subjects 2 and 4, we also found significant feedback information, again in the superficial depth of V1 (Figure 3C). Encoding analysis of

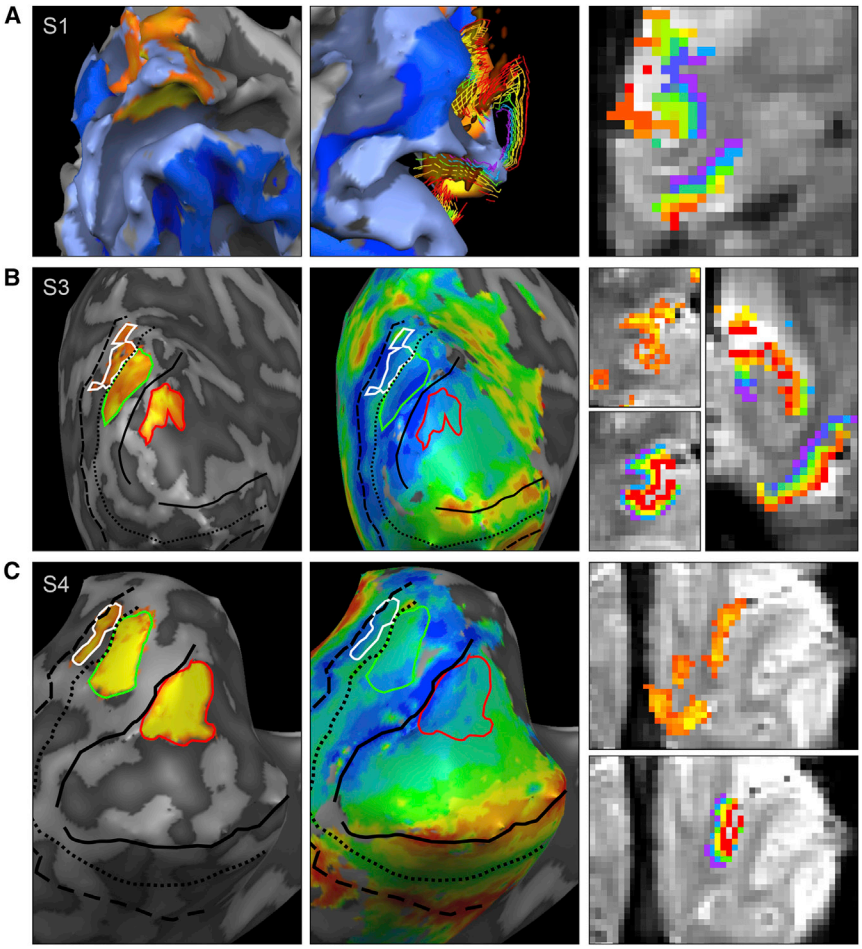


Figure 2. Layer-Specific Regions of Interest

(A) For subject 1: surface reconstruction overlaid with "target > surround" activity map (left); cortical grid lines depicting depth layers (middle); regions of interest in depth layers overlaid onto GE-EP images (right).

(B) For subject 3: inflated surface reconstruction overlaid with map of target responses outlined in red (V1), green (V2), and white (V3); borders between visual areas V1 and V3 are shown by black (dashed) lines (left); inflated surface reconstruction overlaid with polar angle retinotopic mapping data (middle); regions of interest colored in activity (top) or depth overlaid onto GE-EP images (right).

(C) As in (B), but for subject 4. Note: The cortex appears thicker when the slice plane cuts through it at a shallow angle (right column).

3D-GRASE data confirmed the presence of more informative voxels in the superficial, outer depth of V1 (Figure S2).

To investigate how feedback interacts with feedforward processing, whether they carry different information content, and whether they coexist in specific layers, we performed cross-layer decoding. We trained an SVM algorithm on a given depth and a given signal (i.e., either feedforward or feedback) and tested its performance against the same or the other signal, across all depths (Figure 3D). We found that training on feedback and testing on feedforward leads to highest (FDR $q < 0.05$) accuracy (FBxFF > FFxFB $t(575) = 4.67$; >FBxFB $t(575) = 6.85$) except when testing and training on feedforward (FFxFF). Training the model on the coarser feedback information pattern leads to a more general model that allows for accurate decoding of the finer feedforward multivoxel pattern. However, the opposite is not true as the model built on specific feedforward information cannot be generalized to decode the coarser feedback information content.

With the cross-layer cross-signal decoding, we can also demonstrate the following general points: (1) the feedforward signal is homogeneous across layers (i.e., training on depth layers allows to generalize to other depth layers) and (2) although feedforward and feedback activation are very different in amplitude, they can share common information in their patterns of BOLD signal.

Up until now, we have described decoding analyses across cortical depths in V1. We performed identical analyses in V2 and V3. The decoding analysis in V2 revealed that contextual Information was significant in the most inner depth (90%) in two subjects (Figure S4; subjects S2 and S4), and in the most outer two depths (10% or 26%) in another three subjects (S1, S2, and S3). In V3, there was reduced feedforward information compared to V1 and V2. All subjects had significant classification at mid-depth 58%, and three subjects showed significant information at least at two other depths. There was no consistent pattern in the feedback condition in V3 (Figure S4). We propose that both the reduced size of area V3 and its reduced functional specialization for static scene information might explain the decreased decoding performance.

In the spatial shift experiment, we tested the precision of feedback signals by shifting the contextual surround image. We trained the classifier to discriminate two images and then tested it on shifted versions of these two images. Consistent with the first experiment, feedback was most prominent in the outer layers. Furthermore, decoding was only significant when cross-classifying across the smallest shift (2°) in the superficial depth (10%) in subject 1, second and third outermost depths (26% and 42%, respectively) for subject 2, and the two outermost depths (10% and 26%) in subjects 3 and 4 (Figure 4).

DISCUSSION

When primary visual cortex does not receive sensory input, it nonetheless communicates with other brain areas. Such internally generated processing states are not sufficiently understood [18, 19] but are an essential feature of the brain, accounting for 90% of overall energy consumption [20, 21]. We propose that such internal communications include feedback sent across large cortical networks, including back to sensory areas [22–24]. The effects of feedback inputs on sensory neurons are uniquely challenging to study because it requires that feedback inputs be separated from externally induced (feedforward) processing. Access

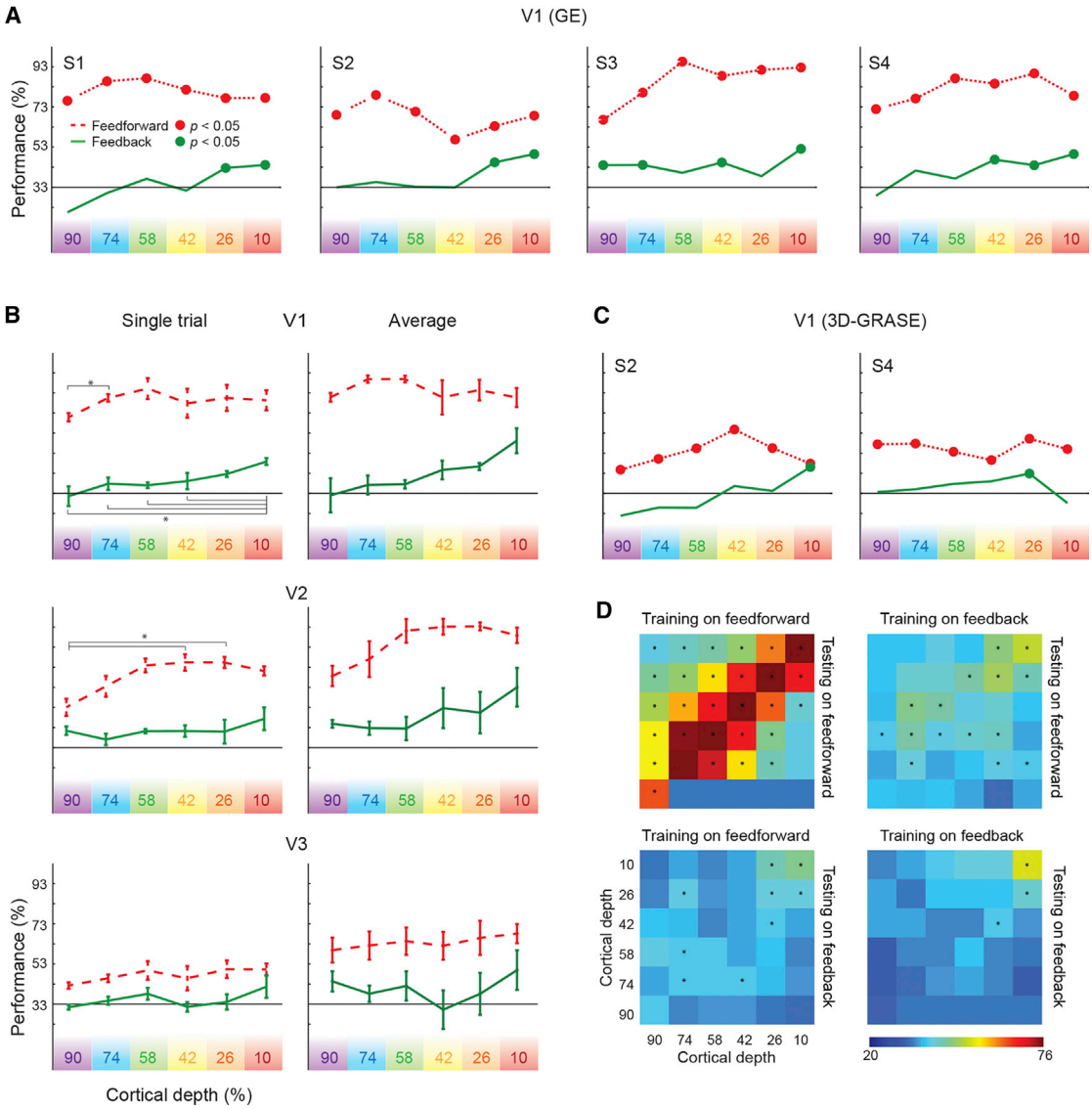


Figure 3. Layer-Specific Information Decoding

(A) For V1 using GE fMRI data, SVM classification performance for all four subjects during feedforward (red dashed) and feedback (green) processing, in individual depths (color-coded purple to red). Chance decoding level was 33%, and significant classification is marked by circles.

(B) For V1, V2, and V3 averaged across subjects using GE fMRI data, SVM classification performance in cortical depths (from white matter 90% depth to superficial depth 10%). Left panels show prediction of single trials in the left out run; right panel shows the averaged condition of the left out run. Significant differences in decoding performance between depths are marked on subject-averaged single run plots (permutation tested); error bars represent SEM (across subjects and leave-one-run-out folds).

(C) As in (A), but for subject 2 and 4's 3D-GRASE fMRI data.

(D) We trained an SVM algorithm on a given cortical depth and a given signal (i.e., either feedforward or feedback) and tested its performance against the same or the other signal across all depths. Asterisks indicate significance against theoretical chance level (FDR $q < 0.05$).

See also Figures S1–S4.

to feedback signals during human visual cognition will explicate century-old theories of inferential brain processing [6]. Taking advantage of the retinotopic organization of V1, we use a partial occlusion paradigm whereby we can drive higher visual areas to feed natural scene information back to regions of V1 that do not receive feedforward scene inputs [14]. We propose that ultra-high-field fMRI (7T) provides a means to investigate the presence of such contextual feedback in distinct cortical layers.

We found contextual feedback information in the superficial layers of V1. Several candidate regions could be sending contextual information to V1, including extrastriate visual (V2–V7) and more distant cortical and subcortical areas. We know from studies of rodent cortex that long-distance cortical feedback and associative thalamo-cortical interactions have dense projections to supragranular layer L1 [1]. Up to 90% of synaptical input in layer 1 is from long-distance sources, with only 10% from local

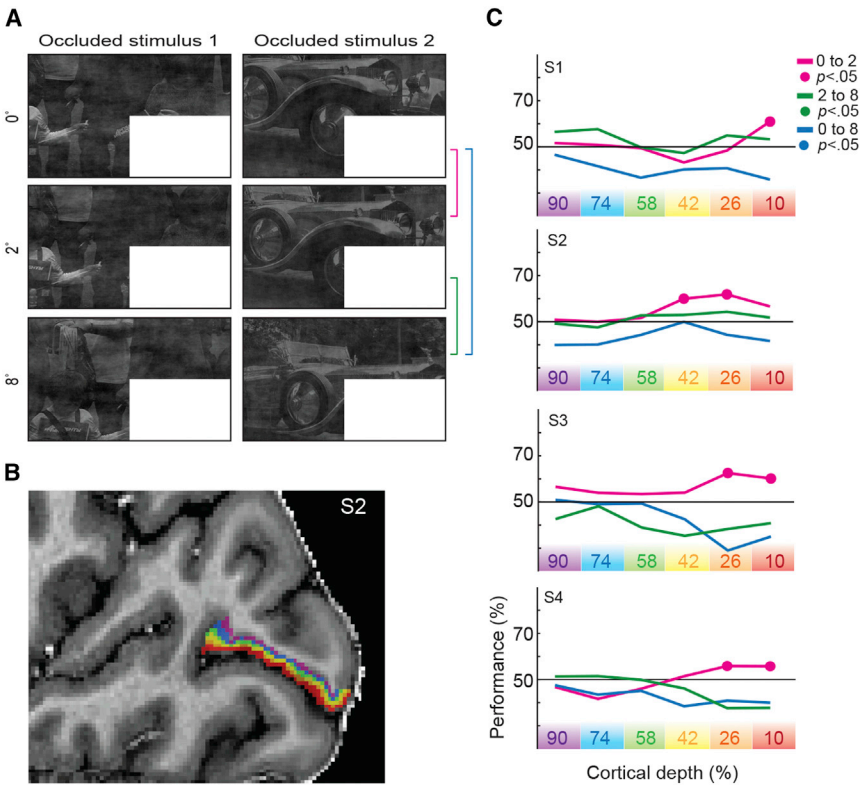


Figure 4. Feedback Sensitivity to Shifts of the Surround

(A) Occluded stimuli used in experiment 2. We shifted the original stimuli (0°) by 2° and 8°.

(B) Cortical depth layers shown in the sagittal plane of subject 2, from deep, inner (purple) to superficial, outer (red).

(C) SVM decoding performance for different depths of V1 when cross-classifying images of different shifts, for all four subjects.

neurons [1, 2]. Larkum proposes a neuronal mechanism that integrates the long-distance top-down projections with bottom-up input, which involves layer 5 pyramidal cells that have distal tuft dendrites in layer 1. David Mumford [10, 25] proposed a conceptual interpretation that cortical feedback from higher areas predicts the most likely feedforward input to the "active blackboard" of V1. The predictive coding framework describes neuronal computations in cortex, which integrate the predictions carried by cortical feedback with sensory inputs carried by feedforward projections [4]. Predicted (or irrelevant) sensory inputs may be dis-amplified by the inhibition of apical amplification [13]. In such a functional architecture, feedback signals from regions with larger receptive fields would code at a more abstract level and at a coarser spatial scale. Consistent with this idea, we found that the information in supragranular layers is similar even when we shifted the surrounding visual context by 2° (visual angle).

Functional MRI is sensitive to neuronal energy consumption, which includes dendritic activity and activity of inhibitory interneurons [26]. The use of ultra-high-resolution fMRI allowed us to focus on the information in activation patterns at different cortical depths. Previous ultra-high-resolution layer-specific fMRI focused on activation amplitude rather than information [27–31]. Our design permits the sampling of top-down dendritic activity that contains scene information in superficial layers of human V1.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2015.08.057.

AUTHOR CONTRIBUTIONS

Conceptualization, L.M. and F.D.M.; Methodology, L.M., F.D.M., E.Y., K.U., and F.W.S.; Software, L.M., L.V., F.W.S., and R.G.; Formal Analysis, L.M., L.V., L.S.P., and F.D.M.; Investigation, L.M., F.D.M., E.Y., L.S.P., and F.W.S.; Writing – Original Draft, L.M. and L.S.P.; Writing – Review & Editing, L.M., L.S.P., L.V., F.D.M., F.W.S., K.U., E.Y., and R.G.; Visualization, L.M., L.S.P., and L.V.

ACKNOWLEDGMENTS

The institutional review board at the University of Minnesota and the ethics committee at the University of Glasgow approved the studies. We thank Jan Zimmermann for support in data acquisition and Jack Gallant, Kendrick Kay, and Elia Formisano for discussion of encoding analysis. R.G. is supported by ERC grant "ColumnarCodeCracking" (ERC-2010-AdG; grant agreement no. 269853). E.S., K.U., and work conducted in CMRR are supported by NIH grants P41 EB015894 (NIBIB) and P30 NS076408 (NINDS). L.M., L.V., and L.S.P. are supported by ERC grant (ERC StG 2012_311751-"Brain reading of contextual feedback and predictions") dedicated to L.M.

Received: June 9, 2015
Revised: July 27, 2015
Accepted: August 25, 2015
Published: October 1, 2015

REFERENCES

1. Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. Trends Neurosci. *36*, 141–151.
2. Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., et al. (2014). Anatomy of hierarchy: feedforward and feedback pathways in macaque visual cortex. J. Comp. Neurol. *522*, 225–259.

3. Sillito, A.M., Cudeiro, J., and Jones, H.E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. Trends Neurosci. *29*, 307–316.
4. Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical microcircuits for predictive coding. Neuron *76*, 695–711.
5. Bullier, J. (2001). Integrated model of visual processing. Brain Res. Brain Res. Rev. *36*, 96–107.
6. Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. *36*, 181–204.
7. Gilbert, C.D., and Sigman, M. (2007). Brain states: top-down influences in sensory processing. Neuron *54*, 677–696.
8. Meyer, K. (2012). Another remembered present. Science *335*, 415–416.
9. Muckli, L., and Petro, L.S. (2013). Network interactions: non-geniculate input to V1. Curr. Opin. Neurobiol. *23*, 195–201.
10. Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol. Cybern. *66*, 241–251.
11. Park, H.J., and Friston, K. (2013). Structural and functional brain networks: from connections to cognition. Science *342*, 1238411.
12. Friston, K. (2005). A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. *360*, 815–836.
13. Phillips, W.A., Clark, A., and Silverstein, S.M. (2015). On the functions, mechanisms, and malfunctions of intracortical contextual modulation. Neurosci. Biobehav. Rev. *52*, 1–20.
14. Smith, F.W., and Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. Proc. Natl. Acad. Sci. USA *107*, 20099–20103.
15. Van de Moortele, P.F., Auerbach, E.J., Olman, C., Yacoub, E., Uğurbil, K., and Moeller, S. (2009). T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. Neuroimage *46*, 432–446.
16. De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., and Goebel, R. (2013). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. PLoS ONE *8*, e60514.
17. Blinder, P., Tsai, P.S., Kaufhold, J.P., Knutsen, P.M., Suhl, H., and Kleinfeld, D. (2013). The cortical angiome: an interconnected vascular network with noncolumnar patterns of blood flow. Nat. Neurosci. *16*, 889–897.
18. Douglas, R.J., and Martin, K.A. (2007). Mapping the matrix: the ways of neocortex. Neuron *56*, 226–238.
19. Markov, N.T., and Kennedy, H. (2013). The importance of being hierarchical. Curr. Opin. Neurobiol. *23*, 187–194.
20. Raichle, M.E. (2006). Neuroscience. The brain's dark energy. Science *314*, 1249–1250.
21. Raichle, M.E. (2011). The restless brain. Brain Connect. *1*, 3–12.
22. Clavagnier, S., Falchier, A., and Kennedy, H. (2004). Long-distance feedback projections to area V1: implications for multisensory integration, spatial awareness, and visual consciousness. Cogn. Affect. Behav. Neurosci. *4*, 117–126.
23. Dehaene, S., and Changeux, J.P. (2011). Experimental and theoretical approaches to conscious processing. Neuron *70*, 200–227.
24. Markov, N.T., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013). Cortical high-density counterstream architectures. Science *342*, 1238406.
25. Mumford, D. (1991). On the computational architecture of the neocortex. I. The role of the thalamo-cortical loop. Biol. Cybern. *65*, 135–145.
26. Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. Nature *453*, 869–878.
27. De Martino, F., Moerel, M., Xu, J., van de Moortele, P.F., Ugurbil, K., Goebel, R., Yacoub, E., and Formisano, E. (2014). High-resolution mapping of myeloarchitecture in vivo: localization of auditory areas in the human brain. Cereb. Cortex Published online July 3, 2014. http://dx.doi.org/10.1093/cercor/bhu150.
28. Goense, J., Merkle, H., and Logothetis, N.K. (2012). High-resolution fMRI reveals laminar differences in neurovascular coupling between positive and negative BOLD responses. Neuron *76*, 629–639.
29. Koopmans, P.J., Barth, M., Orzada, S., and Norris, D.G. (2011). Multi-echo fMRI of the cortical laminae in humans at 7 T. Neuroimage *56*, 1276–1285.
30. Polimeni, J.R., Fischl, B., Greve, D.N., and Wald, L.L. (2010). Laminar analysis of 7T BOLD using an imposed spatial activation pattern in human V1. Neuroimage *52*, 1334–1346.
31. Olman, C.A., Harel, N., Feinberg, D.A., He, S., Zhang, P., Ugurbil, K., and Yacoub, E. (2012). Layer-specific fMRI reflects different neuronal computations at different depths in human V1. PLoS ONE *7*, e32536.

Current Biology

Supplemental Information

Contextual Feedback to Superficial Layers of V1

Lars Muckli, Federico De Martino, Luca Vizioli, Lucy S. Petro, Fraser W. Smith, Kamil Ugurbil, Rainer Goebel, and Essa Yacoub

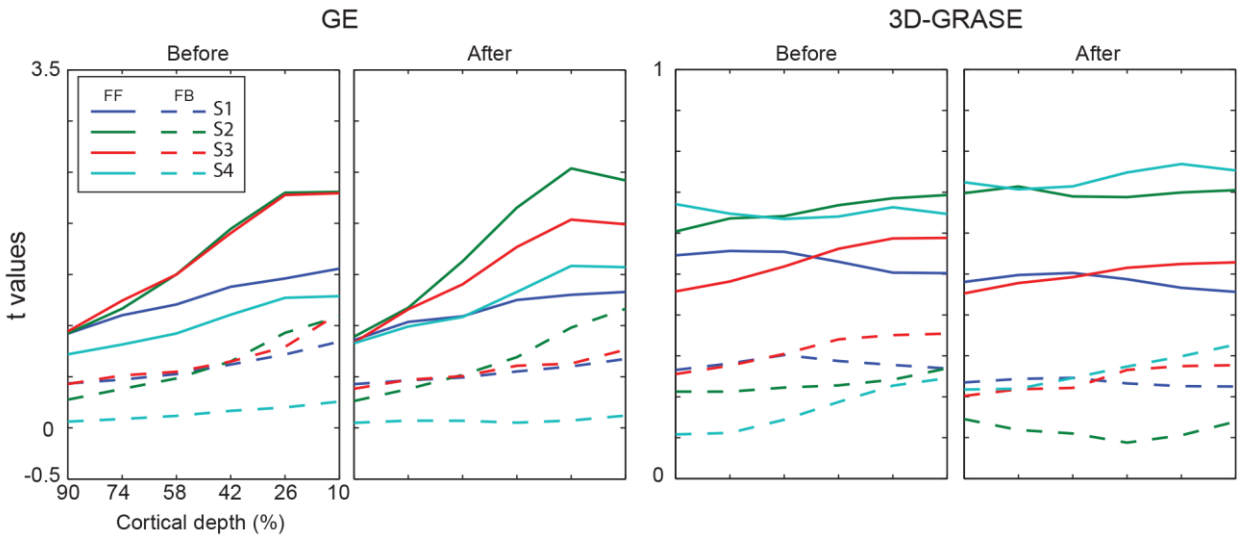


Figure S1 related to Figure 3. Mean t-values (y-axis) across all conditions for the Gradient Echo (GE) and 3D-GRASE sequences before and after voxel selection. Within each sequence, the left and right plots represent the average t-values *before* and *after* voxel selection. Before depicts visual active voxel (contrast: t average feedforward > 0). After depicts retinotopic specific selected voxels (contrast: t target stimulus > t surround ∩ t target stimulus > 0 ∩ t full-stimulus > t occluded-stimulus). Colours denote different subjects. Solid lines show the response to non-occluded (i.e. feedforward) images and dashed lines show the response to the occluded (i.e. feedback) portions of the image. X-axis shows cortical depths from inner (90%) to outer (10%).

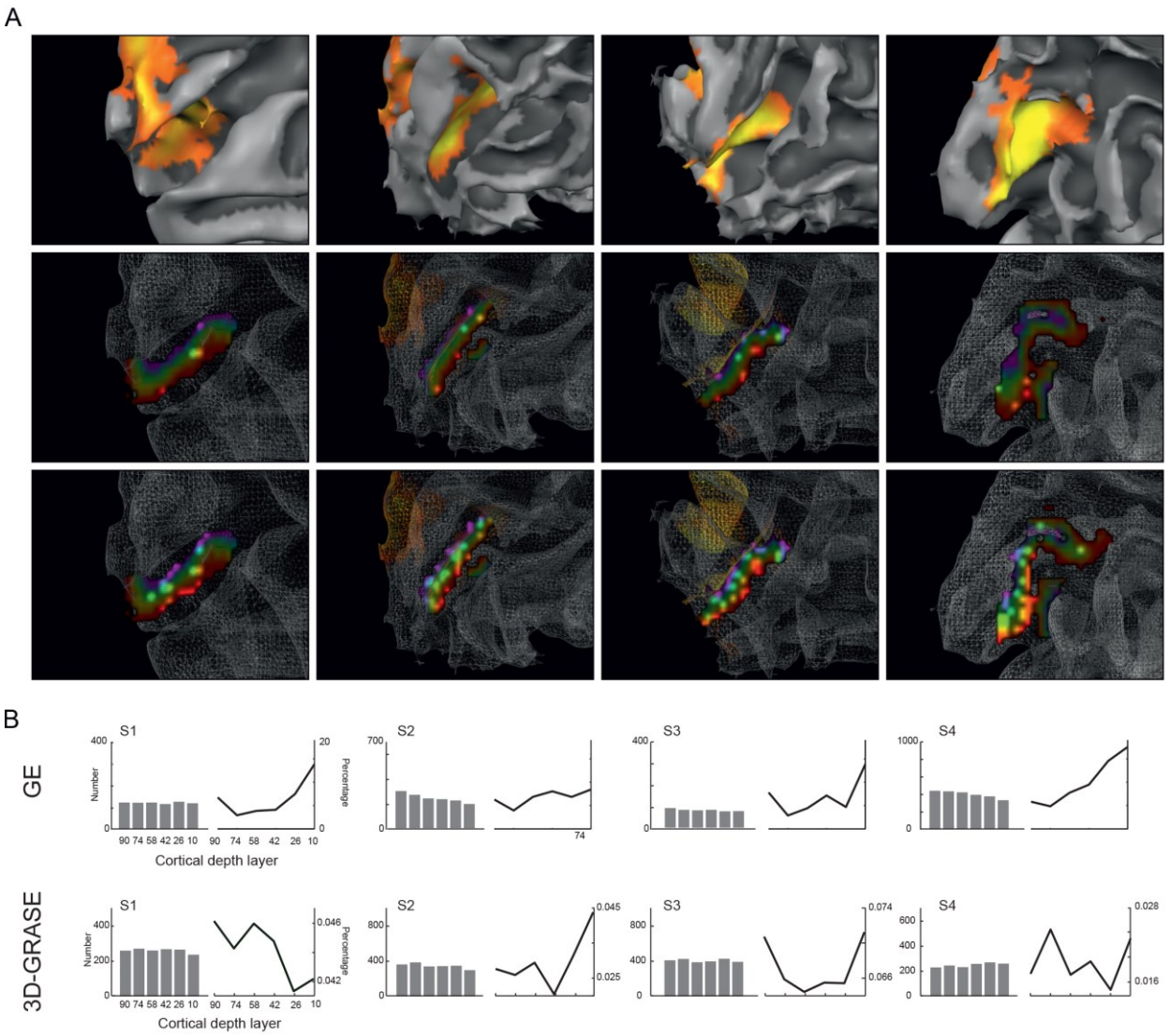


Figure S2 related to Figure 3. Encoding analysis in V1, for four subjects in experiment 1. **A.** The top row shows the location of the target region (target > surround) on the cortical surface reconstruction (along the grey-white matter boundary) of four subjects. The second row shows target-activated voxels (GE data) in one sagittal cut through grey matter. Target-activated voxels are shown in transparent rainbow colours in accordance with their different cortical depths (superficial red 10%, orange 26%, green 42%, blue 58%, dark blue 74%, purple 90%). We show informative voxels as indicated by the encoding analysis (threshold at $p<0.05$) in saturated colours. The third row illustrates the informative voxels (saturated colours) for the GE data as indicated by the encoding analysis (threshold at bootstrapped $p<0.05$) within each cortical depth layer (transparent colours). **B.** The fourth row displays the percent of informative voxels (line plot) within each cortical depth layer for each subject (GE data). The bar plots show the number of voxels in the retinotopic representation of the occluded region. The fifth row portrays the percent of informative voxels (lines) within each cortical depth layer for each subject (GRASE data).

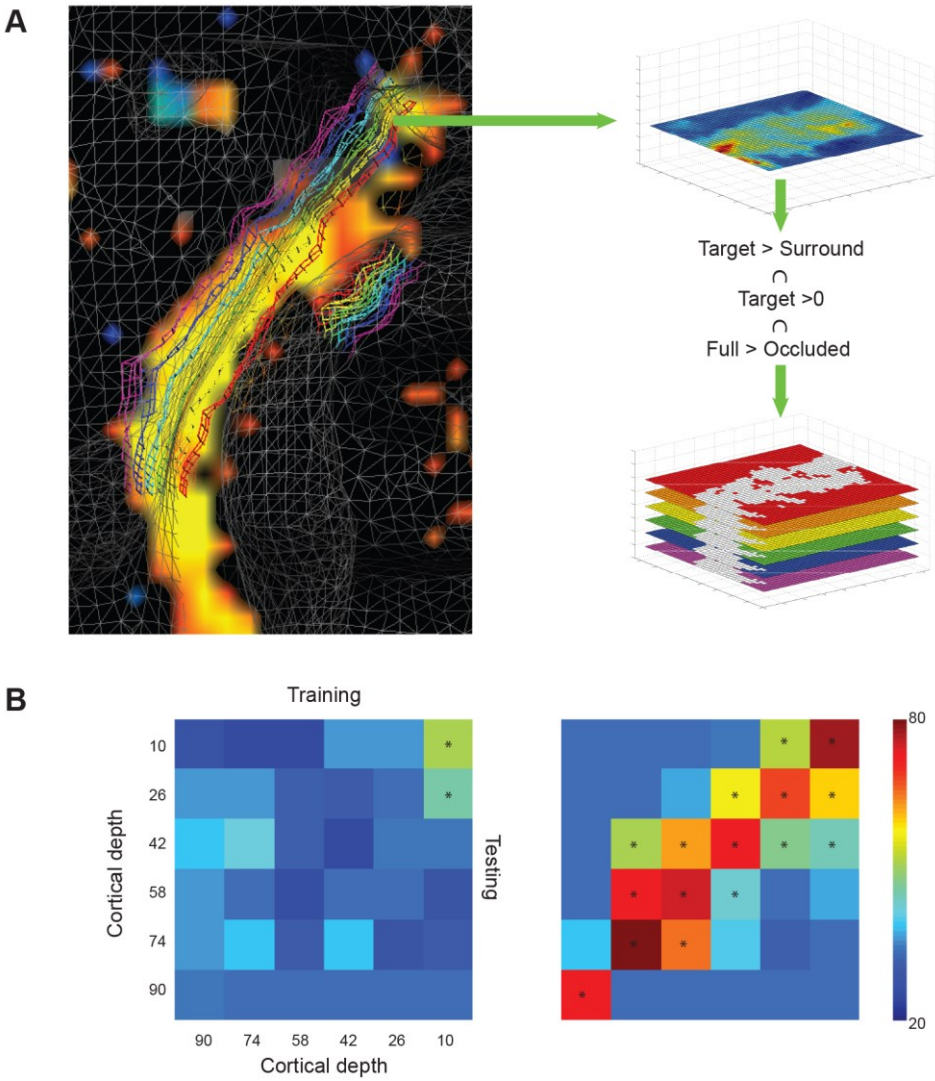


Figure S3 related to Figure 3. Cross-depth support vector machine analysis. **A,** White transparent wireframe shows cortical surface reconstruction along the grey-white matter boundary of subject 2's left occipital cortex (sagittal view). A sagittal section of visual activation map shown in yellow (t target stimulus > t surround ∩ t target stimulus > 0 ∩ t full-stimulus > t occluded-stimulus) and six depth plane-sections (2 gridlines in depths) in purple (inner) to red (outer). Grid-unfolding leads from 3-dimensional voxel space to 2-dimensional unfolded grid space used for subsequent analysis. For a given depth (e.g. depth 58% shown in green top right), we selected the pixels of interest (contrast amongst T-values as indicated). We repeated the same procedure independently for all other cortical depths and combined the resulting pixels subpopulations to produce one population of interest which included only pixels shared across all depths. We projected the resulting pixel subpopulation (shown in white in the bottom right portion of the figure) on all depths. We then trained and tested the classifier on this pixel subpopulation *across* depths and signals (i.e. feedback and feedforward – see supplementary method). **B,** shows subject 2's cross-depth SVM classification performance trained and tested on the activity patterns elicited by feedforward (right matrix) and feedback (left matrix) portions of

the image stimuli. We trained SVM classifiers on a given depth (x-axis) and tested on all other depths (y-axis). Stars indicate significant above chance SVM accuracy.

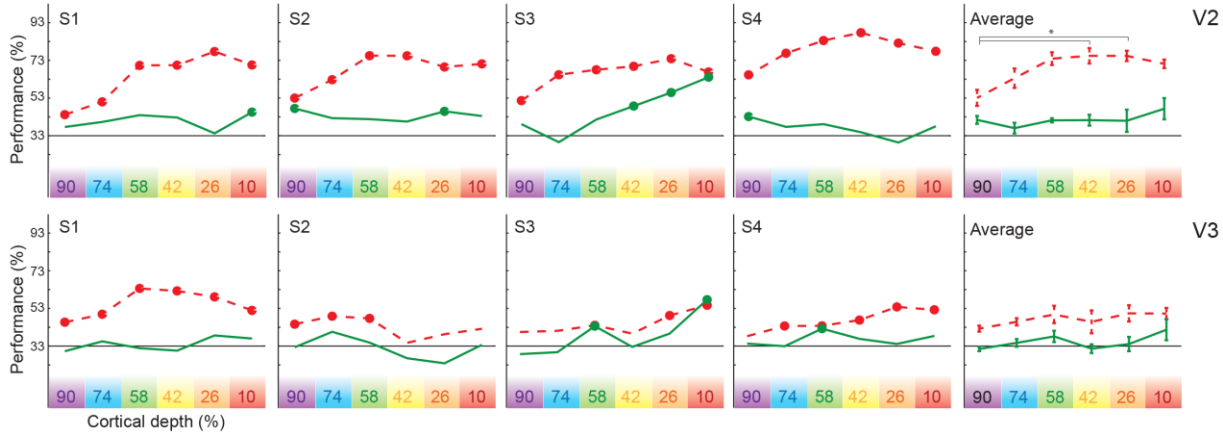


Figure S4 related to Figure 3. SVM single run classification performance for V2 (top row) and V3 (bottom row) using GE fMRI data. Prediction of single trials in the left out run (cross-validation) is shown for all four subjects (columns) and averaged across subjects (in right column). SVM classification performance is shown for cortical depths (from white matter 90% depth to superficial depth 10%), for feedforward conditions in red and feedback conditions in green Significant differences in decoding performance between depths are marked on subject-averaged single run plots (permutation tested), error bars represent standard error of mean (across subjects and leave-one-run-out folds).

Supplemental Experimental Procedures

Subjects

We measured seven subjects; four subjects for experiment 1 at the Center for Magnetic Resonance Research (CMRR, Minneapolis, MN, USA), and three subjects for experiment 2 at the Maastricht brain imaging Centre (Maastricht, Netherlands). One subject participated in both experiments. We ran experiment 1 twice on the same four subjects with different fMRI sequences (see below). All subjects were healthy volunteers with normal or corrected visual acuity. Subjects gave written informed consent and we paid them for participation. The institutional review board for human subject research at the University of Minnesota (Experiment 1), and the ethics committee at the University of Glasgow College of Science and Engineering (CSE01209, Experiment 2) approved the studies.

Stimuli

In experiment 1, subjects viewed three visual scenes (’car on street’, ‘people at market’, ‘ship in harbour’, same as in [S1] (Figure 1)). We controlled the images for global luminance, contrast and energy, using matlab shine-toolbox [S*2*]. Scenes were presented in full (’feedforward’ condition) or masked with an occluder over the lower right visual field (’feedback’ condition). We presented a set of contrast-reversing checkerboard mapping stimuli for ‘target’ and ‘surround’ regions in each run, and in a separate localiser run. The surround checkerboards mapped the outer 2 degrees of the white occluder and the target mapped the remaining inner section of the occluder (figure 1b). The design of the experiment is comparable to our previous study [S1], but with the visual stimuli reduced in size by 20% to fit the smaller MRI bore due to the use of a head gradient insert (see below). We kept the width of the ‘surround’ region at 1 degree of visual angle, and added an additional 1-degree border between the surround and the edge of target region. The second experiment consisted of two occluded images (’crowd of people’, and ‘vintage car’, luminance, contrast and energy controlled) which were presented in the original space and in two spatially shifted versions (2 and 8 degrees visual angle, Figure 4). We presented target and two surround conditions (1 degree, and 2 degree) in each of 5 runs. In both experiments we conducted a separate phase-encoded retinotopic mapping experiment [S3-5]. Stimuli consisted of a wedge-shaped (22.5 degrees) checkerboard rotating slowly (64s for full 360 degree rotation) around the fixation point in the middle of the screen. A white ‘spider web’ configuration was presented in the background to stabilise fixation together with a centre fixation colour change task [S6].

Paradigm

Experiment 1 comprised four functional runs of 350 volumes each. An experimental condition was presented for 6 volumes (12s), and each of 6 experimental conditions was presented in a randomised order within a block followed by 12 volumes (24s) of baseline (6x12s+24=96s per block). Mapping blocks consisting of 2 conditions (’target’, ‘surround’) were presented for 6 volumes (12s) interleaved with 6 volumes (12s) of baseline in between conditions and 12 volumes (24s) at the end of the block (2x12s+12s+24s=60s). A functional run consisted of 6 experimental blocks and 2 mapping blocks and an additional baseline of 2 volumes (4s) at the start of the run (6x96s + 2x48s + 4s = 700s = 350 volumes). Therefore, each experimental condition was presented 24 times across four runs.

The retinotopic mapping run comprised of 12 repetitions of a full rotation lasting 32 volumes (64s), with an extended baseline of 10 volumes (20s) at the beginning and 12 volumes (24s) at the end of the run (resulting in 406 volumes: 12x64s+20s+24s=812s). An additional localiser run comprised 12 repetitions of 'target' and 12 repetition of 'surround' mapping, with 25 baseline periods in-between, all of which lasted for 6 volumes (12s), resulting in 294 volumes ((12+12+25)x12s=588s).

Experiment 2 comprised of a retinotopic mapping run (as above) and five experimental runs. An experimental run had the same temporal structure as experiment 1 only now the six experimental conditions were two occluded scenes each shifted for 0, 2, 8 degrees. The mapping blocks had an additional 'surround-2' condition presented twice for 6 volumes (12s) on and 6 volumes (12s) off, lasting therefore for 360 volumes (740s).

Subjects viewed the visual stimuli on a projection screen mounted to the rear end of the head coil using a head-coil mounted mirror. A video projector combined with a mirror projected the stimuli onto the screen. Stimuli were presented using Presentation software (Neurobehavioral Systems, CA, USA) for the experiments, and for retinotopic mapping with StimulGL (custom-built stimulation software, Maastricht University, Maastricht, NL). We instructed the subjects to keep fixation to the centre of the screen and to perform a colour-change detection task at the centre of the screen, during both the experimental runs and retinotopic mapping.

MRI Acquisition

MRI data for the first experiment was conducted with ultra-high magnetic field (7 Tesla, 90cm bore, Magnex Scientific, Abingdon, UK) at the CMRR in Minneapolis (MN, USA). The MRI was driven with a Siemens console (Erlangen, Germany), and used a head gradient insert with a 6-channel receive array RF coil that covered only the visual areas, and a surface coil with one transmit and six receiving channels (1-Tx, 6-Rx). We recorded the second experiment at the MBIC-/Scannexus in Maastricht (NL) with 7T Magnetom (Siemens, Erlangen Germany), a 32-channel receive head coil, a whole body gradient, and a whole brain coil (1-Tx, 32-Rx; Nova Medical Inc., USA).

Functional scans were recorded using GE-EPI at high resolution (nominal resolution, isotropic 0.8 mm^3, TE = 17 ms, maximum flip angle determined by a flip angle map =85°, slices = 38, TR= 2000 ms, FOV=128 x 128 mm^2, matrix: 160 x 160, IPAT = 2, partial Fourier = 6/8, pixel bandwidth = 1375). We repeated the first experiment with a 3D gradient and spin echo (3D-GRASE) sequence. Details of the sequence and results and advantages of this sequence have been discussed in our previous paper [S7]. Positioning of the reduced 3D-GRASE slab was optimised with previous localiser runs and had 12 slices again with nominal isotropic 0.8mm^3 resolution (TE= 40 ms, TR = 2000 ms, FOV= 25.6 x 204.8 x 9.6 mm^3, slice partial Fourier = 5/8, pixel bandwidth = 1955, matrix: 32 x 256 x 12). The second set of experiments were recorded with a GE-EPI sequence at (nominal resolution, isotropic 0.8 mm^3, TE = 19 ms, FA =70°, slices = 31, TR= 2000 ms, FOV=148 x 148 mm2, matrix: 186 x 186, IPAT = 3, pixel bandwidth = 1120). Anatomical reference scans were recorded at 1mm^3 (experiment 1) and at 0.5mm^3 (experiment 2) with an MPRAGE sequence optimised for T1-weighted contrast, and proton density (PD)-weighted contrast. An additional T2*-weighted anatomical data was acquired in experiment 2.

Anatomical data analysis - cortical depth sampling

All data were analysed with BrainVoyager QX 2.8. Proton density scans with identical slice positioning were used to remove spatial intensity inhomogeneities from T1 scans, by dividing T1 data by PD [S8]. We adjusted manually inner and outer grey matter boundaries along the local intensity values to eliminate pial blood vessels, and to correct for GE-EPI distortions. We used relative cortical depth values to create Laplace-based equipotential grid-lines at six depths (from inner to outer 90%, 74%, 58%, 42%, 16%, and 10% depths, Figure 2). The gridlines were calculated smoothly at a highly up-sampled spatial coordinate system [S9]. In a subsequent step, smooth gridlines are used to assign voxels to a respective cortical depth. Individual voxels were allowed to belong to adjacent depths (Figures 1, 2, 4, S3). The depth grid-lines covered the cortical representation of the occluded image section in the lower right visual field quadrant of retinotopic areas V1d, V2d, and V3d (Figure 2). The layered regions of interest were saved as BrainVoyager QX VOI files (volume of interest) and further processed with Matlab code reading in the BrainVoyager QX volume time courses (vtc) files using the BVQX-Toolbox.

Functional data analysis

We pre-processed the fMRI data using slice scan time correction (GE only, sinc interpolation), 3D rigid body motion correction (sinc interpolation), intra-session alignment to the functional data of the last run, and temporal high pass filtering of 4 cycles. In experiment 2, we used GE EPI images acquired with opposite phase-encoding direction (recorded before the first functional run) to correct for distortions in EPI images [S10]. Functional data were aligned to anatomical data with manual adjustments and iterative optimisation. Activation maps of retinotopic mapping were used to optimise segmentation and alignment. Analysis included general linear model (GLM) estimation of averaged conditions and single trials. Design matrices were generated by convolution of a double gamma function with a "boxcar" function (representing onset and offset of the images stimuli).

Decoding

The linear SVM was implemented using the LIBSVM toolbox [S11], with default parameters (notably C = 1). Note that the activity of each vertex was normalized (separately for training and test data) within a range of −1 to 1 before input to the SVM. All decoding and encoding analyses were performed only on voxels responding to target more than to surround (defined by the contrast t target stimulus > t surround ∩ t target stimulus > 0 ∩ t full-stimulus > t occluded-stimulus). We have described the decoding analyses in more detail previously [1]. In brief, we trained support vector machine classifiers (linear pattern) to map between activation patterns from three scenes (full feedforward images in experiment 1) or between occluded scenes (3 occluded scenes in experiment 1, or 6 occluded images in experiment 2). We tested the trained classifiers on independent data (leave one run out cross validation). In experiment 2, we trained a cross-classifier on two images at one 'shift' and tested on data from a different 'shift'. We measured the classifier performance of each cortical depth independently. For each cortical depth, we tested the single trial classification for significance using permutation testing (1000 iterations of randomly assigned labels). To test the main effects and interaction between signal (feedforward and feedback) and layer depths (1 to 6), we performed a 2 by 6 Linear Mixed Model (Matlab, The Mathworks Inc, 2014). We combined the data from 4 runs and 4 participants (i.e. using 16 data points). Participants were considered as random effect. We estimated our fixed effects coefficients by means of maximum likelihood estimation

Encoding

To estimate whether individual voxels within the retinotopic representation of the occluded region in V1-V3 hold meaningful contextual information in the absence of feedforward thalamic input we designed two simple linear encoding models. We computed GLM design matrices modelling only occluded images and concatenating all runs to produce matrices measuring 1200 (350 volumes x 4 runs) x 31 (3 conditions + 4 constant terms, 1 per run, and discrete cosine transform sets, consisting of 6 predictors per run). In model 1, we kept each image within its original condition, so that the first 3 columns of our design matrix – representing our three stimuli –contained different trials of the *same* image stimulus. In model 2, instead we randomly shuffled the trials, across condition so that each predictor contained trials of images with different contextual information (the first 3 columns of our design matrix). We repeated this procedure 400 times to generate 400 different design matrices. As conditions in model 1 share the same contextual information surrounding the occluders (i.e. the encoded feature), we refer to this model as "high level contextual model" (three conditions, one for each image). On the other hand, in model 2, contextual information differs significantly across trials within a given condition, with the only shared features (i.e. the encoded feature) represented by the low-level visual properties of the occluders. We refer to this model as "the occluder-model" (three randomised occluder conditions). We then performed GLM analysis and computed the percentage of variance explained (defined as $1\text{-}R^2$) by each model independently per voxel. This led to 401 values per voxel (1 for the high level contextual model and 400 for the low level models). In order to estimate whether a given voxel explained significantly more variance in the high-level contextual model compared to the occluder-model, we subtracted the variance explained by the 400 occluder-models to that of the high-level contextual model. We then counted the number of times this computation led to a positive value and considered a voxel as explaining significantly more variance in the high-level contextual model if 95% or more of the instances ($p =< .05$ one tailed) were positive (high-level contextual voxel). We reasoned that as the visual information held within the occluded region is identical across different images, from a purely feed-forward standpoint, no difference in terms of variance explained should be expected between the 2 models (i.e., our H0).

Cross-depth support vector machine analysis

To investigate how feedback interacts with feedforward processing; whether they carry different information content; and if they coexist in specific layers, we performed cross-layer decoding analysis. We trained an SVM algorithm on a given cortical depth and a given signal (i.e. either feedforward or feedback) and tested its performance against the same or the other signal, across all depths. This gave four combinations: (1) training and testing on feedforward; (2) training on feedback and testing on feedforward; (3) training on feedforward and testing on feedback; (4) training and testing on feedback. We first converted our data from voxel to grid space using BrainVoyager QX 2.8. This approach creates 2-dimensional grids of equal number of points at all cortical depths, allowing geometrically accurate sampling of high-resolution (sub-millimetre) functional data [S12, S13]. We defined structural columns as points perpendicularly located across depth grids. To choose the data to include in the cross-layer cross-signal SVM we: independently per depth grid, selected the pixels within the retinotopic representation of the occluded region as defined by the contrast: (t target stimulus > t surround ∩ t target stimulus > 0 ∩ t full-stimulus > t occluded-stimulus). This produced one pixel subpopulation of interest per

depth grid. We then combined these subpopulations across all depths. Only pixels shared by all grid depths were included in the analysis. We trained our SVM classifier on the T values elicited by one signal (e.g. feedforward) within a depth grid, and tested on those elicited by the other signal (e.g. feedback) at all other depths. We repeated this procedure by training and testing on the four signal combinations at all depths. To test the main effects and interaction of different combinations of feedforward and feedback information (FFxFF, FFxFB, FBxFF, FBxFB) and different layers (6x6), we performed a 4 by 36 Linear Mixed-Effect Model in Matlab (The Mathworks Inc, 2014). We combined the data from 4 runs and 4 participants (i.e. using 16 data points) and estimated our fixed effects coefficients by means of maximum likelihood estimation (Subjects as random effect). Paired t-tests were carried out post hoc on significant main effects and interactions.

Supplemental References

S1. Smith, F.W., and Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. Proceedings of the National Academy of Sciences of the United States of America *107*, 20099-20103.

S2. Willenbockel V. *et al.*, (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior research methods* **42**, 671-684.

S3 Petro, L.S., Smith, F.W., Schyns, P.G., and Muckli, L. (2013). Decoding face categories in diagnostic subregions of primary visual cortex. The European journal of neuroscience *37*, 1130-1139.

S4. Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science *268*, 889-893.

S5. Muckli, L., and Petro, L.S. (2013). Network interactions: non-geniculate input to V1. Current opinion in neurobiology *23*, 195-201.

S6. Schira, M.M., Tyler, C.W., Breakspear, M., and Spehar, B. (2009). The foveal confluence in human visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 9050-9058.

S7. De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., and Goebel, R. (2013). Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. PloS one *8*, e60514.

S8. Van de Moortele, P.F., Auerbach, E.J., Olman, C., Yacoub, E., Ugurbil, K., and Moeller, S. (2009). T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. NeuroImage *46*, 432-446.

S9. De Martino, F., Moerel, M., Xu, J., van de Moortele, P.F., Ugurbil, K., Goebel, R., Yacoub, E., and Formisano, E. (2014). High-Resolution Mapping of Myeloarchitecture In Vivo: Localization of Auditory Areas in the Human Brain. Cerebral cortex.

S10. Andersson, J.L., Skare, S., and Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. NeuroImage *20*, 870-888.

S11. Lin, C.-C.C.C.-J. (2011). LIBSVM: A Library for Support Vector Machines http://www.csie.ntu.edu.tw/~cjlin/libsvm. ACM Transactions on Intelligent Systems and Technology *2*, 1.

S12. Zimmermann, J., Goebel, R., De Martino, F., van de Moortele, P.F., Feinberg, D., Adriany, G., Chaimow, D., Shmuel, A., Ugurbil, K., and Yacoub, E. (2011). Mapping the organization of axis of motion selective features in human area MT using high-field fMRI. PloS one *6*, e28716.

S13. Jones, S.E., Buchbinder, B.R., and Aharon, I. (2000). Three-dimensional mapping of cortical thickness using Laplace's equation. Human brain mapping *11*, 12-32.