



University
of Glasgow

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 (46). pp. 20099-20103. ISSN 0027-8424

<http://eprints.gla.ac.uk/46532>

Deposited on: 25 November 2010

Classification: Biological Sciences: Neuroscience.

Non-stimulated early visual areas carry information about surrounding context

Fraser W. Smith^{1,2,*} and Lars Muckli^{1,2}

¹ Centre for Cognitive Neuroimaging, University of Glasgow, Glasgow, UK,
G12 8QB.

² Department of Psychology, University of Glasgow, Glasgow, UK, G12 8QB.

* Now at the Centre for Brain & Mind, & Department of Psychology
University of Western Ontario, Canada.

Correspondence should be addressed to FWS (e-mail: fws252@gmail.com).

Centre for Cognitive Neuroimaging, Department of Psychology,
58 Hillhead Street, University of Glasgow, Glasgow, UK. G12 8QB.

Phone: +44 (0)141 330 5351; Fax: +44 (0)141 330 4606.

Keywords: early visual areas, natural visual scenes, context, linear classifier.

Abstract

Even within the early sensory areas, the majority of the input to any given cortical neuron comes from other cortical neurons. To extend our knowledge of the contextual information that is transmitted by such lateral and feedback connections we investigated how visually non-stimulated regions in primary visual cortex (V1) and V2 are influenced by the surrounding context. We used functional magnetic resonance imaging (fMRI) and pattern classification methods to show that the cortical representation of a non-stimulated quarter-field carries information that can discriminate the surrounding visual context. We further show that the activity patterns in these regions are significantly related to those observed with feed-forward stimulation and that these effects are primarily driven by V1. These results thus demonstrate that visual context strongly influences early visual areas even in the absence of differential feed-forward thalamic stimulation.

Non-stimulated early visual areas carry information about Surrounding context

\body

Introduction

It is well known that the majority of input that arrives to a specific neuron in the early visual system comes from other cortical neurons (either local or long range connections). Such connections provide one way for prior knowledge and context to modulate the responses of neurons in early vision. Yet studies which investigate the role of such mechanisms within natural vision are relatively rare. In the present experiments we devised a novel paradigm in which we analyzed the influence of surrounding context on visually non-stimulated parts of V1 and V2. We set out to investigate whether Blood Oxygen Level Dependent (BOLD) functional Magnetic Resonance Imaging (fMRI) activity in non-stimulated early visual regions carries information about a surrounding visual context. Our goal was to test the hypothesis that lateral- and feedback connections modulate, and possibly prime, regions of visual cortex by transmitting relevant contextual information.

If this hypothesis were true, then we would expect such mechanisms to work most productively when participants are presented with natural visual stimuli, as these contain a multitude of rich contextual associations across space and time (1). Hence we present participants with natural visual scenes in all the experiments reported here. Support for the above hypothesis is provided by recent demonstrations illustrating that V1 receives feedback even in regions that extend beyond the bottom-up stimulated region: for instance, the size of the activated region in retinotopic cortex corresponds to perceived size not absolute size in the context of a size illusion (2), the subjectively perceived apparent motion illusion, moreover, activates non-stimulated V1 on the motion trace (3) and object classification triggers discriminable activity at non-stimulated foveal representations (4). Other work has shown that feature based attention spreads to non-stimulated regions of V2-V4 (5). Hence contextual influences can extend beyond the range of feed-forward stimulated retinotopic space (see also

6-7). In the present work we used natural visual scenes to ask whether visually non-stimulated regions of V1 and V2 contain contextual information that permits decoding of the surrounding visual context.

Our approach consisted in presenting three different natural visual scenes (car scene, boat scene or people scene, see Supp. Fig. 1A; one individual scene of each type) with one quadrant (lower right) occluded by a uniform white field (Fig. 1A: Occluded), in a block design fMRI experiment. Participants maintained fixation and monitored the sequence of images for a color change (whole image; random occurrence within / across blocks). We independently mapped the cortical representation of the occluded quarter-field in early visual areas, **V1 and V2**, making sure to minimize the risk of any spill-over effects from surrounding regions (Fig. 1B, Methods & Supp. Text). We then trained a linear classifier to discriminate between the different scenes presented, based purely on the signal from these non-stimulated early visual regions (8-12). After training we could then use the classifier to decode which scene had been presented to observers, in an independent set of test data. If the classifier can decode above chance, this would provide evidence that activity in non-stimulated early visual regions does indeed discriminate the surrounding context. For the purposes of comparison, we also included blocks where the full visual scene was presented, i.e. the cortical representation of the occluded quarter-field received visual scene stimulation (see Fig. 1A: Control) which we refer to as the natural stimulation (Control) condition.

Results

Figure 2A shows the performance (cross-validated) of the classifiers (SVM – linear Support Vector Machine; LDA – Linear Discriminant Analysis) in decoding which scene was presented when the cortical representation of the target region received visual scene stimulation (Control condition) as a function of the number of randomly selected vertices entering the classifier (pooled here over V1 and V2). The corresponding event related average

of the fMRI signal is shown in Supp. Fig. 1B. The BOLD fMRI signal of 70 vertices was sufficient to correctly diagnose which scene was presented with, on average, 96 \pm 2.2 % accuracy (mean \pm SEM; SVM classifier, average level prediction, $t(5)=28.3$, $p=5.2*10^{-7}$; single block 81 \pm 2.8 %, $t(5)=17.2$, $p=6.11*10^{-6}$), greatly above chance (33%). More importantly, Figure 2B, shows the performance of the classifiers in decoding when the target region received *no* visual scene stimulation (Occluded condition). Performance reaches a maximum of 65 \pm 5.2 % correct here (SVM classifier, average level prediction, $t(5)=6.16$, $p=.0008$; single block 50 \pm 5.1 %, $t(5)=3.22$, $p=.012$), which again is well above chance (33%). Thus the linear classifier can decode reliably which scene has been shown to participants even when no visual scene information is presented to these regions of early visual cortex. We note that the difference in performance between the two types of classifier (SVM Vs LDA) is greatest as the number of vertices approaches the number of training examples (here 54), leading to over-fitting in the case of LDA: a well-known problem in machine learning (13). The SVM classifier, however, is a *regularized* classifier and hence does not suffer from this problem, maintaining high performance as the number of vertices increase (13-14).

Where does this contextual information arise from? In order to investigate the nature of the information driving successful classifier performance in the absence of visual scene stimulation, we explored the relation between the activity patterns observed on the two different trial types presented to observers (Control and Occluded) (11,15). We trained a new classifier on one set of trials (e.g. Control) and tested its generalization performance for the other type of trial (e.g. Occluded). If the classifier could decode above chance in this situation that would be evidence for some degree of commonality in the activity patterns generated in these regions of early visual cortex on trials where visual scene stimulation is present (Control) Vs absent (Occluded). We found our classifiers trained on Control trials could still discriminate above chance when tested with Occluded trials (see Fig 2C), reaching 70 \pm 8 % accuracy with 70 vertices (SVM classifier, average level prediction, $t(5)=4.59$, $p=.0029$;

single block prediction $47 \pm 3.7\%$, $t(5)=3.74$, $p=.0067$; chance 33%). Similarly if we trained on Occluded trials, and tested with Control trials (see Fig 2D), the classifiers could also perform well above chance ($79 \pm 8.9\%$ accuracy, SVM classifier, average level prediction, $t(5)=5.19$, $p=.0017$; $57 \pm 4.1\%$ single block prediction, $t(5)=5.71$, $p=.0012$; chance 33%). The linear classifiers thus show that the activity patterns elicited in the presence Vs absence of visual scene stimulation are similar to a non-trivial degree. The reason we observe higher performance in trial type generalization (Fig 2C-2D) compared to within trial generalization in the Occluded condition (Fig 2B) could well be due to having a greater number of trials to train and test the classifier with in the trial type generalization (72 for training and testing in the latter, 54 and 18 in the former).

We have shown that non-stimulated regions of early visual cortex carry information about surrounding context and further that the activity patterns in these non-stimulated regions are similar on Occluded and Control trials. The previous experiment utilized a block design paradigm, which leads to robust pattern estimates due to averaging over many volumes for each stimulus presentation, but also suffers from the problem that it is unclear how what happens on a block of stimulation (12s here) compares to any one trial of perception (i.e. the standard design in visual cognition). Within the duration of a block, for instance, it is unclear which other cognitive processes the subject engages in therefore it is beneficial if the result can be replicated in single trial conditions. Thus, in order to increase the generalizability of our results, we conducted a second experiment with a rapid-event related design (4s ITI), with a different task (one back repetition detection) and seven new participants. Figures 2E-H show the results of this experiment. On Control trials (Fig 2E), the linear classifier reached $79 \pm 5.5\%$ correct performance (SVM classifier, average level prediction, $t(6)=8.41$, $p=7.71 \times 10^{-5}$; $48 \pm 1.5\%$ single trial, $t(6)=9.76$, $p=3.32 \times 10^{-5}$; chance 33%). Most importantly, the classifier again successfully decoded which scene had been shown on Occluded trials (Fig 2F), reaching performance of $53 \pm 8.4\%$ correct (SVM, average level; $t(6)=2.30$, $p=.031$). Note that the classifier also reliably decoded Occluded trials based purely

on single trial data (39 ± 2.8%, SVM; chance 33%; $t(6)=2.06$, $p=.043$). In order to corroborate this result, the weakest we report in the current manuscript, we conducted a group level permutation test (see Supp. Text) that gave $p = 1 \times 10^{-4}$. Thus with this different fMRI design, task and participants the linear classifier successfully decoded which Occluded scene was presented on the basis of just 1 second of stimulation. We note, however, that the advantage of the SVM over LDA is much reduced in the present experiment (in fact LDA outperforms or is equal to SVM here) compared to Experiment 1. The reasons for this are not clear: one influential factor is likely to be the many more training observations in the present experiment (minimum of 180 here) compared to Experiment 1 and as such over-fitting is much less of a problem for the LDA classifier. A second factor is the greater inter-subject variability for the SVM than LDA in Experiment 2. In any case, taken together we have strong evidence that a linear classifier can indeed discriminate between the different scenes presented even when the target region receives no visual scene stimulation.

We finally used the event related data and tested the performance of the linear classifier in generalizing across trial types, and we again found that the classifiers could generalize successfully (Figs 2G-H): the classifier trained on Occluded trials but tested with Control trials resulted in performance of 69 % ± 12.4 % average level prediction (SVM classifier, chance 33%; $t(6)=2.84$, $p=.015$) and 42 ± 3.3 % single trial prediction (SVM $t(6)=2.75$, $p=.017$). The classifier trained on Control trials but tested with Occluded trials, on the other hand, had 61 ± 13 % (SVM; average level; $t(6)=2.17$, $p=.037$) and 41 ± 3.2 % correct (SVM; $t(6)=2.45$, $p=.025$) in single trial prediction). Thus we have successfully replicated the main pattern of results of Experiment 1 with this new experiment.

We have shown in two independent experiments that non-stimulated regions of early visual cortex do indeed carry information about the surrounding context. Furthermore we have shown that the activity patterns in these non-stimulated regions are similar on Occluded and Control trials. What is the contribution of each visual area to the performance we observe? To explore this question we re-ran our classifier analyses independently for patches

of V1 and V2 (see Supp. Text - Methods). Figure 3A shows the asymptotic performance (defined as performance with the maximum number of vertices) obtained by the SVM classifier (average level prediction) for each visual area (V1, V2 or pooled), for the three types of analysis: Control, Occluded or Generalization of Experiment 1. Figure 3B shows the same for Experiment 2 (see Supplementary Figs 2 & 3 for the full classifier results for both Experiments split by visual area; and Supplementary Tables 1-3 for vertex details). It is clear from these plots that, across both experiments and the different analyses, V1 greatly outperforms V2 – reaching or even exceeding the performance obtained by pooling across V1 and V2, with a smaller number of vertices. Thus, V1 responses are the primary determinant of the ability of our classifiers to successfully discriminate between the three scenes, regardless of the specific trial type considered (Control, Occluded or Generalization analyses).

What kind of mechanism could be responsible for these effects? One plausible explanation would be the presence of an auto-associative memory-based process, active at the level of early visual areas (16-18). As the experiment progresses, relevant early visual neurons would develop a strong expectation of the structure that is present in the occluded region of the stimuli, obtained from presentation of that information on Control trials. Hence on Occluded trials, the three-quarters of the stimulus which is presented would be a very good cue to *retrieve* the remaining, occluded, part of the stimulus.

In order to test this possibility, we conducted a third experiment where participants were only ever presented with Occluded trials (i.e. they never saw the actual visual scene information of the Occluded region). We present the results of this experiment in Figure 4 (A-D). For three out of the four participants clear discrimination can be seen (SVM average level prediction). For the remaining participant the SVM classifier (average level) does not perform well, however, the LDA classifier suggests that discrimination ability is present even within this subject. Thus the effect we report clearly goes beyond auto-associative recall based on memories of the actual visual scene information of the Occluded region.

One additional explanation is that information regarding the distribution of low-level visual features, e.g. luminance & contrast; energy at each spatial frequency and orientation, in the stimulated area is transmitted to the non-stimulated region. To address this possibility we additionally ran two more participants on a modified version of Experiment 3 where we explicitly controlled these low-level image properties within the naturally stimulated region (see Fig 4E & Supp. Text - Low Level Image Control). Performance clearly remained well above chance (Fig 4F-G) despite our control of these low-level image properties, demonstrating that the observed contextual sensitivity is not carried by these specific low-level visual properties.

Finally, to what extent could spreading activity (lateral connections) or spill-over activity account for the present results? We performed an additional analysis investigating the weights from the pattern classification analyses to address this question (see Supp. Text – Weight Analysis). Specifically we correlated the absolute value of the weights, from the SVM and LDA classifiers leading to maximum performance on Occluded trials, with the t values of both the Target and the Surround mapping conditions. The absolute value of the weight at each vertex indicates the relative influence of each vertex to the classifier's solution (e.g. 11, 19). The logic is that if high (absolute) weights are correlated with high Surround t values, that might suggest a possible influence of spill-over signal (or spreading activity). On the other hand, a positive correlation with the Target condition would indicate that the more important vertices in the classifier decision function are those with a strong signal to the Target stimulus, which we might expect to be the case purely on signal to noise considerations. While the results differed between classifiers (LDA & SVM, see Supp. Text – Weight Analysis), crucially the results with the LDA classifier confirmed that it is possible to decode the surrounding context with a set of weights that have no significant positive relationship to the Surround (in fact a suggestive negative or negative correlation) and a significant positive relation to the Target stimulus (Experiments 1 and 2). The results with the

LDA classifier, thus speak against interpretations of our data in terms of simple spill-over activity.

Discussion

Thus, in sum, across four independent experiments we have shown that non-stimulated early visual regions (V1 & V2) carry information about a surrounding visual context. We have shown, moreover, that the activity patterns in the non-stimulated early visual regions are significantly related to the activity observed in those regions under feed-forward stimulation (Experiments 1 and 2). We have further shown that this effect is largely driven by V1 (Experiments 1 and 2), that it does not depend on memory of complete images (Experiment 3) and that it is not carried by several basic low-level visual properties (Experiment 4). Finally with the classifier weight analysis we have provided some reassurance that simple spreading activity or spill-over activation from feed-forward stimulated neighbouring regions is unlikely to account for the results we report here, at least for the LDA classifier.

What type of mechanism could account for all experiments we report? The question concerns how information can be transmitted from regions representing the surrounding context to regions representing the non-stimulated quadrant. Bayesian models of human vision (20-22) offer one possible explanation in which the surrounding visual context would bias cortical feedback to the non-stimulated early visual areas. Such feedback could be spatially precise resulting in something like “filling in”, or spatially diffuse, e.g. leading to general expectation of some property in the target region (i.e. categorical; human bodies or cars etc, see 1,23). Although we did not set out to directly test such predictive accounts here, our paradigm could be adapted in order to test for prediction error signals in the target region of V1 after inducing appropriate contextual expectations in the surround (see 24-25).

The type of visual information that is transmitted to the target region, however, could in principle concern low-level visual features (e.g. distributions of contrast, luminance, energy, spatial frequencies, orientations etc), or as mentioned above, higher-level visual features (such as contours, categories). Although we demonstrated in Experiment 4 that several basic low-level visual features (global luminance & contrast, energy at each spatial frequency and orientation) were not responsible for generating the observed context sensitivity, there are many other low level features that remain to be tested (e.g. local and global distribution of orientations) before we could exclude a mechanism based on surround signals carrying information about distributions of low level visual information (local or global) to the non-stimulated areas. Further, while we have demonstrated a significant degree of “similarity” between the responses present in the target region as a result of feed-forward or surround only stimulation (Experiments 1 & 2), this does not imply that the signals necessarily reflect identical information. Thus the nature of the information that is carried by the surround signal to the non-stimulated area is a question that still requires further study.

One additional explanation of our findings concerns the possible allocation of attention to expected features in the occluded region. An intriguing report (5) showed that feature based attention spreads even to non-stimulated parts of the visual field. The authors showed that attended direction of motion (45 vs. 135 deg.) in the stimulated visual field (Left or Right) could be decoded from ipsi-lateral and hence non-stimulated early visual regions, specifically V2, V3 and V4. Based on this report, one could propose that attention is allocated to certain locations in the occluded region, where participants expect to find object features. These locations would differ across the three different scenes presented to participants, and could hence be responsible for the classification results that we observe. Note that this account is not independent of the predictive accounts given earlier, but shifts the emphasis to neural mechanisms that implement visual attention.

At a neurophysiological level, it is known that most of the input any given V1 neuron receives comes from intra- and inter-areal projections, with the minority coming from

thalamic afferents (26-27). As such, monosynaptic lateral connections (i.e. single synapse intra-cortical projections) might play a role in explaining how contextual information is promulgated from an area of cortex which receives feedforward stimulation to an area that does not (a distance here of 2 deg on the fixation diagonal), possibly by transmitting scene-specific information across the low contrast summation fields of V1 neurons (28-29). These connections, however, are unlikely to account for all the activity in non-stimulated V1 as they are not wide enough to span the entire non-stimulated region. Furthermore, laterally transmitted effects would be predicted to be strongest close to the border of the occluded region which disagrees with the results of the classifier weight analyses (see Supp. Text Weight Analysis). Inter-areal projections (i.e. cortical feedback), hence are likely to be critical in explaining the present effects.

A potential mechanism through which cortical feedback might contribute to the observed context effect would be by direct connections between mirror-symmetric parts of the visual field, as reported in primate MT (30). Analysis of spontaneous BOLD fMRI fluctuations in anaesthetized primates has also revealed correlations between mirror-symmetric parts of the visual field, pronounced along the horizontal meridian (31). Such mirror-symmetric sensitivity in higher visual areas could certainly play a role in explaining where such cortical feedback originates. Although at present we cannot disentangle the contribution from cortical feedback and lateral interaction, we believe that the fact that such information is transmitted at all, from stimulated to non-stimulated regions, highlights the ubiquitous role that context might play in shaping the activity of *all* neurons in early vision. Indeed, several authors have recently pointed out that to increase our understanding of V1, beyond the best current models, it is fundamentally a deeper consideration of the role of context (both spatial and temporal) and how such contextual information is transmitted that is needed (32-34).

In the present experiments we found that V1 was the primary driving force for the context effect. Indeed there was barely any evidence for context sensitivity in V2 (Fig 3).

While in most cases significant classification performance in V1 co-occurs with similar performance in V2 (9, 10 –Ex. 1,15,35,36), these studies typically reflect above threshold stimulation protocols. Interestingly, one exception (10; Ex. 2) involved decoding subjectively invisible stimulus orientation and led to significant decoding only in V1, mirroring the present results. It has been shown, furthermore, that V1 contains more reliable information for binary image reconstruction than V2 (36). Taken together, these latter results make the asymmetry that we found in performance across V1 and V2 less surprising.

Much recent work has demonstrated that a multitude of important visual information can be extracted from the fMRI BOLD response in early visual areas, when taken as a multivariate quantity, allowing for discrimination between visual features (such as orientation and motion direction (9,35)), discrimination between scene categories (37), discrimination of the identity of natural images (38), and even allowing for the reconstruction of arbitrary binary contrast images (36). Here, extending these lines of research, we have shown that discrimination between individual scenes is possible even in non-stimulated regions of primary visual cortex.

Taken together, we believe that our data demonstrate the crucial role of context, carried by the physiological mechanisms of feedback and / or lateral interaction, in driving activity even in non-stimulated early visual areas. We further believe these data to be consistent with theoretical views concerning the importance of predictive codes in the visual system (16,20, 39-40).

Methods

Subjects

Six subjects took part in Exp. 1, seven in Exp. 2, four in Exp. 3 and two in Exp. 4 (Supp. Text).

Paradigm

Participants were presented with occluded natural visual scenes (Fig 1A). There were three such scenes, one instance of each (see Supp. Fig. 1A). In Experiments 1 & 2, the same scenes were also presented non-occluded as a control. Participants had to maintain fixation and either detect a frame color change (Exp. 1, – block design), perform one-back repetition (Exp. 2 – rapid event related) or detect colors change of the fixation marker (Exp. 3 & 4 – block design). Independently, we mapped the cortical representation of the target region in V1/V2 (Fig 1B & Supp. Text) by presenting contrast reversing checkerboards (4 Hz) in either the target or the surround position (Fig 1B) in a standard block design experiment. Vertices were selected from the defined regions of V1 and V2 that met the following criteria: significant effect to Target alone and non-significant effect to the Surround alone (Supp. Text).

Structural and Functional MRI.

MRI was performed at 3Tesla using standard MRI parameters. Anatomical data were transformed to Talairach-space, and the cortical surface was reconstructed. Functional MRI time series were preprocessed using standard parameters (no smoothing) and coaligned to anatomical data-set (Supp. Text). A GLM was used to estimate the activity patterns on each single block (or trial).

Pattern Classification

Independently for occluded and control trials, we trained two linear pattern classifiers (Supp. Text) to learn the mapping between a set of brain activity patterns and the presented scene. We then tested the classifiers on an independent set of test data (leave one run out cross-validation). We chosen input features (vertices) randomly from the set that met our mapping criteria, initially pooled across V1/V2 but also later split by V1/V2 (Supp. Text). We also performed trial type generalization analyses where the classifiers (in Exp. 1 & 2) were trained on one type of trial (e.g. Occluded) and tested on the other (e.g. Control).

Acknowledgements

This work was funded by Biotechnology and Biological Sciences Research Council Grant No. BB/G005044/1. We thank Francisco Pereira for discussion of classifier weight analyses and Chih-Jen Lin (libSVM) for technical assistance.

References

- 1) Bar M (2004) Visual objects in context *Nat Rev Neurosci* 5:617-629.
- 2) Murray SO, Boyaci H, Kersten D (2006) The representation of perceived angular size in human primary visual cortex *Nat Neurosci* 9:429-434.
- 3) Muckli L, Kohler A, Kriegeskorte N, Singer W (2003) Primary visual cortex activity along the apparent-motion trace reflects illusory perception *Plos Biology*, 3:e265.
- 4) Williams MA et al (2008) Feedback of visual object information to foveal retinotopic cortex *Nat Neurosci* 11:1439-1445.
- 5) Serences JT, Boynton, GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron*, 55:301-312.
- 6) Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. *J Neurosci*. 30:2960-6.
- 7) 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* 20:131-139.
- 8) Cox DD, Savoy RL (2003) Functional Magnetic Resonance Imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex *Neuroimage* 19:261-270.
- 9) Kamitani Y, Tong F (2005) Detecting the visual and subjective contents of the human brain *Nat Neurosci* 8:679-685.
- 10) Haynes JD, Rees G (2005) Predicting the orientation of invisible stimuli from activity in human primary visual cortex *Nat Neurosci* 8:686-691.
- 11) Formisano E, De Martino F, Bonte M, Goebel R (2008) “Who” is saying “What”? Brain-based decoding of human voice and speech *Science* 322:970-973.
- 12) Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data *TICS* 10:424-430.

- 13) Duda RO, Hart PE, Stork DG (2001) *Pattern Classification* (John Wiley & Sons, New York, ed 2).
- 14) Hastie T, Tibshirani R, Friedman J (2001) *The Elements of Statistical Learning* (Springer-Verlag, New York Berlin Heidelberg).
- 15) Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas *Nature* 458:632-635.
- 16) Hawkins J, Blakeless S (2004) *On Intelligence* (Times Books).
- 17) McClelland JL, McNaughton BL, Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory *Psych Rev* 102:419-457.
- 18) Nadel L, Samsonovich A, Ryan L, Moscovitch M (2000) Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus* 10:352-368.
- 19) Mourao-Miranda J, Bokde AL, Born C, Hampel H, Stetter M (2005) Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. *Neuroimage* 28:980-995.
- 20) Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex *J Opt Soc Am A* 20:1434-1448.
- 21) Yuille A, Kersten D (2006) Vision as Bayesian inference: analysis by synthesis *TICS* 10:301:308.
- 22) Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference *Annu Rev Psych* 55:271-304.
- 23) Peelen MV, Fei-Fei L, Kastner S (2009) Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460:94-7.
- 24) den Ouden HE, Friston KJ, Daw ND, McIntosh AR, Stephan KE (2009) A dual role for prediction error in associative learning. *Cereb Cortex* 19:1175-85.

- 25) den Ouden HE, Daunizeau J, Roiser J, Friston KJ, Stephan KE (2010) Striatal prediction error modulates cortical coupling *J Neurosci*. 30:3210-9.
- 26) Douglas RJ, Martin KAC (2004) Neuronal circuits of the neocortex. *Ann. Rev. Neurosci* 27:419-451.
- 27) Douglas RJ, Martin KAC (2007) Mapping the matrix: The ways of neocortex. *Neuron* 56:226-238.
- 28) Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons *Prog Brain Res* 154:93-120.
- 29) Angelucci A et al (2002) Circuits for local and global signal integration in primary visual cortex *J Neurosci* 22:8633-8646.
- 30) Maunsell JH, Van Essen DC (1983) The connections of the middle temporal area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3:2563-2586.
- 31) Vincent JL et al (2007) Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* 447:83-86.
- 32) Carandini M et al (2005) Do we know what the early visual system does? *J Neurosci* 25:10577-10597.
- 33) Masland RH, Martin P (2007) The unsolved mystery of vision *Curr Biol* 18:R577-R582.
- 34) Olshausen BA, Field DJ (2005) How close are we to understanding V1? *Neural Comput* 17:1665-1699.
- 35) Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex *Curr Biol* 16:1096-1102.
- 36) Miyawaki Y et al (2008) Visual image reconstruction from human brain activity using a combination of multiscale local image decoders *Neuron* 60:915-929.

- 37) Walther DB, Caddigan E, Fei-Fei L, Beck DM (2009) Natural scene categories revealed in distributed patterns of activity in the human brain *J Neurosci* 29:10573-10581.
- 38) Kay KN, Naselaris T, Prenger RJ, Gallant JL (2008) Identifying natural images from human brain activity *Nature* 452:352-355.
- 39) Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects *Nat Neurosci* 2:79-87.
- 40) Mumford D (1992) On the computational architecture of the neocortex: The role of cortico-cortical loops *Biol Cybern* 66:241-251.

Figure Legends

Fig 1: Experimental Design and Cortical Mapping Procedure.

A) The two conditions of visual stimulation shown in the current experiments. On Occluded trials (left), the lower right quadrant of the image was occluded with a uniform white field. On Control trials (right) there was no such occlusion. The same three individual scenes were shown in both Occluded and Control conditions. The participant always had to fixate the central fixation marker (very small checkerboard) and either detect a frame colour change (Experiment 1) or perform a one back repetition task (Experiment 2). Note that the black bars highlighting the occluded region are for display purposes only. The whole image spanned 22.5 by 18 degrees, with the occluded region spanning approximately 11 by 9 degrees.

B) Cortical mapping procedure used to map the retinotopic representation of the Occluded region. In a standard block design protocol, participants were presented with contrast reversing checkerboards in either the Target (green) or the Surround (red) area of the occluded portion of the stimulus. We defined a patch of V1 and a patch of V2 on the inflated cortical surface (left hemisphere; for one representative participant) from the contrast of Target minus Surround. Any vertex taken as representing the occluded portion of the stimulus had to meet two additional constraints: significant effect to the Target alone (Target > baseline; $t > 1.65$) and, crucially, a non-significant effect to the Surround stimulus alone (Surround = baseline; absolute $t < 1.65$).

Fig 2: Pattern Classification analysis for Experiments 1 & 2.

A-D) Classifier performance for Experiment 1 (Block design). The left panel (A) shows performance (percentage correct) for two linear classifiers (Linear Discriminant – LDA;

Support Vector Machine – SVM) in decoding which scene was presented on *Control* trials as a function of the number of vertices entering the classifier, for both average level (Av.) and single block prediction (SB). Note vertices are pooled across V1 and V2 in this analysis. Performance is averaged across participants (error bars represent one standard error of the mean). Chance performance is indicated by the dark green bar at 33%. The second panel (B) shows the same for *Occluded* trials. The next two panels (C & D) show the results for the trial type generalization analyses (see text).

E-H) Classifier performance for Experiment 2 (Rapid Event Related). Panels are arranged as above though here we report performance for Average (Av.) and Single *Trial* (ST) prediction.

Fig 3: Asymptotic performance of the SVM classifier in Experiments 1 & 2.

Asymptotic classifier performance (average level prediction) for each visual area (V1, V2 and pooled) and each different analysis (Occluded trials, Control trials, and Generalization 1 [G1: train Control, test Occluded], Generalization 2 [G2: train Occluded, test Control]) for Experiments 1 (Panel A) and 2 (Panel B). Note asymptotic is defined as performance for the maximum number of vertices (here 70 for V1 & V2 pooled, 30 for each area considered independently). Error bars show one standard error of mean across participants.

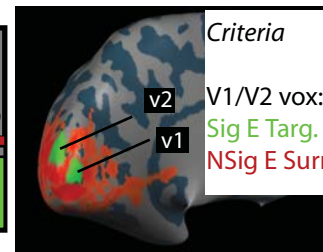
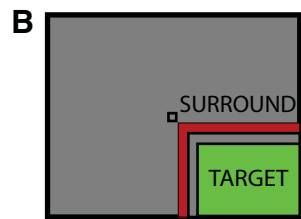
Fig 4: Pattern Classification analysis for Experiment 3 & 4.

A-D) Classifier performance per participant for Experiment 3. Panel (A) shows performance (percentage correct) for two linear classifiers (Linear Discriminant – LDA; Support Vector Machine – SVM) in decoding which scene was presented on *Occluded* trials as a function of the number of vertices entering the classifier, for both average level (Av.) and single block

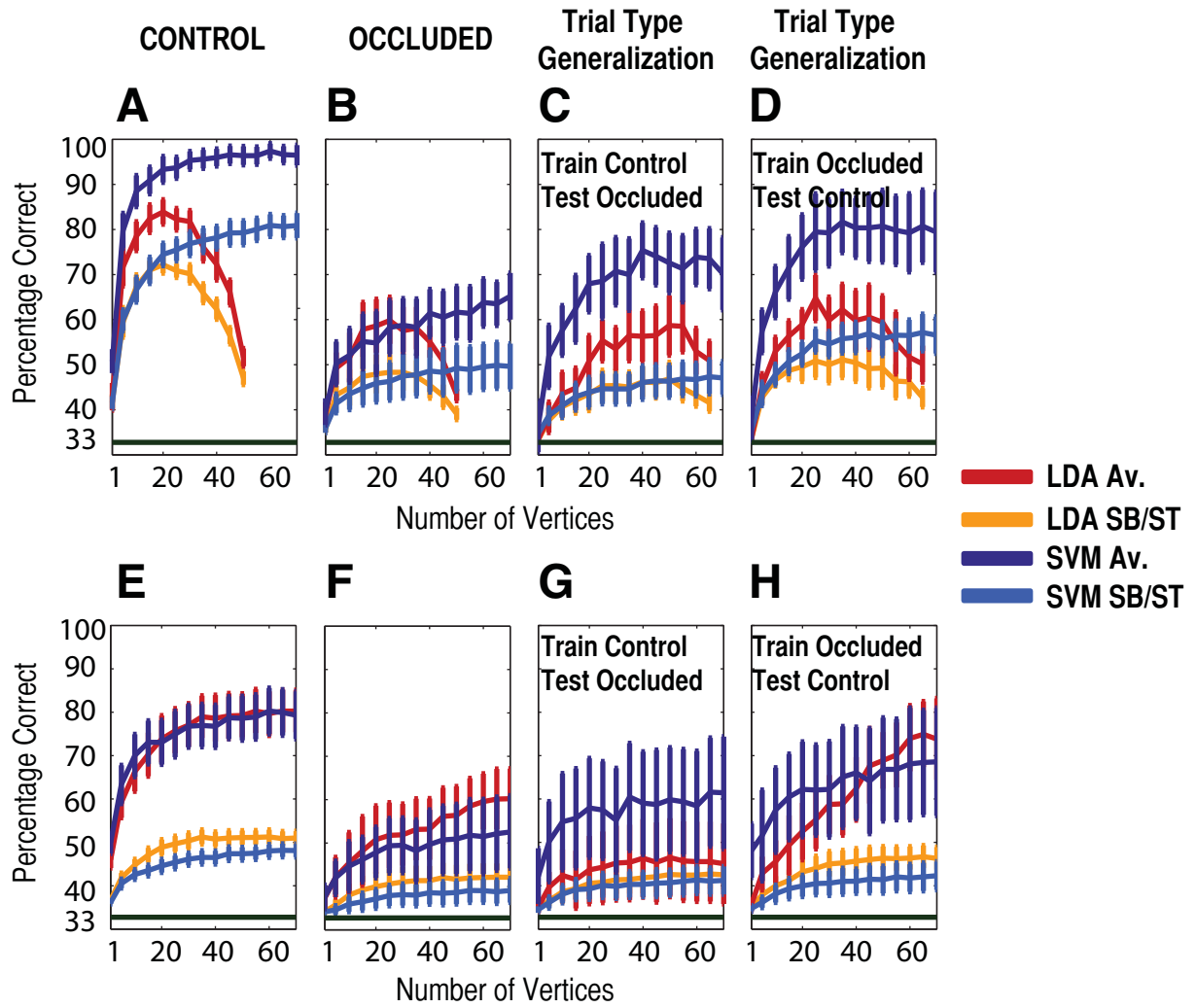
prediction (SB), for subject 1. Note vertices are pooled across V1 and V2 in this analysis. Performance is averaged across sampling iterations and cross-validation cycles (error bars represent one standard error of the mean). Chance performance is indicated by the dark green bar at 33%. Panels B-D show the same for the remaining participants in Experiment 3. Panel (E), second row, shows an example of a scene stimulus shown in Experiment 4 (see Supp. Text. Low Level Image Control). Panels (F-G) show performance for the two participants who took part in Experiment 4 (layout as in A-D).

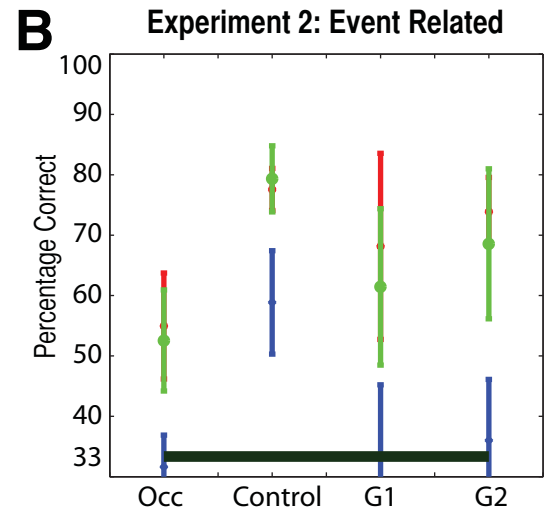
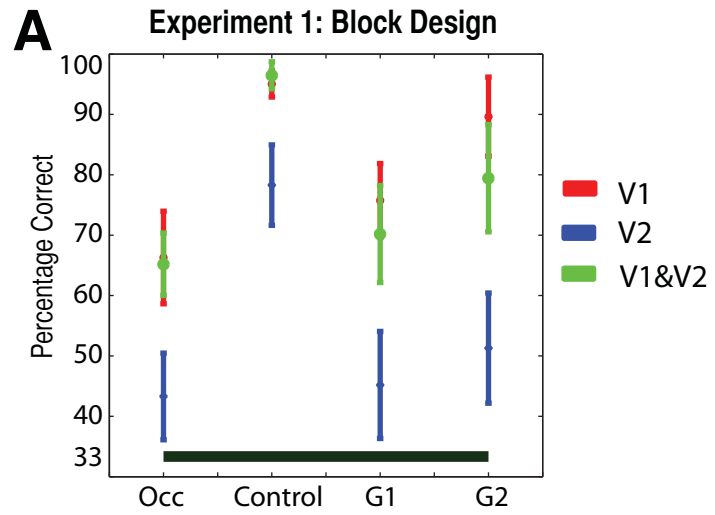
OCCLUDED

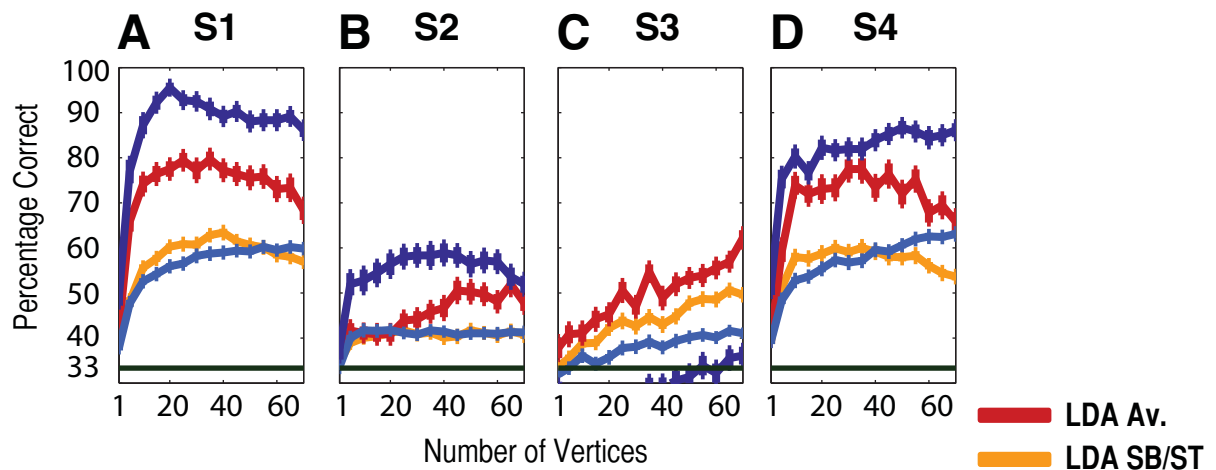
CONTROL



Target - Surround







E



F S1

G S2

