

# Top-down attention switches coupling between low-level and high-level areas of human visual cortex

Naseem Al-Aidroos<sup>a,1</sup>, Christopher P. Said<sup>b</sup>, and Nicholas B. Turk-Browne<sup>a,1</sup>

<sup>a</sup>Department of Psychology, Princeton University, Princeton, NJ 08540; and <sup>b</sup>Center for Neural Science, New York University, New York, NY 10003

Edited by Thomas D. Albright, The Salk Institute for Biological Studies, La Jolla, CA, and approved July 26, 2012 (received for review February 7, 2012)

**Top-down attention is an essential cognitive ability, allowing our finite brains to process complex natural environments by prioritizing information relevant to our goals. Previous evidence suggests that top-down attention operates by modulating stimulus-evoked neural activity within visual areas specialized for processing goal-relevant information. We show that top-down attention also has a separate influence on the background coupling between visual areas: adopting different attentional goals resulted in specific patterns of noise correlations in the visual system, whereby intrinsic activity in the same set of low-level areas was shared with only those high-level areas relevant to the current goal. These changes occurred independently of evoked activity, persisted without visual stimulation, and predicted behavioral success in deploying attention better than the modulation of evoked activity. This attentional switching of background connectivity suggests that attention may help synchronize different levels of the visual processing hierarchy, forming state-dependent functional pathways in human visual cortex to prioritize goal-relevant information.**

category selectivity | functional MRI | goal-directed attention | retinotopic occipital cortex | ventral temporal cortex

The ventral visual stream, the neural substrate of object perception (1), is organized hierarchically. At early stages, occipital cortex decomposes visual images into simple features, such as form and orientation (2). At later stages, ventral temporal cortex combines these features into complex objects, such as faces and scenes (3). Although this hierarchy is hard-wired (4), human vision is flexible: our goals and intentions determine what we see via top-down attention (5). How does top-down attention prioritize goal-relevant information in the ventral visual stream?

The conventional answer is that attention prioritizes certain information by enhancing evoked responses in cortical areas that represent this information. For example, when attending to faces (e.g., when looking for a friend in a crowd, or in our study, monitoring for a repeated face in a stream of composite images that contain both a face and a distracting scene), the response of the fusiform face area (FFA) to faces is enhanced; in contrast, when attending to scenes (e.g., when looking for a restaurant in a new town, or in our study, monitoring for a repeated scene in the composite images), the response of the parahippocampal place area (PPA) to scenes is enhanced (6). This attentional modulation is interpreted as resulting from top-down selection of goal-relevant information and relative inhibition of goal-irrelevant information. Similar effects have been observed throughout visual cortex and with diverse methodologies, including positron emission tomography (7), functional magnetic resonance imaging (fMRI) (6, 8), and single-cell recordings (9, 10). By strengthening representations, top-down attention may ensure that goal-relevant information competes better against goal-irrelevant information and wins the battle for conscious awareness (11).

We propose a complementary but different answer: attention may prioritize information by strengthening the coupling or interaction between goal-relevant areas of human visual cortex. This hypothesis is motivated by theoretical models of cognitive control, such as guided activation theory, in which internal goals specify mappings of weights between posterior brain regions to route

information through goal-relevant pathways (12, 13). Moreover, this hypothesis is prompted by neurophysiological findings from nonhuman primates (14) that attention coordinates neural activity between extrastriate and frontoparietal cortex by coupling neuronal spikes with local field potentials (LFPs) (15, 16) and increasing long-range LFP coherence (17, 18). Here we exploit the spatial coverage and noninvasiveness of fMRI to assess how top-down attention modulates interactions between posterior visual areas of the human brain. We develop an fMRI analysis technique to test the prediction that selective attention to a visual category (faces or scenes) enhances the coupling of occipital areas that process low-level features common to all categories (V1, V2, V3, V4) with the ventral temporal area selective for objects of that category (FFA or PPA, respectively; Fig. 1A).

To measure coupling, we assessed how intrinsic blood oxygen level-dependent (BOLD) activity correlated among brain regions. Such correlations have been used extensively to infer the latent functional architecture of the brain at rest (19, 20)—without external stimulation, the only reason for two regions to express a common signal is if they interact directly or indirectly. Although such approaches have been very successful, they are best suited to studying the resting state rather than specific attentional states. To draw strong conclusions about the relationship of brain interactions to attention requires measuring correlations while subjects perform attention tasks.

The interpretation of BOLD correlations during tasks, however, is complicated by evoked responses (20, 21). Consider our attention task: a subject is instructed to attend to scene information and is shown images containing a face and a scene (Fig. 1B). Our hypothesis predicts that correlations between occipital cortex and the goal-irrelevant FFA should not increase. However, because the FFA responds somewhat to most external stimuli, and this response is time-locked to responses in occipital cortex, occipital-FFA correlations will strengthen (Fig. S1). The challenge for measuring how the goal of attending to scenes modulates connectivity is thus to distinguish intrinsic correlations related to sustaining this goal (state-dependent correlations) from those related to synchronized stimulus-evoked responses (stimulus-driven correlations). Our solution is to project stimulus-evoked responses out of the data with linear regression and measure correlations in the residual spontaneous fluctuations (Fig. 1B and Fig. S2). This extraction is possible because evoked and spontaneous signals are linearly superimposed in human fMRI data (22). The resulting *background connectivity* can be used to assess how the functional architecture of the human brain is influenced by attentional state.

Regions of interest (ROIs) in occipital and ventral temporal cortex were identified in individual subjects. To establish

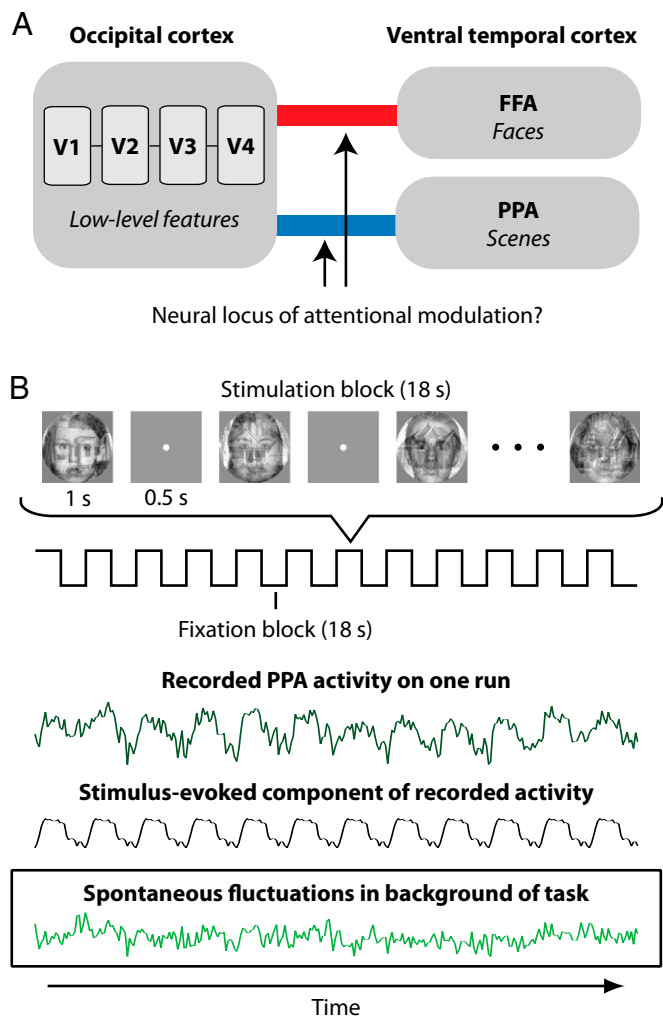
Author contributions: N.A.-A., C.P.S., and N.B.T.-B. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: naseem@uoguelph.ca or ntb@princeton.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1202095109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1202095109/-DCSupplemental).



**Fig. 1.** Study overview. (A) We hypothesize that top-down attention to an object category enhances interactions between retinotopic areas in occipital cortex selective for low-level features and the area of ventral temporal cortex selective for the attended category. (B) Attention was directed to the face or scene component of composite stimuli. Time series for each attentional state were extracted from all ROIs (example subject, dark green). Stimulus-evoked responses (gray) were projected from the time series, isolating spontaneous fluctuations (light green). We analyze whether correlations of these fluctuations between occipital and temporal ROIs depend on attentional state.

a baseline, subjects first completed two *rest* runs with no task. Subjects then completed two *face-attention* runs and two *scene-attention* runs. In both conditions, blocks of composite images containing both a face and a scene were presented centrally (Fig. 1B), with the streams of faces and scenes generated independently with occasional back-to-back repetitions. Subjects detected the face repetitions in face-attention runs and the scene repetitions in scene-attention runs. Because repetitions were always present in both categories, successful behavioral performance required detecting repetitions in one category, while ignoring those in the other. This type of selective task over multiplex stimuli is commonly used to study top-down attention (6–8, 21). Thus, subjects sustained a single attentional state for entire runs, prioritizing one category of information over the other.

Our central prediction is that attention to visual categories will modulate coupling between occipital and ventral temporal cortex. In occipital cortex, we focus on human area V4 because of its proximity to FFA and PPA—in terms of both physical distance

(4, 23) and representational complexity (24)—and because V4 is strongly affected by attention (8, 9, 11, 15). Thus, if attention operates by regulating interactions within visual cortex, then attention to a category should be associated with increased background connectivity between V4 and the goal-relevant region of ventral temporal cortex.

## Results

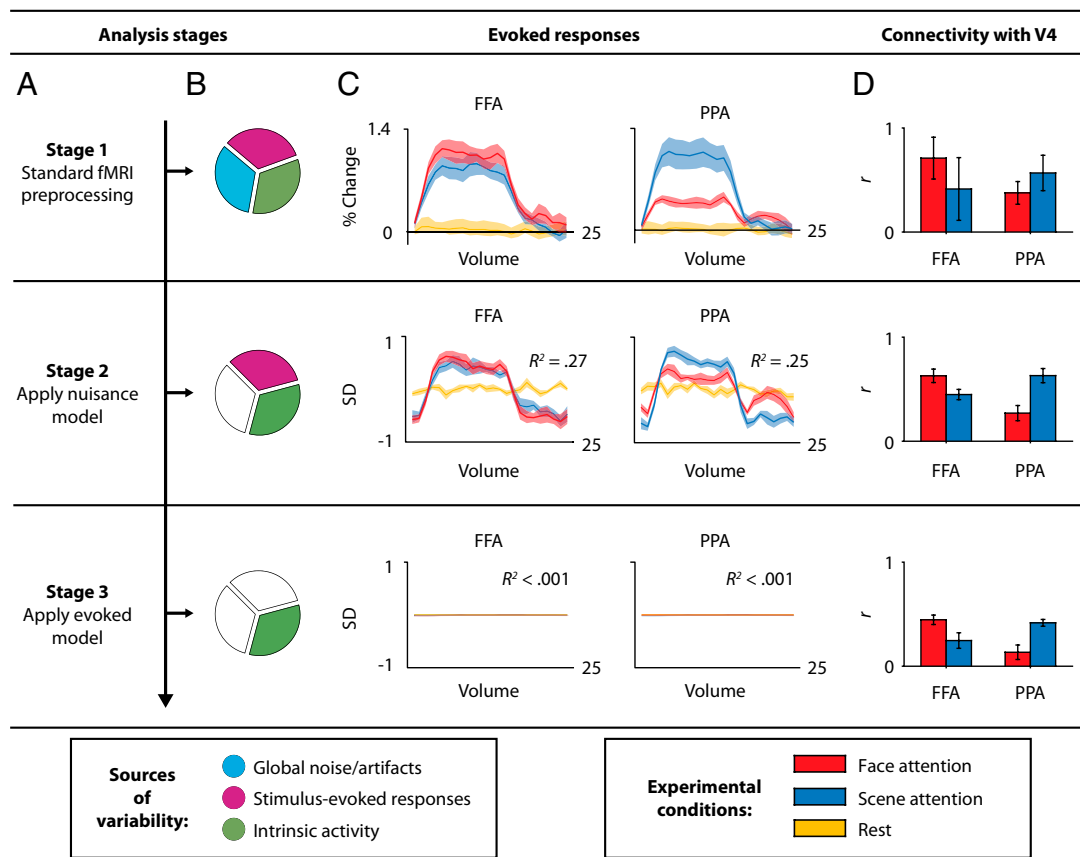
Behavioral performance in the attention task was highly accurate in terms of sensitivity (mean  $d' = 3.6$ ;  $t_6 = 17.46$ ,  $P < 0.001$ ). The hit rate for detecting goal-relevant repetitions was not at ceiling (mean = 85.7%), suggesting—along with the need to selectively attend and respond to stimuli from one category while inhibiting stimuli from the other—that the task was generally demanding of attention. Behavioral performance did not reliably differ by attentional state (face-attention mean  $d' = 3.5$ ; scene-attention mean  $d' = 4.1$ ;  $t_6 = 1.94$ ,  $P = 0.101$ ).

As predicted, however, background connectivity did depend on attentional state (ROI  $\times$  state interaction:  $F_{1,6} = 22.39$ ,  $P = 0.003$ ; Fig. 2A): V4–PPA connectivity was stronger under scene attention than face attention ( $t_6 = 4.22$ ,  $P = 0.006$ ), and V4–FFA connectivity was stronger under face attention than scene attention ( $t_6 = 3.12$ ,  $P = 0.020$ ). This interaction was not limited to area V4, extending to earlier occipital areas (V1–V3: all  $P$  values  $< 0.057$ ; Fig. 2C), nor was it limited to the FFA and PPA, with attention modulating connectivity between V4 and most voxels in occipitotemporal cortex that showed category selectivity for faces or scenes (Fig. S3). Because our stimuli, ROIs, and analyses were identical across attention runs, these modulations of background connectivity can only be attributed to the manipulation of top-down attentional state.

If background connectivity reflects the implementation of attentional goals, then it should have functional significance for subjects' ability to perceive goal-relevant information. Indeed, although our sample size was small, individual differences in attentional modulation of background connectivity between V4 and FFA/PPA correlated strongly with behavioral performance on the attention tasks ( $r = 0.91$ ,  $P = 0.004$ ; Fig. 2B). There was also a moderate but nonsignificant correlation between attentional modulation of evoked responses and behavior ( $r = 0.43$ ,  $P = 0.341$ ). The correlation of V4 connectivity with behavior remained significant, however, after partialing out individual differences in evoked-response modulation ( $r = 0.89$ ,  $P = 0.017$ ). Although residuals are typically treated as error or noise when modeling evoked responses in cognitive neuroscience, they accounted for 79% of the behavioral variance unexplained by evoked responses in our study.

Is attentional modulation of connectivity a consequence of attentional modulation of evoked responses? For example, enhanced FFA and PPA responses under face attention and scene attention, respectively, may have increased interactions with occipital cortex through feedback (26). Our data are inconsistent with this possibility (Fig. 3). First, we observed modulation of connectivity during nonstimulated periods between blocks (ROI  $\times$  state interaction:  $F_{1,6} = 15.08$ ,  $P = 0.008$ ), when there was essentially no evoked response, and thus no attentional modulation of the response (ROI  $\times$  state interaction:  $F < 1$ ). Indeed, modulation of connectivity was not weaker during nonstimulation than stimulation periods (period  $\times$  ROI  $\times$  state interaction:  $F_{1,6} = 2.23$ ,  $P = 0.186$ ), despite a concomitant drop in the evoked modulation ( $F_{1,6} = 24.88$ ,  $P = 0.002$ ). Second, we observed evoked responses in cases when connectivity did not significantly change: goal-irrelevant regions (e.g., FFA during scene attention) showed robust evoked responses ( $F_{1,6} = 16.90$ ,  $P = 0.006$ ) but no increase in background connectivity relative to rest ( $F < 1$ ). This double dissociation shows that evoked responses are neither necessary nor sufficient for modulation of background connectivity.





**Fig. 4.** Analysis stages. (A) Each stage shown in a row, with analyses to the right performed on residuals from that stage. (B) Sources of variance in the residuals of current stage. (C) Mean evoked responses across blocks, with proportions of variance explained by average block response. (D) V4-FFA/PPA correlations. Despite the large reduction in the variance explained by evoked responses across analysis stages, there are negligible changes in the attentional modulation of V4 connectivity (main effect of reduced correlations, but no change in the ROI  $\times$  state interaction; see text). This suggests that attentional modulation of connectivity does not depend on stimulus responses but rather on spontaneous activity commonly used to measure intrinsic brain networks at rest.

frequencies up to 0.3 Hz (Fig. S4B). Additional analyses demonstrated that block-by-block variance around the mean evoked response also could not explain our results (Figs. S2B and S5).

Finally, we can quantify the contribution of evoked responses to background connectivity by measuring correlations between brain areas that could not possibly interact. First, each subject completed two identical runs per condition, and so we correlated the FFA with itself across runs, and the PPA with itself across runs (Fig. S6). Second, all subjects completed the same tasks, and so we correlated the FFA/PPA in each subject with occipital ROIs averaged over all other subjects (Fig. S7). In both cases, strong correlations were observed before, but not after, removing evoked responses, suggesting that background connectivity reflects state-dependent intrinsic correlations.

### Discussion

Together, these results demonstrate that top-down attention can switch background connectivity across multiple regions of visual cortex: attention to faces increased the proportion of intrinsic variance shared between occipital cortex and the FFA, and attention to scenes increased the proportion of intrinsic variance shared between occipital cortex and the PPA. How might attention cause these changes in background connectivity?

One possibility is that attention prioritizes information by directly altering communication between regions of visual cortex. In accordance with recent neurophysiological findings in nonhuman primates, attention could improve the fidelity of communication about attended information by enhancing long-range LFP

coherence (17, 18). This synchronization may, in turn, increase the likelihood of input from one area (e.g., V4) arriving at moments of peak excitability in another area (e.g., FFA or PPA) (27). Although the frequencies assayed with fMRI are at least an order of magnitude slower than other physiological measures, how the modulation of BOLD correlations relates to attentional reductions in low-frequency EEG power (28) and firing rate correlations (29), and to attentional increases in high-frequency LFP coherence (15–18), remains an important open question (30, 31).

Alternatively, the increase in shared variance may reflect coordinated top-down input to occipital and ventral temporal cortex from control regions, such as those in frontoparietal attention networks (32). An exploratory analysis revealed frontal regions whose background connectivity with posterior areas depended on attentional state (Fig. S8), which may provide a useful starting point for future research to investigate the control of connectivity within visual cortex. Note, however, that any top-down input that was time-locked to stimulus presentation would have been removed by our analysis, and thus our results could only be supported by spontaneous or random fluctuations in synchronized input. Whether the observed attentional modulation of background connectivity reflects a direct or indirect change in communication, our findings reveal a unique neural correlate of attention that can strongly predict behavior and is reflected in a component of the BOLD signal that is typically discarded.

These results also emphasize that attention operates in a distributed fashion throughout the ventral visual stream. Conventionally, attention is thought to prioritize goal-relevant

information by modulating the visual area that is most specialized for processing this kind of information, both in terms of its evoked response (3, 6–10) and how it interfaces with frontoparietal cortex (14, 17, 18, 33). The present findings suggest that attention can also influence visual areas not specialized for this information (e.g., retinotopic human area V4 during category-based attention to faces/scenes), increasing their coupling with more specialized areas. In this way, beyond modulating isolated regions at one level of the visual processing hierarchy, top-down attention may help synchronize different levels, forming state-dependent functional pathways to prioritize processing of goal-relevant information.

## Materials and Methods

**Subjects.** Eight naive, right-handed adults with normal or corrected-to-normal vision participated for monetary compensation. Seven subjects completed two fMRI sessions and were included in the analysis (two men, mean age = 26 y); one subject was excluded because they completed the first (retinotopic mapping) session but did not return for the second (experimental) session. The Princeton University Institutional Review Board approved the study protocol, and all subjects provided informed consent.

**Localizer Runs.** We performed phase-encoded retinotopic mapping using standard stimuli and procedures (23). A rotating wedge of a flickering colored checkerboard was used to map polar angle (two runs each of clockwise and counterclockwise rotation), and an expanding and contracting ring of the same checkerboard was used to map eccentricity (one run each of expansion and contraction). We localized FFA/PPA with standard procedures (21). The localizer occurred after all attention runs and used the same face and scene stimuli, now presented individually.

**Attention Runs.** Face-attention and scene-attention runs consisted of an on-off block design, with 18-s blocks of stimulation interleaved with 18-s blocks of fixation (Fig. 1B). Stimulation blocks consisted of 12 1-s presentations of face/scene composite images separated by 500-ms interstimulus intervals, with face and scene identity determined pseudorandomly (21). Stimuli were presented centrally and subtended  $6 \times 6^\circ$  of visual angle. Stimulus onsets were time-locked to the repetition time (TR) and triggered by the scanner. To manipulate attentional state, subjects performed a selective one-back task on either the faces or scenes while ignoring the other category. For example, during face-attention runs, subjects pressed a response button only if the face component of two successive images was identical. During fixation periods, the only stimulus was a central fixation point. Rest runs had the same underlying structure as attention runs, but subjects saw only a continuous fixation point throughout the run. All subjects completed two rest runs, followed by alternating face-attention and scene-attention runs (two each, order counterbalanced across subjects).

**Image Acquisition.** The fMRI data were acquired with a 3T scanner (Siemens, Allegra) using a birdcage volume coil (Nova Medical). An occipital rf surface coil (Nova Medical) was used for four subjects during retinotopic mapping sessions to improve signal-to-noise in occipital cortex. During retinotopic mapping, a high-resolution magnetization-prepared rapid acquisition gradient-echo (MPRAGE) anatomical scan was acquired for spatial registration and surface reconstruction. Additional high-resolution anatomical scans were available for some subjects from past sessions and were included to improve surface reconstruction. Functional images for retinotopic mapping were acquired with a gradient-echo echo-planar imaging (EPI) sequence [TR 2.0 s; echo time (TE) 40 ms; flip angle (FA)  $71^\circ$ ; matrix  $128 \times 128$ ; resolution  $2 \times 2 \times 3$  mm] with 20 slices aligned to the calcarine sulcus. Functional images for attention runs and face/scene localizer runs were acquired with a gradient-echo EPI sequence (TR 1.5 s; TE 28 ms; FA  $64^\circ$ ; matrix  $64 \times 64$ ; resolution  $3.5 \times 3.5 \times 3.5$  mm) with 26 axial slices aligned to the anterior commissure/posterior commissure line. To improve registration, an additional T1 fast low angle shot (FLASH) anatomical scan was acquired at the end of each session, coplanar to the functional scans. To correct B0-field inhomogeneities, phase and magnitude field maps were collected at the end of all sessions, coplanar to the functional scans and with the same resolution.

**Image Preprocessing.** The fMRI data were analyzed using FSL (34), FreeSurfer (35), and Matlab (MathWorks). All images were skull stripped to improve registration. Volumes from the first 6 s of functional runs were discarded. The remaining volumes were corrected for slice-acquisition time and head motion, high-pass filtered (100-s period cutoff), debiased using FSL FAST

(surface-coil runs) or FSL FUGUE (volume-coil runs), spatially smoothed (retinotopic mapping runs, 3 mm FWHM; attention/face-scene localizer runs, 5 mm FWHM), and registered to the coplanar anatomical, high-resolution anatomical, and Montreal Neurological Institute standard brain.

**Regions of Interest.** We drew occipital ROIs manually using FreeSurfer. To visualize occipital areas simultaneously, each subject's cortical surface was segmented along the white matter/gray matter boundary, inflated, and flattened. Functional runs from retinotopic mapping sessions were pre-whitened and phase-decoded to determine the angle and eccentricity of maximal stimulation for each voxel in visual cortex. Phase estimates from clockwise/counterclockwise and expansion/contraction were averaged to account for hemodynamic lag. We projected these values onto the cortical surface and determined occipital ROIs according to the relative locations of polar-angle reversals and foveally stimulated regions (23).

We defined the FFA and PPA ROIs from the face/scene localizer as the peak voxel in right ventral temporal cortex that was most selective for faces (face blocks > scene blocks) and scenes (scene blocks > faces blocks), respectively. BOLD activity was extracted from the FFA and PPA as a weighted average of the surrounding voxels, with weights determined by a 10-mm FWHM Gaussian kernel centered on the peak voxel. The pattern of results was robust to this a priori choice of kernel size (ROI  $\times$  state interaction in V4 for kernel sizes from 0 to 20 mm: all  $P$  values < 0.021). We restricted analysis to right FFA because the right hemisphere is dominant in face processing (36), and to right PPA to ensure comparability with FFA (21). We combined occipital ROIs from both hemispheres because these regions are strongly retinotopic and have no such asymmetries. Secondary analyses confirmed that attentional modulation of connectivity occurred independently for V4 in each hemisphere (all  $P$  values < 0.014).

**Background Connectivity.** We used a recently developed approach to measure interactions: background connectivity. After preprocessing, the BOLD signal from attention runs was scrubbed of nuisance and stimulus-evoked variance using two general linear models (Figs. 1B and 4A and Fig. S2). The first (nuisance) model contained regressors for the global mean BOLD signal, six motion correction parameters obtained from preprocessing, and BOLD activity from four white-matter voxels and four ventricle voxels (20). Residuals from the nuisance model served as input to the second (evoked) model. Each attention run consisted of 12 identically structured blocks of trials, during which a single attentional state was induced. Accordingly, we modeled stimulus-evoked responses with finite impulse response (FIR) basis functions that captured the mean evoked response across blocks. Residuals from the evoked model allowed us to assess functional connectivity associated with each attentional state, independent of correlations attributable to stimulus-evoked responses. The details and merits of this approach are described in Fig. S2 and elsewhere (21, 37).

**Other Analyses.** We replicated the standard attentional modulation of evoked FFA/PPA responses (6) (ROI  $\times$  state interaction:  $F_{1,6} = 27.42$ ,  $P = 0.002$ ). To visualize evoked responses after stage 1 (Fig. 4C), the same FIR model that was used for stage 3 was applied directly to the preprocessed data. Evoked responses after stages 2 and 3 (Fig. 4C) were calculated as the residuals from the nuisance and evoked models, respectively, signal averaged across blocks. These block averages were normalized to the SD of the residuals, allowing comparison of their amplitudes. Twelve repetitions of each average were concatenated and correlated with the original residuals to quantify how much variance they explained.

To determine stimulated and nonstimulated block volumes (Fig. 3), the grand average evoked response was calculated by collapsing the FIR block time course extracted after stage 1 from all conditions (i.e., FFA/PPA and attentional states), and the percent signal change for each timepoint/volume was compared using a  $t$  test with that of the final (i.e., 24th) volume in the block. This approximates the baseline activity just before block onset. The volumes that produced a significant difference ( $P < 0.05$ ) were labeled as stimulated, and the remainder as nonstimulated.

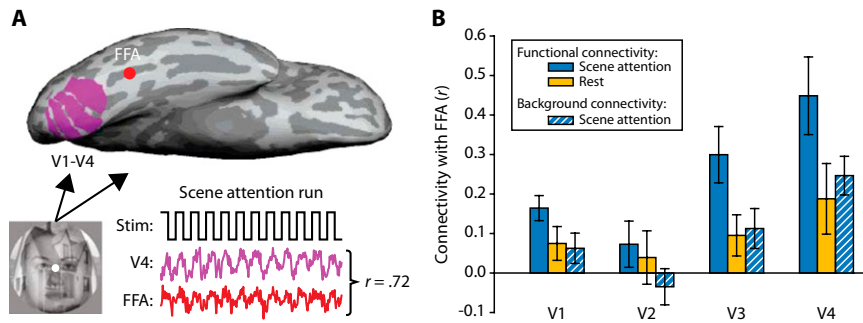
For all statistical analyses, correlations were variance-stabilized using the Fisher transformation. For visualization in figures, correlations and error bars were reconverted using the inverse transform. The threshold for statistical significance was  $P < 0.05$  (all tests two-tailed).

**ACKNOWLEDGMENTS.** We thank M. Arcaro for assistance with retinotopic mapping, A. Tomparny and M. Simon for help with data collection, and U. Hasson and C. Honey for useful discussions. This work was supported by a Natural Sciences and Engineering Research Council of Canada Postdoctoral Fellowship (to N.A.-A.) and National Institutes of Health Grant R01 EY021755 (to N.B.T.-B.).

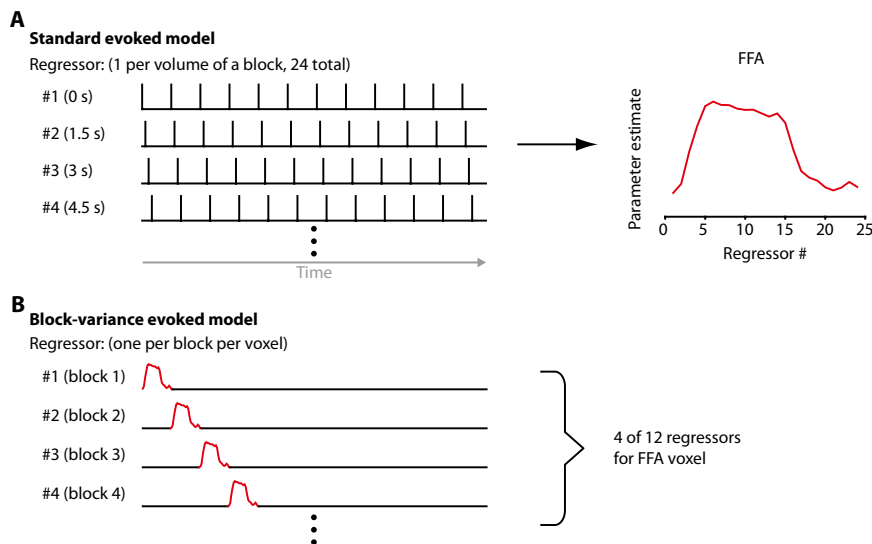
1. Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: Two cortical pathways. *Trends Neurosci* 6:414–417.
2. Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol* 195:215–243.
3. Kanwisher N (2010) Functional specificity in the human brain: A window into the functional architecture of the mind. *Proc Natl Acad Sci USA* 107:11163–11170.
4. Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
5. Chabris CF, Simons DJ (2010) *The Invisible Gorilla: And Other Ways Our Intuitions Deceive Us* (Crown Archetype, New York).
6. O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 401:584–587.
7. Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–1559.
8. Saenz M, Buracas GT, Boynton GM (2002) Global effects of feature-based attention in human visual cortex. *Nat Neurosci* 5:631–632.
9. McAdams CJ, Maunsell JHR (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441.
10. Treue S, Maunsell JHR (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382:539–541.
11. Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. *Annu Rev Neurosci* 18:193–222.
12. Cohen JD, Aston-Jones G, Gilzenrat MS (2004) *Cognitive Neuroscience of Attention*, ed Posner MI (Guilford, New York), pp 71–90.
13. Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
14. Noudoost B, Chang MH, Steinmetz NA, Moore T (2010) Top-down control of visual attention. *Curr Opin Neurobiol* 20:183–190.
15. Fries P, Reynolds JH, Rorie AE, Desimone R (2001) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563.
16. Lee H, Simpson GV, Logothetis NK, Rainer G (2005) Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45:147–156.
17. Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324:1207–1210.
18. Saalmann YB, Pigarev IN, Vidyasagar TR (2007) Neural mechanisms of visual attention: How top-down feedback highlights relevant locations. *Science* 316:1612–1615.
19. Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med* 34:537–541.
20. Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711.
21. Norman-Haignere SV, McCarthy G, Chun MM, Turk-Browne NB (2012) Category-selective background connectivity in ventral visual cortex. *Cereb Cortex* 22:391–402.
22. Fox MD, Snyder AZ, Zacks JM, Raichle ME (2006) Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat Neurosci* 9:23–25.
23. Arcaro MJ, McMains SA, Singer BD, Kastner S (2009) Retinotopic organization of human ventral visual cortex. *J Neurosci* 29:10638–10652.
24. Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *J Neurophysiol* 57:835–868.
25. Morey RD (2008) Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutor Quant Methods Psychol* 4:61–64.
26. Lamme VAF, Supèr H, Spekreijse H (1998) Feedforward, horizontal, and feedback processing in the visual cortex. *Curr Opin Neurobiol* 8:529–535.
27. Fries P (2005) A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends Cogn Sci* 9:474–480.
28. Sauseng P, et al. (2005) A shift of visual spatial attention is selectively associated with human EEG alpha activity. *Eur J Neurosci* 22:2917–2926.
29. Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63:879–888.
30. He BJ, Raichle ME (2009) The fMRI signal, slow cortical potential and consciousness. *Trends Cogn Sci* 13:302–309.
31. Koch SP, Werner P, Steinbrink J, Fries P, Obrig H (2009) Stimulus-induced and state-dependent sustained gamma activity is tightly coupled to the hemodynamic response in humans. *J Neurosci* 29:13962–13970.
32. Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
33. Chadick JZ, Gazzaley A (2011) Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nat Neurosci* 14:830–832.
34. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl 1):S208–S219.
35. Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9:179–194.
36. Yovel G, Tambini A, Brandman T (2008) The asymmetry of the fusiform face area is a stable individual characteristic that underlies the left-visual-field superiority for faces. *Neuropsychologia* 46:3061–3068.
37. Summerfield C, et al. (2006) Neocortical connectivity during episodic memory formation. *PLoS Biol* 4:e128.

# Supporting Information

Al-Aidroos et al. 10.1073/pnas.1202095109



**Fig. S1.** Stimulus-evoked responses complicate the interpretation of functional connectivity. (A) Inflated surface of one subject's right hemisphere, depicting ventral V1–V3, human V4 (each in purple), and fusiform face area (FFA) (red). Time series during rest and tasks can be extracted from these regions of interest (ROIs) and correlated to assess functional connectivity. (B) At rest, functional connectivity reflects intrinsic activity shared between brain areas that have latent direct or indirect connections (yellow). We hypothesize that top-down attention to a visual category (e.g., scenes) will increase the strength of these intrinsic correlations between areas of occipital cortex (especially V4) and the area of ventral temporal cortex selective for that category [e.g., parahippocampal place area (PPA)]. In contrast, we predict no increase, or perhaps a slight decrease, in such connectivity with ventral temporal areas not selective for the attended category (e.g., face-selective FFA). However, widespread correlations might occur in visual cortex whenever any stimulus is introduced, simply because even nonselective areas will exhibit some evoked blood oxygen level-dependent (BOLD) response time-locked to its appearance. For example, when subjects attend to the scene in a composite face/scene image, V4 and FFA show increased functional connectivity relative to rest (compare solid blue with yellow). These correlations—unlike ones with areas selective for the attended category (i.e., PPA; see main text)—are eliminated when evoked responses are removed to isolate background connectivity (striped blue). This suggests that V4–FFA correlations are stimulus-driven and do not reflect enhanced intrinsic interactions related to the current attentional state.

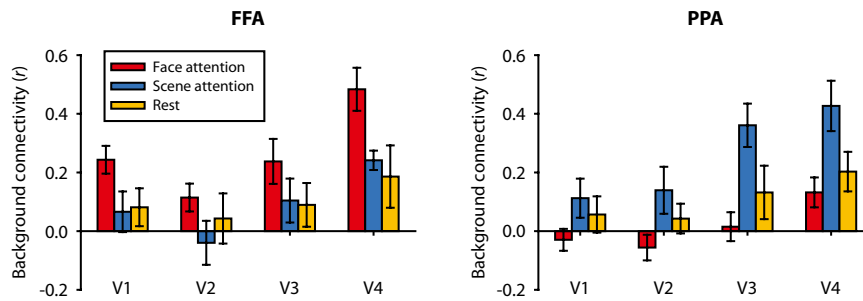


**Fig. S2.** The logic behind the background connectivity approach is that synchronized stimulus-evoked responses in two or more brain areas can produce correlations that are unrelated to connectivity, and thus stimulus-evoked responses should be removed to isolate correlations of intrinsic activity in the background of stimuli/tasks. (A) In the standard model (as used in all analyses except Fig. S5), 24 finite impulse response (FIR) regressors are used to estimate the mean response at each time point across blocks. Each regressor has a constant height at that time point and zeroes elsewhere. Block responses were modeled with FIR functions rather than canonical hemodynamic response functions (HRFs) because the more data-driven FIR approach does not require us to assume that the evoked response has a particular shape, nor that this shape is constant across brain areas. If the evoked response in two or more brain areas differed from an assumed HRF, residual evoked variance would persist and correlate between these areas. The standard model estimates the mean evoked response but does not account for block-by-block variance around the mean. The block-variance model takes the output of the standard model as a template block response for each voxel, generates separate regressors for each block, and then models the original data that were input to the standard model. As a result, the amplitude of the response in each block and voxel is estimated independently, removing any variance in the amplitude across blocks. We found little contribution of this variance to our results (Fig. S5).

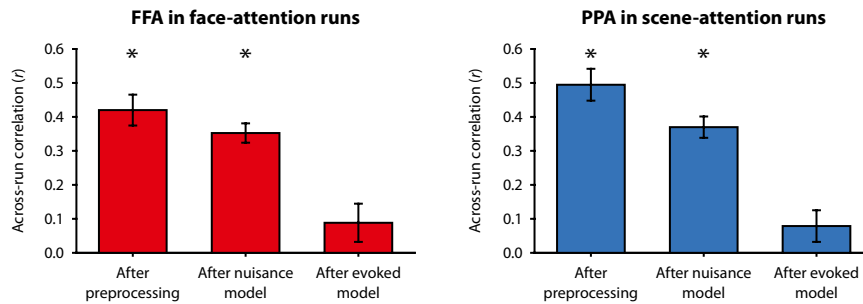




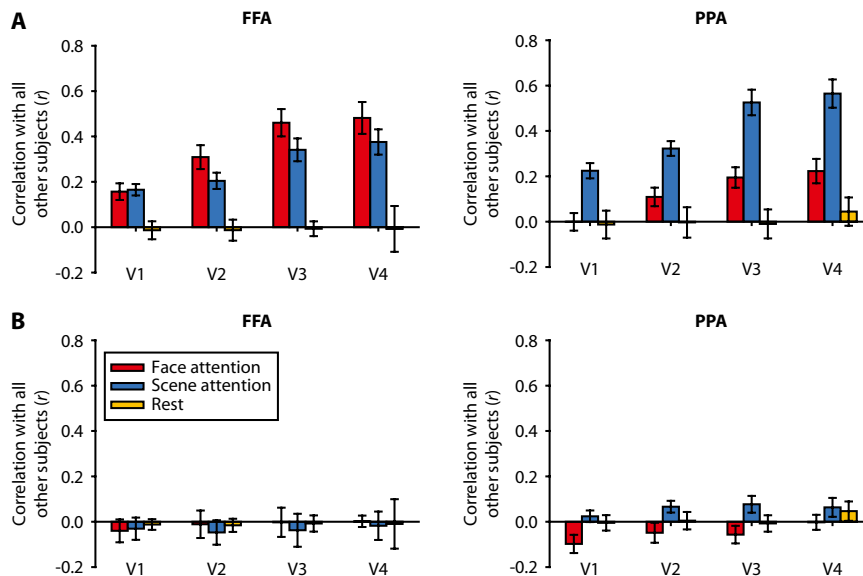




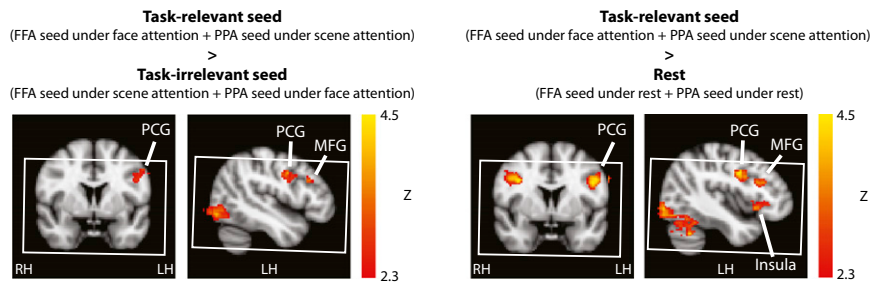
**Fig. S5.** For all background connectivity analyses, stimulus-evoked responses were removed by modeling the mean block activity (Figs. 1B and 4A and Fig. S2A). However, the evoked response of an individual block may have varied around the mean (e.g., if subjects closed their eyes or were especially motivated). If this variance was shared across brain areas, it may contribute to background connectivity. To assess this possibility, we applied a model of the stimulus-evoked response that captured the variance across blocks (Fig. S2B). This more-sophisticated model had little impact on the magnitude of correlations or on their modulation by attention, suggesting that variance around the mean did not meaningfully contribute to background connectivity.



**Fig. S6.** Across-run correlations. All subjects completed two runs of each attention task. Because the stimulus timing of these runs was identical, stimulus-driven correlations should emerge across runs if the runs still contain evoked responses. For example, FFA activity in both face-attention runs should be correlated, despite the fact that intrinsic activity cannot be correlated (i.e., the FFA cannot interact with itself several minutes in the future). As expected, across-run correlations were robust after stage 1 (preprocessing) and stage 2 (the nuisance model) but were eliminated statistically after stage 3 (the evoked model). Thus, background connectivity was not observed in situations where it should not occur (see also Fig. S7).



**Fig. S7.** Across-subject correlations. The attention runs of all subjects used the same stimulus timing, and thus stimulus-driven correlations should emerge across subjects. Correlations were computed using an  $n$ -fold approach: the time series from ventral temporal ROIs of one subject were correlated with the time series from occipital ROIs averaged over all other subjects, and this was repeated for each subject. (A) After stage 2 (the nuisance model), there were strong correlations across subjects. (B) After removing stimulus-evoked responses in stage 3 (the evoked model), these correlations were eliminated. These results further suggest that background connectivity reflects spontaneous fluctuations that were idiosyncratic to particular subjects and moments in time.



**Fig. S8.** Group voxelwise analysis of attentional modulation of background connectivity. Primary analyses focused on how top-down category-selective attention regulates background connectivity within the visual system, with task-relevant areas of ventral temporal cortex coupling more strongly with lower-level areas in occipital cortex. This finding complements past neurophysiological studies using nonhuman primates showing that attention also enhances the coupling between individual areas of visual cortex and control regions of the frontoparietal attention network (1–3). To help connect our findings with these studies, we examined FFA/PPA background connectivity outside of visual cortex. Although our slice prescription (depicted for a typical subject above with a white box) did not include key regions of the frontoparietal attention network—most notably the superior parietal lobe and possibly the frontal eye fields—we performed an exploratory analysis of how attention modulated connectivity with FFA/PPA. For each run, we modeled the background time series in every voxel (i.e., the residuals of the stage 3 evoked model) separately with the FFA time series from the same data, and the PPA time series from the same data. The resulting statistical maps were then collapsed across the two runs of each attention condition, and we performed two contrasts depending on the task-relevance of the seed ROI: (i) task-relevant vs. -irrelevant seed, FFA seed under face attention + PPA seed under scene attention > FFA seed under scene attention + PPA seed under face attention; (ii) task-relevant seed vs. rest, FFA seed under face attention + PPA seed under scene attention > FFA seed during rest + PPA seed during rest. We computed the statistical significance of these contrasts, correcting for multiple comparisons using cluster-mass thresholding ( $P < 0.05$ ; cluster-forming threshold  $Z = 2.3$ ). Despite the loss of power in group voxelwise analyses (as opposed to the ROI analyses reported throughout the article), we observed robust attentional modulation of background connectivity in several lateral frontal regions: left precentral gyrus (PCG; Brodmann area 6) and left middle frontal gyrus (MFG; BA 9) exhibited greater connectivity with ventral temporal ROIs when they were task-relevant vs. task-irrelevant; and bilateral PCG, bilateral MFG, and left insula (BA 13) exhibited greater connectivity with ventral temporal ROIs when they were task-relevant vs. during rest. Thus, similar to previous studies, attention increased the coupling of frontal cortex with task-relevant areas of visual cortex. The identified frontal regions may provide a useful preliminary guide for future investigations of how control regions modulate background connectivity within visual cortex. Namely, PCG, MFG, and/or insula may provide coordinated top-down input to low- and high-level areas of visual cortex.

1. Gregoriou GG, Gotts SJ, Zhou H, Desimone R (2009) High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324:1207–1210.
2. Saalmann YB, Pigarev IN, Vidyasagar TR (2007) Neural mechanisms of visual attention: How top-down feedback highlights relevant locations. *Science* 316:1612–1615.
3. Noudoost B, Chang MH, Steinmetz NA, Moore T (2010) Top-down control of visual attention. *Curr Opin Neurobiol* 20:183–190.