

# Current Biology

## Attention Reduces Spatial Uncertainty in Human Ventral Temporal Cortex

### Highlights

- We developed a population receptive field (pRF) model for ventral temporal cortex (VTC)
- Model accurately predicts responses to novel stimulus positions and sizes
- Attention increases pRF gain, eccentricity, and size in high- but not low-level areas
- Attention improves the quality of spatial representations in VTC

### Authors

Kendrick N. Kay, Kevin S. Weiner, Kalanit Grill-Spector

### Correspondence

kendrick@post.harvard.edu

### In Brief

Kay et al. tackle the problem of modeling the final stage in the ventral visual pathway. They do so with a model of visual space, which is surprising since this stage is traditionally viewed as coding form, not space. They show that attention reduces uncertainty in the peripheral visual field where humans and primates are the worst at seeing.

# Attention Reduces Spatial Uncertainty in Human Ventral Temporal Cortex

Kendrick N. Kay,<sup>1,\*</sup> Kevin S. Weiner,<sup>2</sup>  
 and Kalanit Grill-Spector<sup>2,3</sup>

<sup>1</sup>Department of Psychology, Washington University in St. Louis, St. Louis, MO 63130, USA

<sup>2</sup>Department of Psychology, Stanford University, Stanford, CA 94305, USA

<sup>3</sup>Stanford Neurosciences Institute, Stanford University, Stanford, CA 94305, USA

## Summary

Ventral temporal cortex (VTC) is the latest stage of the ventral “what” visual pathway, which is thought to code the identity of a stimulus regardless of its position or size [1, 2]. Surprisingly, recent studies show that position information can be decoded from VTC [3–5]. However, the computational mechanisms by which spatial information is encoded in VTC are unknown. Furthermore, how attention influences spatial representations in human VTC is also unknown because the effect of attention on spatial representations has only been examined in the dorsal “where” visual pathway [6–10]. Here, we fill these significant gaps in knowledge using an approach that combines functional magnetic resonance imaging and sophisticated computational methods. We first develop a population receptive field (pRF) model [11, 12] of spatial responses in human VTC. Consisting of spatial summation followed by a compressive nonlinearity, this model accurately predicts responses of individual voxels to stimuli at any position and size, explains how spatial information is encoded, and reveals a functional hierarchy in VTC. We then manipulate attention and use our model to decipher the effects of attention. We find that attention to the stimulus systematically and selectively modulates responses in VTC, but not early visual areas. Locally, attention increases eccentricity, size, and gain of individual pRFs, thereby increasing position tolerance. However, globally, these effects reduce uncertainty regarding stimulus location and actually increase position sensitivity of distributed responses across VTC. These results demonstrate that attention actively shapes and enhances spatial representations in the ventral visual pathway.

## Results

### Does a pRF Model Predict Responses in VTC?

To develop a model of how spatial information is encoded in ventral temporal cortex (VTC), we measured fMRI responses (3T, 2-mm voxels) in a series of face-selective regions [13] while subjects fixated centrally and viewed images of faces that varied systematically in position and size (Figure 1A). We used face-selective regions as a model system as they are a highly studied subsystem of VTC [3, 14, 15] with a well-understood functional organization that is anatomically consistent across subjects [13, 16]. After estimating and denoising stimulus-evoked responses [17], we modeled

responses in each voxel using the compressive spatial summation (CSS) model [12]. The CSS model characterizes the population receptive field (pRF) [11] of a voxel and predicts the response to a face by first computing the spatial overlap between the face and an isotropic 2D Gaussian and then applying a compressive nonlinearity (Figure 1B). Cross-validation analyses demonstrate that the CSS model accurately characterizes responses of individual voxels in face-selective regions located on the inferior occipital gyrus (IOG), posterior fusiform gyrus (pFus), and mid-fusiform gyrus (mFus) [13] and successfully predicts responses to faces at novel positions and sizes (Figures 1C and S1A). To assess whether these results are specific to face stimuli, we also performed measurements using phase-scrambled faces. Although phase-scrambled faces evoke weaker responses and produce noisier pRF estimates, pRF properties are largely invariant to stimulus type (Figures S1B and S1C).

### What Is the Nature of pRFs in VTC?

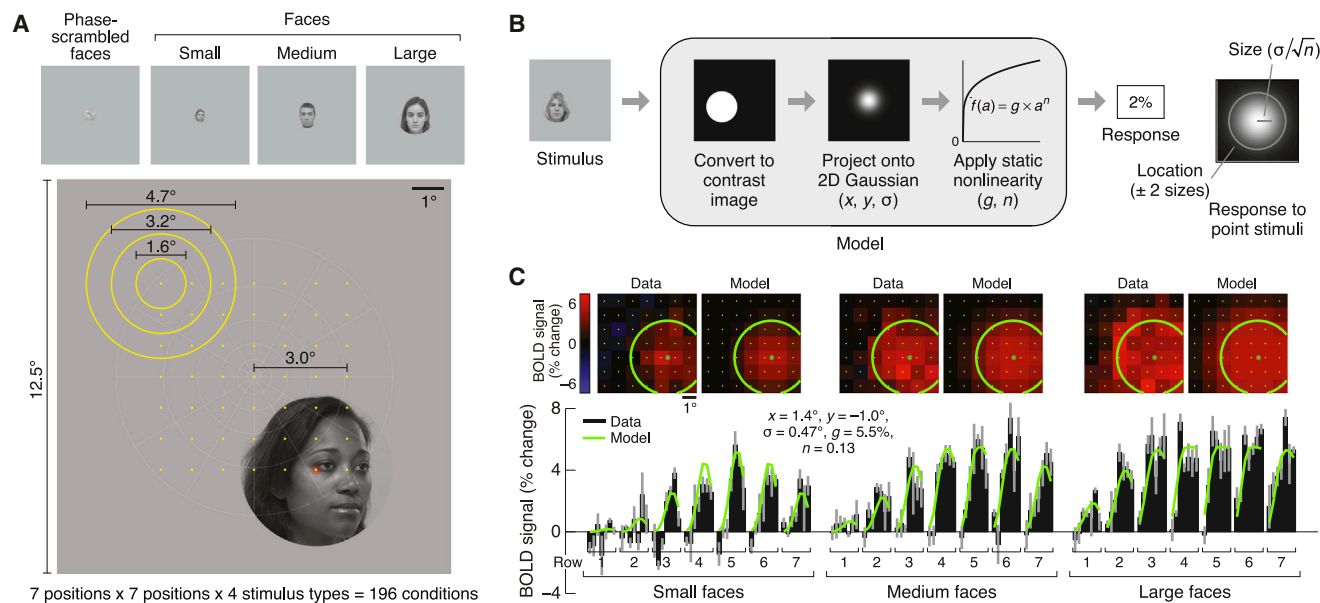
Similar to early and intermediate visual areas [11, 12], pRF size increases with eccentricity in face-selective regions within VTC (Figures 2A and S2B), suggesting that size-eccentricity scaling is a pervasive organizing principle across the ventral visual pathway. However, different from earlier visual areas, pRFs in face-selective regions are quite large compared to their eccentricity. Consequently, these pRFs extend substantially into the ipsilateral visual field (Figure 2B). Also, unlike pRFs in earlier areas, pRFs in face-selective regions are consistently centered near the fovea, producing a representational scheme in which nearly all neural resources are dedicated to the central portion of the visual field (approximately the central 7°; see Figures 2B and S2A). This convergence of spatial coverage is consistent with the foveal bias of face-selective regions [14, 15]. Notably, this organization is different from the distributed tiling of visual space in earlier retinotopic visual regions [18], suggesting unique computational strategies in VTC. Interestingly, pRF properties vary hierarchically across face-selective regions: anterior regions in VTC generally have larger and more foveal pRFs than posterior regions (Figures 2A and S2C), features also observed in monkey inferotemporal cortex (IT) [19–22].

### How Are pRF Properties Affected by Attention?

To understand the contribution of top-down attentional signals to the observed results, we measured pRFs under different attentional states. While maintaining central fixation, subjects performed one of three tasks: digit task (one-back task on rapid serial presentation of digits at fixation), dot task (detection of a red dot appearing on the faces; same as the first experiment), and face task (one-back task on the identity of the faces; see Supplemental Experimental Procedures and Figure S3A). In the digit task, attention is directed toward fixation, whereas in the dot and face tasks, attention is directed toward the faces.

Comparing pRF properties across tasks, we find no substantial changes in pRF properties in early visual areas V1–V3 (Figures 3A and S3C). However, in hV4 and, more substantially, in face-selective regions, voxel responses are strongly modulated by the task (Figure S3B). In these regions,

\*Correspondence: [kendrick@post.harvard.edu](mailto:kendrick@post.harvard.edu)



**Figure 1. Compressive Spatial Summation Accurately Models Responses in VTC**

(A) Stimuli. Subjects viewed faces while fixating centrally. Faces varied systematically in position (centers indicated by yellow dots) and size (sizes indicated by yellow circles). During each trial, face position and size were held constant while face identity and viewpoint were dynamically updated.

(B) Compressive spatial summation (CSS) model. The response to a face is predicted by computing the spatial overlap between the face and a 2D Gaussian and then applying a compressive power-law nonlinearity. The model includes two parameters ( $x, y$ ) for the position of the Gaussian, a parameter ( $\sigma$ ) for the size of the Gaussian, a parameter ( $n$ ) for the exponent of the nonlinearity, and a parameter ( $g$ ) for the overall gain of the predicted responses.

(C) Example voxel (left IOG, subject 1). Top row: responses arranged spatially according to face position. Bottom row: responses arranged serially for better visualization of measurement reliability and goodness of fit of CSS model. Blood-oxygenation-level-dependent (BOLD) response magnitudes (black bars; median across trials  $\pm 68\%$  confidence interval [CI]) are accurately fit by the model (green line). Note that a single set of model parameters accounts for the full range of the data.

See also [Figure S1](#).

pRFs exhibit increased eccentricity, size, and gain when subjects attend to the faces (dot and face tasks) compared to when they attend to fixation (digit task) ([Figures 3A–3C](#)). These effects are consistent with the concept of response enhancement at the attended location [23], and the effects are large in size: for example, in mFus, comparing pRF properties across the digit and face task, respectively, the median pRF eccentricity increases from  $1.3^\circ$  to  $1.9^\circ$ , the median pRF size increases from  $1.8^\circ$  to  $3.4^\circ$ , and the median pRF gain increases from  $0.83\%$  to  $1.32\%$ .

Control experiments reveal that changes in pRF properties are observed even if tasks are interleaved on a trial-by-trial basis ([Figure S3G](#)), indicating that the changes cannot be attributed to variation in general subject arousal across tasks. Furthermore, performing the digit task on digits presented to the left of fixation produces leftward shifts of pRFs in hV4, IOG, and pFus compared to performing the digit task on central digits ([Figure S3H](#)). This indicates that even though attention is drawn away from faces during the digit task, pRF modulations occur in a manner consistent with response enhancement at the attended location, irrespective of the content of the attended stimulus.

Interestingly, attentional effects in face-selective regions are stronger for the face task, which specifically requires perceptual processing of the faces, compared to the dot task ( $p < 10^{-9}$ , two-tailed sign test in each region for each pRF property). Increases in pRF size under the dot and face tasks relative to the digit task are particularly intriguing as they indicate that locally, at the voxel level, attention to the stimulus increases the position tolerance of the neural representation.

### What Is the Benefit of Attentional Modulation of pRFs?

Although we have demonstrated local changes in pRF properties as a result of attention, an open question is whether these attention-induced changes are beneficial to the global, or distributed, representation of the stimulus. Specifically, we ask the following question: does attention affect the ability of a collection of pRFs to discriminate the location of the stimulus? This question cannot be answered through simple summary statistics of pRF properties (such as the ones in [Figure 3A](#)) because discrimination performance depends not only on the properties of individual pRFs but also on how the pRFs collectively tile the visual field. For example, large but overlapping pRFs might discriminate stimulus locations better than small, non-overlapping pRFs [24]. We therefore designed a model-based decoding analysis that quantifies the spatial discrimination performance of a collection of pRFs. In this analysis, we calculate spatial uncertainty, that is, the distance over which changes in stimulus position cannot be well discriminated based on the distributed responses across the pRFs (thus, low spatial uncertainty indicates good discrimination performance). We applied this analysis separately to each region, analyzing the pRFs observed under each task.

As expected from the stability of pRF properties in early visual areas, there is little change in spatial uncertainty in these areas across tasks ([Figure 4A](#), top). In all tasks, spatial uncertainty in V1–V3 is less than  $0.5^\circ$  near the fovea ( $1^\circ$  eccentricity) and less than  $1.5^\circ$  in the periphery ( $5^\circ$  eccentricity) ([Figures 4B](#) and [4C](#)). In contrast, there are large changes in spatial uncertainty in face-selective regions across tasks. In the periphery, spatial uncertainty is substantially reduced under the dot task

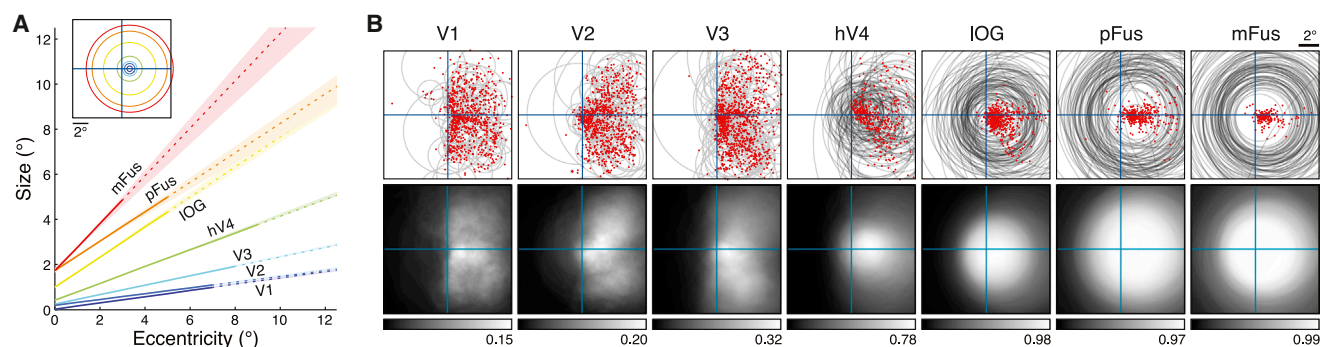


Figure 2. Systematic Organization of pRF Properties across the Ventral Visual Pathway

(A) pRF size versus eccentricity. Each line represents a region (median across voxels  $\pm$  68% CI). Dotted lines indicate eccentricity ranges containing few voxels. The inset shows a schematic of pRF sizes at 1° eccentricity. IOG, inferior occipital gyrus; pFus, posterior fusiform; mFus, mid-fusiform. (B) pRF locations and visual field coverage in the left hemisphere. Top row: pRF centers (red dots) and pRF sizes for 100 randomly selected voxels from each region (gray circles). Bottom row: visual field coverage, computed as the proportion of pRFs covering each point in the visual field. Each image is normalized such that black corresponds to 0, and white corresponds to the maximum value. IOG, pFus, and mFus contain large pRFs centered near the fovea. See also Figure S2.

(1.9-fold reduction, on average, across face-selective regions) and the face task (2.7-fold reduction, on average, across regions) compared to the digit task (Figure 4A, bottom; Figure 4C). For example, in mFus, uncertainty in the periphery is more than 3° under the digit task but only about 1° under the face task (Figure 4C). Importantly, these improvements are not simply due to increased pRF gain: improvements are still observed if pRF gain is held constant and only the task-induced changes in pRF location and size are considered (Figure S4A). These results indicate that attending the stimulus either explicitly (face task) or implicitly (dot task) reduces uncertainty with respect to the location of the stimulus. As a complement to our model-based decoding analysis, we also performed direct decoding of the distributed response patterns evoked by faces with no intervening modeling step. Results are consistent with our model-based analysis: in face-selective regions, there is improved decoding of face position in the periphery under the face task compared to the digit and dot tasks (Figures S4B and S4C).

## Discussion

The experiments in the present study reveal that spatial representations are prevalent in the ventral “what” visual pathway. First, we have shown that responses in VTC are modulated by changes in the position and size of the stimulus. These modulations are systematic and are accurately characterized by a pRF model utilizing spatial summation and a compressive nonlinearity. Second, spatial representations within VTC are actively shaped by top-down task demands. Specifically, attention modulates pRFs in high and intermediate levels of the ventral pathway, but not early visual regions. While prior research has shown that spatial attention shifts receptive fields in the dorsal “where” visual pathway [6–9, 25], as well as intermediate visual areas in the ventral pathway [23, 26–28], we extend these results to high-level areas in the ventral pathway for the first time.

### Attentional Effects in the Ventral Visual Pathway

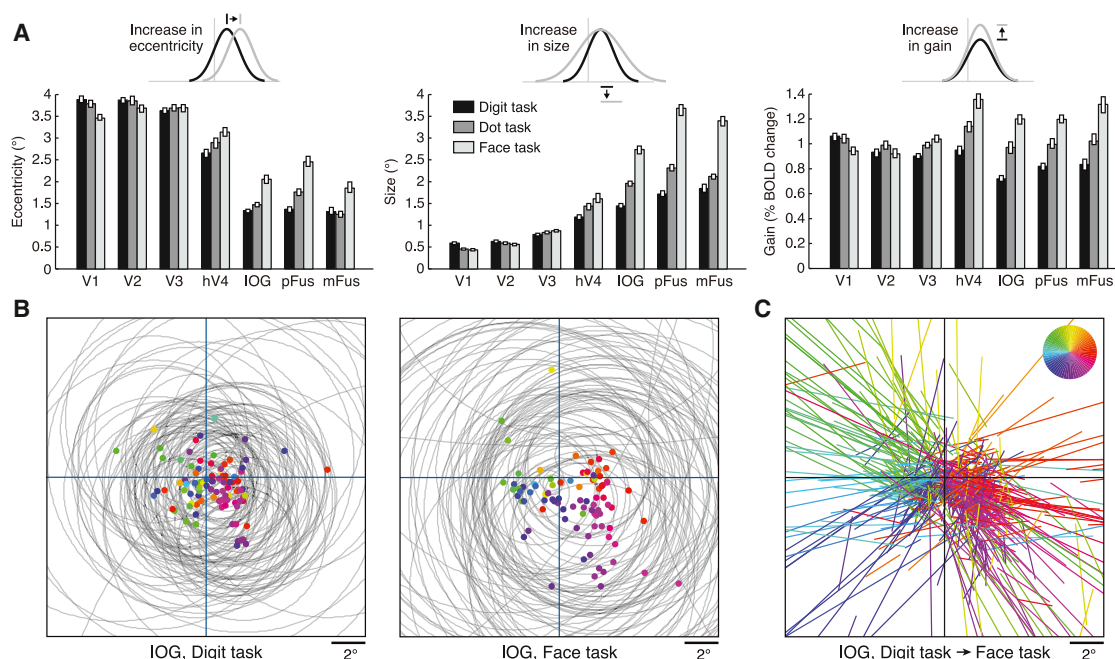
The observed attentional modulations of pRFs are consistent with the theory that neural responses in visual cortex reflect the combination of bottom-up stimulus drive and a top-down

attentional field that enhances responses to stimuli at the current locus of attention [23, 26, 29]. While both implicit (dot task) and explicit (face task) attention toward faces lead to response enhancement, we find that explicit attention toward faces produces larger modulations (see Figure 3A). This suggests that responses in the ventral visual pathway are modulated by both spatial and object-based attention, consistent with recent demonstrations of category-based attentional effects in the ventral pathway [30]. An interesting subject for future work is examining whether the attentional modulations observed here can be quantitatively described as an interaction between a global attentional field [7, 10, 23] and local classical receptive fields. Recent data suggest that the effect of a global attentional field on pRFs depends on pRF size, with larger effects obtained for larger pRFs [10]. Thus, these models predict larger attentional shifts of pRFs at higher stages of the visual processing hierarchy. We facilitate efforts to examine such questions and to further develop attentional models by making our data publicly available (<http://kendrickkay.net/vtcddata/>).

One question that stems from our findings is whether the demonstrated impact of attention on cortical responses has behavioral consequences. We hypothesize that attention-induced changes in the representation of spatial information in VTC may affect behavioral judgments of spatial position. Specifically, reduction of neural spatial uncertainty during the dot and face tasks compared to the digit task suggests that behavioral judgments of face position would be more accurate during the dot and face tasks. This hypothesis can be tested in future behavioral studies.

Another open question is exactly how the attentional modulations measured with fMRI manifest at the level of individual neurons. As prior electrophysiological studies have demonstrated that attention modulates neuronal firing rates in monkey IT [26, 28], we hypothesize that similar attentional modulations of receptive fields (RFs) occur for individual neurons in the ventral visual pathway. Notably, our observation of task-dependent pRFs might explain the variability of previous reports of neuronal RF sizes in monkey IT: RFs were largest during passive viewing [31, 32] and anesthesia [19], whereas RFs were smallest during demanding discrimination tasks near the fovea [33].





**Figure 3. Attention Modulates pRF Properties in VTC**

pRFs were measured under three tasks using the same stimulus. While maintaining fixation, subjects performed a one-back task on centrally presented digits (digit task), detected a small dot superimposed on the faces (dot task), or performed a one-back task on face identity (face task).

(A) Summary of results. Each bar represents a region under a single task (median across voxels  $\pm$  68% CI). In hV4, and more so in IOG, pFus, and mFus, attending to the stimulus (dot task, face task) causes an increase in pRF eccentricity, size, and gain compared to attending to fixation (digit task). These effects are larger for the face task than for the dot task.

(B) Visualization of pRFs for 100 randomly selected voxels from an example region, bilateral IOG (colored dots represent pRF centers; gray circles represent pRF sizes). In the digit task, pRFs are small and cluster near the fovea, whereas in the face task, pRFs are large and spread out into the periphery.

(C) Visualization of pRF shifts for region IOG. For each voxel, a line is drawn that connects the pRF center under the digit task to the pRF center under the face task; color indicates the direction of the shift (see legend), and the same color is used for the corresponding dots in (B). In general, pRFs shift away from the center. Although it appears as if there are many shifts to far eccentricities, the majority (81%) of pRF centers under the face task are actually located within 5° eccentricity.

See also [Figure S3](#).

### Rethinking Position Tolerance in the Ventral Visual Pathway

Position and size tolerance are considered key features of the ventral visual pathway, useful for object and face recognition. Tolerance indicates reduced sensitivity to incidental properties of a stimulus, such as the specific position or size at which it is viewed [34, 35]. Prevailing theories suggest that tolerance is achieved by systematic increase in RF sizes across processing stages in the ventral visual pathway [19, 20, 22, 36, 37]. Intuitively, a large RF implies that a wide range of stimulus positions and sizes drives the neural response [12].

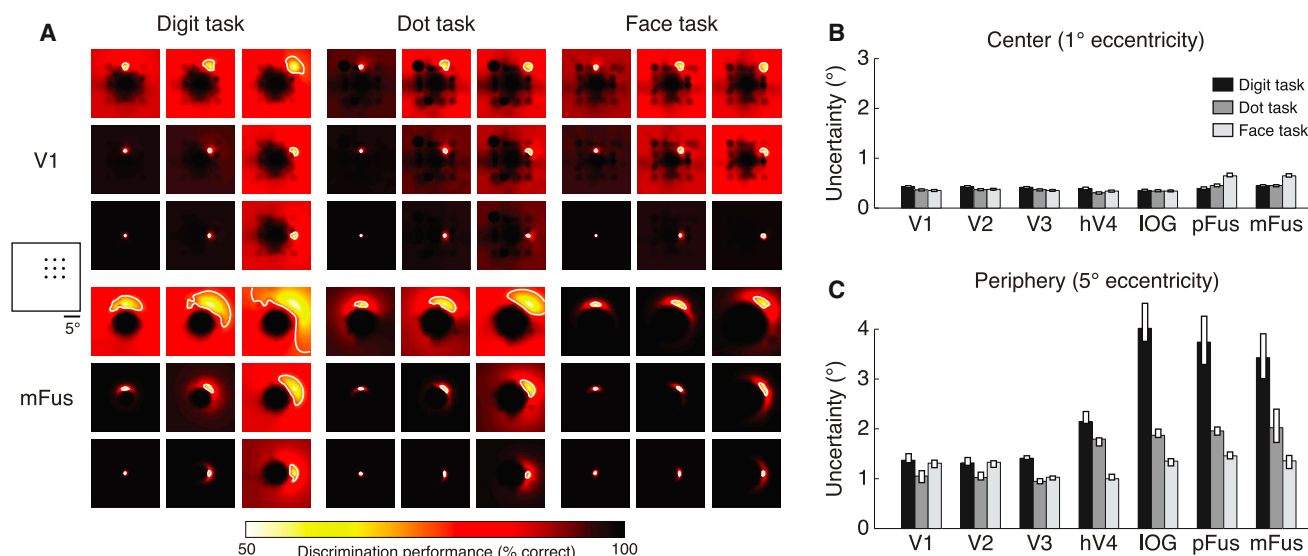
Although our pRF measurements are consistent with this account and reveal a hierarchy of pRF sizes within VTC, there are two aspects of our data that prompt a rethinking of position tolerance in VTC. First, we show that position tolerance at the level of individual voxels is partially the result of top-down attentional mechanisms and not simply due to static RF properties (see also [38]). Specifically, we find that when subjects attend to the stimulus, pRFs enlarge, thereby increasing position tolerance. Second, we show that the common intuition that larger pRFs degrade spatial information may be misleading. Despite the enlargement of pRFs when subjects attend the stimulus, the spatial precision with which the location of the stimulus is represented in VTC improves, rather than worsens.

At first glance, these observations seem inconsistent: how can attention increase spatial tolerance while also increasing

spatial precision? The answer lies in the distinction between the local scale (i.e., information carried by a single voxel) and the global scale (i.e., information carried by distributed responses across voxels). At the local scale, each individual voxel shows reduced sensitivity to stimulus location due to increased pRF size. However, at the global scale, sensitivity to stimulus location improves due to increased pRF coverage and scatter in the periphery (Figures 3B and 3C), which together provide a better tiling of the visual field.

### Conclusions

We have used a model-based approach to understand how attention influences representation in visual cortex. Our approach consisted of measuring responses to a wide range of stimulus conditions [39, 40], developing an encoding model that describes how stimulus information is represented locally [11, 41, 42], and using decoding analyses to quantify the information present in distributed responses [43]. Importantly, although we implemented this approach with fMRI, the approach is general and can be applied to other experimental techniques, such as electroencephalography (EEG), magnetoencephalography (MEG), electrocorticography (ECoG), and electrophysiology. Comparing results from different experimental techniques in a common model-based framework may help elucidate the neural signals measured by different techniques [44] and may help resolve discrepancies in the sizes



**Figure 4. Attention Reduces Spatial Uncertainty in VTC**

For each region and task, we assess the quality of the representation of spatial information using a model-based decoding analysis. This analysis quantifies how well a linear classifier can discriminate stimuli at different visual field positions from a stimulus at a reference position.

(A) Example results for a  $3 \times 3$  grid of reference positions in the upper-right visual field (left inset). Each image is a map of discrimination performance for one reference position (indicated by the relative position of the image). We define spatial uncertainty as the square root of the area of the 75% correct contour (white line).

(B) Uncertainty at 1° eccentricity. Each bar represents uncertainty in a region under a single task (median across angular positions  $\pm$  68% CI). All regions exhibit low uncertainty, irrespective of the task.

(C) Uncertainty at 5° eccentricity. Face-selective regions IOG, pFus, and mFus exhibit high uncertainty under the digit task. However, this uncertainty is dramatically reduced under the dot and face tasks.

See also [Figure S4](#).

of attentional effects found by different techniques [45]. Overall, our study reveals that spatial information is systematically represented in the ventral visual pathway and that attention modifies and enhances this spatial representation. These results provide important insights into how position coding is implemented in the ventral visual pathway.

#### 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.2014.12.050>.

#### Author Contributions

K.N.K. conducted the experiment and analyzed the data. K.S.W. performed localizer experiments and assisted with data collection. K.N.K., K.S.W., and K.G.-S. conceived and designed the experiments. K.N.K., K.S.W., and K.G.-S. wrote the paper.

#### Acknowledgments

We thank T. Naselaris, A. Rokem, J. Winawer, and E. Zohary for helpful discussions. We also thank A. Stigliani for assistance with behavioral experiments; J. Winawer for providing retinotopic mapping data; and N. Withoft for assisting in the collection of face photographs. This work was supported by the McDonnell Center for Systems Neuroscience and Arts & Sciences at Washington University (K.N.K.), NEI grant 1R01EY02391501A1 (K.G.-S.), and NEI grant R01 EY03164 (to Brian Wandell). Computations were performed using the facilities of the Washington University Center for High Performance Computing, which were partially provided through grant NCRR 1S10RR022984-01A1.

Received: September 7, 2014

Revised: November 17, 2014

Accepted: December 18, 2014

Published: February 19, 2015

#### References

1. Ungerleider, L.G., and Mishkin, M. (1982). Two cortical visual systems. In *Analysis of Visual Behavior*, D.J. Ingle, M.A. Goodale, and R.J.W. Mansfield, eds. (MIT Press), pp. 549–586.
2. Goodale, M.A., Milner, A.D., Jakobson, L.S., and Carey, D.P. (1991). Object awareness. *Nature* 352, 202.
3. Schwarzklose, R.F., Swisher, J.D., Dang, S., and Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proc. Natl. Acad. Sci. USA* 105, 4447–4452.
4. Kravitz, D.J., Kriegeskorte, N., and Baker, C.I. (2010). High-level visual object representations are constrained by position. *Cereb. Cortex* 20, 2916–2925.
5. Carlson, T., Hogendoorn, H., Fonteijn, H., and Verstraten, F.A.J. (2011). Spatial coding and invariance in object-selective cortex. *Cortex* 47, 14–22.
6. Silver, M.A., Ress, D., and Heeger, D.J. (2005). Topographic maps of visual spatial attention in human parietal cortex. *J. Neurophysiol.* 94, 1358–1371.
7. Sprague, T.C., and Serences, J.T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nat. Neurosci.* 16, 1879–1887.
8. Szczepanski, S.M., Koenen, C.S., and Kastner, S. (2010). Mechanisms of spatial attention control in frontal and parietal cortex. *J. Neurosci.* 30, 148–160.
9. Saproo, S., and Serences, J.T. (2010). Spatial attention improves the quality of population codes in human visual cortex. *J. Neurophysiol.* 104, 885–895.
10. Klein, B.P., Harvey, B.M., and Dumoulin, S.O. (2014). Attraction of position preference by spatial attention throughout human visual cortex. *Neuron* 84, 227–237.
11. Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660.
12. Kay, K.N., Winawer, J., Mezer, A., and Wandell, B.A. (2013). Compressive spatial summation in human visual cortex. *J. Neurophysiol.* 110, 481–494.

13. Weiner, K.S., and Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* 52, 1559–1573.
14. Levy, I., Hasson, U., Avidan, G., Hendler, T., and Malach, R. (2001). Center-periphery organization of human object areas. *Nat. Neurosci.* 4, 533–539.
15. Yue, X., Cassidy, B.S., Devaney, K.J., Holt, D.J., and Tootell, R.B.H. (2011). Lower-level stimulus features strongly influence responses in the fusiform face area. *Cereb. Cortex* 21, 35–47.
16. Weiner, K.S., Golarai, G., Caspers, J., Chuapoco, M.R., Mohlberg, H., Zilles, K., Amunts, K., and Grill-Spector, K. (2014). The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage* 84, 453–465.
17. Kay, K.N., Rokem, A., Winawer, J., Dougherty, R.F., and Wandell, B.A. (2013). GLMdenoise: a fast, automated technique for denoising task-based fMRI data. *Front Neurosci* 7, 247.
18. 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.
19. Gross, C.G., Bender, D.B., and Rocha-Miranda, C.E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* 166, 1303–1306.
20. Bousssaoud, D., Desimone, R., and Ungerleider, L.G. (1991). Visual topography of area TEO in the macaque. *J. Comp. Neurol.* 306, 554–575.
21. Op De Beeck, H., and Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *J. Comp. Neurol.* 426, 505–518.
22. Issa, E.B., and DiCarlo, J.J. (2012). Precedence of the eye region in neural processing of faces. *J. Neurosci.* 32, 16666–16682.
23. Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. *Neuron* 61, 168–185.
24. Snippe, H.P., and Koenderink, J.J. (1992). Discrimination thresholds for channel-coded systems. *Biol. Cybern.* 66, 543–551.
25. Treue, S., and Maunsell, J.H. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382, 539–541.
26. Moran, J., and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science* 229, 782–784.
27. Connor, C.E., Preddie, D.C., Gallant, J.L., and Van Essen, D.C. (1997). Spatial attention effects in macaque area V4. *J. Neurosci.* 17, 3201–3214.
28. Richmond, B.J., Wurtz, R.H., and Sato, T. (1983). Visual responses of inferior temporal neurons in awake rhesus monkey. *J. Neurophysiol.* 50, 1415–1432.
29. Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
30. Çukur, T., Nishimoto, S., Huth, A.G., and Gallant, J.L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nat. Neurosci.* 16, 763–770.
31. Desimone, R., Albright, T.D., Gross, C.G., and Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* 4, 2051–2062.
32. Tovee, M.J., Rolls, E.T., and Azzopardi, P. (1994). Translation invariance in the responses to faces of single neurons in the temporal visual cortical areas of the alert macaque. *J. Neurophysiol.* 72, 1049–1060.
33. DiCarlo, J.J., and Maunsell, J.H. (2003). Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *J. Neurophysiol.* 89, 3264–3278.
34. DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How does the brain solve visual object recognition? *Neuron* 73, 415–434.
35. Poggio, T., and Ullman, S. (2013). Vision: are models of object recognition catching up with the brain? *Ann. N Y Acad. Sci.* 1305, 72–82.
36. Yamins, D.L.K., Hong, H., Cadieu, C.F., Solomon, E.A., Seibert, D., and DiCarlo, J.J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci. USA* 111, 8619–8624.
37. Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proc. Natl. Acad. Sci. USA* 104, 6424–6429.
38. Olshausen, B.A., Anderson, C.H., and Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.* 13, 4700–4719.
39. Kay, K.N., Naselaris, T., Prenger, R.J., and Gallant, J.L. (2008). Identifying natural images from human brain activity. *Nature* 452, 352–355.
40. Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., and Bandettini, P.A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* 60, 1126–1141.
41. Naselaris, T., Kay, K.N., Nishimoto, S., and Gallant, J.L. (2011). Encoding and decoding in fMRI. *Neuroimage* 56, 400–410.
42. Kay, K.N. (2011). Understanding visual representation by developing receptive-field models. In *Visual Population Codes: Towards a Common Multivariate Framework for Cell Recording and Functional Imaging*, N. Kriegeskorte and G. Kreiman, eds. (MIT Press), pp. 133–162.
43. Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., and Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
44. Winawer, J., Kay, K.N., Foster, B.L., Rauschecker, A.M., Parvizi, J., and Wandell, B.A. (2013). Asynchronous broadband signals are the principal source of the BOLD response in human visual cortex. *Curr. Biol.* 23, 1145–1153.
45. Boynton, G.M. (2011). Spikes, BOLD, attention, and awareness: a comparison of electrophysiological and fMRI signals in V1. *J. Vis.* 11, 12.