progress in neurobiology

Peer Review Overview

Manuscript Title: "Point-spread function of the BOLD response across columns and cortical depth in human extra-striate cortex"

Received	08-Feb-2020
1st Decision	06-Sep-2020
1st Revision Submitted	26-Nov-2020
2nd Decision	22-Jan-2021
2nd Revision Submitted	08-Mar-2021
Accepted	14-Mar-2021

1st Decision Letter

Dear Dr. Fracasso,

Thank you for submitting your manuscript to Progress in Neurobiology.

We have completed our evaluation of your manuscript. The reviewers recommend reconsideration of your manuscript following major revision. We invite you to resubmit your manuscript after addressing the comments below. Please resubmit your revised manuscript by Nov 05, 2020.

When revising your manuscript, please consider all issues mentioned in the reviewers' comments carefully: please outline every change made in response to their comments and provide suitable rebuttals for any comments not addressed. Please note that your revised submission may need to be re-reviewed.

To submit your revised manuscript, please log in as an author at https://www.editorialmanager.com/proneu/, and navigate to the "Submissions Needing Revision" folder.

Progress in Neurobiology values your contribution and we look forward to receiving your revised manuscript.

Kind regards,

Essa Yacoub
Guest Editor

Sabine Kastner
Editor-in-Chief
Progress in Neurobiology

Editor and Reviewer comments:

Reviewer #1: This paper reports an approach to estimating spatial resolution is innovative and valuable. The quality of the anatomical images is truly superb. The authors are to be commended for putting significant effort into illustrating the registration between functional and anatomical data, which is key to all data interpretation (and surprisingly good, especially in view of the fact that it appears as though no distortion compensation was applied to the data).

My primary concern on reading the original version of the manuscript is with the interpretation of the lower limit for blurring in the GE EPI BOLD data and comparison between this number and the blurring associated with the anatomical images. The functional data were acquired with half the resolution as the anatomical data. Does the voxel size not establish a floor for estimated PSF? My rough sense is that a

voxel size of 1 mm (the resolution of the functional data) would be associated with a PSF of ~0.7 mm FWHM, and necessary motion compensation, distortion compensation, and resampling to anatomical space would further blur the data, raising the floor. But I haven't done the appropriate simulation for this dataset. The manuscript reports that the PSF floor for the functional data is about 2X larger than the floor for the anatomical data. But since the resolution of the functional data was about 2X lower, the possible effects of voxel size in shaping the conclusions of the paper should be considered more carefully in a revised manuscript.

My other concerns are rather minor:
- the Introduction is quite long and somewhat redundant with the Discussion. The introduction would benefit from being shortened by removal of material covered in the discussion.
- Methods, page 5, line 7: what was the role of the prisms in the participant's optical path?
- section 2.4: specifying the particular flags used with 3dAllineate would help the reader know whether non-linear distortion compensation was indeed applied to the data and what kind of care was taken in resampling to the volume space to minimize blurring. Since motion- and distortion-compensation increase PSF of functional data, it is possible that the resolution floor is also set by necessary data handling.
- was a GM mask applied to the activation shown in Fig. 2 or are the data truly limited to GM that nicely?
- at several points in the ms it is mentioned that the smoothing done to simulate PSFs of different sizes was done iteratively. Please specify exactly what is meant by iterative smoothing.
- the first paragraph of the Results should be re-written to more clearly explain the estimation of fMRI responses. This paragraph, read by itself, does not explain accurately.
- page 11, line 30, says that the functional features appear more distinct near the WM surface (Fig. 3abcd). However, to my eye, the opposite seems true -- Fig. 3e shows more distinct feature than Fig. 3a. More explanation seems to be in order.
- for Figures 3 and 4, it would be nice to have some labels that indicate where V1 is and what direction is anterior, posterior, dorsal, ventral ...
- at several points, the manuscript refers to 1.0-mm functional voxels as "ultra-high" resolution, but really, this just seems like high resolution. Let's save the use of "utlra-high" for 0.5 mm.
- in the discussion of PSF on pg. 16, it would be nice to acknowledge the heterogeneity of voxel blurring profiles, e.g., Kriegeskorte 2009
- in the Methods, and again in the Discussion (section 4.3.2) it is mentioned that the anatomical data were binarized in a way that assigned an equal number of voxels to "magno" and "parvo" stripes. However, anatomical studies show that the stripes are thick and thin, so forcing the data into equal populations is likely to introduce error.
This decision to ignore the fact that stripes responding to high temporal frequencies are thicker than stripes responding to low temporal frequencies should be justified or reconsidered.

Reviewer #2: The authors estimate the depth-dependent PSF of the BOLD response. To this end, they measure BOLD-fMRI responses to parvo-like and magno-like visual stimuli in area V2, where the cytochrome-oxidase (CO) stripes organization provides a structure with a variation of T1w MRI signal and fMRI responses to the two stimuli. They estimate an "anatomical PSF", a BOLD-fMRI PSF, and a BOLD-fMRI PSF after accounting for the confounding "anatomical PSF." They report that GE BOLD PSF increases approximately linearly along the cortical depth, and that "anatomical PSFs" are on average smaller than functional PSF and remain unchanged with increasing cortical depth.

Comments:
Relying on the CO stripes for estimating the BOLD PSF is a creative idea. However, I have concerns with regard to implementation and terminology.

Main Comments:
(1) Regarding: implementation. The straight forward way to rely on the CO stripes for estimating the BOLD PSF is to first identify the stripes and the boundaries between stripes using the anatomical MRI. Then, to sample the T1w and the BOLD responses across the boundaries between stripes. The authors should take this approach.
(2) Instead, the authors measure fMRI in V2, but it seems that their measurements are quite independent of the CO stripes. They define the anatomical boundaries based on an arbitrary threshold (median T1w signal). It is unclear how this approach delineates the real boundaries between stripes. They define the boundaries between the responses with preferences to the parvo- and magno-like stimuli in a similar manner, based on an arbitrary threshold. It seems that the anatomical boundaries are not directly related to the stripes and that the fMRI boundaries are not related to the defined anatomical boundaries. This makes the interpretation of the results difficult. It questions whether the BOLD-PSF minus the "anatomical PSF" has any interpretable meaning.

(3) Along the same lines, it is also a concern that the anatomical and functional data are obtained in different sessions. This makes the alignment between the two modalities less straight forward relative to measurements done in the same session.
(4) To align the anatomical and functional data from the two sessions, the authors use a Pearson-Correlation base cost function. Since the T1w and the GE-BOLD have different intensity levels and contrasts, using mutual information would be more appropriate.
(5) Regarding: terminology. The BOLD PSF is a well-defined concept. However, I believe that the term 'anatomical PSF' is not meaningful. If the authors were to measure the variation in the T1w signal across the boundaries between the stripes (which apparently, they do not measure, see my previous comment), they could use the term 'gradient' to describe the change in T1w signal (likely reflecting variation in myelin content) across these boundaries. What does the term 'anatomical PSF' mean? The features of the anatomy do not "spread." They vary as a function of space, not necessarily in the same way in different parts of the cortex.

Additional comments:
(6) The authors should include a discussion on how the relatively low spatial resolution they used influences their capacity to measure variations as a function of cortical depth. I expect that the relatively large voxels smooth the depth-dependent measurements and can only show large -scale trends, possibly missing fine structures.
(7) The authors discard the first 13x2.6 sec worth of data to avoid the analysis of data obtained in a non-steady state. This seems over-exaggerated.
(8) The authors report that "GE BOLD PSF increases approximately linearly along the cortical depth." This can be interpreted such that the PSF increases with increasing depth. Instead, I recommend mentioning that the BOLD PSF is maximal in the superficial part of the cortex, and it decreases with increasing cortical depth.

1st Author Response Letter

Editor and Reviewer comments:

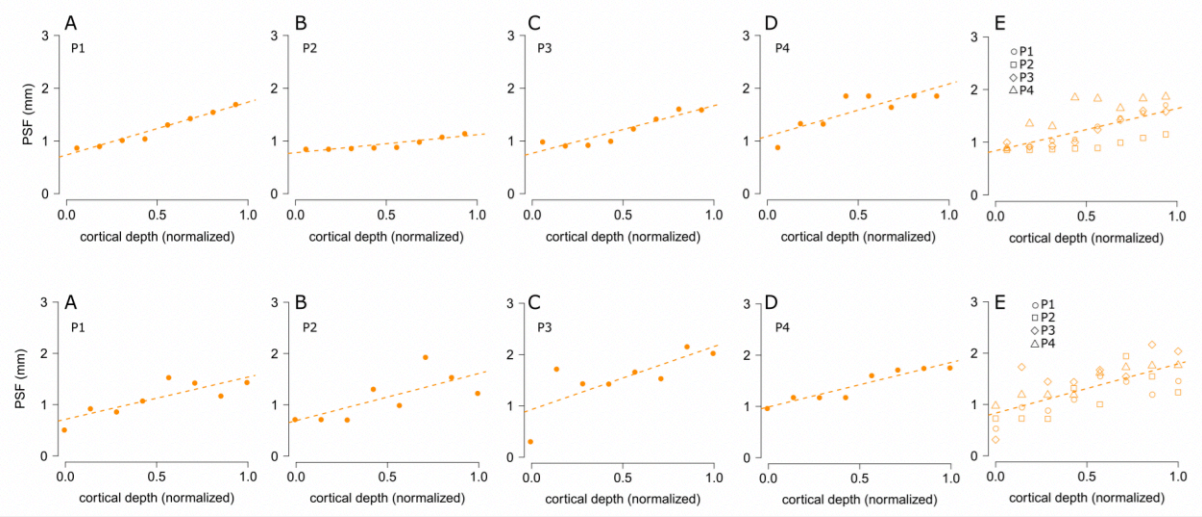
We would like to thank the reviewers for the comments they provided. We believe they have helped us greatly improving the quality of the manuscript.
Major changes in the main manuscript are reported in green.
In this rebuttal, replies are reported in regular Arial font, green. Extracts from the main manuscript are reported in Arial, italic, green.

Reviewer #1: This paper reports an approach to estimating spatial resolution is innovative and valuable. The quality of the anatomical images is truly superb. The authors are to be commended for putting significant effort into illustrating the registration between functional and anatomical data, which is key to all data interpretation (and surprisingly good, especially in view of the fact that it appears as though no distortion compensation was applied to the data).
Thank you

My primary concern on reading the original version of the manuscript is with the interpretation of the lower limit for blurring in the GE EPI BOLD data and comparison between this number and the blurring associated with the anatomical images. The functional data were acquired with half the resolution as the anatomical data. Does the voxel size not establish a floor for estimated PSF? My rough sense is that a voxel size of 1 mm (the resolution of the functional data) would be associated with a PSF of ~0.7 mm FWHM, and necessary motion compensation, distortion compensation, and resampling to anatomical space would further blur the data, raising the floor. But I haven't done the

appropriate simulation for this dataset. The manuscript reports that the PSF floor for the functional data is about 2X larger than the floor for the anatomical data. But since the resolution of the functional data was about 2X lower, the possible effects of voxel size in shaping the conclusions of the paper should be considered more carefully in a revised manuscript.

We have reflected on this concern and the main concern expressed by reviewer 2 (about our PSF estimation method in the previous version of the manuscript). In the current version of the manuscript we decided to modify the BOLD PSF estimation method, to avoid blurring and direct comparison with the anatomical data. The sections about the estimation of the floor PSF have been removed from the current version of the manuscript. We focus our efforts on functionally defined stripes and measuring the GE BOLD PSF. This is performed using the rate of change of phase specified coherence occurring at the border between successive V2 stripes in humans. For this purpose, we measured the geodesic distance from the boundary between successive V2 stripes, defined as 0, and used a modelling approach to estimate GE BOLD PSF isolating components that can be ascribed to global signal changes (see sections 2.6.1, 2.6.2, 2.6.3 and 2.6.4 in the current version of the manuscript). The new method yields similar GE BOLD estimates as those obtained in the previous version of the manuscript:



The upper row shows PSF estimates from each participant across cortical depth obtained with the method adopted in the previous version of the manuscript; the bottom row shows the new estimates (current version of the manuscript). Overall estimates are very similar.

We perform a similar analysis on T1-w data only with the purpose of validating our method and do not directly compare BOLD and T1-w data.

Below we report sections 2.6.1 and 2.6.2 from the current version of the manuscript for a description of the method:

2.6 Data analysis

2.6.1 Determination of the border between two consecutive V2 stripes from phase-specified coherence (GE BOLD)

To estimate the GE BOLD PSF we started by building a map measuring the distance in millimetres from the boundary between two consecutive stripes. This distance is computed using the geodesic distance between each two points over the cortical surface for each surface along cortical depth.
First, we used the phase specified coherence to identify the stripes in human V2 responding to high or low temporal frequencies. Phase specified coherence is ideal for this measurement. In our experimental design, a positive phase specified coherence value indicates a relative preference for low temporal frequency, a negative phase specified coherence value indicates a relative preference for high temporal frequency. A phase specified coherence of 0 indicates the shift between a preference for low and high temporal frequency. Hence, 0 demarcates the border between different stripes in human V2.
We identified the nodes on the surface that prefer high or low temporal frequency ('positive' and 'negative' set of nodes, respectively). We computed the geodesic distance from every node on the positive to every node on the negative group. For each node from the positive group we assigned the smaller distance among all the possible geodesic distances computed towards the negative nodes. We repeated the same operation from the negative to the positive nodes (See Fig. 3). This operation results in a geodesic distance map, indicating the smallest distance from each node to the boundary between successive stripes in V2. A single geodesic distance map was obtained for each surface along cortical depth for human V2. We used these geodesic distance maps as a basis to estimate the GE BOLD PSF across cortical depth.

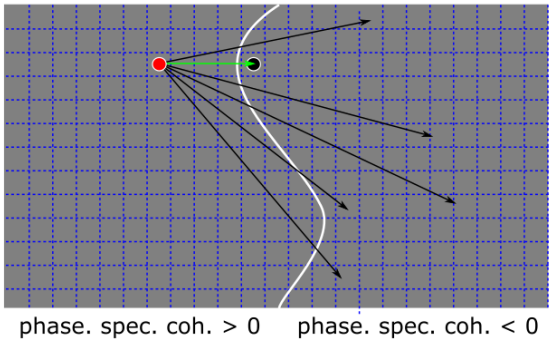


Figure 3. Computing the geodesic distance map (cortical distance)
The grey area depicts a portion of V2 over the cortical surface. Normally, this portion would be a two-dimensional sheet that extends in a three-dimensional space. To simplify, in this example we consider the two-dimensional sheet outlined by the grey area above. The grey area is composed by nodes, represented by the squares delimited by dashed lines. The grey area is divided into two portions (two consecutive stripes), where the phase specified coherence is larger than 0 and smaller than 0, respectively. The border between these two portions is identified by the white line and it divides 'positive nodes' form 'negative nodes'). We computed the geodesic distance from every node in the positive

group to every node in the negative group. For each node from the positive group, we assigned the smaller distance among all the possible geodesic distances computed. In the example above we highlight one node among all the positive nodes (red dot) and several distances to negative nodes (arrows). The smallest distance computed from the red dot is represented by the green arrow. We assign the distance depicted by the green arrow to the red node. To obtain a complete geodesic distance map, we repeat the same operation for all positive nodes and negative nodes.

For each surface across cortical depth and each node we obtained the phase specified coherence (ranging between -1 and 1, see Section 2.4) and the geodesic distance (from now on referred to as: 'cortical distance'), representing the distance from the border between consecutive stripes in human V2 (see Fig. 4). Cortical distance data was binned based on 2% quantile bins. Within each cortical distance bin, we computed the average phase specified coherence as well as cortical distance and plot the former as a function of the latter (see Fig. 4C).

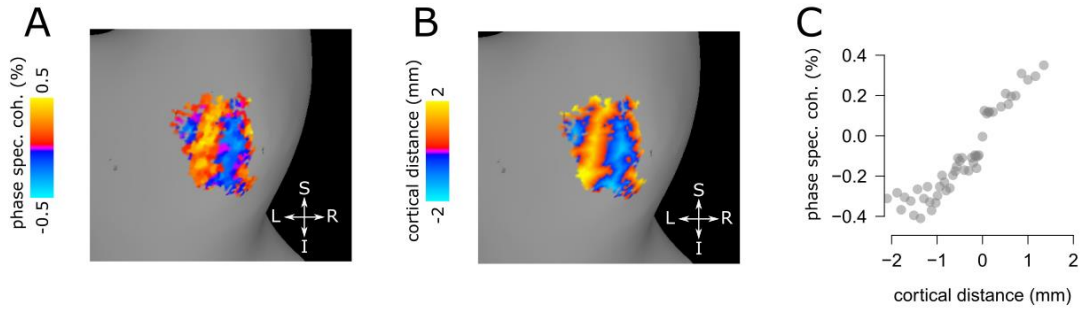


Figure 4. Cortical distance map and phase specified coherence.
We selected a small portion of human V2 (single participant, dorsal V2, left hemisphere), to show the result of our cortical distance measure. Note that the analysis presented in the paper was performed over the entire V2 ROI. Panel A. Phase-specified coherence for a portion of V2. Panel B. Cortical distance derived with the method described in Fig. 3 from the same portion of human V2. The resulting cortical distance ranges between -2 to 2 mm, representing the distance from the border between successive V2 stripes. Panel C. Scatterplot showing the relationship between phase specified coherence and cortical distance. Cortical distance data was binned based on 2% quantile bins and averaged within each bin. Phase specified coherence was averaged within the cortical distance bins. This plot shows the change of phase-specified coherence over cortical distance, across the border between two successive V2 stripes (represented by 0 in the x axis, cortical distance).

2.6.2 Estimation of BOLD PSF

We obtained the binned phase-specified coherence along cortical distance for each surface along cortical depth, for each participant. We fit a model to the measured phase-specified coherence along cortical distance, assuming a Gaussian FWHM.

The model consisted of a modified cumulative gaussian function (CGF). The standard CGF takes the form reported in equation 3 (Eq. 3). The function ranges from 0 to 1 and has location (μ) and sigma (σ) as parameters. We modified the parametrization of the standard CGF. First, we fixed the location parameter (μ) to zero, as we built the distance map placing the 0 at the cross between successive stripes. Second, we added a scaling constant, to account for different amplitudes of the signal. Third, we added a shift constant to accommodate the range of phase specified coherence values, which naturally varies between -1 and 1. Overall we obtained a model with 3 parameters, CGF_m (Eq. 4): sigma, scaling constant and shift constant (see Fig. 5).

Eq. 3:

$$\boldsymbol{CGF = \frac{1}{2}\left[1 + \operatorname{erf}\frac{(x-\mu)}{\sigma\sqrt{2}}\right]}$$

Eq. 4:

$$\boldsymbol{CGF_m = mult * \left(\frac{1}{2}\left[1 + \operatorname{erf}\frac{(x)}{\sigma\sqrt{2}}\right]\right) + shift}$$

Eq. 5:

$$\boldsymbol{\mathrm{FWHM} = 2\sqrt{2\ln 2}\sigma \approx 2.355\sigma}$$

Conveniently, in this parametrization the sigma (s) represents the width of the Gaussian PSF. We multiplied sigma by 2.355 (Eq. 5) to obtain the Gaussian full-width half-max, representing our estimate GE-BOLD PSF.

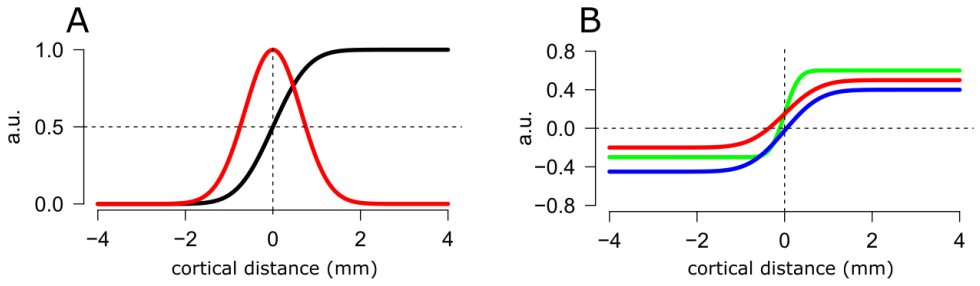


Figure 5, Modelling the GE-BOLD PSF

Panel A. The plot shows a Gaussian PSF (red line). We used the cumulative of the Gaussian (black line, (CGF)) to fit a model to the measured phase-specified coherence along cortical distance, for each surface along cortical depth. Panel B. We modified the parametrization of the standard CGF: we fixed the location parameter (μ) to zero (at the boundary between two successive stripes in human V2); we added a scaling parameter, to account to different amplitudes of the signal. Third, we added a shift parameter to accommodate the range of phase specified coherence values [-1, 1], see Eq. 4. The plot shows three CGFs with different parameters: green line: FWHM=0.6mm, scaling parameter: 0.9, shift parameter: -0.3. blue line: FWHM=1.5mm, scaling parameter: 0.85, shift parameter: -0.45. red line: FWHM=1.5mm, scaling parameter: 0.7, shift parameter: -0.2. Please note that the red and

blue lines appear to have a different rate of change around the zero crossing along cortical distance, however this difference is accounted for by a difference in amplitude, not FWHM (0.85 and 0.7 for the blue and red lines, respectively). The parametrization adopted in our modelling approach allows to disentangle between different components of the trend between phase specified coherence and cortical distance, and to detect differences that could otherwise be wrongly be ascribed to FWHM alone.

But since the resolution of the functional data was about 2X lower, the possible effects of voxel size in shaping the conclusions of the paper should be considered more carefully in a revised manuscript.

The sections about the estimation of the floor PSF have been removed from the current version of the manuscript. We have now added the following section in the discussion, to acknowledge that voxel size per se can drive the estimation of the FWHM parameter on T1w data:

4.3 Anatomical Width

Analysis of T1-w signal was performed for validating our modelling approach. Although we adopt the same modelling strategy for GE BOLD and T1-w signal, it is important to note that the interpretation of the FWHM parameter differs between the two modalities. The fitted FWHM for GE BOLD can be interpreted as an estimate of BOLD PSF, that is the spreading of BOLD signal largely due to the action of draining veins, in response to an otherwise localized neuronal activity (Uğurbil et al., 2003). On the other hand, the same parameter fitted for the T1-w signal cannot be interpreted as a measure of signal spread, as anatomical features per se do not spread and are not affected by the presence of large draining veins. However, the fitted FWHM for T1-w signal is a useful measure that represents the anatomical variability ('width') of the T1-w signal along successive V2 stripes in the absence of venous spread.

We find that anatomical width is about 0.4mm. It is important to highlight that this parameter remains flat along cortical depth, thus providing a validation for our analysis pipeline. Furthermore, the anatomical variation of about 0.4mm might reflect the acquisition resolution of the anatomical data (0.5mm isotropic), so it is possible that voxel size drives the estimates of anatomical width.

My other concerns are rather minor:

- the Introduction is quite long and somewhat redundant with the Discussion. The introduction would benefit from being shortened by removal of material covered in the discussion.

We have shortened the introduction as requested. Specifically, we shortened the section regarding the problems introduced by neuronal receptive field size and scatter when measuring PSF using retinotopy.

- Methods, page 5, line 7: what was the role of the prisms in the participant's optical path?

The prisms were used to reflect the image from the screen located at the back of the bore towards participants eyes. We have now added this information in the main manuscript.

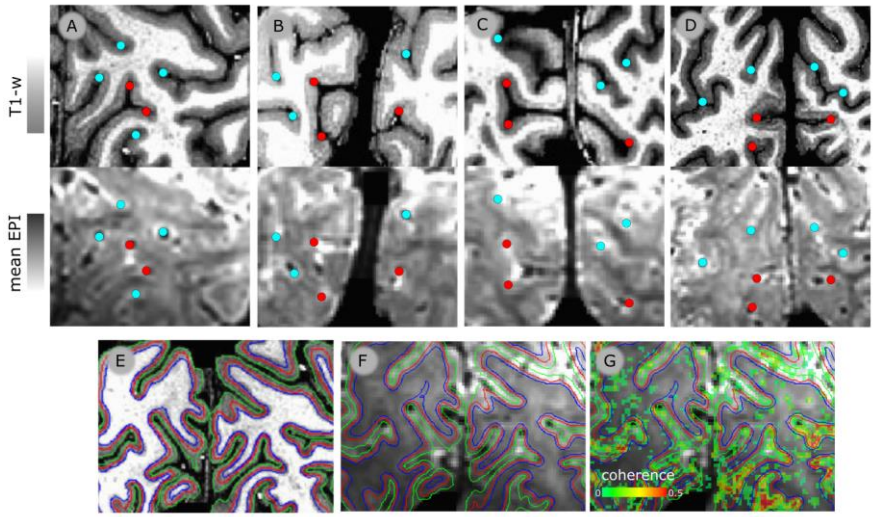
- section 2.4: specifying the particular flags used with 3dAllineate would help the reader know whether non-linear distortion compensation was indeed applied to the data and what kind of care was taken in resampling to the volume space to minimize blurring. Since motion- and distortion- compensation increase PSF of functional data, it is possible that the resolution floor is also set by necessary data handling.

The first step in the co-registration was to manually align the EPI to the anatomical data using the 'nudge' tool in AFNI.

After this first step we ran an automated co-registration algorithm (3dAllineate), using local Pearson Correlation as a cost function. We set the following option to the 3dAllineate algorithm: first we limited the rotation and the shift to 3 mm (-maxrot 3 maxshf 3), and we imposed a single pass only (-onepass). We have now added this information in the current version of the manuscript.

- was a GM mask applied to the activation shown in Fig. 2 or are the data truly limited to GM that nicely?

Yes, the map was limited to GM. In the current version of the manuscript we changed the panel, showing voxels that survived the thresholding within the EPI mask, without limiting it to GM.



Here is the current Figure 2, panel G shows the activation within the EPI mask, without limiting it to GM.

- at several points in the ms it is mentioned that the smoothing done to simulate PSFs of different sizes was done iteratively. Please specify exactly what is meant by iterative smoothing.

In the current version of the manuscript we changed the method to estimate the PSF.

In the previous version of the manuscript we simulated PSFs of different size iteratively, that is, we

selected a bank of possible kernel sizes. On each iteration we smoothed the binarized map with the given kernel size and compared the similarity (deviance) between the binarized/smoothed map and the observed map (empirical map).

- the first paragraph of the Results should be re-written to more clearly explain the estimation of fMRI responses. This paragraph, read by itself, does not explain accurately.

We have changed the first paragraph in the results section:

We recorded fMRI responses while participants viewed concentric gratings with contrast reversing at slow or fast rate (1.5 Hz or 7.5 Hz, respectively). We derived a measure of the fMRI response frequency and phase in response to the stimuli, adjusted for the hemodynamic delay (the phase-specified coherence). The values ranged between −1 and 1; positive values reflect stronger responses to the slow-rate stimulus (1.5 Hz) and negative values reflect stronger responses to the fast-rate stimulus (7.5 Hz).

- page 11, line 30, says that the functional features appear more distinct near the WM surface (Fig. 3abcd). However, to my eye, the opposite seems true -- Fig. 3e shows more distinct feature than Fig. 3a. More explanation seems to be in order.

We wrote this sentence having in mind the PSF results, scaling linearly with cortical depth. We have removed this sentence from the current version of the manuscript as it introduces an unwanted element of subjectivity.

- for Figures 3 and 4, it would be nice to have some labels that indicate where V1 is and what direction is anterior, posterior, dorsal, ventral ...
Now added

- at several points, the manuscript refers to 1.0-mm functional voxels as "ultra-high" resolution, but really, this just seems like high resolution. Let's save the use of "utlra-high" for 0.5 mm.

We have removed the wording ‘ultra-high’ in the current version of the manuscript

- in the discussion of PSF on pg. 16, it would be nice to acknowledge the heterogeneity of voxel blurring profiles, e.g., Kriegeskorte 2009
We now acknowledge the suggested paper in the current version of the manuscript

- in the Methods, and again in the Discussion (section 4.3.2) it is mentioned that the anatomical data were binarized in a way that assigned an equal number of voxels to "magno" and "parvo" stripes.

However, anatomical studies show that the stripes are thick and thin, so forcing the data into equal populations is likely to introduce error.

We agree with the reviewer, anatomical data does not allow for this clear demarcation between the stripes, and we would always need to include a somewhat arbitrary criteria to define the border between stripes. For this reason, we focus our efforts on functionally defined stripes and limited the analysis of T1-w data for validation purposes (see the current version of the manuscript, section 2.6.1, also reported above).

This decision to ignore the fact that stripes responding to high temporal frequencies are thicker than stripes responding to low temporal frequencies should be justified or reconsidered.

Regarding the number of identified stripes using phase specified coherence, we have now added a dedicated section in the discussion addressing this point:

4.2 V2 stripes

Phase-specified coherence allows us to identify two subdivisions in human V2 (see Figure 8). On the other hand previous neurophysiology studies distinguish three types of stripes (Horton, 1984; Livingstone and Hubel, 1982; Tootell et al., 1983). In a previous publication (Dumoulin et al., 2017) we interpret our stripe-based subdivisions as collapsing two stripe types into one, speculating that our subdivisions distinguish 'thick' versus 'thin' and 'pale' stripes combined. This speculation is based on previous non-human histological observations suggesting that 'thick' stripes in V2 receive a larger projection from layer 4B in V1, dominated by magnocellular input (Shipp and Zeki, 2002; Sincich and Horton, 2005a).

Reviewer #2: The authors estimate the depth-dependent PSF of the BOLD response. To this end, they measure BOLD-fMRI responses to parvo-like and magno-like visual stimuli in area V2, where the cytochrome-oxidase (CO) stripes organization provides a structure with a variation of T1w MRI signal and fMRI responses to the two stimuli. They estimate an "anatomical PSF", a BOLD-fMRI PSF, and a BOLD-fMRI PSF after accounting for the confounding "anatomical PSF." They report that GE BOLD PSF increases approximately linearly along the cortical depth, and that "anatomical PSFs" are on average smaller than functional PSF and remain unchanged with increasing cortical depth.

Comments:
Relying on the CO stripes for estimating the BOLD PSF is a creative idea. However, I have concerns with regard to implementation and terminology.

Main Comments:

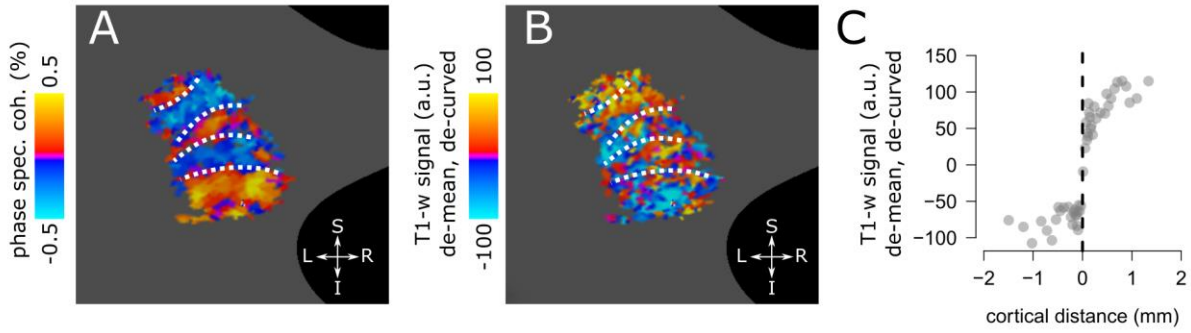
(1) Regarding: implementation. The straight forward way to rely on the CO stripes for estimating the BOLD PSF is to first identify the stripes and the boundaries between stripes using the anatomical MRI. Then, to sample the T1w and the BOLD responses across the boundaries between stripes. The authors should take this approach.

(2) Instead, the authors measure fMRI in V2, but it seems that their measurements are quite independent of the CO stripes. They define the anatomical boundaries based on an arbitrary threshold (median T1w signal). It is unclear how this approach delineates the real boundaries between stripes. They define the boundaries between the responses with preferences to the parvo- and magno-like stimuli in a similar manner, based on an arbitrary threshold. It seems that the anatomical boundaries are not directly related to the stripes and that the fMRI boundaries are not related to the defined anatomical boundaries. This makes the interpretation of the results difficult. It questions whether the BOLD-PSF minus the "anatomical PSF" has any interpretable meaning.

We considered this concern and the main concern expressed by reviewer 1. We decided to modify the PSF estimation method using the rate of change in phase specified coherence occurring at the border between successive V2 stripes in humans.

Phase-specified coherence allows us to characterize V2 stripes in humans. In our experimental design, a positive phase specified coherence value indicates a relative preference for low temporal frequency, a negative phase specified coherence value indicates a relative preference for high temporal frequency. A phase specified coherence of 0 indicates the shift between a preference for relatively low to relatively high temporal frequency. Hence, we can take 0 as the border between different stripes in human V2.

Phase-specified coherence in V2 is correlated with the underlying T1-w signal but the correlation is not perfect (~0.20). We report the correlation between phase-specified coherence and T1-w signal a previous paper (Dumoulin et al., 2017). While the borders between consecutive V2 stripes identified with phase-specified coherence and T1-w signal appear to be close, they do not coincide perfectly. We illustrate this point in the figure below:

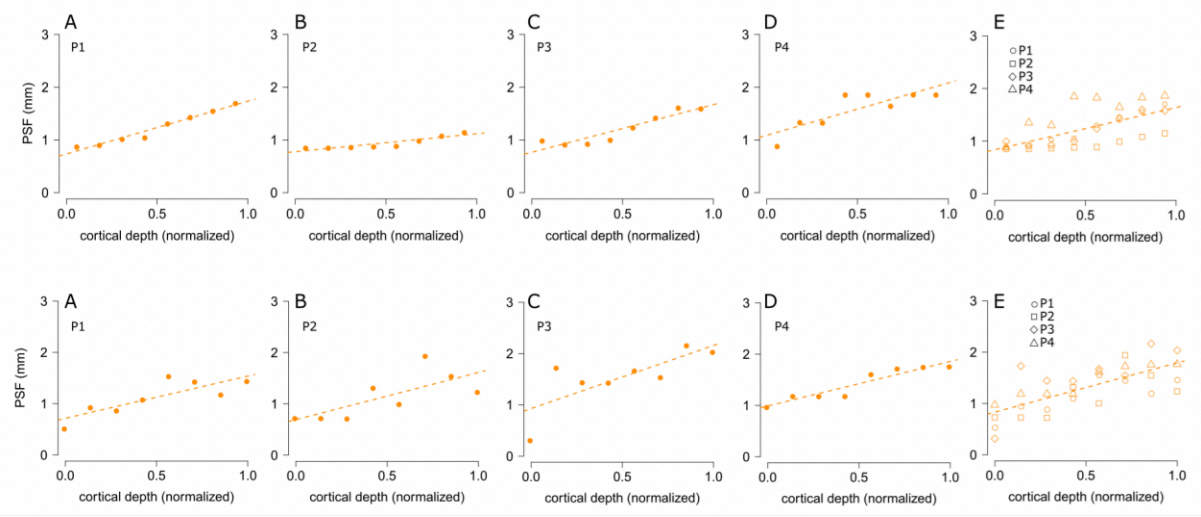


Cortical distance map and T1-w anatomy.

Here we selected a small portion of human V2 to compare between phase specified coherence and T1-w data. Panel A. Phase-specified coherence for a portion of V2 for one participant. Dashed lines represent manually defined borders between successive stripes based on phase specified coherence map. Panel B. The same area, T1-w signal (de-meaned, de-curved). While phase specified coherence and T1-w signal are related, the correlation is not perfect. Visual inspection suggests that the borders between consecutive V2 stripes identified with phase-specified coherence and T1-w anatomy appear to be close, but do not coincide perfectly. Panel C. Scatterplot showing the relationship between T1-w signal and cortical distance. Given the misalignment between T1-w data and phase-specified coherence, in Panel C cortical distance was computed based on T1-w signal (de-meaned, de-curved), not phase specified coherence. Data was binned based on 2% quantile bins and averaged within each bin. T1-w signal (de-meaned, de-curved) was averaged within the cortical distance bins.

Given the misalignment between T1-w data and phase-specified coherence, in the current version of the manuscript we identify successive V2 stripes using phase-specified coherence. We measured the geodesic distance from the boundary between successive V2 stripes, defined as 0, and used a modelling approach to estimate GE BOLD PSF, isolating components that could potentially be ascribed to global signal changes (see sections 2.6.1, 2.6.2, 2.6.3 and 2.6.4 in the current version of the manuscript).

It is important to note that the new method yields similar GE BOLD estimates as those obtained in the previous version of the manuscript:



The upper row shows PSF estimates from each participant across cortical depth obtained with the method adopted in the previous version of the manuscript; the bottom row shows the new estimates (current version of the manuscript). Overall, estimates are very similar.

In the current version of the manuscript we perform a similar analysis on T1-w data, only with the purpose of validating our method and do not directly compare BOLD and T1-w data.

Below we report sections 2.6.1 and 2.6.2 from the current version of the manuscript for a description of the method:

2.6 Data analysis

2.6.1 Determination of the border between two consecutive V2 stripes from phase-specified coherence (GE BOLD)

To estimate the GE BOLD PSF we started by building a map measuring the distance in millimetres from the boundary between two consecutive stripes. This distance is computed using the geodesic distance between each two points over the cortical surface for each surface along cortical depth.
First, we used the phase specified coherence to identify the stripes in human V2 responding to high or low temporal frequencies. Phase specified coherence is ideal for this measurement. In our experimental design, a positive phase specified coherence value indicates a relative preference for low temporal frequency, a negative phase specified coherence value indicates a relative preference for high temporal frequency. A phase specified coherence of 0 indicates the shift between a preference for low and high temporal frequency. Hence, 0 demarcates the border between different stripes in human V2.
We identified the nodes on the surface that prefer high or low temporal frequency ('positive' and 'negative' set of nodes, respectively). We computed the geodesic distance from every node on the positive to every node on the negative group. For each node from the positive group we assigned the smaller distance among all the possible geodesic distances computed towards the negative nodes. We repeated the same operation from the negative to the positive nodes (See Fig. 3). This operation results in a geodesic distance map, indicating the smallest distance from each node to the boundary between successive stripes in V2. A single geodesic distance map was obtained for each surface along cortical depth for human V2. We used these geodesic distance maps as a basis to estimate the GE BOLD PSF across cortical depth.

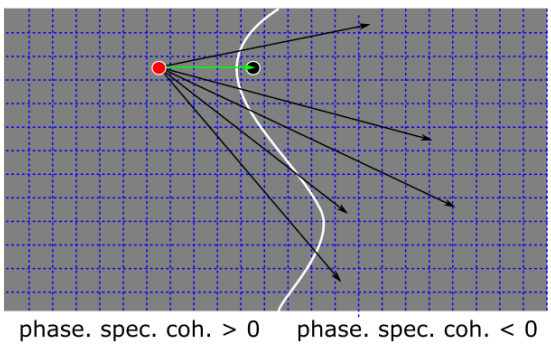


Figure 3. Computing the geodesic distance map (cortical distance)

The grey area depicts a portion of V2 over the cortical surface. Normally, this portion would be a two-dimensional sheet that extends in a three-dimensional space. To simplify, in this example we consider the two-dimensional sheet outlined by the grey area above. The grey area is composed by nodes, represented by the squares delimited by dashed lines. The grey area is divided into two portions (two consecutive stripes), where the phase specified coherence is larger than 0 and smaller than 0, respectively. The border between these two portions is identified by the white line and it divides 'positive nodes' form 'negative nodes'). We computed the geodesic distance from every node in the positive group to every node in the negative group. For each node from the positive group, we assigned the smaller distance among all the possible geodesic distances computed. In the example above we highlight one node among all the positive nodes (red dot) and several distances to negative nodes (arrows). The smallest distance computed from the red dot is represented by the green arrow. We assign the distance depicted by the green arrow to the red node. To obtain a complete geodesic distance map, we repeat the same operation for all positive nodes and negative nodes.

For each surface across cortical depth and each node we obtained the phase specified coherence (ranging between -1 and 1, see Section 2.4) and the geodesic distance (from now on referred to as: 'cortical distance'), representing the distance from the border between consecutive stripes in human V2 (see Fig. 4). Cortical distance data was binned based on 2% quantile bins. Within each cortical distance bin, we computed the average phase specified coherence as well as cortical distance and plot the former as a function of the latter (see Fig. 4C).

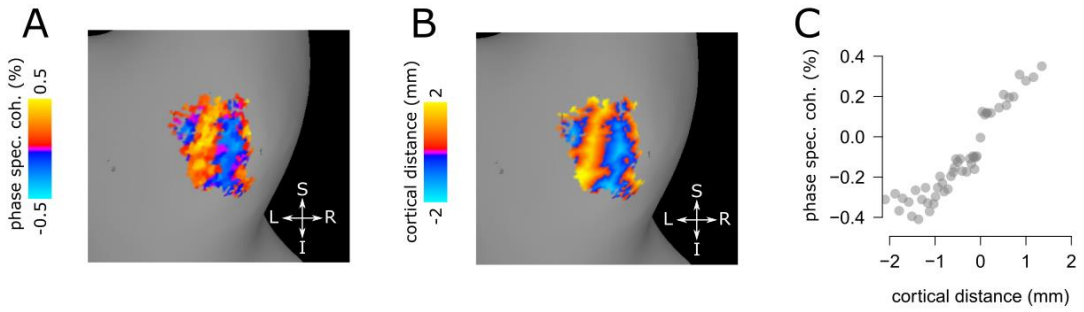


Figure 4. Cortical distance map and phase specified coherence.
We selected a small portion of human V2 (single participant, dorsal V2, left hemisphere), to show the result of our cortical distance measure. Note that the analysis presented in the paper was performed over the entire V2 ROI. Panel A. Phase-specified coherence for a portion of V2. Panel B. Cortical distance derived with the method described in Fig. 3 from the same portion of human V2. The resulting cortical distance ranges between -2 to 2 mm, representing the distance from the border between successive V2 stripes. Panel C. Scatterplot showing the relationship between phase specified coherence and cortical distance. Cortical distance data was binned based on 2% quantile bins and averaged within each bin. Phase specified coherence was averaged within the cortical distance bins. This plot shows the change of phase-specified coherence over cortical distance, across the border between two successive V2 stripes (represented by 0 in the x axis, cortical distance).

2.6.2 Estimation of BOLD PSF

We obtained the binned phase-specified coherence along cortical distance for each surface along cortical depth, for each participant. We fit a model to the measured phase-specified coherence along cortical distance, assuming a Gaussian FWHM.

The model consisted of a modified cumulative gaussian function (CGF). The standard CGF takes the form reported in equation 3 (Eq. 3). The function ranges from 0 to 1 and has location (μ) and sigma (σ) as parameters. We modified the parametrization of the standard CGF. First, we fixed the location parameter (μ) to zero, as we built the distance map placing the 0 at the cross between successive stripes. Second, we added a scaling constant, to account for different amplitudes of the signal. Third, we added a shift constant to accommodate the range of phase specified coherence values, which naturally varies between -1 and 1. Overall we obtained a model with 3 parameters, CGF_m (Eq. 4): sigma, scaling constant and shift constant (see Fig. 5).

Eq. 3:

$$\boldsymbol{CGF = \frac{1}{2}\left[1 + \operatorname{erf}\frac{(x-\mu)}{\sigma\sqrt{2}}\right]}$$

Eq. 4:

$$\boldsymbol{CGF_m = mult * \left(\frac{1}{2}\left[1 + \operatorname{erf}\frac{(x)}{\sigma\sqrt{2}}\right]\right) + shift}$$

Eq. 5:

$$\boldsymbol{\mathrm{FWHM} = 2\sqrt{2\ln 2}\sigma \approx 2.355\sigma}$$

Conveniently, in this parametrization the sigma (s) represents the width of the Gaussian PSF. We multiplied sigma by 2.355 (Eq. 5) to obtain the Gaussian full-width half-max, representing our estimate GE-BOLD PSF.

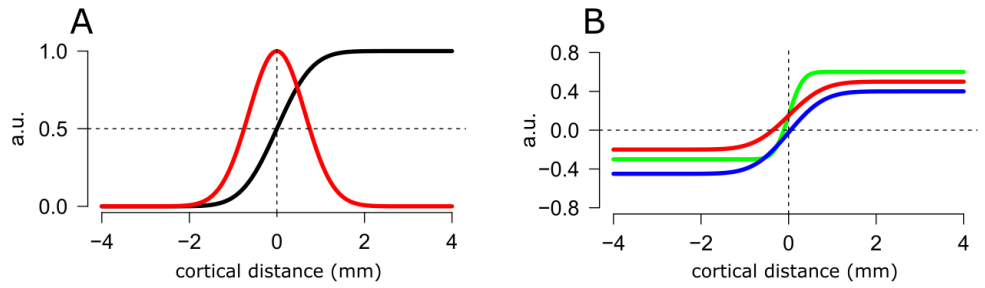


Figure 5, Modelling the GE-BOLD PSF

Panel A. The plot shows a Gaussian PSF (red line). We used the cumulative of the Gaussian (black line, (CGF)) to fit a model to the measured phase-specified coherence along cortical distance, for each surface along cortical depth. Panel B. We modified the parametrization of the standard CGF: we

fixed the location parameter (μ) to zero (at the boundary between two successive stripes in human V2); we added a scaling parameter, to account to different amplitudes of the signal. Third, we added a shift parameter to accommodate the range of phase specified coherence values [-1, 1], see Eq. 4. The plot shows three CGFs with different parameters: green line: FWHM=0.6mm, scaling parameter: 0.9, shift parameter: -0.3. blue line: FWHM=1.5mm, scaling parameter: 0.85, shift parameter: -0.45. red line: FWHM=1.5mm, scaling parameter: 0.7, shift parameter: -0.2. Please note that the red and blue lines appear to have a different rate of change around the zero crossing along cortical distance, however this difference is accounted for by a difference in amplitude, not FWHM (0.85 and 0.7 for the blue and red lines, respectively). The parametrization adopted in our modelling approach allows to disentangle between different components of the trend between phase specified coherence and cortical distance, and to detect differences that could otherwise be wrongly be ascribed to FWHM alone.

Furthermore, as explained above, we used the T1-w data only with the purpose of validating our method and do not directly compare BOLD and T1-w data. Given the robust but relatively low correlation between T1-w data and phase-specified coherence, for this purpose we computed the border between two consecutive V2 stripes from T1-w data itself, following an identical procedure as the one described in the current version of the manuscript, section 2.6.1, this time using T1-w data instead of phase-specified coherence. We have now added the following section in the discussion

4.2 Anatomical Width

Analysis of T1-w signal was performed for validating our modelling approach. Although we adopt the same modelling strategy for GE BOLD and T1-w signal, it is important to note that the interpretation of the FWHM parameter differs between the two modalities. The fitted FWHM for GE BOLD can be interpreted as an estimate of BOLD PSF, that is the spreading of BOLD signal largely due to the action of draining veins, in response to an otherwise localized neuronal activity (Uğurbil et al., 2003). On the other hand, the same parameter fitted for T1-w signal cannot be interpreted as a measure of signal spread, as anatomical features per se do not spread and are not affected by the presence of large draining veins. However, the fitted FWHM for T1-w signal is a useful measure that represents the anatomical variability ('width') of T1-w signal along successive human V2 stripes in the absence of venous spread,

We find that anatomical width is about 0.4mm. It is important to highlight that this parameter remains flat along cortical depth, thus providing a validation for our analysis pipeline. Furthermore, the anatomical variation of about 0.4mm might reflect the acquisition resolution of the anatomical data (0.5mm isotropic), so it is possible that voxel size drives the estimates of anatomical width.

(3) Along the same lines, it is also a concern that the anatomical and functional data are obtained in different sessions. This makes the alignment between the two modalities less straight forward

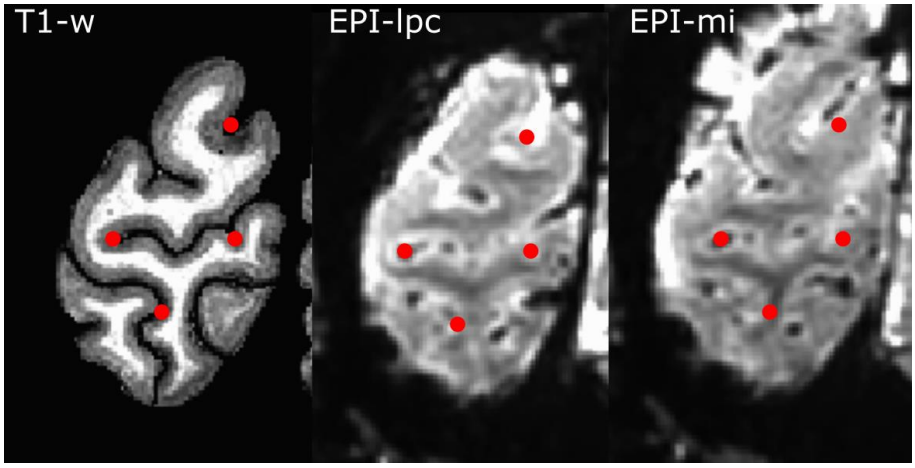
relative to measurements done in the same session.

In the revised version of the manuscript we do not directly compare functional and anatomical data and any potential mis-alignment does not influence the results. However, the quality of the alignment between anatomical and functional modalities can be appreciated in Figure 2 (main manuscript) and in the figure below (see point 4).

(4) To align the anatomical and functional data from the two sessions, the authors use a Pearson- Correlation base cost function. Since the T1w and the GE-BOLD have different intensity levels and contrasts, using mutual information would be more appropriate.

In general, we agree that it is not an easy task to align high resolution anatomical and functional data from two sessions. In our experience using a local-Pearson correlation based cost function (lpc) yields better results than mutual information cost functions. This is shown in the literature for 'standard' acquisition resolution at 3T (Saad et al., 2009).

For high-resolution acquisition there is no published data indicating that lpc performs better than other cost functions. Qualitatively we can report some images, visually comparing alignments using lpc and mutual information cost function.



The T1-w image on the left has red markers on relevant points to assess the quality of alignment. In the middle the EPI-lpc (local Pearson Correlation) shows good agreement with the anatomy. On the right, the EPI-mi (mutual information) shows some agreement, but overall, we regard it as a lower-quality alignment compared to lpc, especially comparing the global shape of the EPI data to the T1-w data, thus we used lpc as the cost function to perform our alignment.

Saad, Z.S., Glen, D.R., Chen, G., Beauchamp, M.S., Desai, R. and Cox, R.W., 2009. A new method for improving functional-to-structural MRI alignment using local Pearson correlation. *Neuroimage*, *44*(3), pp.839-848.

(5) Regarding: terminology. The BOLD PSF is a well-defined concept. However, I believe that the term 'anatomical PSF' is not meaningful. If the authors were to measure the variation in the T1w signal across the boundaries between the stripes (which apparently, they do not measure, see my previous comment), they could use the term 'gradient' to describe the change in T1w signal (likely reflecting variation in myelin content) across these boundaries. What does the term 'anatomical PSF' mean? The features of the anatomy do not "spread." They vary as a function of space, not necessarily in the same way in different parts of the cortex.

In the revised version of the manuscript we do not refer to anatomical PSF anymore. We used T1-w data only with the purpose of validating our method and do not directly compare BOLD and T1-w data. When referring to the T1-w data analysis we refer to the computed full-width-half-max as 'anatomical width' not PSF. We clarify our interpretation of the T1-w data in section 4.3 of the discussion (reported above, in the current version of the manuscript: Discussion, section, ***4.3 Anatomical Width*** in the main manuscript, also reported below).

4.3 Anatomical Width

Analysis of T1-w signal was performed for validating our modelling approach. Although we adopt the same modelling strategy for GE BOLD and T1-w signal, it is important to note that the interpretation of the FWHM parameter differs between the two modalities. The fitted FWHM for GE BOLD can be interpreted as an estimate of BOLD PSF, that is the spreading of BOLD signal largely due to the action of draining veins, in response to an otherwise localized neuronal activity (Uğurbil et al., 2003). On the other hand, the same parameter fitted for the T1-w signal cannot be interpreted as a measure of signal spread, as anatomical features per se do not spread and are not affected by the presence of large draining veins. However, the fitted FWHM for T1-w signal is a useful measure that represents the anatomical variability ('width') of the T1-w signal along successive V2 stripes in the absence of venous spread.

We find that anatomical width is about 0.4mm. It is important to highlight that this parameter remains flat along cortical depth, thus providing a validation for our analysis pipeline. Furthermore, the anatomical variation of about 0.4mm might reflect the acquisition resolution of the anatomical data (0.5mm isotropic), so it is possible that voxel size drives the estimates of anatomical width.

Additional comments:

(6) The authors should include a discussion on how the relatively low spatial resolution they used influences their capacity to measure variations as a function of cortical depth. I expect that the relatively large voxels smooth the depth-dependent measurements and can only show large -scale trends, possibly missing fine structures.

We understand the concern raised by the reviewer. It is indeed true that, by itself, an acquisition resolution of about 0.9mm does not give us the capacity to derive measures across cortical depth. At

best we would be able to measure 2-3 different compartments, assuming a 2mm / 3mm cortical thickness, respectively.
However, the folded nature of cortical surface helps us. We have added the following section in the discussion, last paragraphs of the section: ***4.5.1 The equi-volume model.***

It is important to note that an acquisition resolution of 0.9mm isotropic would at best give us the capacity to measure 2-3 different compartments, assuming a 2mm / 3mm cortical thickness, respectively. However, in Figure 8, we report eight different GE BOLD PSF estimates along normalized cortical depth. We can obtain more that 2-3 different PSF estimates along cortical depth thanks to the folded nature of the cortex. When measuring the fMRI signal along cortical depth, we are using a regular grid (the FOV of the acquisition resolution, 0.9x0.9x1mm) to sample signal from an irregular (folded) portion of the cortex. Thus, the centre of a voxel on the regular grid could be located close to the WM surface, close to the CSF surface or anywhere between the two. This comes at the cost of partial volume effects that can especially influence voxels located at the edge of the grey matter (close to the WM surface, close to the CSF surface), and we acknowledge this limitation. From an ROI, taking advantage of the irregular cortical folding, we can sample continuously along cortical depth, obtaining more than 2-3 signal estimates along cortical depth.

(7) The authors discard the first 13x2.6 sec worth of data to avoid the analysis of data obtained in a non-steady state. This seems over-exaggerated.
The first 13*2.6 of each run did not contain visual stimulation, thus we removed them from the subsequent analysis, we have added the following lines to the main manuscript:

During the first 13 time-frames in each functional run no visual stimuli was presented, and were discarded to ensure that the signal had reached steady-state

(8) The authors report that "GE BOLD PSF increases approximately linearly along the cortical depth." This can be interpreted such that the PSF increases with increasing depth. Instead, I recommend mentioning that the BOLD PSF is maximal in the superficial part of the cortex, and it decreases with increasing cortical depth.
We changed the phrasing following the reviewer's suggestion.

2nd Decision Letter

Dear Dr. Fracasso,

Thank you for submitting your manuscript to Progress in Neurobiology. We have received comments from reviewers on your manuscript. Your paper should become acceptable for publication pending suitable minor revision and modification of the article in light of the appended reviewer comments.

When resubmitting your manuscript, please carefully consider all issues mentioned in the reviewers' comments, outline every change made point by point, and provide suitable rebuttals for any comments not addressed.

To submit your revised manuscript go to https://www.editorialmanager.com/proneu/ and log in as an Author where you will see a menu item called 'Submission Needing Revision'.

Please resubmit your manuscript by Mar 23, 2021.

We look forward to receiving your revised manuscript.

Kind regards,

Essa Yacoub
Guest Editor

Sabine Kastner
Editor-in-Chief
Progress in Neurobiology

Comments from the Editors and Reviewers:

Reviewer #1: My concerns from the previous round of review have been addressed.

Reviewer #2: The authors addressed the comments by the two reviewers with regard to the analysis pipeline. The analysis pipeline makes more sense now.

Remaining comments:

Abstract
The authors conclude that "We postulate that GE BOLD measurements can resolve the underlying spatial organization of stripes with an accuracy below 1mm deeper in the cortex."
This statement is based on circular logic. The quantification of the PSF is based on the assumption that the authors resolved the stripes, and based on the estimated PSF they conclude that they can resolve the stripes.

Part of the current modeling and analysis procedures, including a differential analysis to define a boundary between two conditions, estimating the responses as a function of distance to the boundary, and fitting a cumulative Gaussian function to the sampled responses as a function of distance for estimating a Gaussian PSF, were introduced by Shmuel et al. (2007). The authors should compare their modeling and analysis procedures to those introduced by Shmuel et al. (2007).

Highlights
The authors mention that: "We used the borders of the V2 stripes to compute the BOLD point-spread function (PSF) across cortical depth." 'borders' should be modified to 'functionally defined borders'.

2nd Author Response Letter

Thank you for the comments provided. We believe they have helped us further improving the quality of the manuscript. Changes in the main manuscript are reported in green.
In this rebuttal, replies are reported in Arial font, green. Extracts from the updated version of the manuscript are reported in Arial, italic, green.

Comments from the Editors and Reviewers:

Reviewer #1: My concerns from the previous round of review have been addressed.

Thank you.

Reviewer #2: The authors addressed the comments by the two reviewers with regard to the analysis pipeline. The analysis pipeline makes more sense now.

Remaining comments:

Abstract
The authors conclude that "We postulate that GE BOLD measurements can resolve the underlying spatial organization of stripes with an accuracy below 1mm deeper in the cortex." This statement is based on circular logic. The quantification of the PSF is based on the assumption that the authors resolved the stripes, and based on the estimated PSF they conclude that they can resolve the stripes.

Thanks for pointing this out. We did not realize it while writing the rebuttal, and we agree with the reviewer that this sentence is problematic. We have now removed this sentence from the abstract

Part of the current modeling and analysis procedures, including a differential analysis to define a boundary between two conditions, estimating the responses as a function of distance to the boundary, and fitting a cumulative Gaussian function to the sampled responses as a function of distance for estimating a Gaussian PSF, were introduced by Shmuel et al. (2007). The authors should compare their modeling and analysis procedures to those introduced by Shmuel et al. (2007).

We have now added the following paragraphs in the methods and discussion, comparing the analysis reported in the current manuscript with the one adopted in Shmuel et al., 2007.

Section 2.6.2:

The model followed a similar logic to the analysis introduced by (Shmuel et al., 2007). For a comparison between the current analysis procedure and the one in (Shmuel et al., 2007) see the end of this section and Discussion, section 4.5.3.

…

There are several differences between the current approach and the analysis of Shmuel and colleagues (2007) that are important to point out. In the current approach, we used functionally defined borders of V2 stripes, and used phase specified coherence to identify stripes in human V2, responding to high or low temporal frequencies. This is different from Shmuel and colleagues, which took advantage of the underlying retinotopic organization of human primary visual cortex and used an estimate of BOLD response based on the 1st principal component from the activity profiles sampled from the border.

From the modelling perspective, we fit a cumulative Gaussian model with varying width (FWHM), scaling and shift parameter. We allowed the shift parameter to vary to accommodate for negative phase-specified coherence values. On the other hand, Shmuel and colleagues tested two versions of the cumulative Gaussian model. Using three parameters to accommodate Gaussian centre, width, and scaling; or using two free parameters, assuming the edge of the responding region coincides with the edge of the stimulated region.

Lastly, in the current approach we estimate phase-specified coherence and cortical distance at the surface level (for each node on the surface mesh). This, combined with the sub-millimetre resolution acquisition and modelling, allows us to estimate PSFs along cortical depth. Shmuel and colleagues estimated the PSF in voxel space which encompassed a flat part of cortex in human V1.

Discussion

4.5.3 Comparison between the current modelling approach and (Shmuel et al., 2007).

The method adopted in (Shmuel et al., 2007) inspired us to implement the modelling analysis used in the current manuscript.

While there are several differences between our model and Shmuel and colleagues (see section 2.6.2), these are introduced to account for the different dependent variable adopted in this study and be able to obtain an estimate of PSF along cortical depth. The main difference between our approach and Shmuel et al., is in the nature of the boundary adopted to measure PSF. In both cases we are in the presence of a functionally defined boundary. However, Shmuel and colleagues take advantage of the retinotopic organization in human primary visual cortex, whereas in the current investigation we take advantage of a different organizational principle, the stripe-based arrangement present in human V2, and validate the model using anatomical data. It is important to point out that there is a discussion as to whether vascularisation is similar between V1 and V2. (Duvernoy et al., 1981; Weber et al., 2008).

Weber, B., Keller, A.L., Reichold, J., Logothetis, N.K., 2008. The microvascular system of the striate and extrastriate visual cortex of the macaque. Cerebral cortex (New York, N.Y. : 1991) 18:2318-30.

Duvernoy, H.M., Delon, S., Vannson, J.L., 1981. Cortical blood vessels of the human brain. Brain research bulletin 7:519-79.

Highlights
The authors mention that: "We used the borders of the V2 stripes to compute the BOLD point-spread function (PSF) across cortical depth." 'borders' should be modified to 'functionally defined borders'.

Changed