



RESEARCH ARTICLE

**REVISED** Comparison of human population receptive field estimates between scanners and the effect of temporal filtering [version 2; peer review: 2 approved]

Catherine Morgan <sup>1-3</sup>, D. Samuel Schwarzkopf<sup>3,4</sup>

<sup>1</sup>School of Psychology and Centre for Brain Research, University of Auckland, Auckland, New Zealand

<sup>2</sup>Brain Research New Zealand, Auckland, New Zealand

<sup>3</sup>School of Optometry & Vision Science, University of Auckland, Auckland, New Zealand

<sup>4</sup>Experimental Psychology, University College London, London, UK

**v2** First published: 24 Sep 2019, 8:1681 (<https://doi.org/10.12688/f1000research.20496.1>)  
 Latest published: 26 Feb 2020, 8:1681 (<https://doi.org/10.12688/f1000research.20496.2>)

**Abstract**

**Background:** Population receptive field (pRF) analysis with functional magnetic resonance imaging (fMRI) is an increasingly popular method for mapping visual field representations and estimating the spatial selectivity of voxels in human visual cortex. However, the multitude of experimental setups and processing methods used makes comparisons of results between studies difficult.

**Methods:** Here, we compared pRF maps acquired in the same three individuals using comparable scanning parameters on a 1.5 and a 3 Tesla scanner located in two different countries. We also tested the effect of low-pass filtering of the time series on pRF estimates.

**Results:** As expected, the signal-to-noise ratio for the 3 Tesla data was superior; critically, however, estimates of pRF size and cortical magnification did not reveal any systematic differences between the sites. Unsurprisingly, low-pass filtering enhanced goodness-of-fit, presumably by removing high-frequency noise. However, there was no substantial increase in the number of voxels containing meaningful retinotopic signals after low-pass filtering. Importantly, filtering also increased estimates of pRF size in the early visual areas which could substantially skew interpretations of spatial tuning properties.

**Conclusion:** Our results therefore suggest that pRF estimates are generally comparable between scanners of different field strengths, but temporal filtering should be used with caution.

**Keywords**

population receptive fields, site comparison, replicability, functional MRI

**Open Peer Review**

**Reviewer Status**

	Invited Reviewers	
	1	2
<b>version 2</b> (revision) 26 Feb 2020		
<b>version 1</b> 24 Sep 2019	 report	 report

- Noah Benson** , New York University, New York, USA
- Paola Binda** , University of Pisa, Pisa, Italy

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding author:** Catherine Morgan ([c.morgan@auckland.ac.nz](mailto:c.morgan@auckland.ac.nz))

**Author roles:** **Morgan C:** Data Curation, Investigation, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing;

**Schwarzkopf DS:** Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** This work was supported by a European Research Council Starting Grant (310829, WMOSPOTWU) and start-up funding from the Faculty of Medical and Health Sciences at the University of Auckland.

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2020 Morgan C and Schwarzkopf DS. This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Morgan C and Schwarzkopf DS. **Comparison of human population receptive field estimates between scanners and the effect of temporal filtering [version 2; peer review: 2 approved]** F1000Research 2020, 8:1681 (<https://doi.org/10.12688/f1000research.20496.2>)

**First published:** 24 Sep 2019, 8:1681 (<https://doi.org/10.12688/f1000research.20496.1>)

**REVISED Amendments from Version 1**

Changes in version 2 of the manuscript are in response to the reviewer's comments and suggestions, namely:

Adding short discussion on how increases in TSNR might alter pRF size.

Clarifying which HRF function was used.

Creating new group figures (individual data now in supplementary section)

Adding comment to highlight same scanner manufacturer at both sites.

Adding short discussion on how sequential vs random presentation of pRF stimuli might impact results.

Adding 2 new references to support above points.

**Any further responses from the reviewers can be found at the end of the article**

## Introduction

Population receptive field analysis with functional magnetic resonance imaging (fMRI) has become a popular method in the toolbox of visual neuroscience. It has not only been used for mapping cortical organization<sup>1-3</sup>, but also to study spatial integration in the visual cortex<sup>4,5</sup>, reveal the effects of attention on visual processing<sup>6-10</sup>, show differences in patients and special populations<sup>11-15</sup> and for reconstructing the neural signature of perceptual processes<sup>16-20</sup>. The most wide-spread technique involves fitting a two-dimensional symmetric Gaussian model of the pRF to the time series of each voxel in visual cortex responding to a set of stimuli<sup>1</sup>. Estimates of pRF position and size reflect an aggregate of the position preferences and sizes of thousands of neuronal receptive fields of the cells within the imaging voxel, and also incorporates extra-classical receptive field interactions. Further, in fMRI, as underlying neuronal activity is inferred through neurovascular coupling, pRF measurements are affected by hemodynamic factors (although it has been shown that pRFs estimated from fMRI data have a close correspondence with receptive field properties in electrophysiological experiments<sup>1,21,22</sup> but see also 23).

The indirect nature of estimating pRFs from fMRI data suggests therefore that there could be considerable variability in derived measurements. Direct test-retest evaluations with the same experimental setup have shown that pRF mapping experiments are robust and repeatable<sup>24-26</sup>. However, for *different* experimental setups, for instance, in terms of the magnetic field strength and the particular pulse sequence used to acquire fMRI data the comparability has not been assessed. The signal-to-noise ratio of MRI is proportional to voxel volume and the strength of the static magnetic field<sup>27</sup>, and hence pRF measurements at higher magnetic field strength might be more accurate. However, the temporal resolution (or repetition time, TR, of image acquisition), directly affects the contribution of different noise frequencies and the contribution of physiological nuisance factors like respiration or cardiac pulsation. Noise in fMRI data therefore has multiple contributions (physiological, thermal and system related) and the relationship between the temporal signal to noise ratio (TSNR) in a fMRI time course has a

non-linear relationship with static SNR<sup>28</sup>, i.e. gains in static SNR, from for example increased field strength, may not translate to proportional gains in TSNR due to a limit where physiological noise dominates. If there *were* gains in TSNR to be had from scanning at a higher field strength, one could posit a stronger response amplitude could potentially produce a larger pRF size, because visual field locations farther from the pRF centre respond to the stimulus. However, theory suggests that the pRF model is scale-invariant, that is, the pRF parameters are estimated by correlation between predicted and observed time series and only the pRF *shape* matters to size estimation, *not* amplitude.

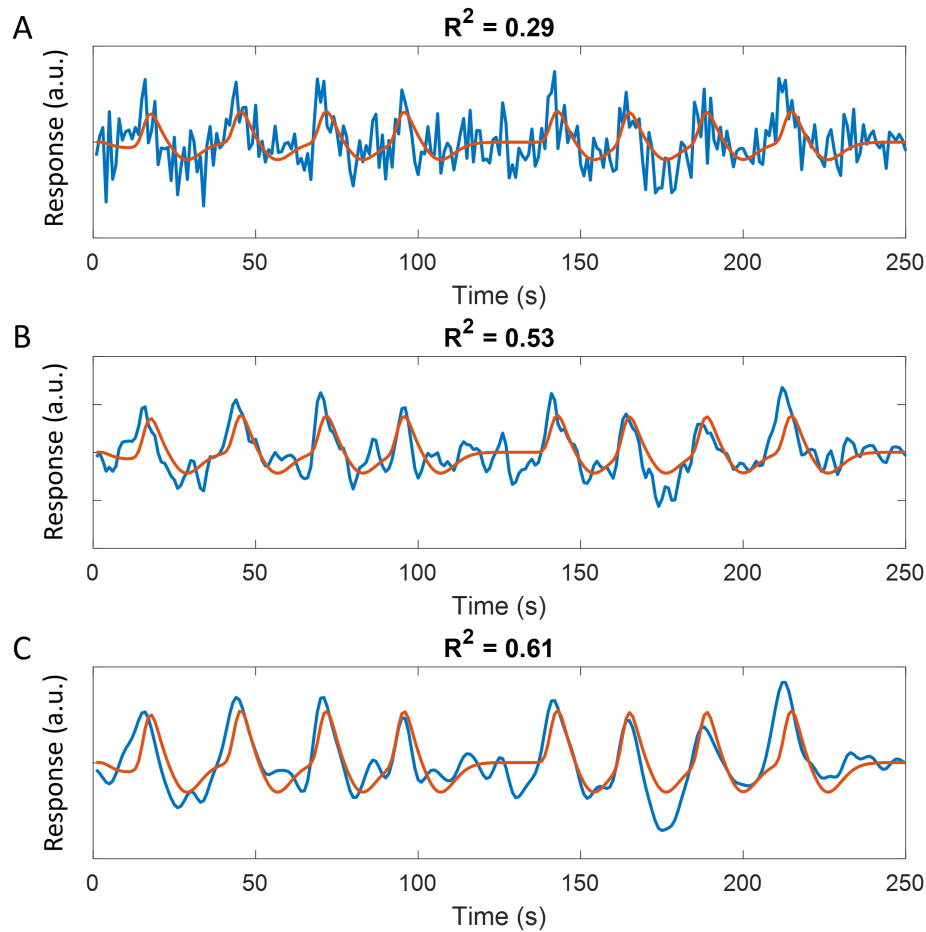
Here, we conducted pRF mapping on three individuals using identical TR and voxel size on a 1.5 Tesla (1.5T) scanner in London, United Kingdom and a 3 Tesla (3T) scanner in Auckland, New Zealand. Critically, our aim was not to test the effect of magnetic field strength alone, but rather to compare pRF estimates from typical methods used at each scanning facility that also *includes* a different field strength. Thus, our findings should reflect a relatively conservative estimate of the variability of pRF parameters under comparable conditions at the two sites.

Typical pRF studies evaluate the goodness-of-fit of the pRF model by means of the coefficient of determination ( $R^2$ ) of the correlation between the observed and predicted time series. This measure, however, strongly depends on the temporal resolution of the signal acquisition. In our recent experiments we have used an accelerated multiband sequence with 1 s TR<sup>25,29-32</sup>. At this temporal resolution, there is a considerable contribution of high-frequency noise to the signal (Figure 1). Since our analysis does not typically involve any temporal filtering beyond linear detrending (essentially, a wide-bandwidth high-pass filter removing only very low frequencies that are attributed to slow drifts in the signal), this probably explains why the overall  $R^2$  in our studies is comparably low compared to those reported by others who use more standard fMRI TRs of 2-3 s. We therefore conducted an analysis comparing pRF parameters obtained using our standard analysis (minimal filtering), to filtering with two low-pass filters that approximate longer TR acquisition.

## Methods

### Participants

We recruited three healthy adult volunteers (2 female, 1 left-handed, aged 30, 39 and 47) from our pool of repeat participants in functional brain imaging studies in London, United Kingdom. An important criterion for their participation was that they would also visit Auckland, New Zealand, within a few months after the first scan in London. Generally, we could only include participants with normal or corrected-to-normal visual acuity (contact lenses only), and no history of eye disease or neurological, neurodevelopmental or neuropsychiatric disorders. Moreover, they could have no other contraindications to magnetic resonance imaging (e.g. claustrophobia, metal implants). All participants gave written informed consent to take part and henceforth referred to as P1, P2, and P3. All procedures were approved by local ethics review boards at University College



**Figure 1. Low-pass filtering the time series improves goodness-of-fit quantified by  $R^2$ .** The same predicted time series (red curves) is shown overlaid on the unfiltered time series (A), and low-pass filtered time series with kernel 1 s (B) or 2 s (C).  $R^2$  is substantially larger for the filtered time series because high-frequency noise has been removed.

London (fMRI/2012/007) and the University of Auckland (017477).

### Procedure

Participants were scanned twice with a pRF mapping protocol. The first scan took place inside a MAGNETOM Avanto 1.5T MRI scanner (Siemens Healthcare, Erlangen, Germany) at the Birkbeck/UCL Centre for NeuroImaging (BUCNI) in the Experimental Psychology department of University College London, United Kingdom (henceforth referred to as London site). The second scan took place several months later in a MAGNETOM Skyra 3T MRI scanner (Siemens Healthcare, Erlangen, Germany) at the Centre for Advanced Magnetic Resonance Imaging (CAMRI) in the Faculty of Medical & Health Sciences of the University of Auckland, New Zealand (henceforth referred to as Auckland site). In both scans, participants were scanned with six runs for pRF mapping lasting 4 min 20 s each during which functional echo-planar images were acquired. Moreover, at both centres a structural T1 weighted brain image was acquired although the structural image from the second scan

(at 3T) was not used in any further analysis. During the scans, participants were instructed to remain as still as possible and fixate continuously on a small dot in the centre of the screen. They were instructed to press a button whenever the fixation dot changed colour.

### Stimuli

Stimuli were generated and presented using **MATLAB** (Mathworks; Version R2017a, 9.2.0.538062) and **Psychtoolbox** (Version 3.0.14)<sup>33</sup> at a resolution of 1920 \* 1080. At both sites, the screen subtended 34° by 19° of visual angle. Stimuli were presented on a screen at the back of the bore via a mirror mounted on the head coil. At the London site, stimuli were projected onto a screen in the back of the bore while in the Auckland site stimuli were presented on an MRI compatible 32" widescreen LCD (Cambridge Research Systems Ltd).

At both sites, the stimuli generated were matched, and comprised bars containing a dynamic high-contrast ripple pattern as used in previous studies<sup>6,13,15,22</sup> on a uniform grey

background. Bars traversed the visual field in a regular sequence (e.g. from the bottom to the top), jumping by  $0.38^\circ$  every second. Each sweep of the bar lasted 25 s and so there were 25 jumps of the bar. Each run started with a sweep from the bottom to the top and then the sweep direction was rotated by  $45^\circ$  clockwise on the next sweep. There were thus eight sweeps covering a complete rotation. After the fourth and eighth sweep, a 25 s baseline period (no bars) was presented.

Bars were always  $0.53^\circ$  wide and at the longest (when crossing the centre of the visual field) subtended the full screen height, but because they were presented only within a circular region (diameter:  $19^\circ$ , i.e. the height of the screen) they were accordingly shorter at the start and end of each sweep. The outer edge of the ripple stimulus was smoothed by ramping it down to zero over  $0.1^\circ$ . Similarly, the stimulus contrast ramped down to zero from  $0.53^\circ$  to  $0.43^\circ$  eccentricity, thus creating a blank hole around fixation.

A blue dot (diameter:  $0.09^\circ$ ) was present in the centre of the screen throughout each scanning run. The whole run was divided into 200 ms epochs. At each epoch, there was a 0.01 probability that the dot could change colour to purple with the constraint that no such colour changes could occur in a row. Participants were instructed to fixate the dot and press a button on a magnetic resonance-compatible button box whenever it changed colour. In addition, a radar screen pattern comprising low-contrast radial and concentric lines around fixation was presented at all times to aid fixation compliance (see 25). The fixation dot and radar screen pattern were also presented during the baseline period.

Prior to each run, we collected 10 dummy volumes to allow ample time for steady-state magnetisation to be reached. During this time, only the fixation dot was presented on a blank grey screen.

### Scanning parameters

At both sites, we used a 32-channel head coil where we removed the front elements because it impeded the view of the stimulus. This resulted in 20 effective channels covering the back and the sides of the head. Six pRF mapping runs of 260 T2\*-weighted image volumes were acquired (including the 10 dummy volumes). We used 36 transverse slices angled to be approximately parallel to the calcarine sulcus (planned using the T1-weighted anatomical image). At both sites, we used an accelerated multiband sequence<sup>34,35</sup> at 2.3 mm isotropic voxel resolution, field of view 96x96, and a TR of 1 s. At the London site, the scan had an echo time (TE) of 55 ms, flip angle of  $75^\circ$ , and a multiband/slice acceleration factor of 4 and rBW 1628 Hz/pixel. At the Auckland site, the scan had a TE of 30 ms, flip angle of  $62^\circ$ , a multiband/slice acceleration factor of 3, an in-plane/parallel imaging acceleration factor of 2 and rBW was 1680 Hz/Px. After acquiring the functional data at the London site, the front portion of the coil was put back on to ensure maximal signal-to-noise levels for collecting a structural scan (a T1-weighted anatomical magnetization-prepared rapid acquisition with gradient echo scan with a 1 mm isotropic voxel size and full brain coverage).

### Data preprocessing

Functional data were preprocessed in *SPM12* (Wellcome Centre for Human NeuroImaging; Version 6906). The first 10 dummy volumes were removed. Then we performed mean bias intensity correction, realignment and unwarping of motion-induced distortions, and coregistration to the structural scan acquired at the London site using default parameters in *SPM12*. Using *FreeSurfer* (Version 6.0.0) we further used the structural scan for automatic segmentation and reconstruction as a three-dimensional surface mesh of the pial and grey-white matter boundaries<sup>36,37</sup>. The grey-white matter surface was then further inflated into a smooth model and a spherical model.

All further analysis was conducted using our custom *SamSrf 6 toolbox*<sup>38</sup>. Functional data were projected from volume space to the surface mesh by finding the nearest voxel located halfway between each vertex in the pial and grey-white matter surface mesh. The time series at each vertex was then linearly detrended and z-standardized before being averaged across the six runs at each scanning site.

For the temporal filtering analysis, we then further convolved the time series of each vertex with a Gaussian filter with standard deviation of 1 or 2 s, respectively.

### pRF modelling

We modelled pRFs as a symmetric, two-dimensional Gaussian defined by  $x_0$  and  $y_0$ , the Cartesian coordinates of the pRF centre in visual space, and the standard deviation of the Gaussian,  $\sigma$ , as a measure of the pRF size. The pRF model predicted the neural response at each TR of the scan by calculating the overlap of the mapping stimulus with this Gaussian pRF profile. A binary mask of 100-by-100 pixels indicated where the stimulus appeared on the screen for each time point. The neural responses were then determined by multiplying each frame of the stimulus mask with the pRF profile and summing over the 10,000 pixels. Subsequently, the time series was convolved with a canonical hemodynamic response function (HRF) determined from previous empirical data<sup>6</sup> and z-standardized. The same HRF function was used for all participants given that the effect of individualized HRF on pRF parameters is expected to be small<sup>25</sup>.

The pRF parameters at each vertex were fit using a two-stage procedure. First, we applied a coarse fit which involved an extensive grid search by correlating the actually observed time series against a set of 7650 predicted time series derived from a combination of  $x_0$ ,  $y_0$ , and  $\sigma$  covering the plausible range for each parameter (see 25,29). The parameters giving rise to the maximal correlation were then retained for the second stage, the fine fit, provided the squared correlation,  $R^2$ , exceeded 0.01. The fine fit entailed an optimization procedure<sup>39,40</sup> to refine the three pRF parameters to further maximize the correlation between observed and predicted time series. Subsequently, we used linear regression between the observed and predicted time series to fit the amplitude,  $\beta_1$ , and the baseline intercept,  $\beta_0$ . Up until this point, all analyses used raw data without any smoothing or interpolation. However, we then applied a Gaussian smoothing kernel (full width at half maximum = 3 mm)

to the final pRF parameter maps on the spherical surface mesh. These smooth maps were used for visualization purposes and delineation of the visual areas. Moreover, they are necessary for calculating the local cortical magnification factor (CMF; Harvey and Dumoulin<sup>41</sup>) because it requires a smooth visual field map without gaps and scatter. For this we determined the cortical neighbours of each vertex and calculated the area subtended by the polygon formed by their pRF centres. The same procedure was used to determine the cortical surface area (calculation performed by FreeSurfer). To calculate the CMF, we then divided the square root of cortical area by the square root of visual area.

In addition, for the data comparing the London and Auckland sites we also calculated the noise ceiling as an estimate of the maximum goodness-of-fit that could theoretically be achieved from the data of each voxel. For this, we split the six pRF mapping runs from each site into even and odd numbered runs and averaged them separately. We then calculated the Pearson correlation between these split time series,  $r_{obs'}$ , and used the Spearman-Brown prophecy formula<sup>42,43</sup> to determine,  $r_{obs}$ , the expected reliability for the average of all six runs:

$$r_{obs} = \frac{2r_{obs'}}{1 + r_{obs'}}$$

The theoretical maximum observable correlation between the predicted and observed time series,  $\rho_o$ , for each voxel is given by the square root of this correlation<sup>44</sup> because

$$\rho_o = \rho_h \sqrt{r_{pred} r_{obs}}$$

where  $r_{pred}$  is the reliability of the predictors, which is 1, and we assume the hypothetical correlation between the observed and predicted time series,  $\rho_h$ , also to be 1. The noise ceiling, the maximum observable goodness-of-fit  $\rho_o^2$  is therefore equal to  $r_{obs}$ .

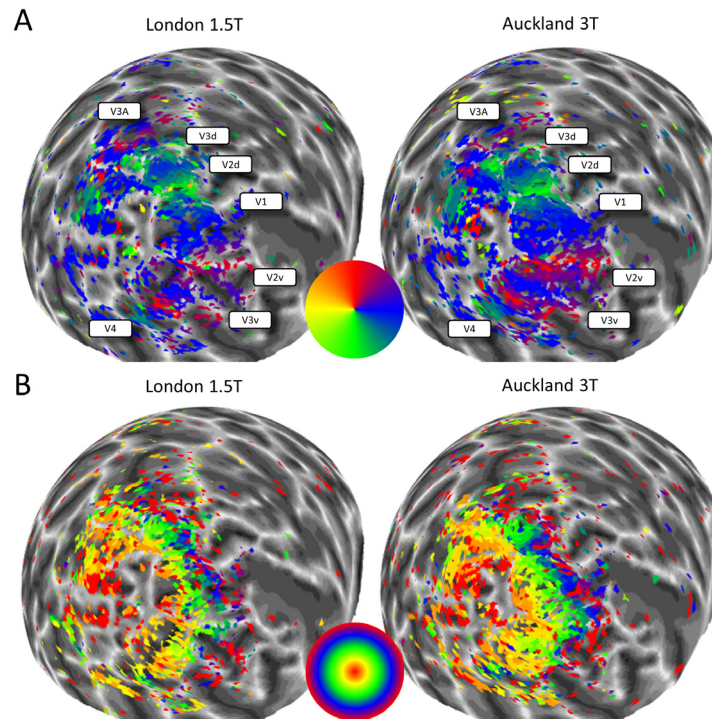
### Further analysis

We manually delineated visual areas V1-V4 and V3A using smoothed maps from the London site by determining the borders between regions from the polar angle reversals<sup>45</sup> and then also applied these delineations to the maps generated in Auckland. For each region of interest (ROIs) we then extracted pRF data and binned them into eccentricity bands  $1^\circ$  in width, starting from  $0.5^\circ$  and increasing up to  $9.5^\circ$ . For each bin, we then calculated the mean pRF size, and the median CMF,  $R^2$ , noise ceiling  $\rho_o^2$ , or the normalized goodness-of-fit, that is  $R^2$  divided by  $\rho_o^2$ . To bootstrap the dispersion of these summary statistics, we resampled the data in each bin 1000 times with replacement and determined the central 95% of this bootstrap distribution as a confidence interval. To avoid the inclusion of artifactual data, for pRF size and CMF we only included data from voxels where  $R^2$  exceeded 0.15.

## Results

### Comparison between scanning sites

We first compared visual field maps from the two sites, London 1.5T and Auckland 3T, by visual inspection. **Figure 2** shows polar angle and eccentricity maps of the left hemisphere of one participant from both sites. It is immediately



**Figure 2.** Polar angle (**A**) and eccentricity (**B**) maps of one participant from the London 1.5T site (left) and the Auckland 3T site (right) displayed on a spherical model of the left hemisphere. Colour wheels denote the pseudo-colour code for visual field maps. No smoothing or interpolation was applied to the mapping data. Voxels were thresholded at  $R^2 > 0.1$ . In **A**, the position of the visual regions have been labelled for reference. The greyscale indicates the cortical folding pattern.

apparent that more voxels survive statistical thresholding ( $R^2 > 0.1$ ) for the 3T data. The cortical territory occupied by visual field maps is somewhat more extensive and more complete. Nevertheless, the orderly organization of V1-V4, V3A, as well as regions in the LO complex is clearly visible in both scans, and their borders are very similar. We further quantified the map similarity by calculating correlations across all voxels above threshold in both scans (circular correlation for polar angle, Spearman's  $\rho$  correlation for eccentricity). This showed that the polar angle maps were well correlated for two participants (P1: 0.71; P2: 0.71) although the correlation was not as strong in the final participant (P3: 0.44). Eccentricity maps were strongly correlated in all three participants (P1: 0.70; P2: 0.75; P3: 0.67).

Next, we compared pRF size using  $1^\circ$  wide eccentricity bands to bin the data from each site and then calculating the mean for each bin. We observed that pRF size consistently increased across eccentricities and also along the visual pathway as expected. But crucially, while pRF size varied somewhat between sites, this difference was not systematic across the five regions of interest (Figure 3A). In V4 and V3A there was somewhat greater variance at some eccentricities, likely due to the smaller size of these regions compared to V1-V3. Overall, pRF sizes were, however, very similar between sites. Similarly, the median cortical magnification factor (CMF) showed the expected exponential decrease from the central to the peripheral visual field. CMF curves were very comparable for the two sites (Figure 3B), except for somewhat greater CMF at 3T in V1 and V2 in the very central visual field of participants P2 and P3<sup>46</sup> (see supplementary figure 1A and B<sup>46</sup>) for individual participants' pRF size and CMF respectively).

We then repeated this analysis for the goodness-of-fit of the pRF model across eccentricities (Figure 3C). This showed that in V1 and V2 goodness-of-fit was notably greater for the Auckland 3T site than the London 1.5T site. In higher extrastriate regions, the pattern of results was less clear, although at least for P3 see supplementary figure 2 for individual participants data fitting results<sup>46</sup>) goodness-of-fit was greater also in V3 and V4 (although note that in V4 for P1 very little data with very low model fits was present beyond  $6^\circ$  eccentricity in the Auckland data). In V3A, the model fits (see Figure 3C last column) at both sites were similar, but generally lower than in the other regions and with greater variability. However, this was presumably largely driven by the overall signal-to-noise ratio. When we normalized model fits relative to the noise ceiling,  $\rho_0^2$ , the maximum goodness-of-fit that could theoretically be achieved given the data from each site, we found no systematic difference in goodness-of-fit between sites (Figure 3D). The curves mostly overlapped for P1 and P2 except in V1 where the London data outperformed the Auckland data (see Figure 2B). In contrast, for P3 normalized model fits for the Auckland site outperformed the London site in all regions except V3A. The noise ceiling itself was consistently higher for the Auckland data than the London data (Figure 4E).

### Effect of temporal filtering

Goodness-of-fit of the modelled time course derived from pRF estimates and the measured fMRI time course are

typically quantified by  $R^2$ . However, the pRF model does not account for the high frequency noise observed in fast temporal resolution acquisitions like the 1 s TR we used here, which therefore likely results in lower  $R^2$  values. Theoretically, high frequency signals could be removed from the data and thus the model fits artifactually improved (Figure 1). We therefore reanalysed the data from the 3T site after temporal low-pass filtering the time series of each voxel by convolving it with a Gaussian kernel of either 1 s or 2 s standard deviation (e.g. Figures 1B and 1C, respectively).

Figure 4 shows visual field maps from the left hemisphere of one participant using the raw data and those after low-pass filtering. A very similar map structure is apparent in all three images. There are, however, also a lot of noise voxels surviving thresholding ( $R^2 > 0.1$ ) outside the general visual field maps. Conversely, while filtering filled in a few missing voxels within the maps, filtering did not affect the overall structure and did not increase the cortical territory of the responsive region.

Next, we quantified mean pRF size and compared this for the three data sets (Figure 5A). In the early areas V1-V3, pRF size is consistently greater for the filtered time series, especially at the longer kernel of 2 s. In V4 and V3A, pRF size remained relatively similar between conditions. We also quantified the median goodness-of-fit of the pRF model (Figure 5B) and observed consistently greater model fits for both the low-pass filtered time series in V1 to V4. A similar pattern could be seen in V3A but results were generally more variable, especially in P1 (see supplementary figure 3 for individual participant results<sup>46</sup>).

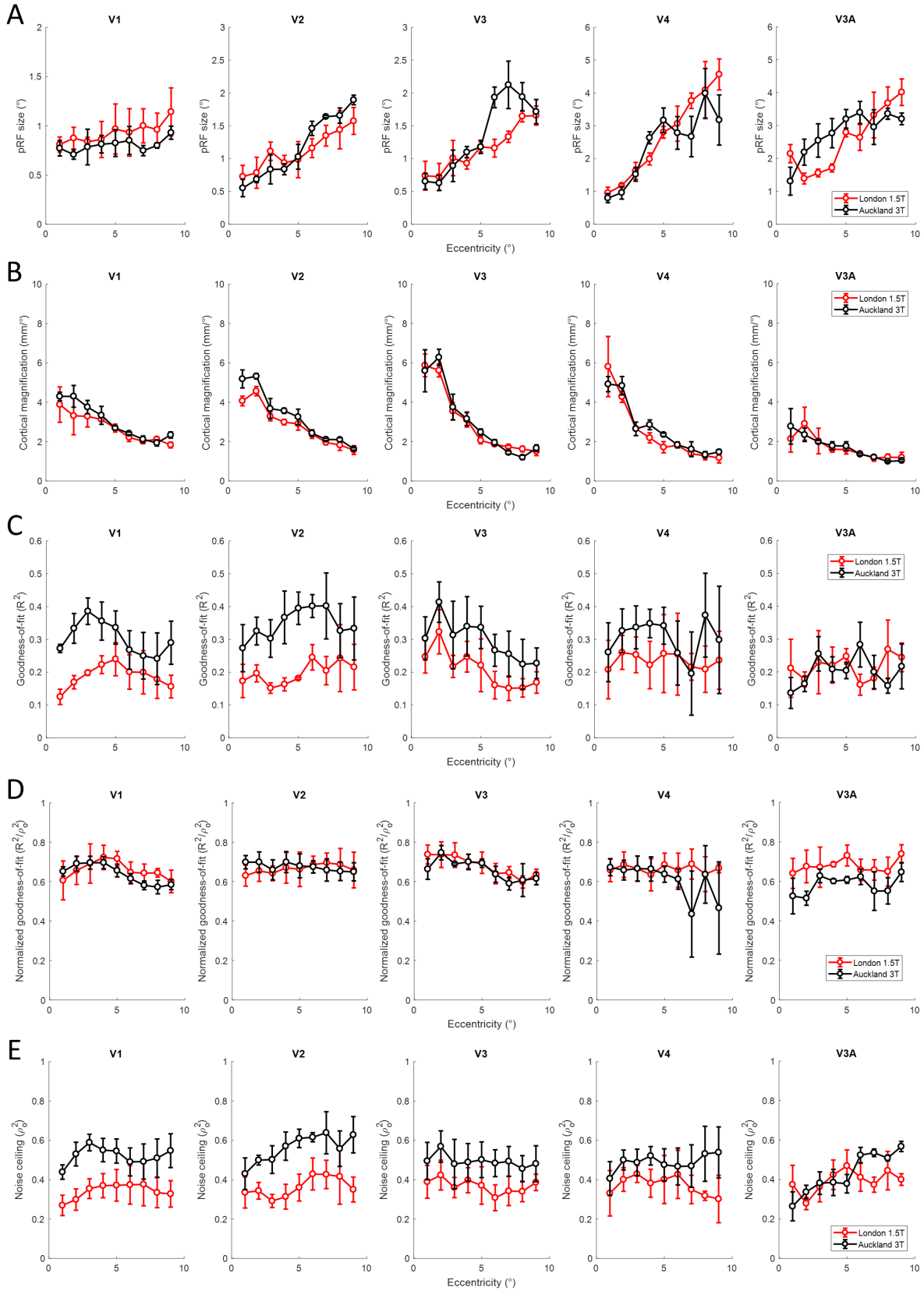
Raw results are available from Open Science Framework as *Underlying data*<sup>47</sup>.

### Discussion

Here, we compared pRF mapping data acquired using the same stimuli and participants and under comparable conditions at two sites, a 1.5T scanner in London, United Kingdom, and a 3T scanner in Auckland, New Zealand. Generally, pRF model fits were better in the Auckland data, probably in large part because of the approximately double signal-to-noise ratio from the greater static magnetic field. However, despite this difference in accuracy of data fitting, we found that visual field maps and actual estimates of pRF size and local cortical magnification were very similar between the two sites.

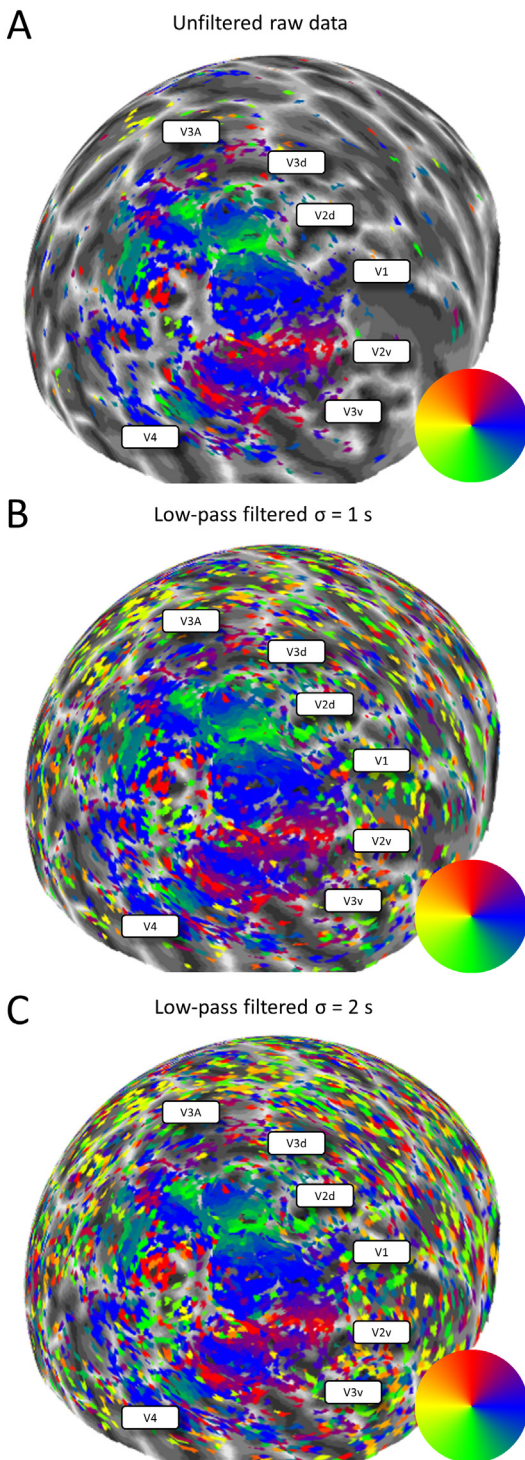
Firstly, it is worth noting that both scanners used were manufactured by the same vendor. While this means our results are not currently generalizable to other platforms, it removes manufacturer as an additional source of variance between sites.

Although the London data were always acquired first, this is unlikely to explain our findings. All our participants were familiar with the fMRI environment, including undertaking previous pRF mapping studies. Therefore, it is unlikely that there should have been substantial training effects that could for example have changed the amount of head motion or fixation compliance between sessions. The exact parameters of the pulse sequences used and general differences in the image



**Figure 3.** Mean population receptive field (pRF) size (A) and median cortical magnification factor CMF (B), median goodness-of-fit (C), normalized goodness-of-fit (D), and noise ceiling (E) binned into eccentricity bands for the London 1.5T site (red) and the Auckland 3T site (black) and averaged across the three participants. Columns show different visual regions. Error bars denote  $\pm 1$  standard error of the mean across participants.





**Figure 4.** Polar angle maps of one participant from the Auckland 3T site displayed on a spherical model of the left hemisphere. Colour wheels denote the pseudo-colour code for visual field maps. No smoothing or interpolation was applied to the mapping data. Voxels were thresholded at  $R^2 > 0.1$ . The position of the visual regions have been labelled for reference. The greyscale indicates the cortical folding pattern. **A.** Unfiltered data. **B.** Low-pass filtered with kernel 1 s. **C.** Low-pass filtered with kernel 2 s.

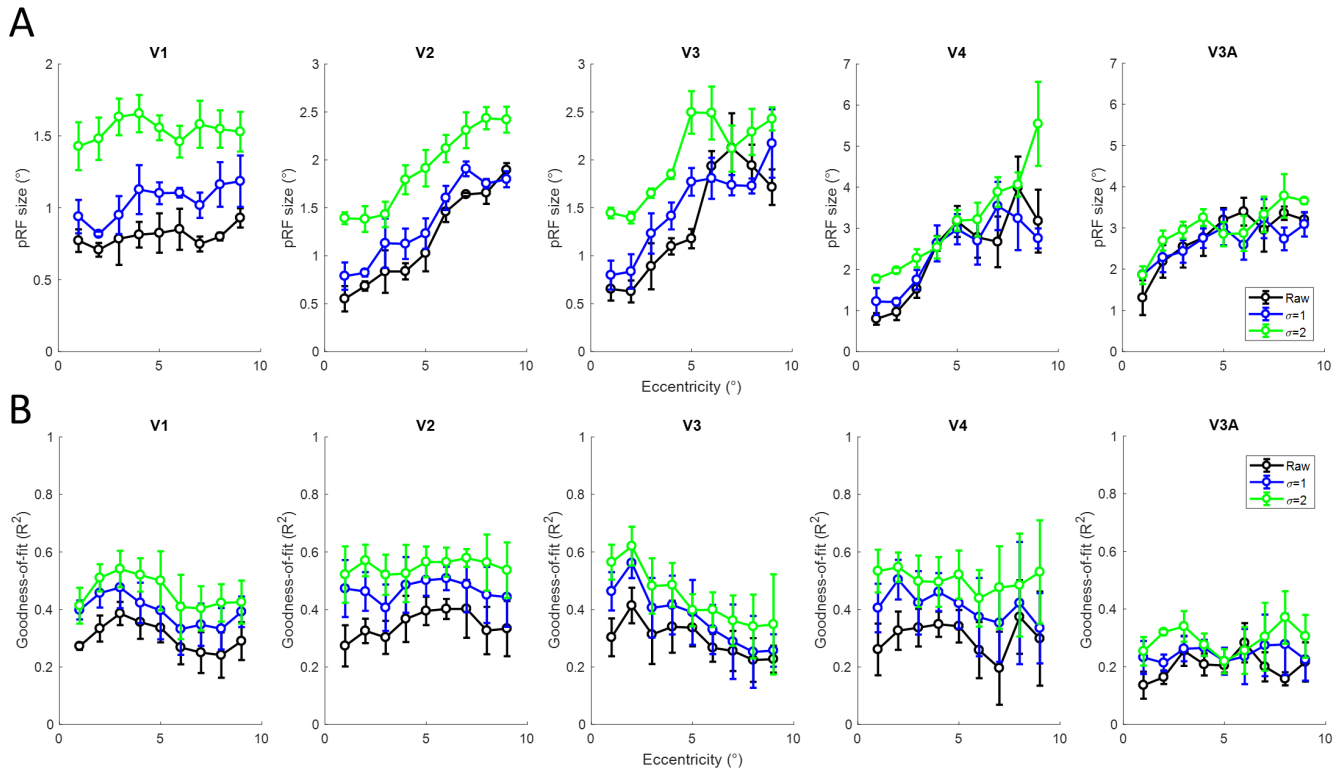
quality of the two scanners could of course also be contributing factors to inter-site differences. We did not apply any correction of distortions caused by the static magnetic field<sup>48</sup> to either site data. Moreover, the exact slice positioning, and thus how the voxel grid resolved the grey matter, may also have differed between sites. Relevant to this, in our previous work quantifying the test-retest reliability of pRF maps we found that reliability was greater for scanning sessions in close succession on the same day than for sessions on different days<sup>25</sup>. This could have been caused by fluctuations in the scanner hardware itself but also relate to differences in head position at set up. While in both cases participants were removed from the scanner between the repeat sessions, it is highly likely that positioning was more similar for the sessions conducted on the same day.

While we took painstaking steps to match the visual displays at the two sites as closely as possible, due to constraints of the experimental setup there may have been small differences between the stimuli presented. In London, images were projected onto a screen and this necessitated focusing and scaling the projected image to be of the exact size. The image in London may have been somewhat blurrier and the viewing angle more variable than in Auckland where we used a clear liquid crystal display that was always placed at the exact same position at the back of the bore. Naturally, the exact viewing angle of the stimuli also depends on the viewing distance and there may have been subtle variation between the sites although this is unlikely to have produced any systematic differences.

Nevertheless, the general extent of the visually responsive area of cortex containing clear retinotopic maps was very comparable across the sites, as were estimates of spatial tuning and cortical magnification. Therefore, we conclude that pRF estimates are very robust across scanning sites, even when using different magnetic field strengths.

At the short TR of 1 s as used in our study, noise contributes high-frequency signals to the fMRI response that are irrelevant to the sluggish blood oxygen level dependent activity the pRF model seeks to characterize. This explains why the model fits in most of our studies are relatively low compared to those reported in the literature using more conventional TRs. We therefore also sought to test what effect temporal low-pass filtering had on pRF parameter estimates. Unsurprisingly, low-pass filtering enhanced the goodness-of-fit of pRF models. However, while this filled in a few gaps in the maps it largely boosted the number of noise voxels outside the visually responsive cortex.

Crucially, low-pass filtering also generally enlarged estimates of pRF size in the early visual regions V1-V3 while data in higher regions were mostly unaffected. The pRF size estimates for the unfiltered data accord well with previous research and what one would expect from electrophysiological recordings<sup>1,21,22,41</sup>. Therefore, the pRF sizes of the filtered data presumably reflect an overestimate that seems unrealistically high (e.g. for P2 mean pRF size in central V1 was close to  $2^\circ$  for the most heavily filtered data).



**Figure 5.** Mean population receptive field (pRF) size (**A**) and median goodness-of-fit (**B**) binned into eccentricity bands for the Auckland data without filtering (black), and low-pass filtering with kernel 1 s (blue) and 2 s (green), averaged across participants. Columns show different visual regions. Error bars denote  $\pm 1$  standard error of the mean across participants.

The reason why filtering increased pRF sizes in early areas is probably the fact that filtering blurred together signals from stimuli presented close in time (proximal bar positions were presented only a few seconds apart). In the early regions where pRFs are small this would therefore be modelled as a larger pRF, whereas in higher areas in which pRFs are large enough to encompass adjacent bar positions this should not affect estimates of pRF size. A slower stimulus design where each bar position is stimulated for 2–3 s, or a random instead of an ordered stimulus sequence, could probably counteract this problem but this would come at the expense of longer scanning durations. In any case, for our standard design using 1 s TR low-pass filtering clearly has no practical advantages because it does not appear to fundamentally improve map quality and skews pRF size estimates in the early visual areas. Of course, the skew we observed in pRF size is related to the sequential presentation of our visual stimuli (sweeping bars, a stimuli commonly used for pRF). One could hypothesise that temporal filtering would *not* bias for pRF sizes if the location of stimuli used for retinotopic mapping was presented randomly. However, recent studies<sup>49,50</sup> have shown that pRF size estimates differ between random and ordered designs (irrespective of temporal filtering), probably due to additional factors. The choice of mapping stimulus should therefore also be taken into consideration when comparing results across studies.

Since we only tested three participants in this study, our analyses were not designed to detect subtle differences between the

scanner environments or small effects of temporal filtering. Retinotopic maps and pRF data are highly replicable<sup>25</sup> and can also be used for case studies on individuals. We were therefore most interested in how consistent results are across all three participants. Our descriptive analysis of the data included bootstrapped confidence intervals, which permit an inference about the statistical evidence in each participant, and thus each individual participant essentially constitutes a replication.

In summary, our results show that pRF mapping with fast 1 s TR produces similar and reliable results on different scanners, even using different magnetic field strengths. It would be interesting to conduct a similar comparison between 3T and 7T scans as these are becoming increasingly commonplace.

## Data availability

### Underlying data

Open Science Framework: pRF data comparison between London 1.5T & Auckland 3T. <https://doi.org/10.17605/OSF.IO/WX4HP><sup>17</sup>.

This project contains Lon-Akl.zip, which itself contains underlying data for each participant in folders 30fl, 39mr and 47fr. These files each contain the following underlying data:

- anatomy (the model of the subject's brain anatomy).

- prf\_bucni (functional MRI data (pRF maps) from London).
- prf\_camri (functional MRI data (pRF maps) from Auckland).
- ROIs (delineation (FreeSurfer label files) for the visual areas).

### Extended data

Supplementary Figure 1.

<https://osf.io/chtra>

Mean population receptive field (pRF) size (A), median cortical magnification factor CMF (B) and response amplitude (beta) (C) binned into eccentricity bands for the London 1.5T site (red) and the Auckland 3T site (black). Panels in columns show different visual regions. Panels in rows show the three participants. Error bars denote 95% confidence intervals based on bootstrapping.

Supplementary Figure 2.

<https://osf.io/chtra>

Median goodness-of-fit (A), normalized goodness-of-fit (B), and noise ceiling (C) binned into eccentricity bands for the London 1.5T site (red) and the Auckland 3T site (black).

Panels in columns show different visual regions. Panels in rows show the three participants. Error bars denote 95% confidence intervals based on bootstrapping.

Supplementary Figure 3.

<https://osf.io/chtra>

Mean population receptive field (pRF) size (A) and median goodness-of-fit (B) binned into eccentricity bands for the Auckland data without filtering (black), and low-pass filtering with kernel 1 s (blue) and 2 s (green). Panels in columns show different visual regions. Panels in rows show the three participants. Error bars denote 95% confidence intervals based on bootstrapping.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

### Acknowledgements

We would like to thank the staff of the Centre for Advanced Magnetic Resonance Imaging (CAMRI) and the Birkbeck-UCL Centre for NeuroImaging (BUCNI) for technical support and the Centre for eResearch, University of Auckland, for computing support.

### References

- Dumoulin SO, Wandell BA: **Population receptive field estimates in human visual cortex.** *NeuroImage*. 2008; **39**(2): 647–660.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Amano K, Wandell BA, Dumoulin SO: **Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex.** *J Neurophysiol*. 2009; **102**(5): 2704–2718.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Winawer J, Horiguchi H, Sayres RA, *et al.*: **Mapping hV4 and ventral occipital cortex: the venous eclipse.** *J Vis*. 2010; **10**(5): 1.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dumoulin SO, Hess RF, May KA, *et al.*: **Contour extracting networks in early extrastriate cortex.** *J Vis*. 2014; **14**(5): 18.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Harvey BM, Dumoulin SO: **Visual motion transforms visual space representations similarly throughout the human visual hierarchy.** *NeuroImage*. 2016; **127**: 173–185.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- de Haas B, Schwarzkopf DS, Anderson EJ, *et al.*: **Perceptual load affects spatial tuning of neuronal populations in human early visual cortex.** *Curr Biol*. 2014; **24**(2): R66–67.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Klein BP, Harvey BM, Dumoulin SO: **Attraction of position preference by spatial attention throughout human visual cortex.** *Neuron*. 2014; **84**(1): 227–237.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Klein BP, Fracasso A, van Dijk JA, *et al.*: **Cortical depth dependent population receptive field attraction by spatial attention in human V1.** *NeuroImage*. 2018; **176**: 301–312.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kay KN, Weiner KS, Grill-Spector K: **Attention reduces spatial uncertainty in human ventral temporal cortex.** *Curr Biol*. 2015; **25**(5): 595–600.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Vo VA, Sprague TC, Serences JT: **Spatial Tuning Shifts Increase the Discriminability and Fidelity of Population Codes in Visual Cortex.** *J Neurosci*. 2017; **37**(12): 3386–3401.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Levin N, Dumoulin SO, Winawer J, *et al.*: **Cortical maps and white matter tracts following long period of visual deprivation and retinal image restoration.** *Neuron*. 2010; **65**(1): 21–31.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Hoffmann MB, Kaule FR, Levin N, *et al.*: **Plasticity and stability of the visual system in human achiasma.** *Neuron*. 2012; **75**(3): 393–401.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Schwarzkopf DS, Anderson EJ, de Haas B, *et al.*: **Larger extrastriate population receptive fields in autism spectrum disorders.** *J Neurosci*. 2014; **34**(7): 2713–2724.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Clavagnier S, Dumoulin SO, Hess RF: **Is the Cortical Deficit in Amblyopia Due to Reduced Cortical Magnification, Loss of Neural Resolution, or Neural Disorganization?** *J Neurosci*. 2015; **35**(44): 14740–14755.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Anderson EJ, Tibber MS, Schwarzkopf DS, *et al.*: **Visual Population Receptive Fields in People with Schizophrenia Have Reduced Inhibitory Surrounds.** *J Neurosci*. 2017; **37**(6): 1546–1556.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kok P, de Lange FP: **Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex.** *Curr Biol*. 2014; **24**(13): 1531–1535.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kok P, Bains LJ, van Mourik T, *et al.*: **Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback.** *Curr Biol*. 2016; **26**(3): 371–376.  
[PubMed Abstract](#) | [Publisher Full Text](#)
- Kok P, van Lieshout LL, de Lange FP: **Local expectation violations result in global activity gain in primary visual cortex.** *Sci Rep*. 2016; **6**: 37706.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ekman M, Kok P, de Lange FP: **Time-compressed preplay of anticipated events in human primary visual cortex.** *Nat Commun*. 2017; **8**: 15276.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Senden M, Emmerling TC, van Hoof R, *et al.*: **Reconstructing imagined letters from early visual cortex reveals tight topographic correspondence between visual mental imagery and perception.** *Brain Struct Funct*. 2019; **224**(3): 1167–1183.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Winawer J, Rauschecker AM, Parvizi J, *et al.*: **Population receptive fields in human visual cortex measured with subdural electrodes.** *J Vis*. 2011; **11**: 1196.  
[Publisher Full Text](#)
- Alvarez I, de Haas BA, Clark CA, *et al.*: **Comparing different stimulus configurations for population receptive field mapping in human fMRI.** *Front Hum Neurosci*. 2015; **9**: 96.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Keliris GA, Li Q, Papanikolaou A, *et al.*: **Estimating average single-neuron visual**

- receptive field sizes by fMRI.** *Proc Natl Acad Sci U S A.* 2019; **116**(13): 6425–6434.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Senden M, Reithler J, Gijzen S, *et al.*: **Evaluating population receptive field estimation frameworks in terms of robustness and reproducibility.** *PLoS One.* 2014; **9**(12): e114054.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
25. van Dijk JA, de Haas B, Moutsiana C, *et al.*: **Intersession reliability of population receptive field estimates.** *NeuroImage.* 2016; **143**: 293–303.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Benson NC, Jamison KW, Arcaro MJ, *et al.*: **The Human Connectome Project 7 Tesla retinotopy dataset: Description and population receptive field analysis.** *J Vis.* 2018; **18**(13): 23.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Edelman WA, Glover GH, Hardy CJ, *et al.*: **The intrinsic signal-to-noise ratio in NMR imaging.** *Magn Reson Med.* 1986; **3**(4): 604–618.  
[PubMed Abstract](#) | [Publisher Full Text](#)
28. Murphy K, Bodurka J, Bandettini PA: **How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration.** *NeuroImage.* 2007; **34**(2): 565–574.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Moutsiana C, de Haas B, Papageorgiou A, *et al.*: **Cortical idiosyncrasies predict the perception of object size.** *Nat Commun.* 2016; **7**: 12110.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Moutsiana C, Soliman R, de Wit L, *et al.*: **Unexplained Progressive Visual Field Loss in the Presence of Normal Retinotopic Maps.** *Front Psychol.* 2018; **9**: 1722.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. de Haas B, Schwarzkopf DS: **Spatially selective responses to Kanizsa and occlusion stimuli in human visual cortex.** *Sci Rep.* 2018; **8**(1): 611.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Hughes AE, Greenwood JA, Finlayson NJ, *et al.*: **Population receptive field estimates for motion-defined stimuli.** *NeuroImage.* 2019; **199**: 245–260.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
33. Brainard DH: **The Psychophysics Toolbox.** *Spat Vis.* 1997; **10**(4): 433–6.  
[PubMed Abstract](#) | [Publisher Full Text](#)
34. Breuer FA, Blaimer M, Heidemann RM, *et al.*: **Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging.** *Magn Reson Med.* 2005; **53**(3): 684–691.  
[PubMed Abstract](#) | [Publisher Full Text](#)
35. Moeller S, Yacoub E, Olman CA, *et al.*: **Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI.** *Magn Reson Med.* 2010; **63**(5): 1144–1153.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Dale AM, Fischl B, Sereno MI: **Cortical surface-based analysis. I. Segmentation and surface reconstruction.** *NeuroImage.* 1999; **9**(2): 179–194.  
[PubMed Abstract](#) | [Publisher Full Text](#)
37. Fischl B, Sereno MI, Dale AM: **Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system.** *NeuroImage.* 1999; **9**(2): 195–207.  
[PubMed Abstract](#) | [Publisher Full Text](#)
38. Schwarzkopf DS, de Haas B, Alvarez I: **SamSrf 6 - Toolbox for pRF modelling.** *Open Science Framework.* 2018.  
[Publisher Full Text](#)
39. Nelder JA, Mead R: **A Simplex Method for Function Minimization.** *Comput J.* 1965; **7**(4): 308–313.  
[Publisher Full Text](#)
40. Lagarias J, Reeds J, Wright M, *et al.*: **Convergence properties of the Nelder—Mead simplex method in low dimensions.** *SIAM J Optim.* 1998; **9**(1): 112–147.  
[Publisher Full Text](#)
41. Harvey BM, Dumoulin SO: **The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture.** *J Neurosci.* 2011; **31**(38): 13604–13612.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
42. Brown W: **Some Experimental Results in the Correlation of Mental Abilities 1.** *Brit J Psychol.* 1910; **3**(3): 296–322.  
[Publisher Full Text](#)
43. Spearman C: **Correlation Calculated from Faulty Data.** *Brit J Psychol.* 1910; **3**(3): 271–295.  
[Publisher Full Text](#)
44. Spearman C: **The Proof and Measurement of Association between Two Things.** *Am J Psychol.* 1904; **15**(1): 72–101.  
[Publisher Full Text](#)
45. Sereno MI, Dale AM, Reppas JB, *et al.*: **Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging.** *Science.* 1995; **268**(5212): 889–893.  
[PubMed Abstract](#) | [Publisher Full Text](#)
46. Schwarzkopf DS, Morgan C: **Supplementary Figures.** 2020.  
[Reference Source](#)
47. Schwarzkopf D, Morgan C: **pRF data comparison between London 1.5T & Auckland 3T.** 2019.  
<http://www.doi.org/10.17605/OSF.IO/WX4HP>
48. Hutton C, Bork A, Josephs O, *et al.*: **Image distortion correction in fMRI: A quantitative evaluation.** *NeuroImage.* 2002; **16**(1): 217–240.  
[PubMed Abstract](#) | [Publisher Full Text](#)
49. Infanti E, Schwarzkopf DS: **Mapping sequences can bias population receptive field estimates.** *bioRxiv.* 2019.  
[Publisher Full Text](#)
50. Binda P, Thomas JM, Boynton GM, *et al.*: **Minimizing biases in estimating the reorganization of human visual areas with BOLD retinotopic mapping.** *J Vis.* 2013; **13**(7): 13.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 1

Reviewer Report 12 December 2019

<https://doi.org/10.5256/f1000research.22528.r56563>

© 2019 Binda P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Paola Binda** 

Faculty of Medicine, University of Pisa, Pisa, Italy

pRF mapping provides a quantitative description of cortical representations, including the spatial representation in visual cortical areas. While pRF parameters are extensively used to quantitatively compare across subjects and conditions, we are well aware that such parameters depend on both neural factors and on nuisance variables that are related to the dynamics of BOLD. However, we lack a precise understanding of how such nuisance variables might affect each of the pRF parameters. Morgan & Schwarzkopf make an important effort in this direction and test the impact of static SNR (affected by MR field strength) and temporal SNR (affected by temporal filtering) on spatial pRF parameters in visual cortical areas.

I have no major concerns with the study and only offer few suggestions that hopefully will increase the usability of the reported results:

1. It would be useful to explicitly describe one (or more) models that predict changes of pRF parameters with SNR.

I feel one intuitive (and probably wrong) model invokes a non-linear thresholding effect. The logic is, approximately, as follows: when BOLD responses are stronger, each voxel responds to a larger range of stimulus position and the pRF size is consequently increased. Is this one of the hypotheses that the study intends to test (and falsify)?

This particular model would predict that beta values (indexing BOLD response amplitude) and pRF size estimates correlate positively. Since both indices were collected here, could the authors measure this (lack of) correlation?

In addition, this model would incorrectly predict larger pRF size estimates at higher MR fields, under the assumption that response amplitude is positively related to SNR. Could the authors report whether this assumption (larger beta-values in the higher SNR, i.e. 3T, dataset) is correct for their specific dataset?

Also, I wonder whether one could also expect variations of the other pRF parameter (pRF position),

e.g. in the form of an attraction towards the fovea / screen borders / horizontal and vertical meridians. The text reports relatively high correlations between pRF center parameters at 1.5 and 3T, but this does not exclude systematic shifts in eccentricity or polar angle. The same applies to the effects of temporal filtering: can the authors exclude systematic shifts of pRF positions?

2. I think that the figures could be more incisive if they were more compact. While I understand the importance of reporting single-participant's data, I feel the need for gathering the data in fewer more illustrative plots. One possibility would be to use the format of scatter-plots comparing pRF parameters at 1.5 vs. 3T (so that the three subjects could be represented as different symbols within a single plot) - perhaps accompanied by a grand-average showing the expected relationship between pRF parameters and visual areas or eccentricity.
3. It might be worth pointing out that the bias introduced by temporal filtering is linked to the regular sequence of visual presentations. Presumably, temporal filtering would have had no effect on pRF size if the visual stimulus used for retinotopic mapping varied its position randomly - instead of sweeping across the visual field.
4. Sharing the raw data is sometimes very difficult, but clearly this is encouraged by many in the field. I wonder whether the authors could provide a justification for not providing .nii files (rather than the analyzed datasets, which are certainly sufficient for reproducing the figures but do not allow for reusing the dataset for novel investigations).

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Yes

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly

**Are the conclusions drawn adequately supported by the results?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Physiology of the visual system

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 15 Feb 2020

**Catherine Morgan**, University of Auckland, Auckland, New Zealand

We'd like to thank the reviewer for their considered questions and useful suggestions on this manuscript. In our response we have reproduced the report, and commented immediately after each point.

"pRF mapping provides a quantitative description of cortical representations, including the spatial representation in visual cortical areas. While pRF parameters are extensively used to quantitatively compare across subjects and conditions, we are well aware that such parameters depend on both neural factors and on nuisance variables that are related to the dynamics of BOLD. However, we lack a precise understanding of how such nuisance variables might affect each of the pRF parameters. Morgan & Schwarzkopf make an important effort in this direction and test the impact of static SNR (affected by MR field strength) and temporal SNR (affected by temporal filtering) on spatial pRF parameters in visual cortical areas.

I have no major concerns with the study and only offer few suggestions that hopefully will increase the usability of the reported results:

It would be useful to explicitly describe one (or more) models that predict changes of pRF parameters with SNR.

I feel one intuitive (and probably wrong) model invokes a non-linear thresholding effect. The logic is, approximately, as follows: when BOLD responses are stronger, each voxel responds to a larger range of stimulus position and the pRF size is consequently increased. Is this one of the hypotheses that the study intends to test (and falsify)?

This particular model would predict that beta values (indexing BOLD response amplitude) and pRF size estimates correlate positively. Since both indices were collected here, could the authors measure this (lack of) correlation?"

>>Thank you for this important question. A stronger response amplitude could potentially appear as a larger pRF because visual field locations farther from the pRF centre respond to the stimulus. However, this is unlikely to manifest as a positive correlation with the response amplitude (beta) parameter because the pRF model is scale-invariant, that is, the pRF parameters are estimated by correlation between predicted and observed time series and thus theoretically only the *pRF shape* matters to estimation.

In fact, the opposite relationship, that pRF size estimates are larger with low beta parameters, is actually more realistic: response amplitudes are linked to signal-to-noise and thus goodness-of-fit. Under low noise conditions, a small pRF would show a clean response time series with very clearly defined peaks. In a noisier pRF, this time series would however be contaminated and so the peaks of the response would be less clearly defined. In order to fit a pRF model, the algorithm will therefore produce an artifactual bias on pRF size estimates.

We have already seen this relationship in previous work. However, for the sake of completeness we now conducted the suggested correlation analysis between beta parameters and pRF size at the single-vertex level. This confirms that the second model is indeed correct: across participants and visual regions of interest there is a negative correlation between pRF size estimates and beta parameters (Spearman's rho ranging between -0.17 and -0.5 for most cases, with the exception of

V3A where the small amount of data may skew the results).

In other words, when response amplitude is low, pRFs tend to be estimated as *larger*. Unfortunately, this analysis is complicated by several factors: the data are highly heteroscedastic (there is far more variability in response amplitude for small pRFs than large ones), a small but considerable proportion of beta parameters is negative (we believe these are related to artifactual pRF estimates with negative BOLD responses due to large blood vessels or edge-of-stimulus artifacts), and there is also a strong but incomplete ceiling artifact as many pRF sizes cluster at a small size. Nevertheless, we believe due to its consistency and robustness to thresholding this relationship probably correct.

We now briefly discuss this issue in the introduction of the revised manuscript. Please note that because we find no consistent difference in pRF size between scanner sites, there is no evidence that this link has any bearing on the conclusions of our present study. Therefore, and because of their complexity, we chose not to present and discuss the results of this control analysis in the revised manuscript. However, the analysis script for this is now included in our online repository.

“In addition, this model would incorrectly predict larger pRF size estimates at higher MR fields, under the assumption that response amplitude is positively related to SNR. Could the authors report whether this assumption (larger beta-values in the higher SNR, i.e. 3T, dataset) is correct for their specific dataset?”

>> As the reviewer correctly points out, in MRI, the signal-to-noise ratio (SNR) is directly proportional to main magnetic field strength (and voxel volume) (Edelstein, Glover, Hardy, & Redington, 1986). Therefore power to detect BOLD activation theoretically increases with static magnetic field (or larger voxel sizes), or where this is not possible/desirable, longer scan times. However, noise in fMRI data has multiple contributions (physiological, thermal and system related) and therefore the relationship between the temporal signal to noise ratio (TSNR) in a fMRI time course has a non-linear relationship with static SNR. This has been demonstrated theoretically and empirically (Murphy, Bodurka, & Bandettini, 2007) showing that any gains in static SNR, from e.g. increased field strength, will not translate to proportional gains in TSNR due to a limit where physiological noise dominates.

Since response amplitude (beta) is related to goodness-of-fit, we assume that betas are indeed greater at 3T. We now conducted the suggested analysis and this confirms that prediction: betas are larger at 3T at least in the early visual areas showing a similar pattern as goodness-of-fit. We added this figure to what is now Supplementary Figure S1.

But importantly, as we explained above, larger betas do not result in larger pRF size estimates but rather the opposite. Despite this, we found no consistent differences in pRF size estimates between the two scanner sites.

“Also, I wonder whether one could also expect variations of the other pRF parameter (pRF position), e.g. in the form of an attraction towards the fovea / screen borders / horizontal and vertical meridians. The text reports relatively high correlations between pRF center parameters at 1.5 and 3T, but this does not exclude systematic shifts in eccentricity or polar angle. The same applies to the effects of temporal filtering: can the authors exclude systematic shifts of pRF



positions? “

We would not predict any systematic relationship between pRF position and magnetic field strength but this is an interesting suggestion. There does not appear to be any strong evidence of such a systematic bias in the maps at least (please refer to the example maps we included – an interested reader may be able to reproduce them using our shared analysis scripts). The biggest differences in eccentricity maps between the two sites appear to be additional voxels within the foveal region in the 3T data, rather than an obvious shift in the map. We however understand that this visual inspection of the maps is only a very coarse analysis.

A quantitative analysis would be to plot visual field positions directly against one another. For reasons that are outside the scope of this discussion, it is imperative that this comparison must be done using an independent criterion for determining which data points we compare between the two sites. We therefore determined the cortical distance of each vertex in the spherical surface model from a point chosen near the occipital pole as a proxy of ground truth eccentricity. We then quantified the mean eccentricity within 5 mm cortical distance bins for the two sites. This revealed no obvious systematic differences in eccentricity between the two sites. We believe this analysis is too complex to be included in the main publication but we included the code for it in our public repository and made this figure publicly available: <https://osf.io/k4dcr/>

“I think that the figures could be more incisive if they were more compact. While I understand the importance of reporting single-participant’s data, I feel the need for gathering the data in fewer more illustrative plots. One possibility would be to use the format of scatter-plots comparing pRF parameters at 1.5 vs. 3T (so that the three subjects could be represented as different symbols within a single plot) - perhaps accompanied by a grand-average showing the expected relationship between pRF parameters and visual areas or eccentricity.”

>> Thank you for this observation. We agree both you and the other reviewer make a good point and we have now consolidated the individual plots into group average plots, and moved the individual figures into the supplementary material.

We also like your idea of scatter plots but we believe they would also be somewhat complex as we would ideally also need to denote the different eccentricities (perhaps with a colour code). Especially for measures that do not change linearly with eccentricity (like beta or goodness-of-fit) this would make for very dense plots.

It might be worth pointing out that the bias introduced by temporal filtering is linked to the regular sequence of visual presentations. Presumably, temporal filtering would have had no effect on pRF size if the visual stimulus used for retinotopic mapping varied its position randomly - instead of sweeping across the visual field.

>> This is a very good suggestion and we now discuss this in the revised manuscript. However, please also keep in mind that other work by ourselves (<sup>1</sup>, currently under peer review, but a preprint is available here: <https://doi.org/10.1101/821918>) and of course also your own work <sup>2</sup> has shown that pRF size estimates differ between random and ordered designs. This issue was also discussed in a study on tonotopic pRF mapping<sup>3</sup>. It is possible that temporal filtering would not show the effect we report here for a random design but there are probably additional factors linking pRF size with randomisation. For instance, we showed in Infanti & Schwarzkopf that whether or not

pRF size is over- or underestimated in ordered relative to random designs also depends on the stimulus parameters.

Sharing the raw data is sometimes very difficult, but clearly this is encouraged by many in the field. I wonder whether the authors could provide a justification for not providing .nii files (rather than the analyzed datasets, which are certainly sufficient for reproducing the figures but do not allow for reusing the dataset for novel investigations).

>> please refer to our response to reviewer 1 to answer this point. In short, there are both legal and ethical barriers to our sharing of raw data files. However, we want to stress that all of the time series information necessary for reproducing the pRF analysis, reanalysing it with a different software, or conducting additional novel analyses is available in our public repository. For anyone seeking to conduct analysis that are not possible with the shared data we are willing to make these data available upon request provided the participants consent to this use.

#### Response References

1. Infanti, E. & Schwarzkopf, D. S. Mapping sequences can bias population receptive field estimates. *bioRxiv* 821918 (2019) doi:10.1101/821918.
2. Binda, P., Thomas, J. M., Boynton, G. M. & Fine, I. Minimizing biases in estimating the reorganization of human visual areas with BOLD retinotopic mapping. *J. Vis.* **13**, 13–13 (2013).
3. Thomas, J. M. *et al.* Population receptive field estimates of human auditory cortex. *NeuroImage* **105**, 428–439 (2015).

**Competing Interests:** No competing interests were disclosed.

Reviewer Report 25 October 2019

<https://doi.org/10.5256/f1000research.22528.r54262>

© 2019 Benson N. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Noah Benson** 

Department of Psychology, Center for Neural Science, New York University, New York, NY, USA

Population receptive field (pRF) modeling is an incredibly important and common technique in visual and perceptual neuroscience. PRF mapping is also complex, and, despite its popularity, few examinations of how the many incidental complexities of collecting pRF maps affect the maps themselves have been reported. This lack makes comparison and interpretation of pRF results across sites and studies difficult. In this article, Morgan and Schwarzkopf provide a straightforward examination of two of these incidental complexities of retinotopic pRF mapping: the field-strength of the scanner and the choice to low-pass filter the BOLD time-series. Encouragingly, they find that the effect of the field-strength is minimal; however, the effects of the filtering are non-trivial.

I have no major concerns regarding the methods of this paper, but I do have a few suggestions. First, it is

not clear if the authors used the same hemodynamic response function (HRF) for all subjects or if a single canonical function was fit to empirical data from each subject individually, nor do the authors explain which canonical HRF was used; I would like to see the HRF form(s) reported explicitly. Secondly, the use of linearly-spaced eccentricity bands for the various further analyses is a somewhat questionable choice given that the number of vertices between 8.5° and 9.5° will be substantially smaller than the number of vertices between 0.5° and 1.5°. This effect may partly explain why the errors in some figures (e.g. Figure 6) are quite large at higher eccentricities. A principled way to obtain similarly-sized eccentricity bins might be to use Horton and Hoyt's 1991 equation <sup>1</sup> for cortical magnification [ $CMF(e) = (17.3 / (0.75 + e))^2$  (mm/deg)<sup>2</sup> for eccentricity  $e$ ], integrate it over the visual field to obtain an estimate of cortical surface area in terms of eccentricity [ $SAF(e)$  (mm<sup>2</sup>)], then to invert this [ $SAF^{-1}(a)$  (deg) for cortical surface area  $a$ ] and choose eccentricity values that correspond to values of  $SAF^{-1}(a)$  for linearly-spaced surface-area values of  $a$ .

Despite the above suggestions, the results of the paper are quite clear: (1) pRF models are robust to changes in the field strength of the scanner, and (2) temporal filtering inflates the  $R^2$  and size parameter of the pRF fits without substantially changing the topographic structure of the retinotopic maps. Both of these insights are useful additions to the field, especially given the rapid rate of change in MRI technology and the frequency with which temporal smoothing is applied to pRF mapping data. More work is needed throughout the field to quantify these kinds of effects, and this paper is a useful contribution to that effort. One minor suggestion is that the authors might explicitly state in the discussion that the two scanners share a manufacturer; this is both a limitation of the comparison, in that the results may not generalize across manufacturers, and one of its strengths, in that one can be more confident of the findings with respect to field-strength because the manufacturer was not an additional variable across sites.

My biggest concerns about the paper center on the supplemental data. Although the authors have provided enough data to reproduce the figures and aspects of the study, the data provided are incomplete and fail to conform to contemporary standards of reusability and interoperability. For one, all of the supplemental data files could have been provided in well established community-vetted formats (such as NIFTI files), yet *none* of the data are provided in such formats (in fact many of the files cannot be easily read without an expensive MATLAB license). Additionally, only some of the data are provided: FreeSurfer reconstructions, anonymized anatomical images, and various raw data are not provided. These data really ought to be part of the dataset, and ideally should be provided in accordance with the BIDS spec <sup>2</sup>. A few examples of things that cannot be examined with the current dataset: concerns about field distortions, questions about partial voluming artifacts, concerns about signal attenuation near vessels, questions about how the results relate to the surface area of the visual areas, and questions about how the results generalize to frontal cortex. To be clear, none of these examples are likely to impact the conclusions of the study, but I point them out as examples of how the current dataset fails to meet the evolving standards of data-availability in the field.

Beyond these comments, I have only minor concerns that are largely aesthetic in nature:

The use of the prime symbol (') and the single quote (') is inconsistent and, due to the occasional adjacency of variable symbols with subscripts next to commas, can be a bit confusing.

Figures 3, 4, and 6 contain a lot of information, and I feel that a panel that summarizes or combines the data across subjects in some way would be useful and easier to interpret.

The eccentricity colormap used in Figure 2 is circular, which causes some ambiguity between very different values.

Figure 1 seems to have a larger range (along the y axis) for the filtered data than for the unfiltered data. It is not clear to me how the filtering described would produce a signal with nearly twice the range of values

as the original signal. This might be a plotting error.

### References

1. Horton JC, Hoyt WF: The representation of the visual field in human striate cortex. A revision of the classic Holmes map. *Arch Ophthalmol*. 1991; **109** (6): 816-24 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline JB, Rokem A, Schaefer G, Sochat V, Triplett W, Turner JA, Varoquaux G, Poldrack RA: The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data*. 2016; **3**: 160044 [PubMed Abstract](#) | [Publisher Full Text](#)

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Yes

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly

**Are the conclusions drawn adequately supported by the results?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Data science / visual neuroscience

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Author Response 15 Feb 2020

**Catherine Morgan**, University of Auckland, Auckland, New Zealand

We'd like to thank the reviewer for their considered questions and useful suggestions on this manuscript. In our response we have reproduced the report, and commented immediately after each point.

"Population receptive field (pRF) modeling is an incredibly important and common technique in visual and perceptual neuroscience. PRF mapping is also complex, and, despite its popularity, few examinations of how the many incidental complexities of collecting pRF maps affect the maps themselves have been reported. This lack makes comparison and interpretation of pRF results across sites and studies difficult. In this article, Morgan and Schwarzkopf provide a straightforward examination of two of these incidental complexities of retinotopic pRF mapping: the field-strength

of the scanner and the choice to low-pass filter the BOLD time-series. Encouragingly, they find that the effect of the field-strength is minimal; however, the effects of the filtering are non-trivial.

I have no major concerns regarding the methods of this paper, but I do have a few suggestions. First, it is not clear if the authors used the same hemodynamic response function (HRF) for all subjects or if a single canonical function was fit to empirical data from each subject individually, nor do the authors explain which canonical HRF was used; I would like to see the HRF form(s) reported explicitly.”

>> For all subjects we used the same canonical HRF based on previous data (de Haas et al. 2014). The reviewer is right that the HRF used can have an effect on the obtained pRF parameters but this effect is typically very small (Dumoulin & Wandell, 2008). In our previous study quantifying the test-retest reliability of pRF parameters on the same scanner we found only minimal effects of obtaining subject-specific HRF fits. While it is indeed possible that the HRF differs between our scanners in the present study, this effect is also unlikely to alter our conclusions considering the already very similar pRF parameters we observed. We thank you for pointing this ambiguity out and will edit the manuscript accordingly.

“Secondly, the use of linearly-spaced eccentricity bands for the various further analyses is a somewhat questionable choice given that the number of vertices between 8.5° and 9.5° will be substantially smaller than the number of vertices between 0.5° and 1.5°. This effect may partly explain why the errors in some figures (e.g. Figure 6) are quite large at higher eccentricities. A principled way to obtain similarly-sized eccentricity bins might be to use Horton and Hoyt's 1991 equation 1 for cortical magnification  $[CMF(e) = (17.3 / (0.75 + e))^2 \text{ (mm/deg)}^2 \text{ for eccentricity } e]$ , integrate it over the visual field to obtain an estimate of cortical surface area in terms of eccentricity  $[SAF(e) \text{ (mm}^2)]$ , then to invert this  $[SAF^{-1}(a) \text{ (deg)} \text{ for cortical surface area } a]$  and choose eccentricity values that correspond to values of  $SAF^{-1}(a)$  for linearly-spaced surface-area values of  $a$ .”

>> The reviewer raises a very important point and we have previously debated this in the lab and also with other researchers. Linearly scaled eccentricity bands is the standard procedure used by most pRF studies . It is true that a logarithmic scaling of eccentricity bands would ensure approximately equal amounts of data per eccentricity band, and thus ensure relatively even error margins. It however seems to us unlikely that this will alter our conclusions in any way or that this will reveal additional information. Practically, it would reduce the number of bins which could also be suboptimal for comparing the findings by visual inspection. As an example, we included the equivalent to the figure for the pRF size from the two scanning sites but using logarithmic bands online: <https://osf.io/nyj4z/>

To our eyes this does not alter what we can interpret from this plot. Note also that the most peripheral band still contains less data because edge artifacts are not always consistent and thus we cannot ensure equal surface area for this band. We therefore decided to go with tradition and stick with linear bins.

“Despite the above suggestions, the results of the paper are quite clear: (1) pRF models are robust to changes in the field strength of the scanner, and (2) temporal filtering inflates the  $R^2$  and size parameter of the pRF fits without substantially changing the topographic structure of the retinotopic maps. Both of these insights are useful additions to the field, especially given the rapid rate of change in MRI technology and the frequency with which temporal smoothing is applied to pRF

mapping data. More work is needed throughout the field to quantify these kinds of effects, and this paper is a useful contribution to that effort. One minor suggestion is that the authors might explicitly state in the discussion that the two scanners share a manufacturer; this is both a limitation of the comparison, in that the results may not generalize across manufacturers, and one of its strengths, in that one can be more confident of the findings with respect to field-strength because the manufacturer was not an additional variable across sites.”

>> We'd like to thank the reviewer for highlighting this very important point and will add it to the discussion.

“My biggest concerns about the paper center on the supplemental data. Although the authors have provided enough data to reproduce the figures and aspects of the study, the data provided are incomplete and fail to conform to contemporary standards of reusability and interoperability. For one, all of the supplemental data files could have been provided in well established community-vetted formats (such as NIFTI files), yet *none* of the data are provided in such formats (in fact many of the files cannot be easily read without an expensive MATLAB license). Additionally, only some of the data are provided: FreeSurfer reconstructions, anonymized anatomical images, and various raw data are not provided. These data really ought to be part of the dataset, and ideally should be provided in accordance with the BIDS spec [2](#). A few examples of things that cannot be examined with the current dataset: concerns about field distortions, questions about partial voluming artifacts, concerns about signal attenuation near vessels, questions about how the results relate to the surface area of the visual areas, and questions about how the results generalize to frontal cortex. To be clear, none of these examples are likely to impact the conclusions of the study, but I point them out as examples of how the current dataset fails to meet the evolving standards of data-availability in the field.”

>> We are very supportive of the idea of open science in general, for data integrity and reproducibility. However, data transparency must be traded off against ethical and data protection concerns. Any brain imaging data is by nature not fully de-identifiable, especially if it contains anatomical/morphological information. This is a particular issue with a small sample size like this (which is not uncommon for retinotopic mapping studies). Moreover we also have reason to believe that under New Zealand laws (specifically the *Privacy Act (1994)* and the Health Information Privacy Code (HIPC); a set of legally-binding rules for applying this act) we cannot share our full imaging data (including volumetric EPI data and FreeSurfer reconstructions) publicly without any custodianship as it can only be shared for research and for the same purpose as our original ethical approval. We are currently in the process of seeking legal guidance on this point and determine a consistent policy on data sharing within our institution.

However, even in the absence of all the raw data, the data we do share are fully transparent and theoretically permit anyone to reproduce our findings. As opposed to the reviewer's concern, our data do not require a MATLAB license but can be read using the free software Octave and FreeSurfer. Octave misses some of the functionality of MATLAB and we are currently working on opening this up but critically the data are fully accessible. All the pre-processed time series for individual voxels/vertices are also included which would allow an interested research to carry out their own pRF analysis on these data.

Regarding the format, at present we always share our data in the MATLAB format because that is what our SamSrf toolbox uses. It is convenient because it allows us to store all the functional data per hemisphere in one file, in a relatively straightforward data structure, with maps and time series

information of all vertices of the cortical reconstruction in a single 2D matrix. We however see the merit in a ready exchange of data between different platforms and are currently considering and debating adding support for standard formats that allow this, but in our view this is outside the scope of this revision..

“Beyond these comments, I have only minor concerns that are largely aesthetic in nature: The use of the prime symbol (') and the single quote (') is inconsistent and, due to the occasional adjacency of variable symbols with subscripts next to commas, can be a bit confusing.”

>> We thank the reviewer for pointing this out and will change this in our revision.

“Figures 3, 4, and 6 contain a lot of information, and I feel that a panel that summarizes or combines the data across subjects in some way would be useful and easier to interpret.”

>> This is an interesting point. Considering our sample size is 3 we believe presenting individual data is important. However, the group results are also very clear and we therefore decided to present only group averages in the main manuscript now, thank you for the suggestion, and instead relegate the individual plots to the Supplementary Information.

“The eccentricity colormap used in Figure 2 is circular, which causes some ambiguity between very different values.”

>> The choice of the same colour scheme for polar angle and eccentricity is in part a leftover from old days in which eccentricity maps were based on phase-encoded analysis. We agree that this could be somewhat confusing, although our rendering ensures that this is not the case. The fovea is represented by red whereas the most peripheral eccentricities are shown in purple. The colour scheme *does not cycle back*. Instead, all eccentricities above the outer edge of the stimulus (9.5 degrees) are set to purple.

“Figure 1 seems to have a larger range (along the y axis) for the filtered data than for the unfiltered data. It is not clear to me how the filtering described would produce a signal with nearly twice the range of values as the original signal. This might be a plotting error.”

>> We thank the reviewer for pointing this out. This is an artifact in this illustration and the Y-axis should really be in arbitrary units. We will change this in our revision.

>> Response References

de Haas B, Schwarzkopf DS, Anderson EJ, et al.: Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Curr Biol*. 2014; 24(2): R66–67.

Dumoulin SO, Wandell BA: Population receptive field estimates in human visual cortex. *NeuroImage*. 2008; 39(2): 647–660.

**Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**