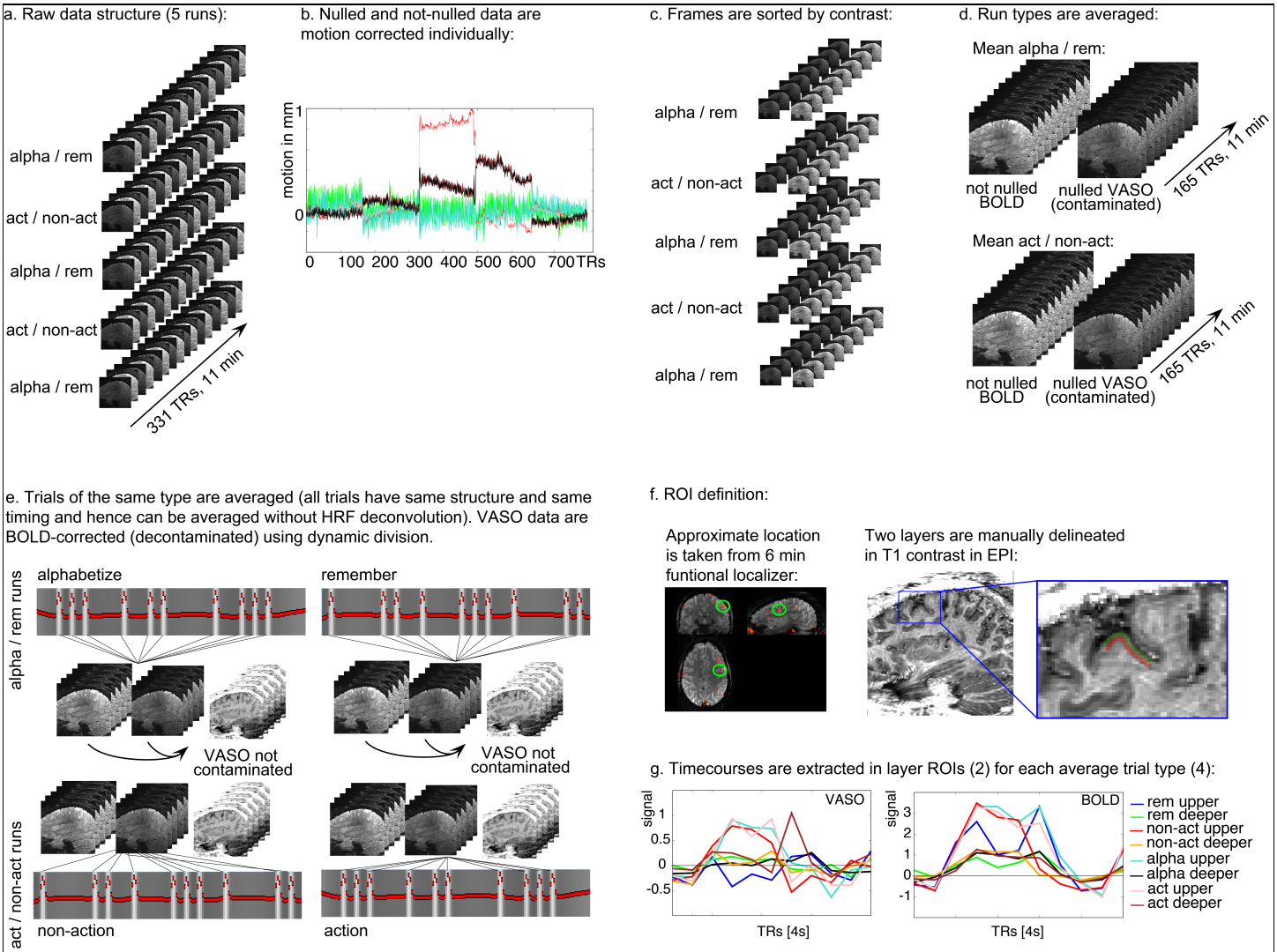
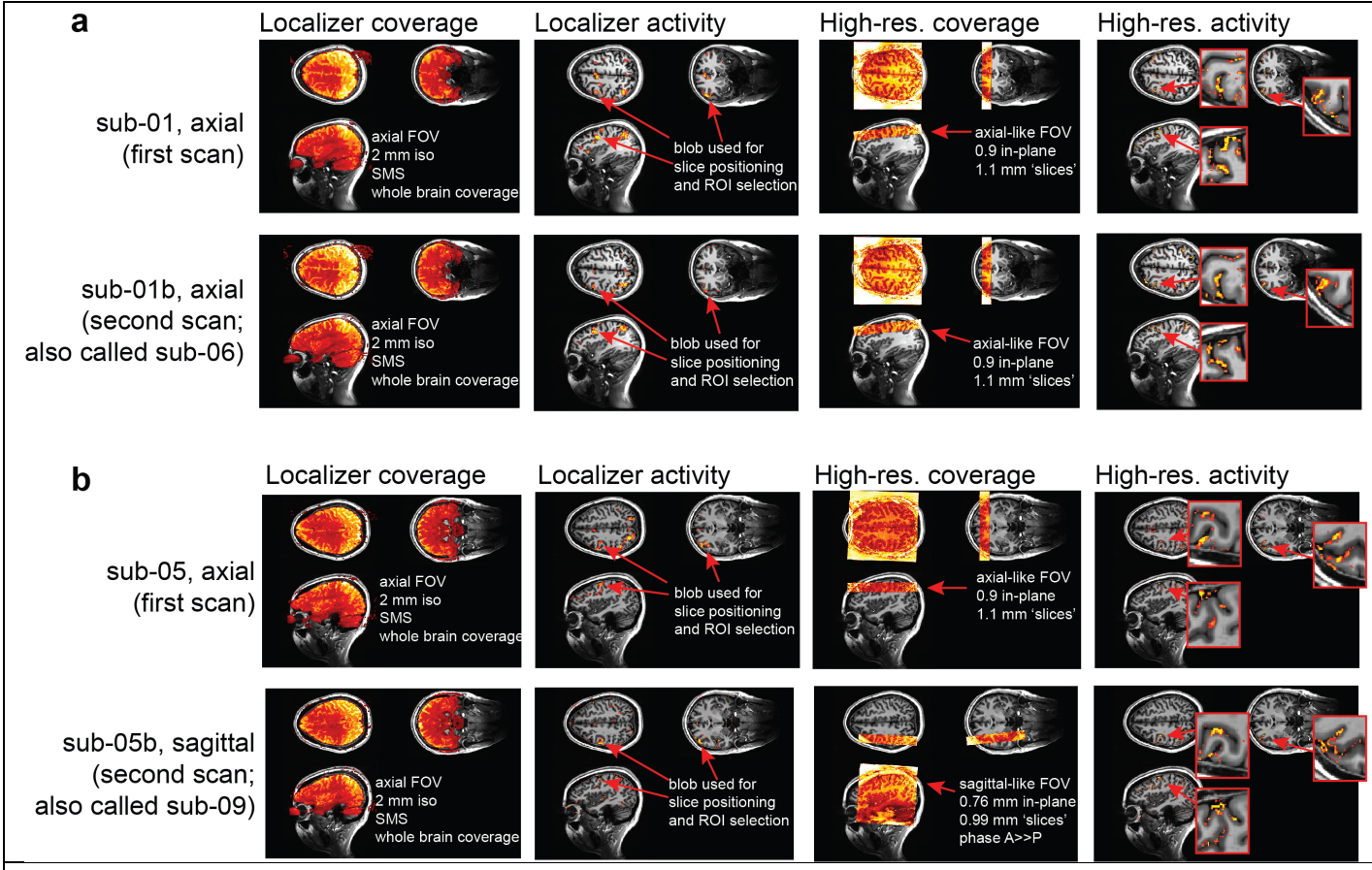


Supplementary Figure 1

Graphical depiction of the analysis pipeline

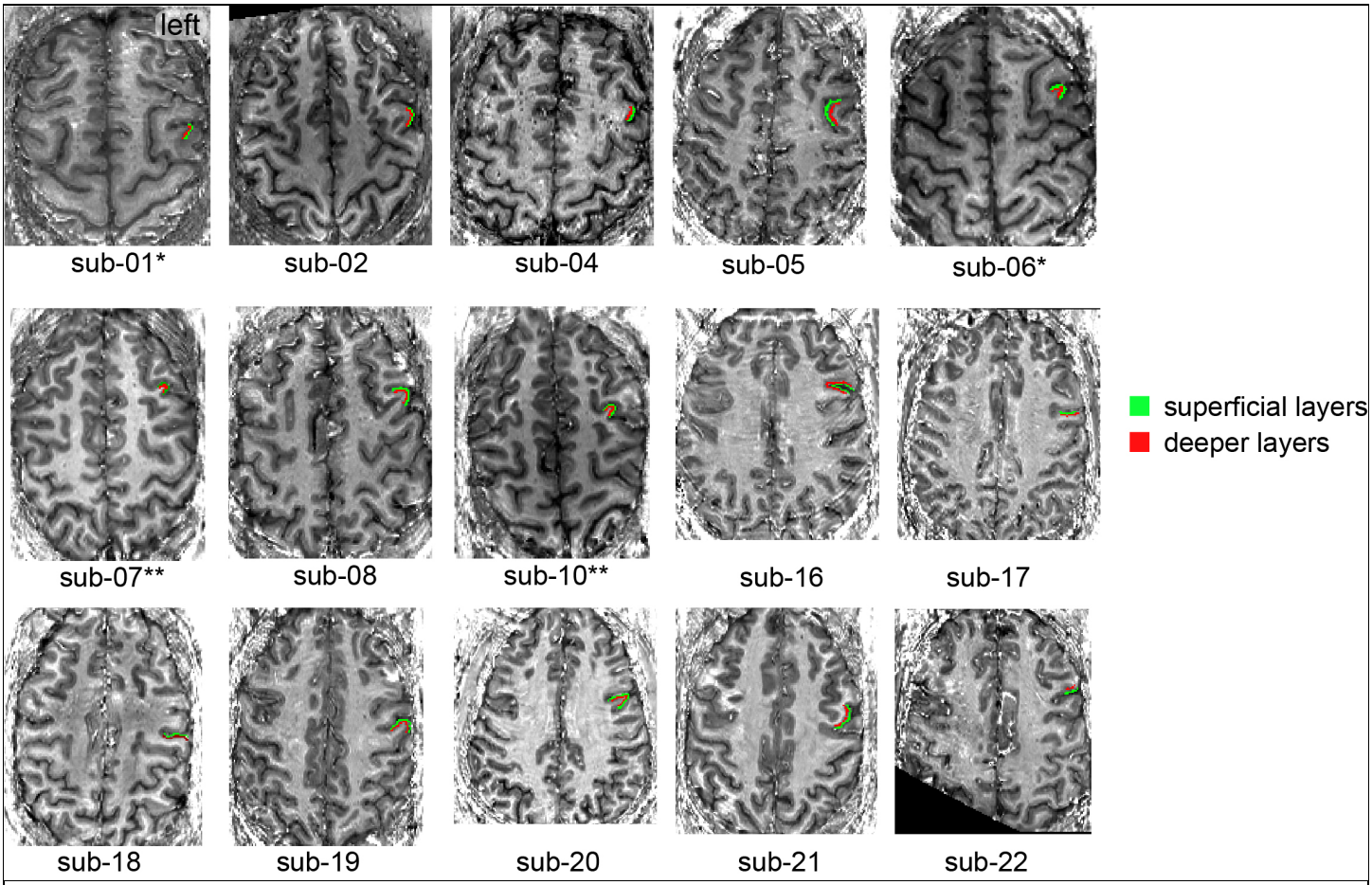
(A) Multiple runs are collected per session with two different types (‘alpha / rem’, or manipulation versus maintenance, and ‘act / non-act’, or action versus non-action). Each run consists of interleaved images with (VASO) and without (BOLD) blood nulling. (B) Nulled and not-nulled images are motion corrected separately for the whole session. It is manually checked that the motion traces of nulled and not-nulled images are matching. (C) The time series are sorted by imaging contrast. (D) Runs of the same type are averaged within contrast. (E) Time series from all trials are averaged based on their task condition. In the 2 × 2 design used here, this results in 4 average trial types. Nulled images are corrected for BOLD contamination with a time-wise division of not-nulled images to provide a clean VASO contrast. (F) Based on the approximate location of activation in the low-resolution functional localizer, two layer ROIs are manually drawn on the T1-EPI anatomical images. (G) Time series for all four task conditions are extracted from the ROI of superficial and deeper layers for BOLD and VASO contrasts, resulting in 8 time series for each contrast, or 16 total time series.



Supplementary Figure 2

Reproducibility of activation loci in two subjects across two different days

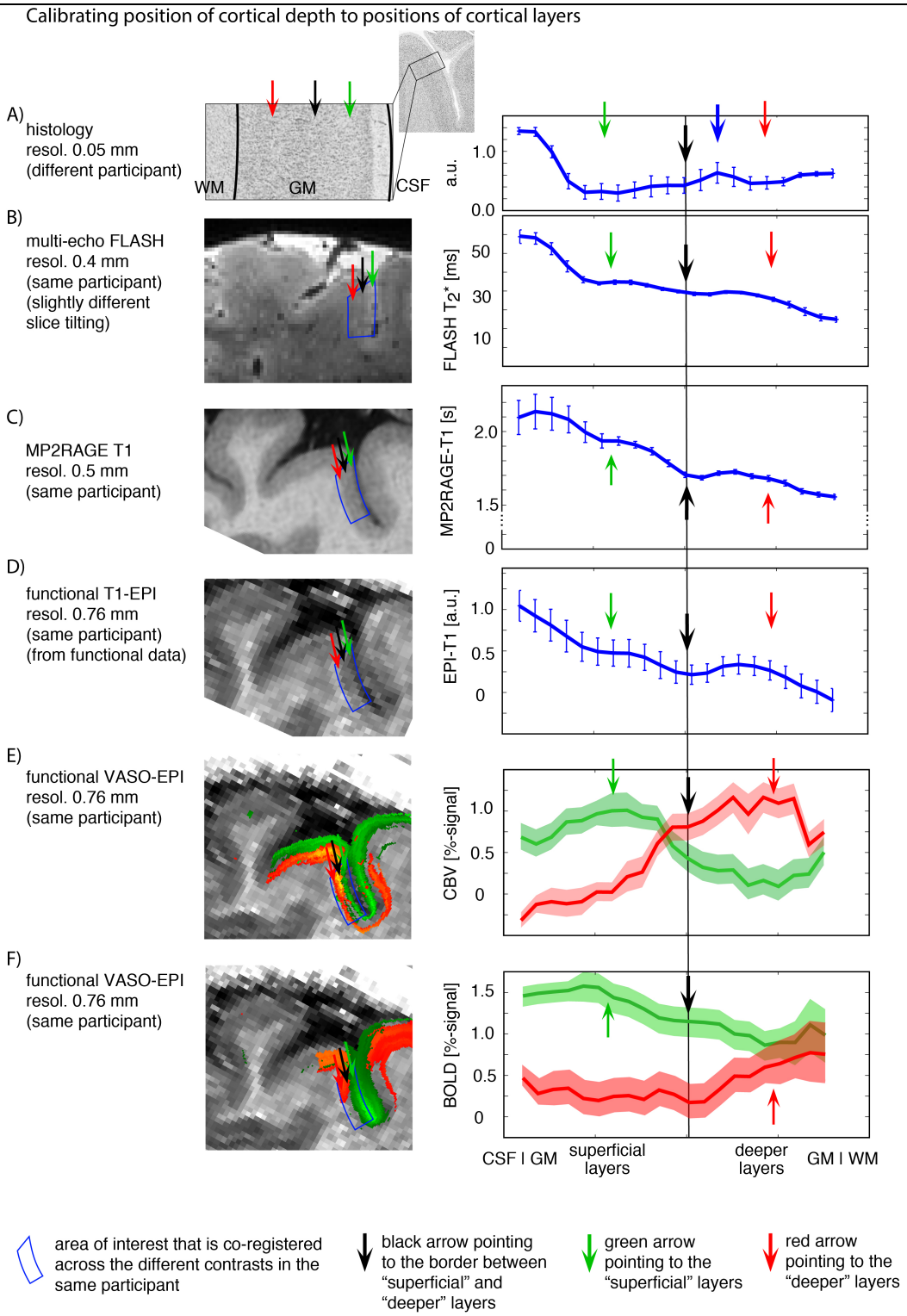
(A) Reproducibility of activation loci in one subject scanned with the axial imaging protocol on two different days. The same parts of gray matter are engaged in the task across the same acquisition protocol on two different days (first session, top row; second session, bottom row). The consistency of the locus of activation is within the accuracy of the registration quality. (B) Reproducibility of activation loci in one subject across the axial and sagittal imaging protocols acquired on two different days. The same parts of gray matter are engaged in the task across the two different acquisition protocols (axial-like, top row; and sagittal-like, bottom row) acquired on two different days. The consistency of the locus of activation is within the accuracy of the registration quality.



Supplementary Figure 3

Layer ROIs for individual subjects and sessions

Masks of superficial and deeper layers for each subject/session that were used to extract the timecourses in Figs. 2, 3, and S5-S7. The grayscale background contrast refers to the T1-contrast in functional VASO data. Here the T1-contrast in the functional data is used to identify superficial and deeper voxels (manually drawn green and red masks). Asterisks (* and **) are used to denote test-retest scans from the same participant (i.e., sub-06 is the same as sub-01; sub-10 is the same as sub-07).

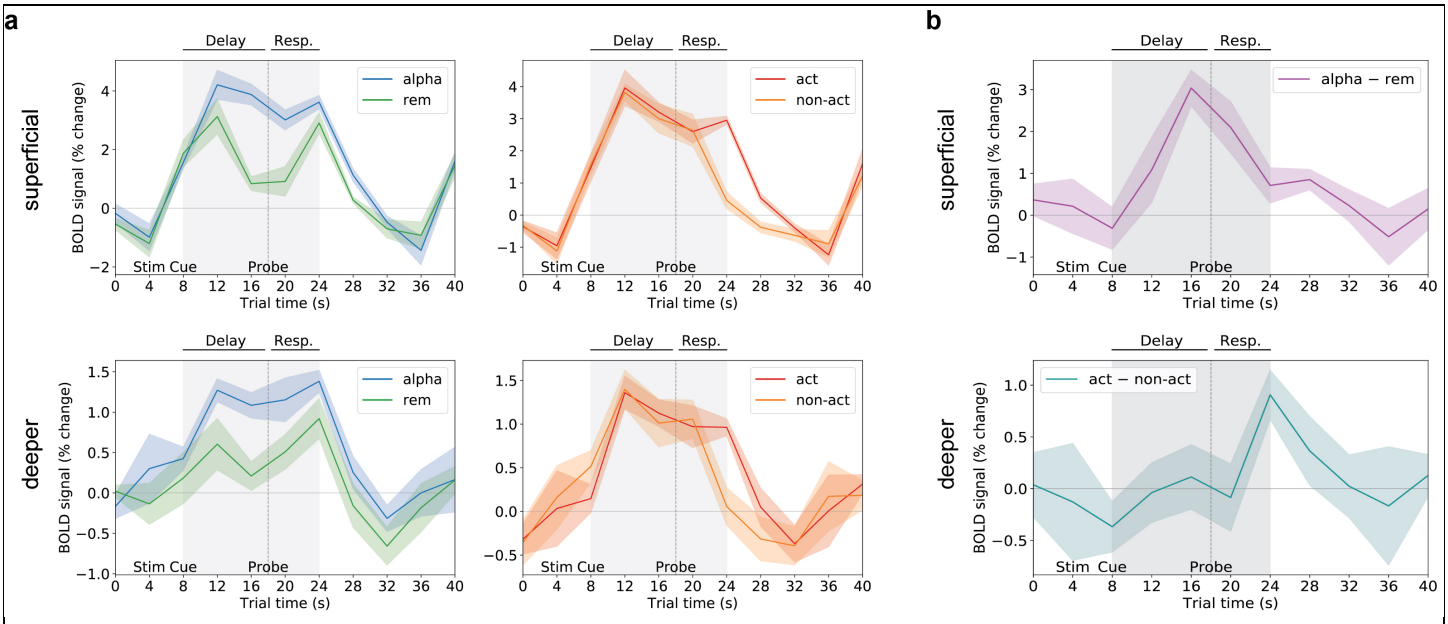


Supplementary Figure 4

Comparison of cortical depth with cytoarchitectonic cortical layers

(A) Grayscale Nissl-stained images of the dlPFC from the Ding Atlas (ref. 67 in Methods) were converted into single-slice NIFTI format. Borderlines of the GM-CSF surface and GM-WM surface are manually drawn in the same space. Layers are extracted with LAYNII as

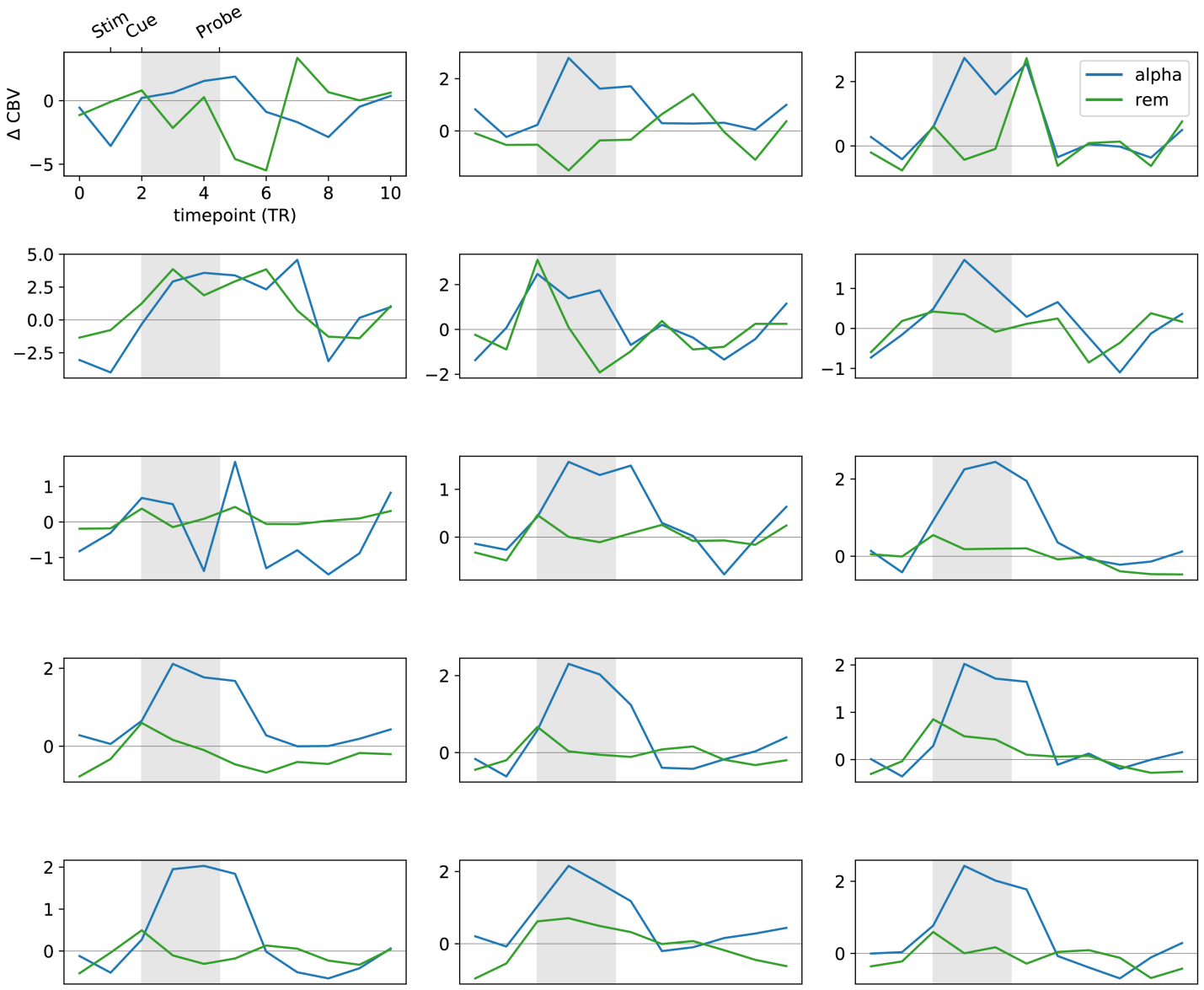
described in Methods subsection “Layering for sagittal protocol”. The position of the internal pyramidal layer is visible as a small bump (blue arrow). Just superior to it is the outer band of Bailager (black arrow), which is taken as a landmark to identify the approximate transition area between layer III (green arrow) and layer V (red arrow). This layer has an elevated myelin content, which comes along with reduced T_1 values in MRI contrast (ref. 68 in Methods); as such, this landmark is also visible in in-vivo T_1 and T_2* weighted profiles (see black arrows in panel B-D). The reason to show the histological slice in panel A) is to confirm that the transition area between later III and V is approximately 40-60% of the cortical thickness away from WM as visible in myelin-sensitive MRI data. This confirmation is helpful because the relative depth of cytoarchitectonic cortical layers may differ across brain areas. Based on this depth calibration, we believe it is appropriate to pool fMRI signal (E-F) and time courses from two independent subsets of voxels, superficial voxels (superficial ≈ 30-40% of the cortex) and deeper voxels (deeper ≈ 30-40%) as shown in Fig. S2A. This dip between cytoarchitectonic layers III and IV is used here as the border between so-called “superficial layers” and “deeper layers” which show different responses to the different functional task contrasts (E-F). (Data in panels A-D come from single subjects: data in panel A are from the Ding Atlas, which is based on a single adult female brain, and data in panels B-F come from the same subject in the present study. Error bars in panel A-D refer to the standard error across columns within the area of interest [without layer smoothing]. In panel A this represents the variability across a distance of 2 mm sampled at 7270-8949 points. In panel B, this represents the variability of across a distance of 17 mm sampled at 29-45 points.) The data in panel E-F) refer to the manipulation (alphabetize) versus. remember runs. 10 trial of each condition where acquired per run. The data shown here are extracted after averaging across 4 runs with (20 trials each). This totals to 80 trials. The green contrast refers to the time average across the delay period, whereas the red contrast refers to the time average across the response period (Fig. 1b). As in Fig. 4, the green contrast refers to the signal difference of alphabetizing minus remember trials ($(S_{alpha}-S_{remember})/S_{rest}$) and the red refers to the signal difference of the response period minus the rest period ($(S_{response}-S_{rest})/S_{rest}$) taken from the inter stimulus interval. For more subjects, see Fig. 4. The uncertainty area in panel E-F refers to the standard error of the mean across the ten respective trials (upon averaging across four runs).



Supplementary Figure 5

Activity timecourses and statistical contrasts for two layers and all four trial types using the BOLD contrast

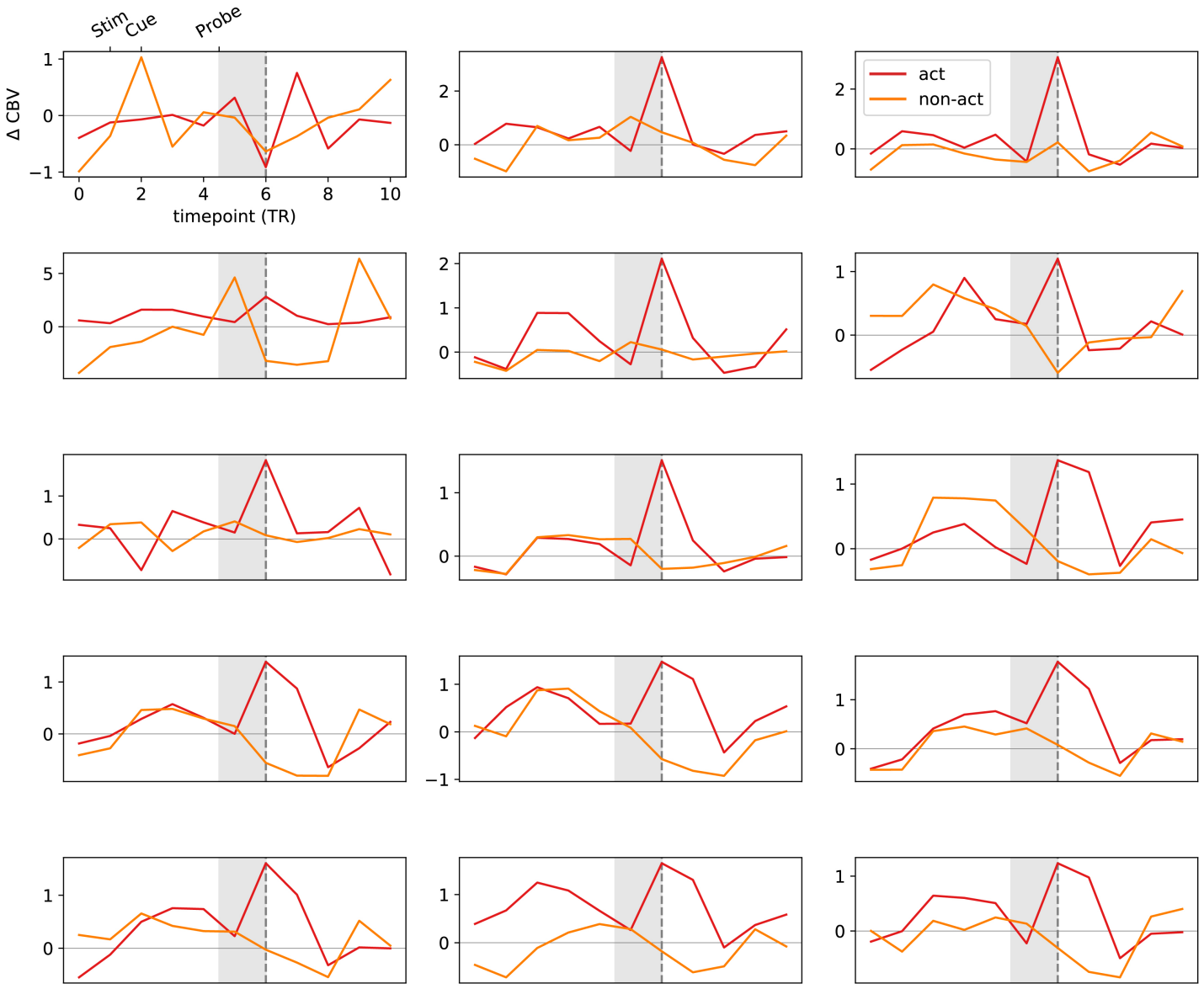
(A) Left panel: mean BOLD percent signal change in superficial layers (top) and deeper layers (bottom) for the first contrast, which consisted of trial types manipulation ('alpha') and maintenance ('rem'). Right panel: mean BOLD percent signal change in superficial layers (top) and deeper layers (bottom) for the second contrast, which consisted of trial types action ('act') and non-action ('non-act'). Lines represent mean and shaded area represents 95 percent confidence intervals for the mean (determined via bootstrapping with 1,000 iterations) across n = 15 sessions (13 unique subjects). While BOLD is more sensitive than VASO, it is less spatially specific. Discrepancies between BOLD timecourses shown here and VASO timecourses shown in Fig. 2—for example, it appears from BOLD that superficial-layer activity is higher at the time of the response in action versus non-action trials (panel A, top right), but this difference is absent in VASO (see Fig. 2a, top right)—are likely due to asymmetrical layer cross-talk, in which signal from deeper layers "leaks" into superficial layers because of draining-vein artifacts (but not the other way around). Thus BOLD results must be interpreted with caution. (B) Top: Superficial-layer BOLD activity during maintenance ('rem') trials subtracted from activity during manipulation ('alpha') trials [purple line]. The largest difference can be seen during the delay period. Bottom: Deeper-layer BOLD activity during non-action ('non-act') trials subtracted from activity during action ('act') trials [teal line]. The largest difference can be seen during the response period. Lines represent mean and shaded area represents 95 percent confidence intervals for the mean (determined via bootstrapping with 1,000 iterations) across n = 15 sessions (13 unique subjects).



Supplementary Figure 6

Average VASO timecourses for each individual subject-session for manipulation and maintenance in superficial layers

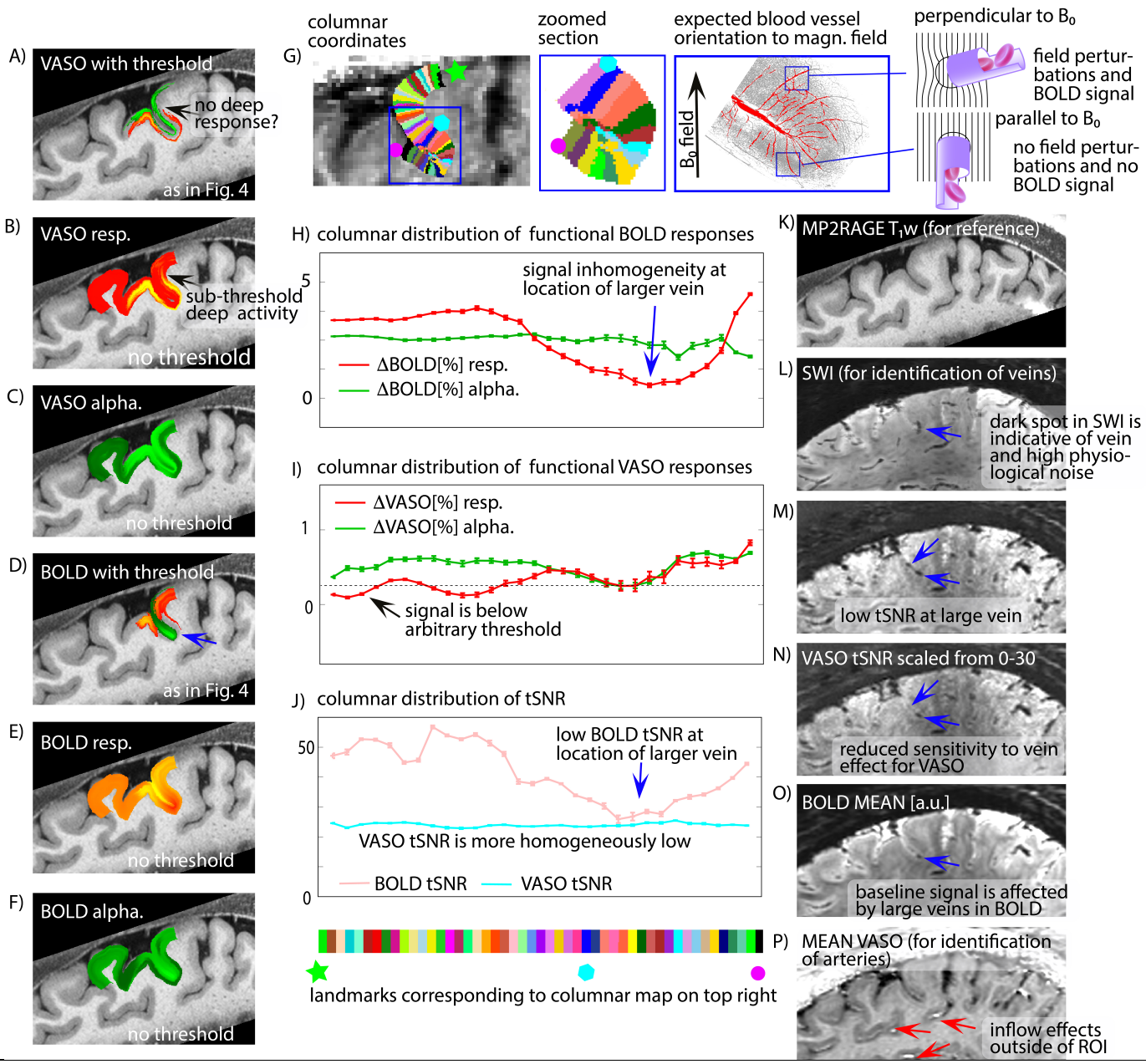
Data from n = 15 individual subject-sessions are shown. In the majority of subjects, it can be clearly seen that activity is higher for manipulation ('alpha', blue lines) over maintenance ('rem', green lines) specifically during the delay period (gray shading).



Supplementary Figure 7

Average VASO timecourses for each individual subject-session for action and non-action in deeper layers

Data from n = 15 individual subject-sessions are shown. In the majority of subjects, an increase in activity can be clearly seen at the timepoint when response-related signal would be expected to peak (gray dashed line) in action ('act', red lines) but not non-action ('non-act', orange lines) trials. This peak occurs at the end of the response period (gray shading), or approximately 6 seconds after the appearance of the probe (accounting for behavioral and hemodynamic delay).



Supplementary Figure 8

Discussion of columnar variance

Layer-dependent activity maps in Fig. 4 exhibit some variance across the columnar dimension. As such, there is no response activity for VASO at the location of the green arrow in panel A and there is no response activity at the location of the blue arrow in panel D). Panels B-C) and E-F) depict how the layer-dependent activity maps depend on the arbitrary threshold chosen. It can be seen in panel A) that here is indeed sub-threshold VASO response activity at the location of the green arrow. Panel G) depicts columnar coordinates that are used to unfold the fMRI signal along the cortical ribbon. The estimation of the columnar coordinates and the unfolding is done in LAYNII with the program LN_COLUMNAR_DIST. (https://layerfmri.com/quick-example-of-cortical-unfolding-in-laynii/). Note that the batch of cortex investigated here is highly folded with respect to the external magnetic field. This means that the BOLD signal change can be substantially variable dependent on the columnar position along the sulcus (refs. 63, 64 in Methods). At the location of the blue arrow in panels J,L-O, it can be seen that at the location of larger draining veins the BOLD sensitivity is reduced and the functional responses exhibit larger inhomogeneities.