Multiple face- and limb-selective activations are consistent over a span of three years, voxel sizes, and visualizations

We scanned subjects in the same localizer experiment longitudinally while varying the voxel size and visualization to examine whether the organization we report is visible with other resolutions and visualizations. To relate the data we show in **Figure 6** to volume-based and inplane visualizations, as well as to data acquired with other voxel sizes, we illustrate data from the same subject (S3) across visualizations.

Face-selective regions

In **Supplementary Figure 1**, we show visualizations of face-selective voxels from the statistical contrast of faces vs. all categories ($t > 3$, voxel level) over four different time periods: time 1 (1.5 mm isotropic voxels), time 2 (4 months later with 1.5 x 1.5 x 3mm voxels), time 3 (8 months later with a more standard resolution of 3.125 x 3.125 x 3mm), and time 4 (3 years later with 1.5 x 1.5 x 3mm voxels). We use an axial slice because this brain volume view allows the visualization of the ventral and lateral face-selective regions (mFus-faces, pFus-faces, and IOG-faces). As illustrated in both the left (top) and right (bottom) hemispheres, the anatomical loci of activations (indicated by arrows in **Supplementary Figure 1**) are consistent over time and resolutions. The reader should also take note that activations are restricted to gray matter. When not restricting to gray matter, fMRI activations extend into white matter, but do not form a single contiguous FFA (**Supplementary Figure 6a**). The only situation in our measurements that produces a single FFA is when the data are spatially smoothed on an unrestricted brain volume (**Supplementary Figure 6a,** top right). In relation to this latter point, when comparing the Talairach coordinates from the original FFA study (**Figure 3a**) to the

average of mFus- and pFus-faces in the present study, the average of these two regions are nearly identical to the original FFA coordinates (especially in the anterior-posterior dimension; **Supplementary Table 1**). Such a finding suggests that the combination of methodological factors (large functional voxels, spatial smoothing, unrestricted volume visualizations) used in the original study caused mFus-faces and pFus-faces to merge together.

Although the anatomical loci are consistent across functional resolutions and time, the reader should be aware of issues with fMRI measurements affecting the definition of face- and body part-selective regions. For example, IOG-faces is present in both hemispheres in all scans, except time 3 where we did not detect a face-selective activation on the right IOG (**Supplementary Figure 1a**, bottom, third column). We draw attention to this because other researchers may encounter a situation of a missing activation. In this case, the lack of the right IOG activation is likely a consequence of signal dropout from a susceptibility artifact produced by the transverse sinus, which exacerbates with larger voxels (**Supplementary Figure 5**; Winawer et al. 2010). Since our procedure of labeling activations is based on anatomical location and relation to other activations, the remaining fusiform activations are labeled accordingly as mFus-faces and pFus-faces in this scan with no IOG-faces. (Note that trying to fit these data to an inaccurate template of a single 'FFA' and a single 'OFA' would have resulted in inconsistent labeling of face-selective activations in this right hemisphere across scans. In times 1, 2, and 4, the fusiform activations would have been merged together and labeled the FFA (as in other studies; **Figure 3**), and the remaining IOG labeled as the OFA,

whereas in time 3, the fusiform activations would have been separated into FFA and OFA due to the absence of an activation on the IOG).

Another factor that may influence the apparent organization is the statistical contrast used to define the functional activations. To address this concern, we: (1) generated separate contrast maps comparing faces to each of the non-face categories in our original study at time 1 (**Supplementary Figure 2**) and 2) scanned three of the original subjects on a new experiment three years later during which they viewed different stimuli (legs, torsos, headless bodies, hands, houses, chairs, and faces) than those used in the original study and generated face-selective contrast maps using these different face and non-face stimuli (**Supplementary Figures 1b** and **2**).

In each subject, we first defined face-selective voxels using a contrast in which all stimuli changed (faces > hands, torsos, legs, headless bodies, and houses, $t > 3$, voxel-level). As illustrated in **Supplementary Figure 1b**, even with this completely different comparison, we replicate the prior result of three face-selective regions located anatomically on the IOG, pFus, and mFus, respectively. Second, when visualizing contrasts comparing faces to each of the nonface categories at either time 1 (**Supplementary Figure 2**, top) or time 3 (**Supplementary Figures 1b** and **Supplementary Figure 2**, bottom), we again replicate the prior results of three face-selective activations on distinct anatomical loci (two on the fusiform and one on the IOG). Nevertheless, varying the contrast does not yield identical maps. There is some variability across contrasts in the distribution, extent, and selectivity (*t*-value) of face-selective voxels, as one would expect when changing all stimuli.

Taken together, these analyses indicate that mFus-, pFus-, and IOG-faces are stable across a span of three years, resolutions, and are not contingent on the view used to visualize the data. However, spatial smoothing functional data on the original brain volume can affect the apparent organization because it can make distinct activations appear as a single cluster (where this effect is exacerbated when not restricting data to gray matter; **Supplementary Figure 6a**; **Figure 8**). Further, researchers should take caution when a functional activation is not visible in a given subject because it may be a consequence of a poor measurement rather than a lack of activation.

Limb-selective regions

Implementing the same approach as with the face-selective activations in **Supplementary Figure 1**, we first illustrate example longitudinal measurements of limb-selective activations on the volume in **Supplementary Figure 3** (see Weiner and Grill-Spector 2011 for additional examples using standard resolution fMRI as well as left hemisphere data). As a guide to readers using brain volume-based visualizations, the sagittal view allows the clearest illustration of the limb-selective OTS in conjunction with the three LOTC limb-selective activations (LOS/MOG, ITG, and MTG). As illustrated in **Supplementary Figure 3**, the series of limb-selective regions forms a crescent organization in LOTC, surrounding hMT+ (**Supplementary Figure 3**, blue) and an additional limb-selective activation is present on the OTS. This pattern of response is reproducible over a span of three years as measured with 1.5mm isotropic voxels (top row) and 1.5 x 1.5 x 3mm voxels (bottom row).

We next tested how changing the statistical contrast affects the organization of limb-selective voxels. In **Supplementary Figure 4a** (right), we compare maps resulting from contrasting different body parts (headless bodies, hands, legs, or torsos) to faces, houses, and chairs (different exemplars than those used at time 1) to those ROIs defined in the original study (**Supplementary Figure 4a**, left; body, body part, and chair images were acquired from the Zohary and Downing labs and were used in the recent Downing et al. 2007 and Orlov et al. 2010 studies). First, irrespective of the statistical contrast used for localization, body part-selective activations surround and largely do not extend into hMT+. Second, there is a correspondence between the location of the original limb-selective activations (shown in the colored outlines in **Supplementary Figure 4a**) and the location of voxels that prefer different parts of the body compared to a baseline of faces, chairs, and houses. That is, voxels that are selective for both the whole body and specific body parts overlap the original ROIs. Third, there are additional voxels that arise from these new contrasts as one would expect when changing both the body part and control exemplars (Orlov et al. 2010). For example, the hand-selective and body-selective contrasts overlap with the limb-selective OTS defined at time 1, while the torso- and leg-selective contrasts do not, suggesting that the OTS may illustrate a preference for different body part stimuli as previously reported (Taylor et al. 2007). Furthermore, when more specific contrasts are used, such as hands vs. bodies, torsos, legs, faces, chairs, and houses, we extend prior results reporting focal activations selective for images of hands on the MTG (Supplementary Figure 4 from Weiner and Grill-Spector 2011). However, when contrasting activations to specific body parts relative to non-body images (**Supplementary Figure 4a**, right), a constellation of activations surrounding hMT+ is

produced rather than a singular focal activation. Finally, using the original ROIs defined based on a limb-selective contrast, we extracted timecourses from the new experiment and calculated selectivity indices 3 years later (average voxel *t*-value) within each ROI for different body parts relative to faces, houses, and chairs. As illustrated in **Supplementary Figure 4b,** each ROI has a positive selectivity to body parts. However, preference for images of specific body parts varies across ROIs, replicating recent findings (Bracci et al. 2010; Orlov et al. 2010).

Taken together, a consistent factor in our measurements is that voxels selective for images of the human body, body parts, and limbs in LOTC largely surround and do not encroach into the center of hMT+. That is, our data do not support a contiguous EBA highly overlapping hMT+. Further, the observed organization in **Figure 6** with multiple limb-selective LOTC regions rather than one EBA is consistent over a span of three years, statistical contrasts, and surface- and volume-based visualizations. Finally, similar to the methodological point raised in the prior section, factors such as spatial smoothing and restricting data to gray matter have a profound effect on determining the spatial organization of these activations (**Supplementary Figure 6b**). Specifically, when comparing common EBA coordinates in stereotaxic space to the average of our LOS and MTG coordinates (**Supplementary Table 1**), the coordinates are nearly identical. Such a finding indicates that though the center of the MTG- and LOS-limbs activations in our measurements are nearly 15 mm apart, when using large voxels, spatial smoothing, and unrestricted volume visualizations, these regions inaccurately merge together to form one EBA. Furthermore, when comparing these average Talairach coordinates of the LOS- and MTG-limbs to the average coordinates of hMT+, they are also nearly identical. These

analyses suggest that prior studies reporting significant overlap between hMT+ and the EBA (Downing et al. 2007; Peelen et al. 2006) are likely a direct result of methodological factors such as large functional voxels and spatial smoothing on the brain volume that inaccurately merged these regions together.

Supplementary Table 1. Comparison of averaged mFus and pFus coordinates to FFA coordinates and averaged MTG and LOS coordinates to EBA and hMT+ coordinates

Face-selective	*Right*			*Left*		
	x	y	z	x	y	z
Average of mFus- and pFus-faces	35(4)	-54(6)	-16(2)	-35 (4)	-57(4)	-17(4)
FFA (Kanwisher et al. 1997)	40	-55	-10			
Limb-selective						
Ave of LOS- and MTG-limbs	43(3)	-66(3)	3(4)	-44(3)	-68(6)	4(3)
EBA (Taylor and Downing)	47	-67	-1	-48	-68	3
hMT+	40(3)	-66(3)	2(4)	-42(4)	-68(4)	3(4)

The original EBA study (Downing et al. 2001) included MNI coordinates, which slightly vary compared to Talairach coordinates. Therefore, we used the Talairach coordinates provided in a recent review by Taylor and Downing 2011). Numbers reflect mean (std).

REFERENCES

Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *J Neurophysiol, 103*(6), 3389-3397.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science, 293*(5539), 2470-2473.

Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *J Neurosci, 27*(1), 226-233.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci, 17*(11), 4302-4311.

Orlov, T., Makin, T. R., & Zohary, E. (2010). Topographic representation of the human body in the occipitotemporal cortex. *Neuron, 68*(3), 586-600.

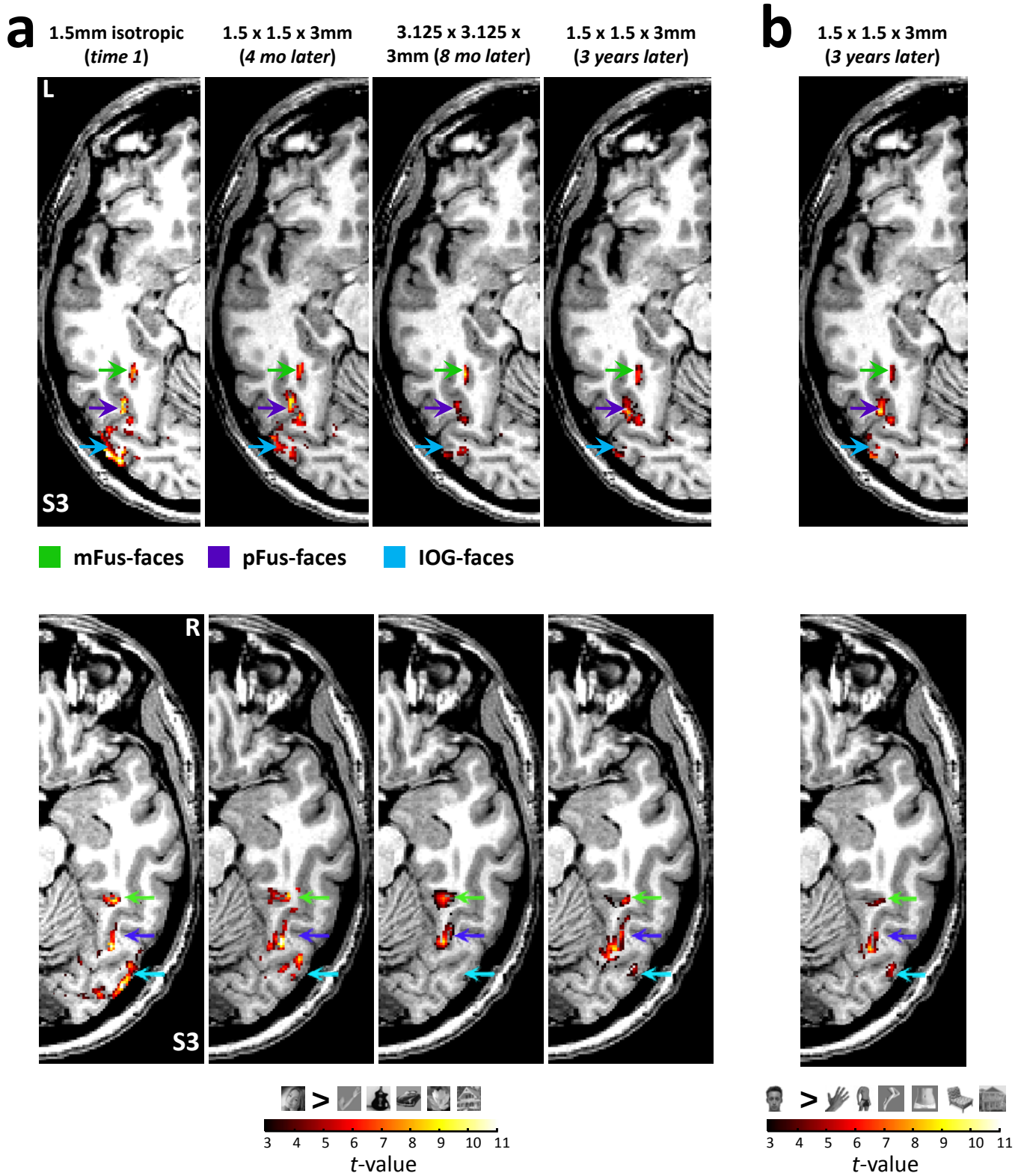
Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron, 49*(6), 815-822.

Taylor, J. C., & Downing, P. E. (2011). Division of Labor between Lateral and Ventral Extrastriate Representations of Faces, Bodies, and Objects. *J Cogn Neurosci*.

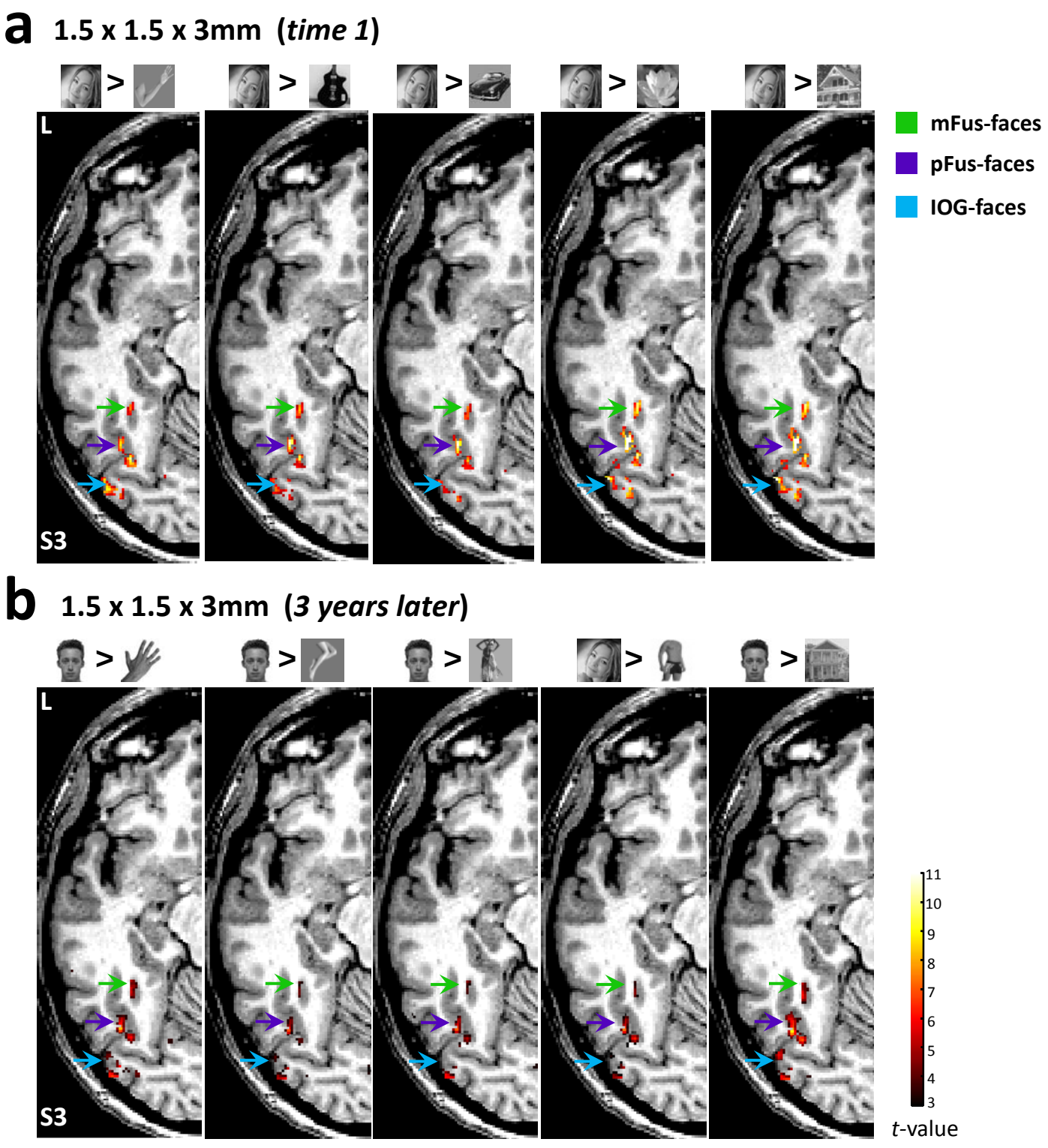
Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *J Neurophysiol, 98*(3), 1626-1633.

Weiner, K. S., & Grill-Spector, K. (2011). Not one extrastriate body area: Using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *Neuroimage, 56*, 2183-2199.

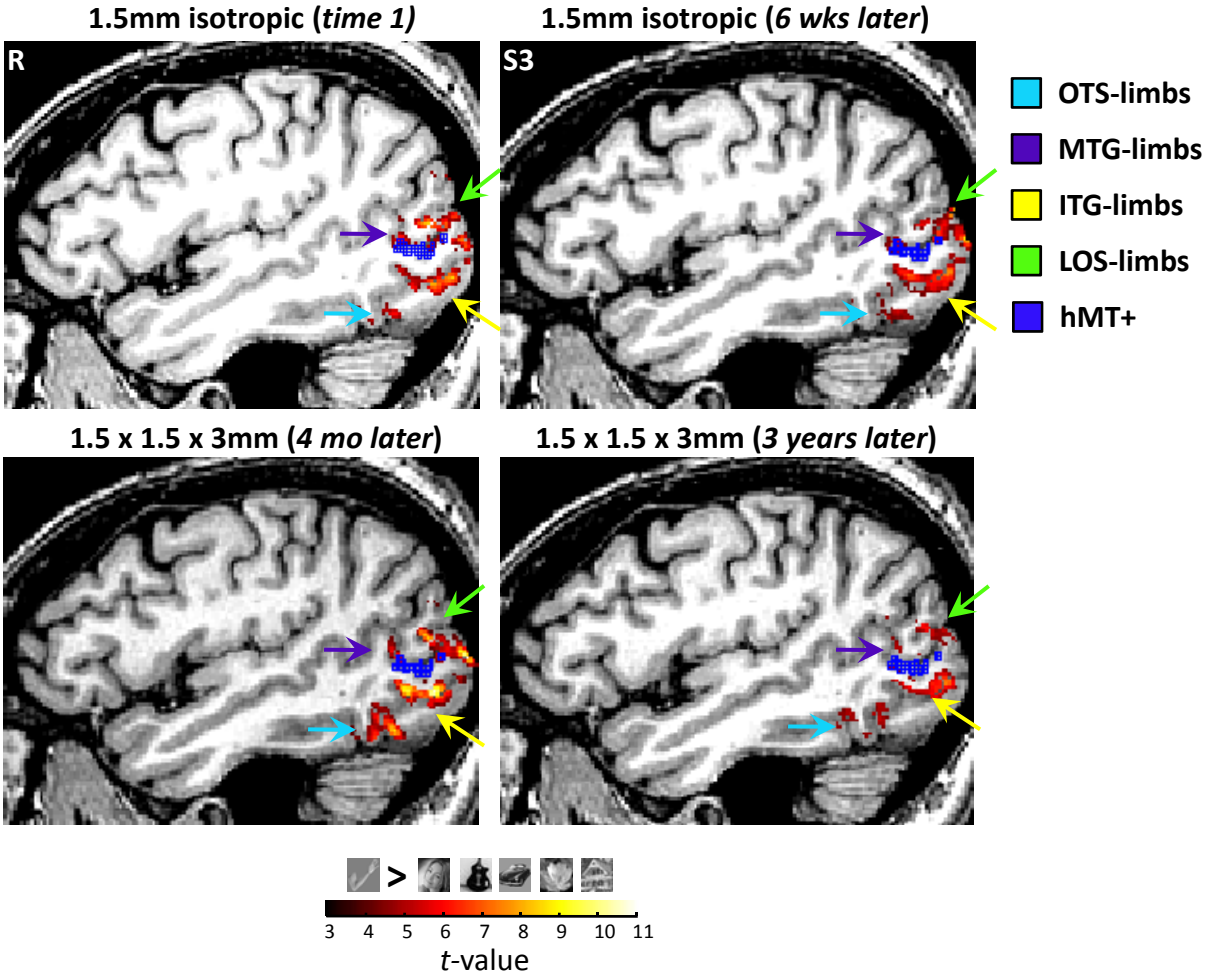
Winawer, J., Horiguchi, H., Sayres, R. A., Amano, K., & Wandell, B. A. (2010). Mapping hV4 and ventral occipital cortex: the venous eclipse. *J Vis, 10*(5), 1.



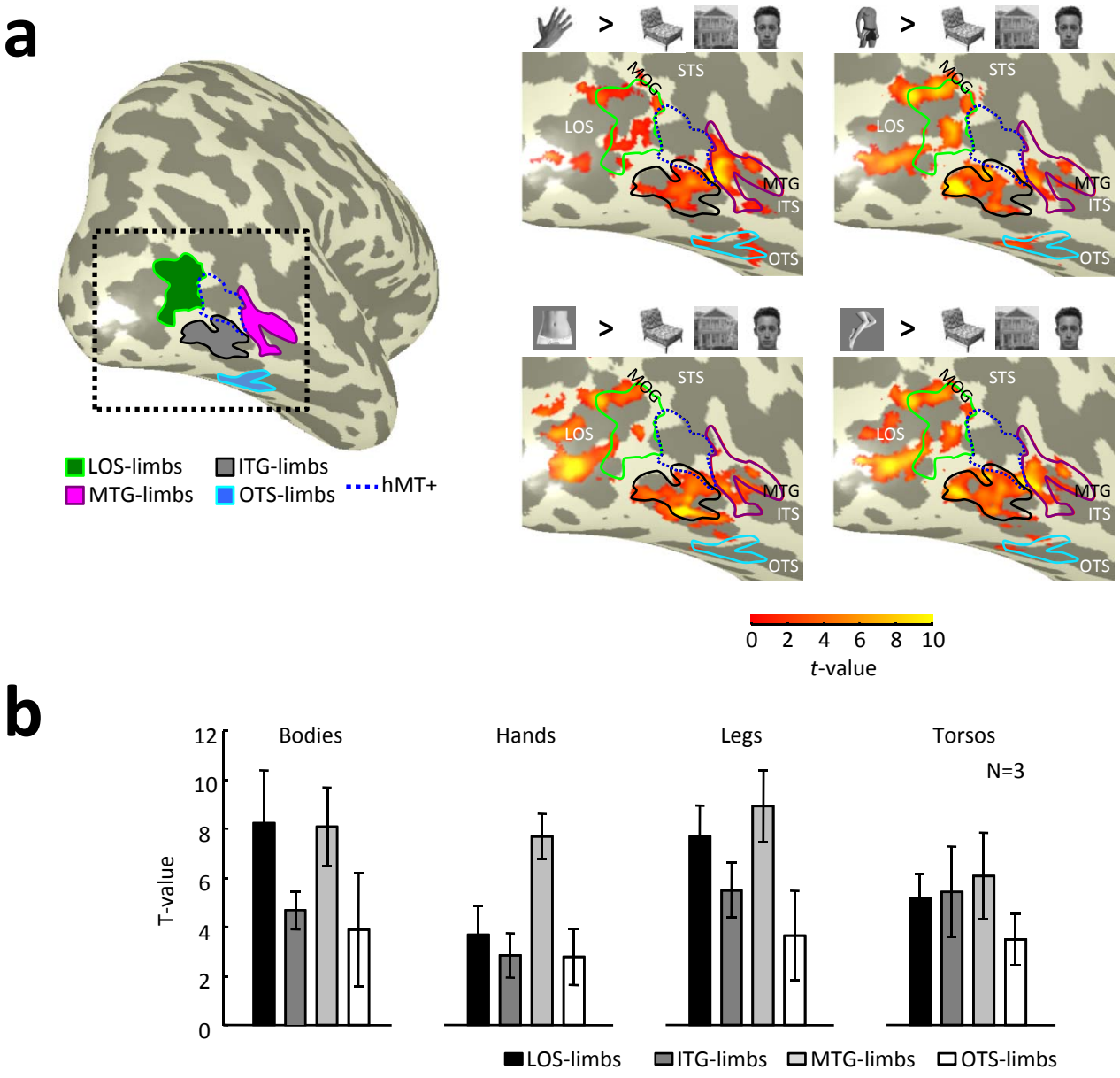
Supplementary Figure 1. The anatomical foci of bilateral face-selective regions is consistent over a span of three years and different scanning resolutions. (a) Statistical contrast of BOLD responses to images of faces vs. images of limbs, cars, houses, guitars, and flowers ($t > 3$, voxel level, no spatial smoothing) from subject S3 on the same axial slice across four different time points over a span of three years. Activations are restricted to gray matter. From left to right: Time 1 (1.5mm isotropic voxels), Time 2 (4 months later; 1.5 x 1.5 x 3mm), Time 3 (8 months later; 3.125 x 3.125 x 3mm), and Time 4 (3 years later; 1.5 x 1.5 x 3mm). (b) Face-selective responses in subject S3 localized three years later using different face images as well as different control images (hands, headless bodies, legs, torsos, chairs, and houses). *Top*: Left hemisphere. *Bottom*: Right hemisphere. Activation foci are anatomically consistent across time, scanning resolution, and statistical contrast. Arrows indicate the location of the functional region labeled with its anatomical location. *Green*: mid-fusiform sulcus, mFus-faces; *Purple*: posterior fusiform gyrus, pFus-faces; *Cyan*: inferior occipital gyrus, IOG-faces.



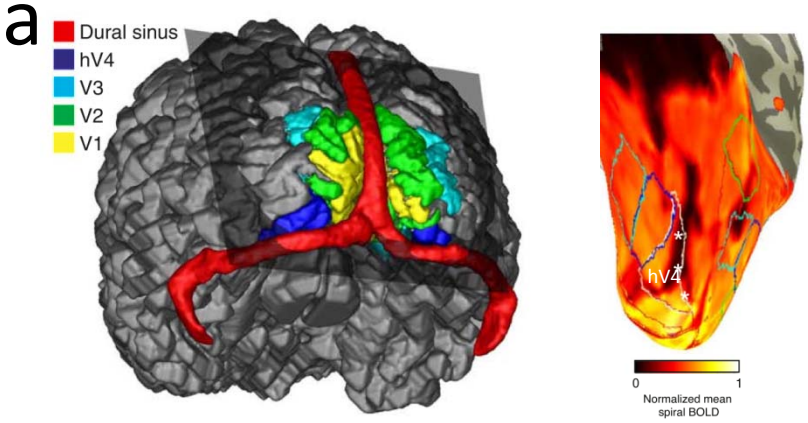
Supplementary Figure 2. Anatomical foci of functional activations are preserved when using different statistical contrasts across time and exemplars. Statistical contrast of BOLD responses to images of faces vs. control images ($t > 3$, voxel level, no spatial smoothing) from subject S3 on the same axial slice across five different statistical contrasts. Activations are restricted to gray matter. (a) Face-selective responses at time 1 with a resolution of 1.5 x 1.5 x 3mm. Contrasts illustrated above each image. (b) Face-selective responses 3 years later with a resolution of 1.5 x 1.5 x 3mm. Body part images in the first three columns are stimuli acquired from Orlov and colleagues, while the headless bodies in the fourth column are those from Downing and colleagues. Third and fourth columns from the left are both contrasts of faces > bodies using different face and body images. Arrows indicate the location of the functional activation labeled with its anatomical location. For each contrast in (a) and (b), note that we find face-selective activations consistently in anatomical foci on the mid-fusiform (mFus) sulcus, posterior fusiform (pFus) gyrus, and inferior occipital gyrus (IOG), though the exact pattern of the response across the voxels may differ for different contrasts and exemplars. *Green*: mFus-faces; *Purple*: pFus-faces; *Cyan*: IOG-faces.

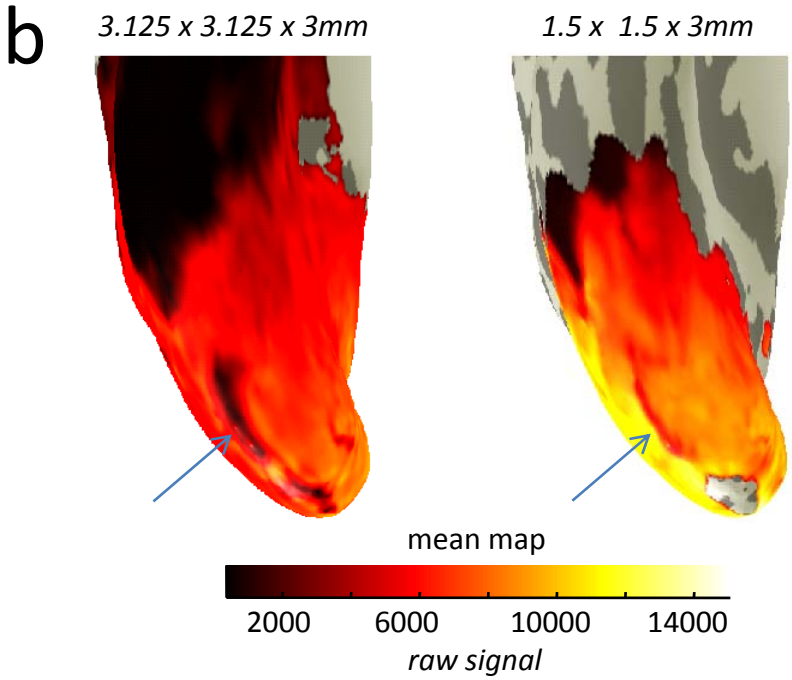


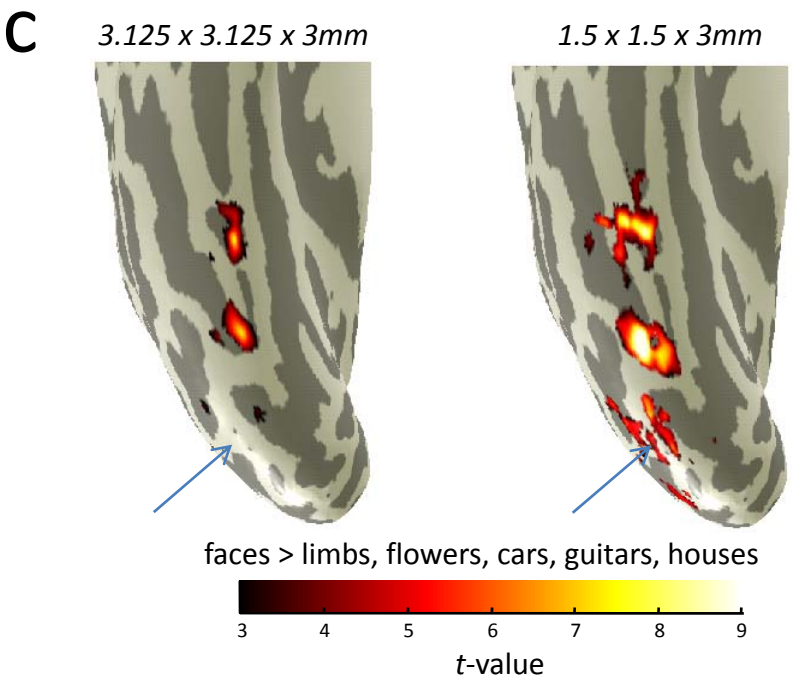
Supplementary Figure 3. The anatomical foci of lateral occipitotemporal (LOTC) limb-selective regions surrounding hMT+ are consistent over a span of three years. Statistical contrast of BOLD responses to images of limbs vs. images of faces, cars, houses, guitars, and flowers ($t > 3$, voxel level) relative to hMT+ (in blue) from subject S3 on an example sagittal slice from a high-resolution volume across four different time points over a span of three years. Each image represents the same statistical contrast acquired at a different time without spatial smoothing. Each row represents data acquired with the same functional resolution. Activations are restricted to gray matter. *Top,* left to right: Time 1 (1.5mm isotropic) and Time 2 (6 weeks later; 1.5mm isotropic); *Bottom*, left to right: Time 3 (8 months later; 1.5 x 1.5 x 3mm) and Time 4 (3 years later; 1.5 x 1.5 x 3mm). Similar to the longitudinal face-selective measurements illustrated in **Supplementary Figure 2**, the anatomical foci of LOTC limb-selective regions are consistent across time. Arrows indicate the location of the functional activation labeled with its anatomical location. *Cyan*: occipitotemporal sulcus, OTS-limbs; *Purple*: middle temporal gyrus, MTG-limbs; *Green*: lateral occipital sulcus, LOS-limbs; *Yellow*: inferior temporal gyrus, ITG-limbs. See Weiner and Grill-Spector 2011 for longitudinal measurements in the left hemisphere, as well as for a comparison with data acquired with larger functional voxels.



Supplementary Figure 4. Effect of varying the contrast on the identification of body part-selective voxels. Different body part-selective contrasts produce slightly different patterns of activation, however, the crescent organization surrounding hMT+ is consistent over time and contrasts**.** (a) *Left*: Lateral view of the inflated right hemisphere of subject S3 illustrating the limb-selective ROIs defined at time 1 (see Figure 6). *Right*: Zoomed view of LOTC indicated by the dotted black rectangle in the left panel showing activations to different contrasts in S3 three years later. hMT+ indicated in dotted blue outline, and limb-selective LOS/MOG, ITG, MTG, and OTS defined from time 1 outlined in green, black, magenta, and cyan, respectively. *Top (from left to right):* Hands > faces, houses, and chairs ($t > 3$, voxel level); Headless Bodies > faces, houses, chairs ($t > 3$, voxel level); *Bottom (from left to right):* Torsos > faces, houses, and chairs ($t > 3$, voxel level); Legs > faces, houses, chairs ($t > 3$, voxel level). Headless bodies, torsos, legs, and chair stimuli were from Downing et al. 2007, Taylor et al. 2007, and Orlov et al., 2010. (b) Using the ROIs from time 1, the *t*-values for each body part stimulus were extracted relative to faces, houses, and chairs in 3 subjects. All regions illustrate substantial preference for body parts compared to control stimuli. Note that the different ROIs show a different profile of response, which both replicates Orlov et al., 2010 and Bracci et al., 2010, as well as extends these results by illustrating that these body part-selective voxels radiate around hMT+.

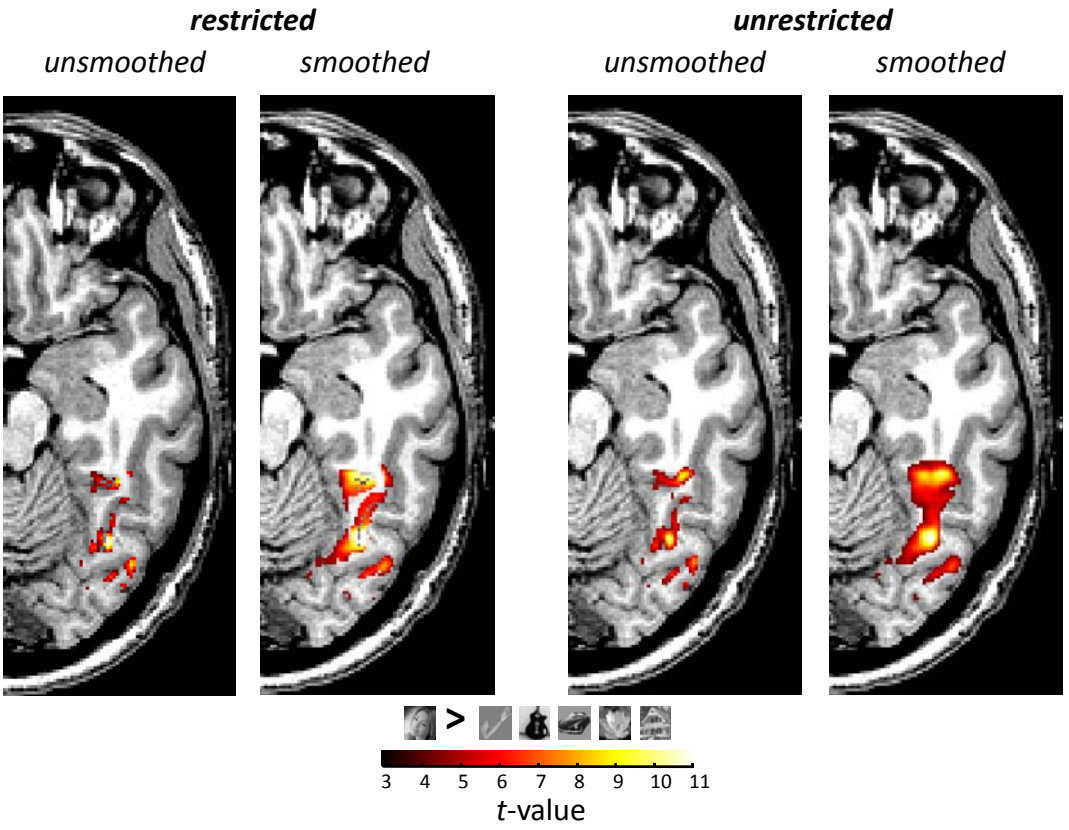






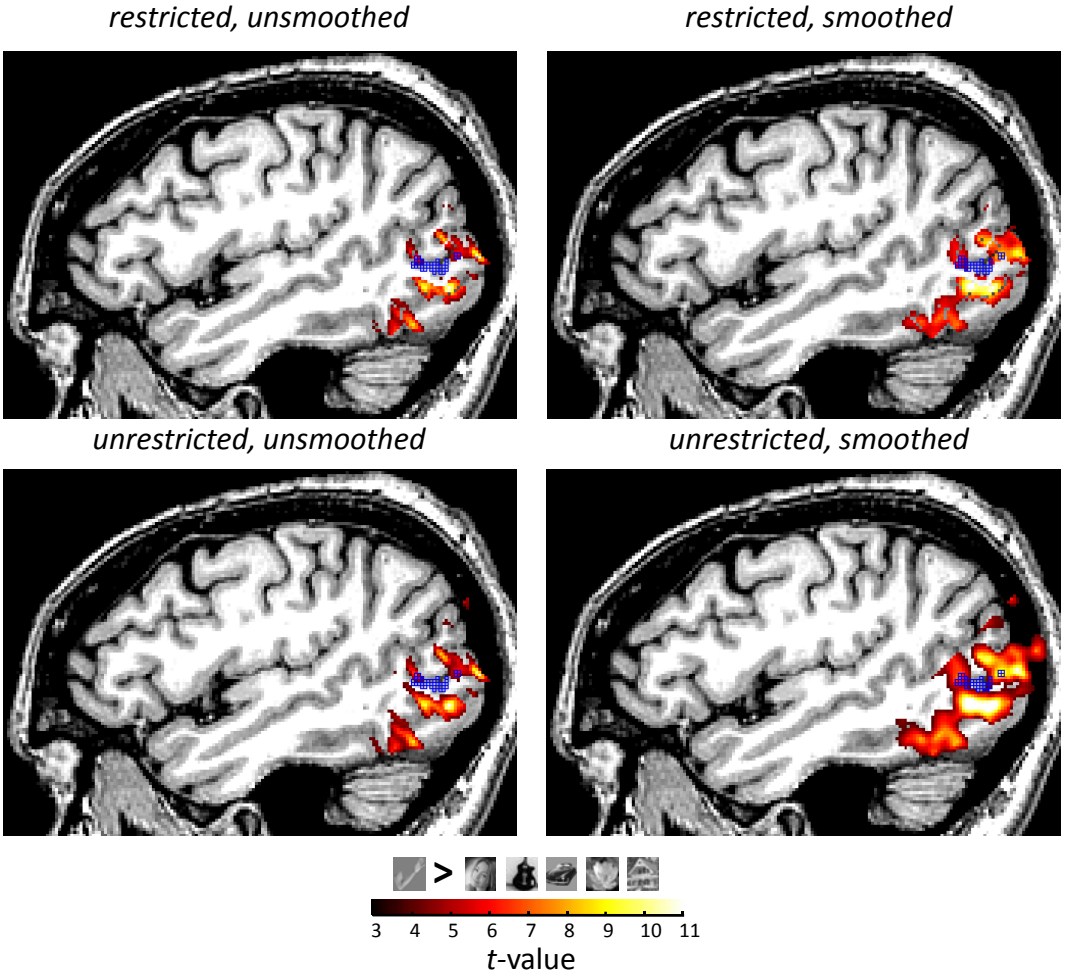
Supplementary Figure 5. Smaller functional voxels improve measurements near the transverse sinus. Smaller voxels improve measurements located near the signal dropout (dark colors in a and b) caused by the transverse sinus (Winawer et al., 2010). (a) *Left*: Anatomical location of the transverse sinus (in red) relative to early visual field maps (V1-hV4). The sinus typically runs up against the posterior portion of the lateral fusiform gyrus extending into the inferior occipital gyrus, which not only affects the measurements of visual field maps in VTC and LOTC, but also nearby face-selective and limb-selective regions. *Right*: Inflated cortical surface from one example subject illustrating the effect of the transverse sinus on the mean map of the fMRI signal. Dark regions indicate signal drop out. Colored outlines indicate visual field maps where white asterisks identify the lateral edge of the hV4 map within the dropout caused by the transverse sinus. Image adapted from Winawer et al., 2010. (b) The inflated cortical right hemisphere of one example subject with mean map (in raw scanner units) projected to the surface with two different functional measurements (3.125 x 3.125 x 3mm and 1.5 x 1.5 x 3mm). Light blue arrows identify the location of the transverse sinus. Notice that with smaller voxels (right), the map is brighter (as compared to larger voxels in the left image) in the posterior fusiform and extending into LOTC and also, that the measurements are not as affected by the signal dropout in the transverse sinus (which is much less dark compared to the image on the left with larger voxels). (b) Face-selective regions in the same subject in each of these scans. Notice that with the improved mean signal with smaller voxels illustrated in *b (*right column*)*, this produces more face-selective voxels in LOTC in c (right column; thresholded with the same *t*-value) especially around the location of the transverse sinus. With the larger voxels and larger drop-out (left column in c), there are barely any face-selective voxels around the IOG (left column in c). Determining the conditions that produce the most reliable measurements are an active topic of current research. It is likely that methodological factors beginning with raw data acquisition influence functional measurements much more than the statistical contrast.

a



Supplemental Figure 6a. Spatial smoothing is more detrimental for precise localization than restricting measurements to gray matter. (a) Statistical contrast of BOLD responses to images of faces vs. images of limbs, cars, houses, guitars, and flowers ($t > 3$, voxel level) from subject S3 on the same axial slice across four different conditions (from left to right): (1) restricted to gray matter and unsmoothed, (2) restricted to gray matter and smoothed with a 4mm kernel, (3) unrestricted to gray matter and unsmoothed, and (4) unrestricted to gray matter and smoothed with a 4mm kernel. Note that the difference between the activation patterns in the restricted and unsmoothed condition (first image) with the unrestricted and unsmoothed condition (third image) is less extensive than with either of the smoothed conditions (2nd and 4th images, respectively), indicating that spatial smoothing is more detrimental for precise localization than restricting to gray matter.

b



Supplemental Figure 6b. Spatial smoothing is more detrimental for precise localization than restricting measurements to gray matter. (b) Statistical contrast of BOLD responses to images of limbs vs. images of faces, cars, houses, guitars, and flowers ($t > 3$, voxel level) relative to hMT+ (in blue) from subject S3 on an example sagittal slice across four different conditions. *Top row:* (1) Restricted to gray matter and unsmoothed (*left)* and (2) restricted to gray matter and smoothed with a 4mm kernel (right). *Bottom row:* (1) Unrestricted to gray matter and unsmoothed (*left*) and (2) unrestricted to gray matter and smoothed with a 4mm kernel (*right).* hMT+ was defined from unsmoothed data restricted to gray matter and is used as an anchor point across the four conditions. Functional scans in both (a) and (b) acquired with 1.5 x 1.5 x 3mm voxels. Note that again, the most imprecise localization occurs in the 4th condition with smoothed data on a volume unrestricted to gray matter.