# Supporting Online Material for

## A Cortical Region Consisting Entirely of Face-Selective Cells

Doris Y. Tsao,* Winrich A. Freiwald, Roger B. H. Tootell, Margaret S. Livingstone

*To whom correspondence should be addressed. E-mail: doris@nmr.mgh.harvard.edu

**This PDF file includes:**

Materials and Methods

SOM Text

Figs. S1 to S7

References

**Supporting Online Material**

**Materials and Methods**

All procedures used in this study conformed to local and NIH guidelines. Two male rhesus macaques (M1: 7 kg, 6 years old, M2: 4 kg, 3 years old) were implanted with Ultem headposts, trained via standard operant conditioning techniques to maintain fixation on a spot for a juice reward, and then scanned in a 3 T horizontal bore magnet to identify face-selective regions. Both animals had a prominent face-selective region located ~ 6 mm anterior to the ear bars (the "middle face patch") which was targeted for single-unit recordings. See (*1, 2*) for details on fMRI procedures, and (*3*) for details on single-unit recording procedures.

**Alert monkey fMRI**. All fMRI procedures for identifying the face patches were identical to those described in (*4*) with one exception: we obtained a field map during each session and used a multi-echo sequence (EPI; TR=3 sec; TE = 30 ms; 64 x 64 matrix; 1.25 x 1.25 x 1.25 mm voxels; 21 coronal slices) which allowed subsequent epi undistortion based on the fieldmap (*5, 6*). The use of this sequence facilitated registration of functional data to high resolution anatomical data. MION contrast agent was used in both animals to improve signal/noise ratio. In total, we obtained 142,020 functional volumes during 30 scan sessions in monkey M1 and 22,333 volumes during 10 scan sessions in monkey M2 (each scan session was ~ 3 hours long). Face selective regions were identified as those regions responding significantly more to faces than to bodies, fruits, gadgets, hands, and scrambled patterns. Monkey M1 has been scanned repeatedly for face patches, and the cortical location of those patches has remained consistent over

the past four years (see Fig. S5 in (*4*)). In monkey M2, the location of the face patches was stable across the 10 scan sessions.

**Targeting a face patch for single-unit recording.** In each monkey, fMRI was used to locate the face patches, and the stereotaxic coordinates of the middle face patch were determined. Then a CILUX chamber was implanted stereotaxically, roughly aimed at the face patch. In a subsequent scan session, a high resolution anatomical scan was obtained under ketamine anesthesia (3D-MPRAGE; 256 x 256 matrix; 128 slices; 0.5 x 0.5 x 0.5 mm$^3$ voxel size). During this scan, a set of oil-filled markers were placed inside a recording grid fixed inside the chamber, allowing us to determine which grid holes would precisely target the center of the fMRI face patch.

**Single-unit recording.** We recorded extracellularly with fine electropolished tungsten electrodes coated with vinyl lacquer. Extracellular signals were amplified, bandpass filtered (500Hz-2 kHz) and fed into a dual-window discriminator. The spike train was recorded at 1 ms resolution. Only well-isolated single units that showed a refractory period were studied. LFP responses were recorded by filtering between 3 and 90 or 0.7 and 170 Hz. Subsequently, the LFP responses were further filtered with a notch filter removing frequencies between 70 to 90 Hz. Eye position was monitored with an ISCAN primate infrared eye tracking system.

**Visual stimuli.** During fMRI scanning, visual stimuli were projected with an LCD projector (1024 x 768 pixels, 75 Hz refresh rate), onto a screen 53 cm in front of the monkey's eyes. The display spanned 28º laterally and 21º vertically. Visual stimuli were generated in MATLAB using the Psychophysics Toolbox extensions. The size of the images for the fMRI experiments was 11° x 11°.

For single-unit recordings, all visual stimuli were written using Microsoft Visual C/C++, and presented at a 60 Hz monitor refresh rate and 640 x 480 resolution on a BARCO

ICD321 PLUS monitor.  The monitor was positioned 53 cm in front of the monkey's eyes.  Stimuli were 7° x 7° and presented in the center of the screen.

**Data analysis.**  Single-unit data were analyzed using custom programs written in MATLAB.  For both fMRI and single-unit experiments, only data points during which the monkey maintained fixation within a $\pm 2.5°$ window were used.  *Fig. 2*: Cells were considered visually responsive if they gave a response above one standard deviation of the baseline for 20 consecutive bins, starting at a time point less than 300 ms post-stimulus onset, where the baseline was defined as the mean response across all images over the first 50 ms.  For each visually responsive cell, we determined the response latency by computing the average response to each of the 6 categories between 100 and 300 ms; we took the latency as the time at which the response of the best category exceeded 3 standard deviations above baseline.  *Fig. 3:* The analyses in panel A and panel B (top and middle rows) were based on average activity over the stimulus duration (200 msec), beginning at the response latency of the cell; we did not consider the temporal structure of the response time course (which could provide additional information for identification).  A distance matrix was calculated for each of the 5 possible permutations of (test, template) trial identities.  The 5 distance matrices were averaged to yield the distance matrix shown in panel A, which was then used for the identification and categorization analysis.

**Supporting Online Text**

**Reduced response to inverted faces**

Compared to other objects, faces are much more easily recognized when upright, compared to when they are inverted (*7*), even though there is no difference in the distinctness of low-level features between upright and inverted faces. This is interpreted as evidence for holistic processing of faces in humans. When we compared single-unit responses to upright versus inverted versions of the same faces, we observed three types of effects (Fig. S7A): the response to inverted faces was 1) reduced, 2) delayed, or 3) more transient (with different combinations in different cells). These results are consistent with previous studies of the effect of face inversion on face cell responses (*8*). Fig. S7B shows the population response, averaged across 69 cells, to 64 upright faces and 64 inverted versions of the same faces. In the population average, all three effects can be seen; the average response magnitude to upright faces was 2.2 times as large as that to inverted faces. This is a larger effect than has been found in fMRI studies of the human FFA, where the reported ratio of upright to inverted face activation ranges from 1.0 (*9*) to 1.4 (*10*). This single-unit face-inversion effect suggests that face selectivity in the middle face patch is probably not due to low level image differences between faces and other objects (which are identical for upright and inverted faces).

**Fig. S1.** Examples of stimuli used to assess the selectivity of single units (Figs. 2, 3).

**Fig. S2.** Location of face patches in five monkeys. **(A)** Activation to human faces versus non-face objects (headless human bodies, fruits, and gadgets) in 5 monkeys. Functional activation is shown overlaid on raw EPI slices. MION contrast agent was used to improve signal/noise ratio. Numbers indicate the location of each coronal section anterior (+) or posterior (-) to the interaural line, in mm. Slices were shifted to obtain best possible alignment between the three animals. The most prominent and robust patch (the "middle face patch"), found in all five monkeys, was located ~ A6. The anterior patch actually consists of several distinct sub-patches, located inside the STS, on the inferior temporal gyrus and the lateral bank of the anterior middle temporal sulcus. The white arrows in the first two panels indicate the patch in which electrophysiology was performed in monkeys M1 and M2. **(B)** Time course of fMRI activation within the middle face patch in monkeys M1-M5. Activation to three cycles of (faces, grid-scrambled faces, objects, grid-scrambled objects) within face-selective voxels located between A4 and A9 were averaged. The ratio of activation to faces versus objects was 3.1, 2.2, 11.3, 2.6, and 1.7 in monkeys M1-M5, respectively.

**Fig. S3.** Face-selective units in the middle face patch respond to a wide variety of face stimuli. **(A)** Responses of a single-unit from monkey M1 to the 96 screening stimuli (left), 45 monkey faces (middle), and 112 highly impoverished cartoon faces (right). Examples of the stimuli are shown below each response map. The animal had not seen the monkey or cartoon faces prior to the experiment. All responses are plotted with the same color scale. **(B)** Bar graph showing response to 64 human faces, averaged across first 49 face-selective units tested. Faces 1-16 were the highly familiar faces from the screening set, while faces 17-64 were seen only during this experiment, and were therefore unfamiliar. Familiarity did not produce the face selectivity observed; responses

to familiar faces were actually slightly smaller on average than responses to novel faces (t-test, $p < 1.9 * 10^{-6}$).

**Fig. S4.** Evoked local field potentials in the middle face patch are face selective. We recorded the LFP by low-pass filtering the same signal that we used to record spiking activity. **(A)** Average evoked LFP responses to the six stimulus categories recorded at three different sites in monkeys M1 (left) and M2 (right). The red trace is the LFP response to faces. At almost all recording sites in the face patch, the LFP showed an initial non-stimulus-specific component at 100 ms, followed by two face-selective components, at 130 and 200 ms. Evoked LFP traces are shown over two stimulus cycles. Since image order was randomized, differences between the 6 traces in the response to the second stimulus cycle are presumably due to adaptation (apparent in sites 2 and 3 in monkey M1, and site 3 in monkey M2). **(B)** Average evoked LFP responses, sorted by category, across 70 recording sites in monkey M1 and 63 recording sites in monkey M2. **(C)** Average single-unit responses to the same categories (see also Fig. S6), obtained by averaging single-unit responses from all visually responsive neurons recorded in each monkey. The average single-unit response to faces shows a peak in firing at a time corresponding to the first trough in the LFP and shows a decrement in firing rate at a time corresponding to the second trough in the LFP. If we assume the first trough corresponds to feedforward inputs and the second to feedback and local inputs, then the feedback/local inputs must contain more inhibition than the feedforward inputs.

**Fig. S5.** Location of single-unit recordings within the recording grid. The color-coded response profiles of all the visually-responsive cells in monkeys M1 (182 cells) and M2 (138 cells) are here grouped by grid hole (3 grid holes were used in each monkey). The

number of independent penetrations into each grid hole is given below each response histogram. The locations of the sampled grid holes are indicated by the colored squares. Conventions as in Fig. 2.
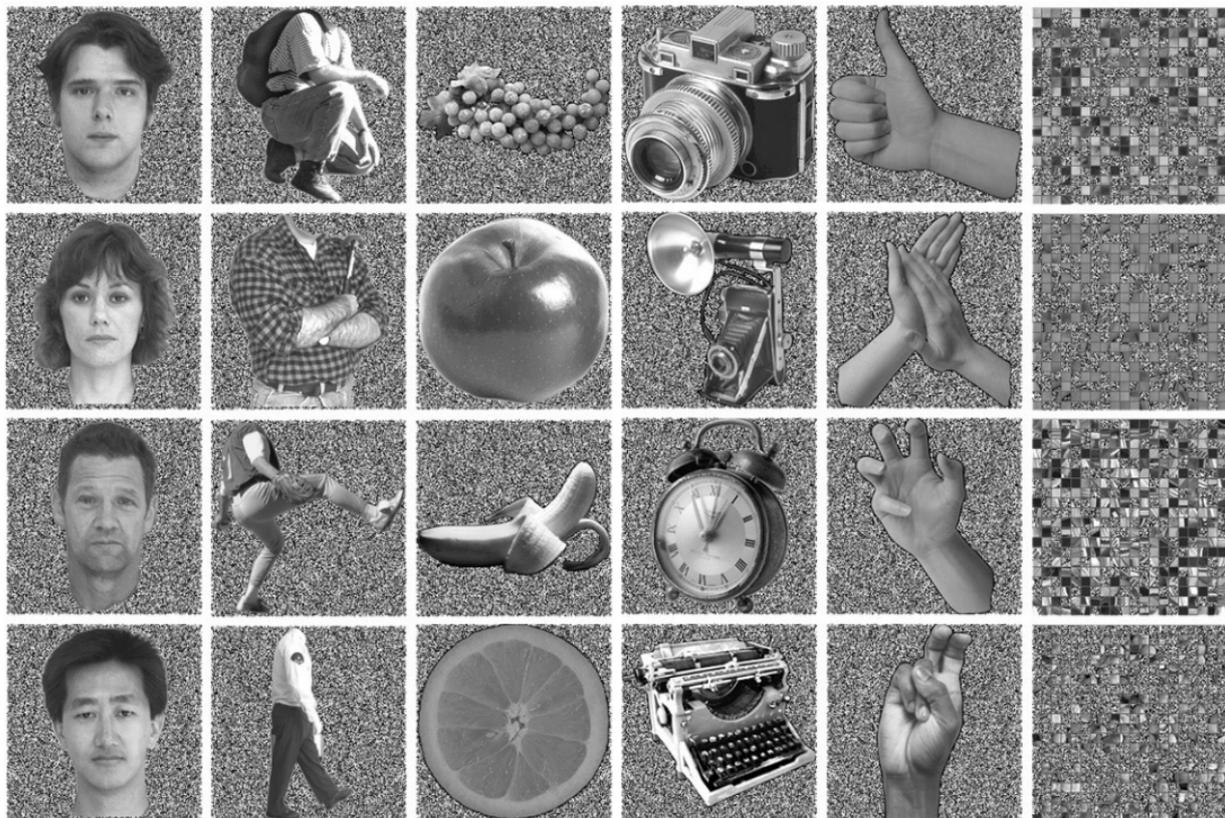
**Fig. S6.** Average spike density histogram, from monkeys M1 and M2, to 96 images of faces, bodies, fruits, gadgets, hands, and scrambled patterns. Individual spike density histograms from the 182 visually responsive single units in monkey M1 and 138 single units in monkey M2 were averaged. The most salient phenomenon is the selective responsiveness to faces (images 1-16). In addition, four more subtle features are evident: 1) a weak response to particular non-face objects which were round (see Fig 2 for pictures of these objects), 2) a weak non-stimulus specific response at 100 ms, 3) in monkey M1, inhibition to many non-face objects, and in monkey M2, inhibition following brief excitation to many non-faces objects, and 4) an OFF response occurring at ~ 300 ms, which, in monkey M1, was stronger on average than the ON-response for non-face objects. These phenomena demonstrate that the responses of face cells in the middle face patch can change over time. Thus fMRI responses observed in both humans (*11*) and macaques (*4, 12*) to non-face objects may not be due to stimulus-specific excitation, but may instead be attributed to non-specific transient responses, off responses, or inhibition.


**Fig. S7.** Effect of face inversion. **(A)** Top: Mean response of a single face-selective cell to 64 upright and 64 inverted faces. Bottom: Mean response across all upright and all inverted faces. This cell was inhibited by inverted faces. **(B)** Response to face inversion averaged across 69 cells. The mean response to inverted faces was delayed and
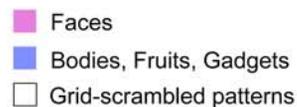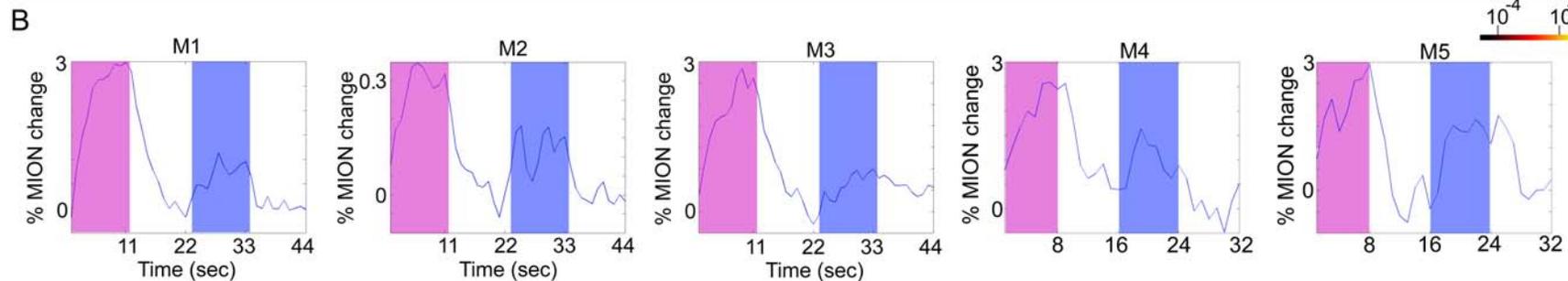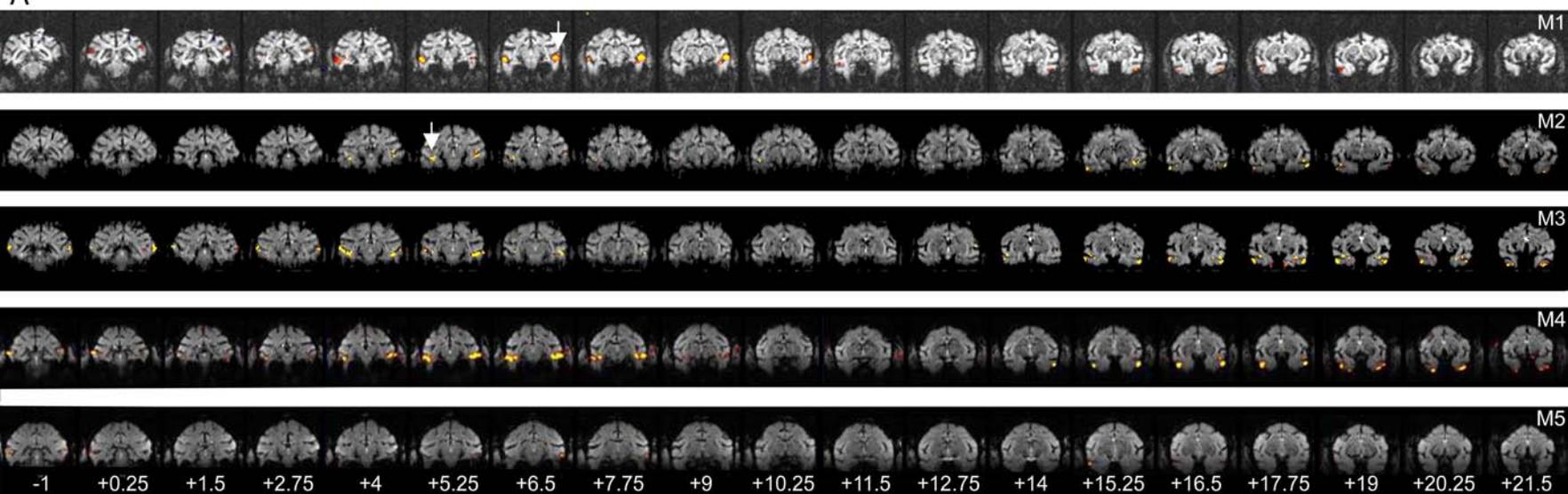
diminished compared to the mean response to upright faces during all response phases;

the overall upright/inverted response ratio was 2.2.

## References

1.      W. Vanduffel *et al.*, *Neuron* 32, 565 (Nov 20, 2001).

2.      D. Y. Tsao *et al.*, *Neuron* 39, 555 (2003).

3.      D. Y. Tsao, B. R. Conway, M. S. Livingstone, *Neuron* 38, 103 (2003).

4.      D. Y. Tsao, W. A. Freiwald, T. A. Knutsen, J. B. Mandeville, R. B. Tootell, *Nat. Neurosci.* 6, 989 (2003).

5.      H. Zeng, R. Constable, *Magn. Res. in Med* 48, 137 (2002).

6.      R. Cusack, M. Brett, K. Osswald, *Neuroimage* 18, 127 (2003).

7.      R. Yin, *J. Exp. Psychol.* 81, 141 (1969).

8.      D. I. Perrett *et al.*, *Proc. R. Soc. Lond. B Biol. Sci.* 223, 293 (1985).

9.      J. Haxby *et al.*, *Neuron* 22, 189 (1999).

10.     N. Kanwisher, F. Tong, K. Nakayama, *Cognition* 68, 1 (1998).

11.     J. Haxby *et al.*, *Science* 293, 2425 (2001).

12.     M. A. Pinsk, K. Desimone, T. Moore, C. G. Gross, S. Kastner, *Proc. Natl. Acad. Sci.* 102, 6996 (2005).

Supplementary Figure 1

M1

M2

M3

M4

M5

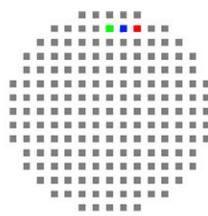| -1 | +0.25 | +1.5 | +2.75 | +4 | +5.25 | +6.5 | +7.75 | +9 | +10.25 | +11.5 | +12.75 | +14 | +15.25 | +16.5 | +17.75 | +19 | +20.25 | +21.5 |

$10^{-4}$  $10^{-8}$

B

M1

M2

M3

M4

M5

% MION change

Time (sec)

Faces

Bodies, Fruits, Gadgets

Grid-scrambled patterns
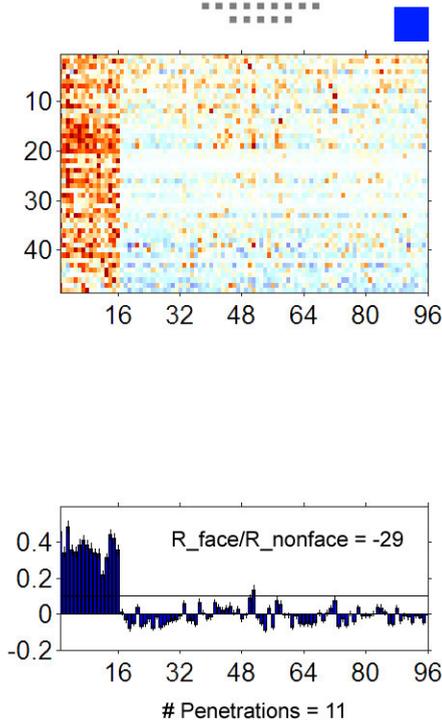
Supplementary Figure 2

Supplementary Figure 3

Supplementary Figure 4

# M1

grid spacing = 1 mm

R_face/R_nonface = -322
# Penetrations = 19

R_face/R_nonface = -29
# Penetrations = 11

R_face/R_nonface = --407
# Penetrations = 16

# M2

R_face/R_nonface = -10
# Penetrations = 2

R_face/R_nonface = 43
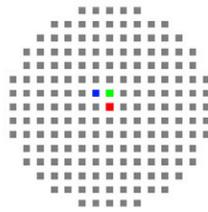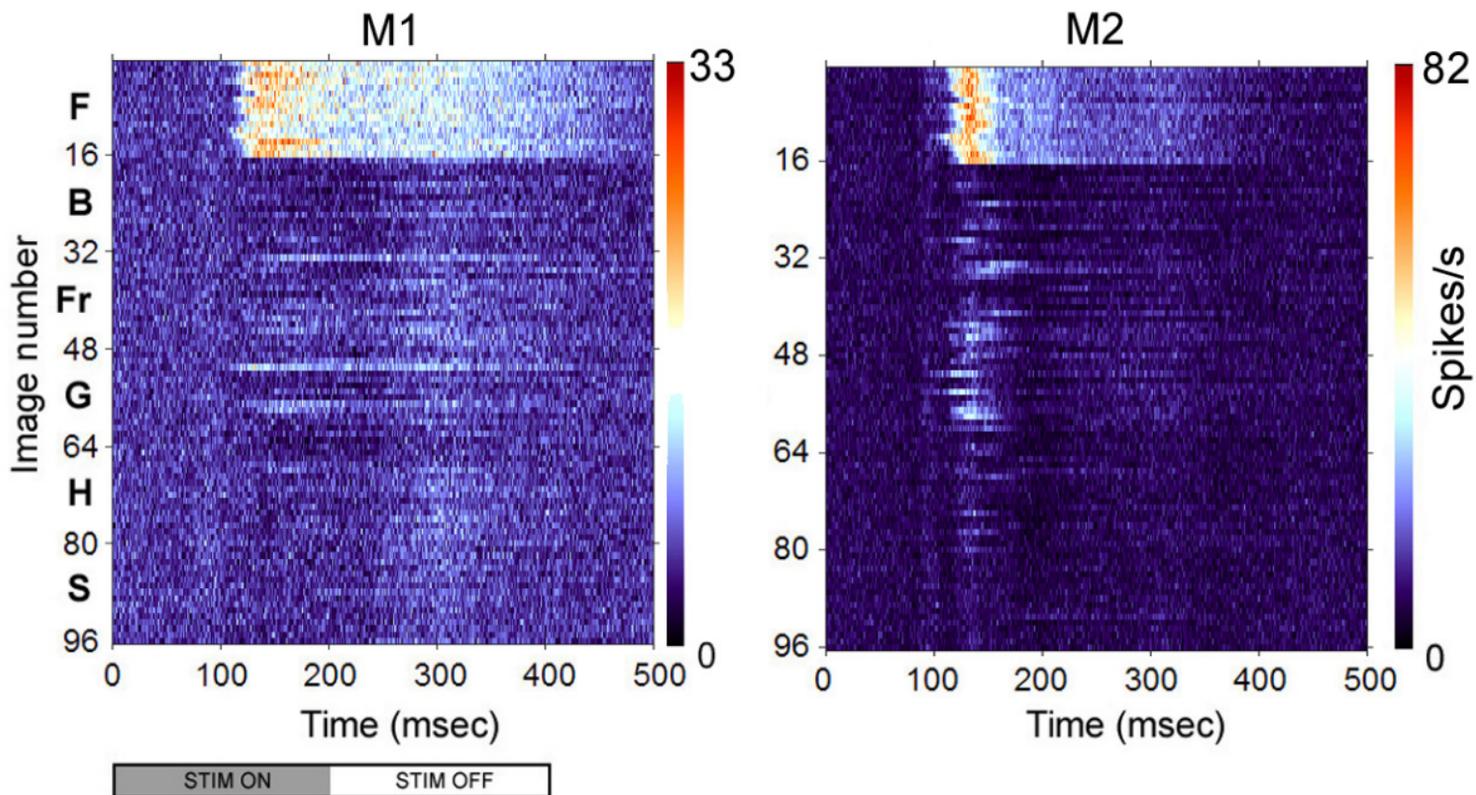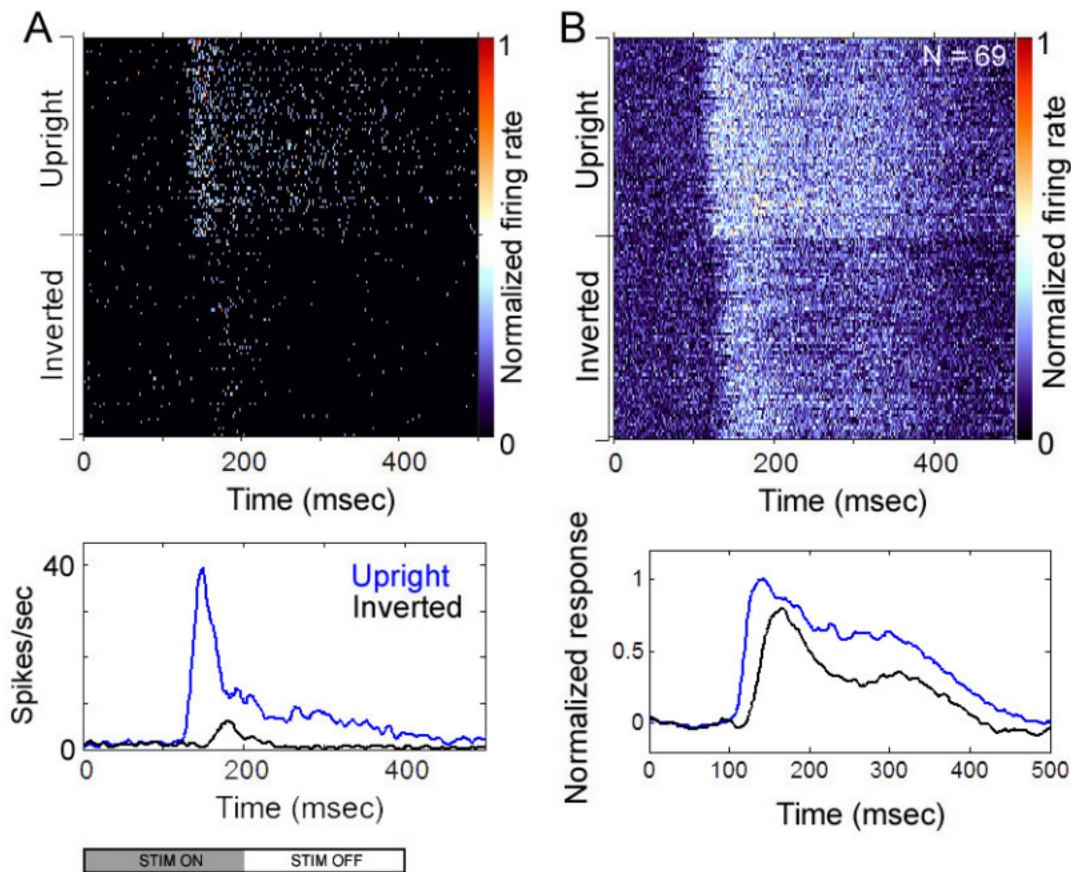# Penetrations = 16

R_face/R_nonface = 34
# Penetrations = 10
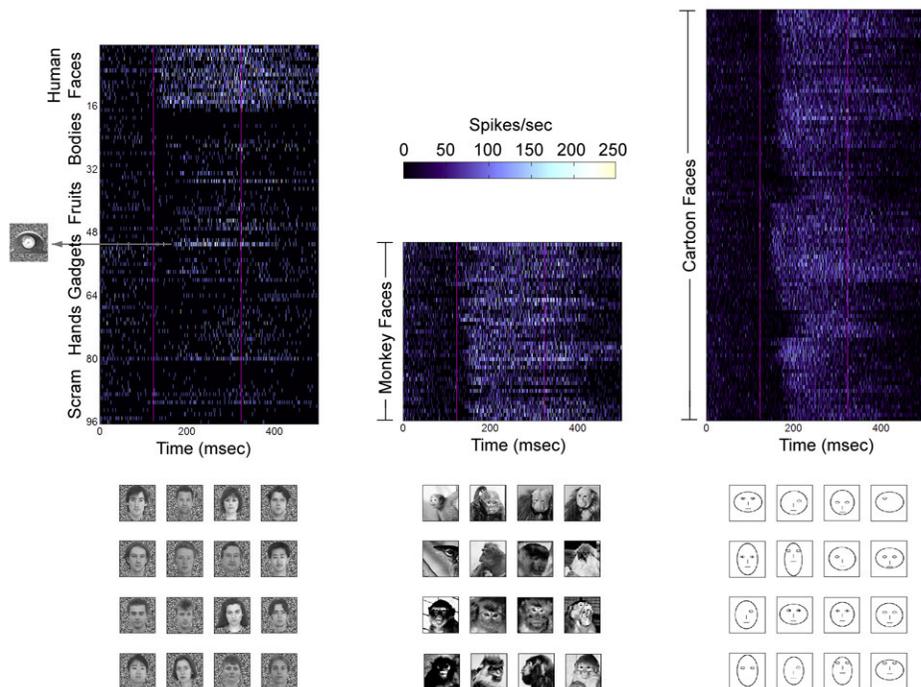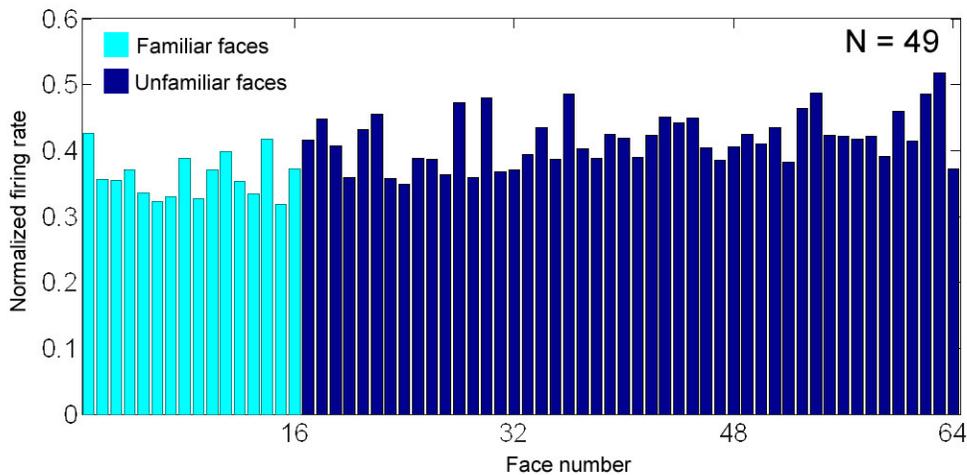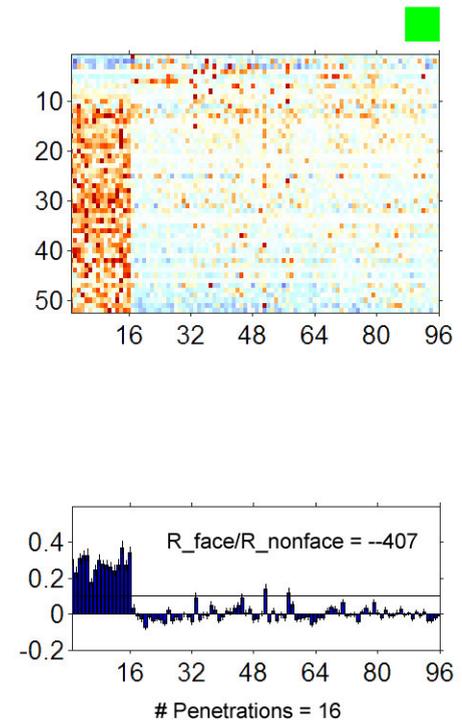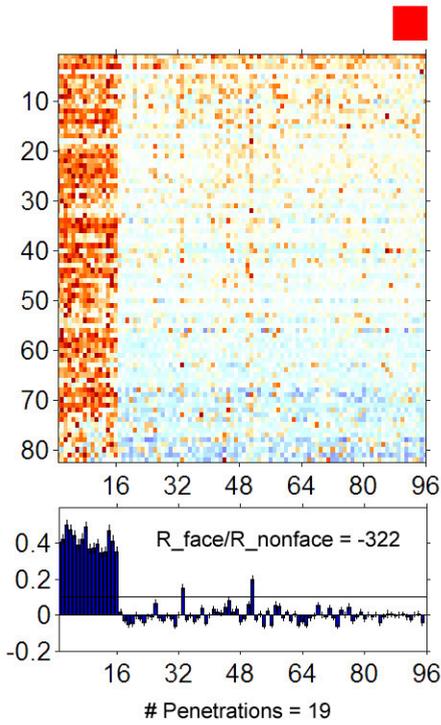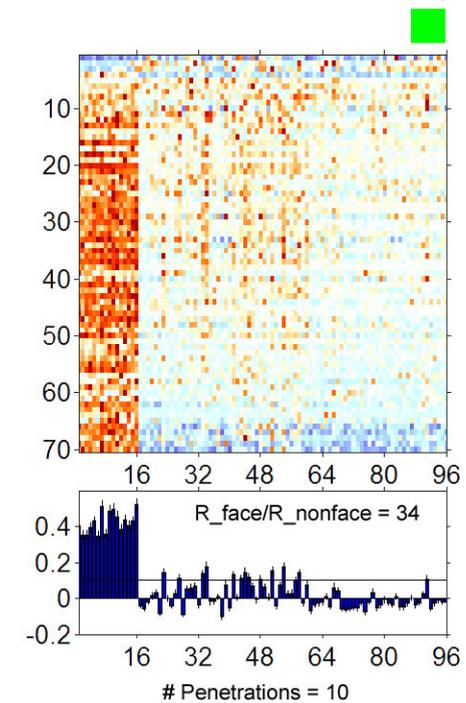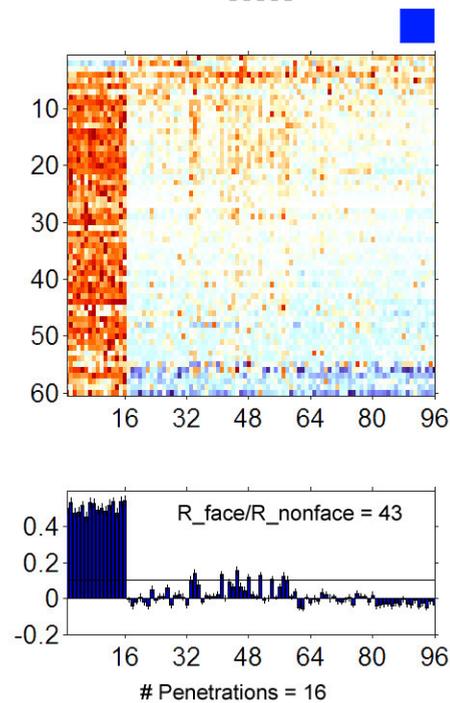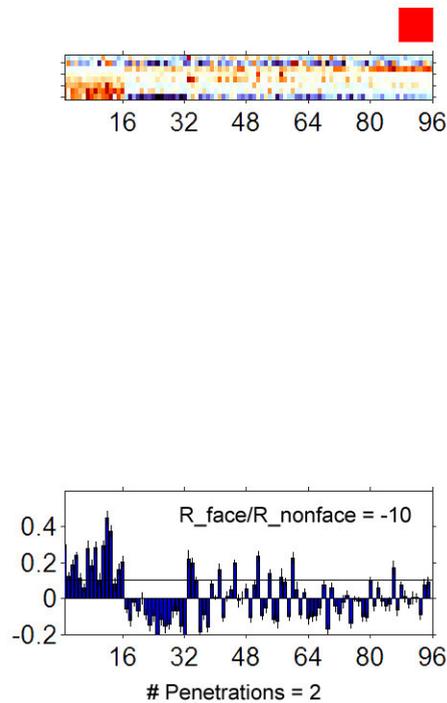
Supplementary Figure 5

Supplementary Figure 6

Supplementary Figure 7