The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

1860-11

Borsellino College 2007. Spike Trains to Actions: Brain Basis of Behavior

3 - 14 September 2007

Visual Perception
(Microstimulation of inferotemporal cortes influences face categorization)

Hossein ESTEKY
School of Medicine, Shaheed Beheshti Univ. and School of Cognitive Sciences, IPM, Tehran, Iran

Microstimulation of inferotemporal cortex influences face categorization

Seyed-Reza Afraz[1,2], Roozbeh Kiani[1,2], Hossein Esteky [1,2,3]

1. Research Center for Brain and Cognitive Sciences, Shaheed Beheshti University of Medical Sciences, Tehran, Iran
2. School of Cognitive Sciences, Institute for Studies in Theoretical Physics and Mathematics, Tehran, Iran
3. Neuroscience Research Center, Shaheed Beheshti University of Medical Sciences Tehran, Iran

Address correspondence to:

Dr. Hossein Esteky
IPM School of Cognitive Sciences, Bahonar Sq. Niavaran, Tehran, Iran
E-mail: esteky@ipm.ir
Tel: (+98) 21 2229 4035
Fax: (+98) 21 2228 0352

The inferior temporal cortex (IT) of primates is thought to be the final visual area in the ventral stream of cortical areas responsible for object recognition[1,2]. Consistent with this hypothesis, single IT neurons respond selectively to highly complex visual stimuli such as faces[3,4,5,6]. However, a direct causal link between the activity of face selective neurons and face perception has never been demonstrated. In the present study we artificially activated small clusters of IT neurons by means of electrical microstimulation while monkeys performed a categorization task, judging whether noisy visual images belonged to "face" or "non-face" categories. Microstimulation of face-selective sites, but not other sites, strongly biased

the monkeys' decisions toward the face category. The magnitude of the effect depended upon the degree of face selectivity of the stimulation site, the size of the stimulated cluster of face-selective neurons, and the exact timing of microstimulation. Our results establish for the first time a causal relationship between the activity of face-selective neurons and face perception.

We trained two adult macaque monkeys to perform a face-nonface categorization task upon viewing single images from one or the other category that were systematically degraded by varying amounts of noise. We chose the noise levels to create a range of difficulties spanning psychophysical threshold: categorization was easy on some trials and difficult on others (Fig. 1A). On each trial, the monkey was presented briefly (54ms) with a face or non-face image degraded by noise. Subsequently, the monkey was required to make a saccadic eye movement to one of two targets to indicate whether the image was a face or non-face. Each correct response was rewarded by a drop of juice. For pure noise stimuli (Fig. 1A, "100%"), the monkey was rewarded randomly with a probability of 0.5.

Our central experimental question was whether electrical microstimulation of clusters of face-selective IT neurons would bias the monkeys' choices toward the

face category. Because of its relatively precise temporal and spatial characteristics, microstimulation is a particularly powerful tool for establishing causal relationships between physiologically characterized neurons and behavioral performance[7,8,9,10]. Even weak microstimulation pulses excite many neurons simultaneously[11,12,13]; successful use of extracellular microstimulation therefore relies on structural regularities within the cortex, such as the presence of cortical columns[14,15]. Face selective neurons are found in relatively large clusters in IT[16,17,18] making them an optimal target for microstimulation.

In each experimental session, we assessed the face selectivity of multiunit clusters of neurons at regular intervals (minimum steps of 150μm) through a single electrode penetration in IT cortex. At each recording site, selectivity was determined by presenting a large number of face and non-face images while the monkey passively fixated a small fixation point on the monitor screen. Face/non-face stimulus selectivity of multiunit responses was quantified with a d' index. A d' value of zero indicates indistinguishable responses to faces and non-faces. Increasingly positive d' values indicate progressively better selectivity for faces.

After recording from several sites within a track (mean number of recorded sites in each track=4), the electrode was positioned in between the recorded sites and neural response selectivity was determined again. Altogether, we assessed stimulus selectivity at 348 recording sites in 86 electrode penetrations in two

monkeys (46 and 40 in monkey FR and KH, respectively). We conducted microstimulation experiments at 31 face-selective sites and 55 non-selective sites, while the monkey performed the object categorization task. Selectivity for faces was defined as having a d' value > 1.

Microstimulation consisted of bipolar current pulses of 50μA delivered at 200 Hz[19,20]. The stimulation pulses were biphasic with the cathodal pulse leading. Each pulse was 0.2ms in duration with a 0.1ms between the cathodal and anodal phase. Each experiment contained three microstimulation conditions differing in the exact time of stimulation delivery as well as an un-stimulated control condition. Stimulating pulses were delivered for 50ms in one of three time periods following onset of the visual stimulus: zero to 50ms, 50 to 100ms or 100 to 150ms. The first period was prior to the earliest visual responses normally observed in IT, and the latter two periods correspond to the earliest and later IT responses, respectively[21,22,23]. The three stimulation conditions and the control trials were randomly interleaved in each experiment.

To reveal the impact of microstimulation on behavior, the monkey's performance in the categorization task was plotted as the proportion of "face" choices as a function of the visual stimulus signal for face and non-face images (Fig. 2). We used positive visual signal values for faces and negative values for non-faces to create a continuum. Logistic regression analysis was used to determine whether

the three microstimulation time conditions caused a significant shift in the psychometric functions compared to the non-stimulated condition. A leftward shift in the psychometric function would indicate an increased tendency to choose faces on trials in which microstimulation was applied.

Figure 2 illustrates results obtained in two typical microstimulation experiments. In both experiments, microstimulation during the 50-100ms interval shifted the monkeys' choices significantly toward the face category (Fig. 2A, logistic regression, $p<0.001$; Fig. 2B, $p<0.01$). Microstimulation during the 100-150ms interval biased choices significantly in the experiment of Fig. 2A ($p<0.001$), but not in the other experimental session depicted in Fig. 2B ($p=0.283$). Microstimulation resulted in a significant leftward shift of the psychometric function in at least one of the stimulation conditions for 19 of 31 face selective sites (61%; 9 in right hemisphere and 10 in left hemisphere) and a significant leftward shift in one non-face site with $d'=0.94$ (see Fig. 3). No significant rightward shift was ever observed.

The impact of microstimulation on perceptual decisions increased as a function of the neural face selectivity of the stimulated site. The scatter plots of figure 3 show the correlation between the degree of face selectivity of the stimulated sites and the shift of the psychometric function in different microstimulation conditions. The strongest correlation was found for microstimulation at 50-100ms after image

onset ($r=0.643$, $p<0.0001$). Microstimulation at 100-150ms also showed a significant correlation ($r=0.539$, $p<0.0001$). No significant correlation was observed for 0-50ms ($r=0.147$, $p=0.18$).

Analysis of the change in reward rate received by monkey following microstimulation confirmed the logistic regression results. In face selective sites ($d'>1$) the reward rate increased significantly in stimulated conditions compared to non-stimulated condition in face presented trials and decreased significantly in non-face presented trials (ANOVA, $F=15.1$ and 10.9; p for both tests <0.001).

The impact of microstimulation on perceptual decisions was much larger when current was injected into larger clusters of face-selective neurons. Recall that we measured stimulus selectivity at recording sites adjacent to the stimulation site as well as at the stimulation site itself. When adjacent sites exhibit selectivity similar to that of the recorded site, we may infer that the cluster of physiologically homogenous neurons is larger, at least along the dimension of our electrode track.

Figure 4A summarizes the effect of stimulating clusters of different sizes. Stimulation effects were substantially more pronounced for larger clusters of face selective neurons (Fig. 4A; black bars) as compared to smaller clusters (Fig. 4A; gray bars). On average, there was no significant shift in the psychometric function following microstimulation of cortical clusters lacking face selectivity (Fig.

4B). A two-way ANOVA showed a significant effect of both neighborhood ($F(1,87)=11.248$, $p=0.001$) and stimulation timing ($F(2,87)=6.092$, $p=0.003$) on the averaged values of the psychometric function shifts across all face-selective sites (sites with d'>1). No such significant effect was observed in non-selective sites. The averaged d' of neighboring sites was correlated with the effect of microstimulation: the correlation coefficients for 0-50, 50-100 and 100-150ms stimulation conditions are: $r=0.12$, $p=0.31$; $r=0.49$, $p<0.001$; $r=0.44$, $p<0.001$, respectively.

To prevent the monkeys to memorize specific exemplars we used a large image bank making it unlikely that the monkey could memorize the specific examples. To further examine the unlikely event of whether the monkeys simply memorized all the images in the image set a behavioral experiment was conducted after the completion of the training in each monkey. In these experiments 40 novel images (20 faces and 20 non-face objects) were intermixed with 40 familiar face and non-face images (randomly chosen from the learned image bank). The stimuli were presented to the monkey without any noise. Face/non-face discrimination performance of the monkeys was measured in several behavioral sessions. In each session we used a new set of novel images. Monkeys' performance for novel stimuli was as good as it was for familiar stimuli (both above 95%) from the very beginning of the behavioral sessions. Furthermore, to prevent the monkeys

to use a general rule other than face/non-face categorization, such as detection of living versus non-living objects, we included non-face animate object images (ex., human and animal bodies from back view or with the face cut out) in our non-face set and included artifact faces (ex., masks and sculpture faces) in the face set. The possibility of familiarity being a factor was also controlled by using images of human and monkey bodies (which are presumably as familiar objects as faces) in the non-face set.

Our findings demonstrate a causal relationship between IT neural activity and visual object perception and categorization. While a general role for IT cortex in object perception has been demonstrated previously in cortical ablation experiments[24,25], our data extend causality to a much finer spatial scale. In addition, our data demonstrate that single neuron response properties provide important clues to the functional role of neurons in perception, even for highly complex stimuli such as faces. The functional role of face selective neurons in behavior has been hotly debated, but our data clearly shows that this role includes, at the very least, categorization of objects into faces and non-faces.

Methods

Behavioral tasks

Each session started with a passive fixation task in which monkeys were required to maintain fixation in a 4ºx4º window at the center of the screen. Following

300ms of fixation, a sequence of visual stimuli (7ºx7º in size) was presented to the monkey. Each image was presented for 200ms without blank intervals between images[26,27,23], 7-10 times pseudorandomly.

The images were grayscale photographs of 30 face objects and 60 non-face objects chosen randomly from an image bank of 600 images.

In the second phase of the experiment, monkeys performed a face-nonface categorization task. The monkey started a trial by fixating on the fixation spot for 300ms. Then a noisy image was presented for 54ms, followed immediately, by two small response targets presented 10 degrees to the left and right of the screen center. The left and right targets represented face and non-face responses, respectively. The monkey was required to make a saccade to the correct target no later than 660ms after the onset of targets.

To minimize monkey's behavioral choice bias, we used a correction scheme[7]. The monkey entered a set of correction trials if he made three consecutive errors within a single category (face or non-face). Upon entering a correction trial, images from the neglected category were presented until the monkey chose that category correctly. All data collected from correction trials were discarded from the analysis. The monkey entered a correction trial in 32 of the 86 sessions, resulting in exclusion of 5.7% of trials in those sessions.

In the categorization task 6 or 5 noise levels were used in monkey Fr and 5 noise levels were used in monkey Kh. Each noise level was generated by assigning a uniformly distributed grayscale value to X percent of image pixels, where X is the noise level. Noisy face and nonface images create a continuum of task relevant visual signal extending from noiseless faces (100) to completely noisy images (0) to noiseless non-faces (-100). For each noise level, 16 face and 16 non-face images were randomly selected from the image bank.

Electrophysiology

Recordings were made on an evenly spaced grid, with 1-mm intervals between penetrations over a wide region of lower bank of STS and TEa cortices[23] (left hemisphere; A-P position 14 to 21 in FR and right hemisphere A-P position 14 to 20 in KH). The recording positions were determined stereotaxically by referring to magnetic resonance images acquired before the surgery. Multiunit neural responses were recorded through tungsten microelectrodes (0.4-1.0 mega-ohm; FHC). Neural selectivity of neighboring sites within ±500μm from the stimulated site along each recording track was determined as the electrode was advanced. The recorded positions were separated by at least 150μm (mean=296μm). After determining the neighborhood selectivity the electrode tip was positioned in the middle of the recorded area and remained there through the rest of the

experiment. The neural selectivity in this site was verified before starting the categorization task.

Data Analysis

Mean multiunit discharge for each stimulus was measured in a period 70-200ms after the image onset. The degree of selectivity of each cortical site for face vs. non-face images was measured by a d' index[28,29] based on the following formula:

$d` = [M(f)-M(nf)] / \sqrt{([\sigma^2(f)+ \sigma^2(nf)]/2)}$

Where M(f) and M(nf) are the mean multiunit response to face and non-face images respectively, $\sigma^2(f)$ and $\sigma^2(nf)$ are the variance of the distributions of neural responses to face and non-face images respectively.

To calculate leftward shift in psychometric function, logistic curves were fit to the monkey's responses in the categorization task based on the following formula:

$$P(x) = \frac{1}{1+e^{-(\alpha+\beta x+\lambda_1 I_1+\lambda_2 I_2+\lambda_3 I_3)}}$$

Where *x* is the visual signal and *P(x)* is the probability of face response. I_1, I_2 and I_3 indicate the presence or absence of microstimulation in the three periods. α, β and λs are free parameters which were fit using the maximum likelihood fitting procedure[30]. The fit was performed separately for all of the behavioral data obtained in each experimental session (86 fits). Microstimulation effect in each site was considered significant if λ_i was significantly different from zero ($p<0.05$).

Leftward shift of the psychometric function in each stimulation condition was defined as the change in the visual signal that would have induced a behavioral effect comparable to that of the microstimulation. This is equal to λ_i/β in the logistic fit. Similar methods have been used in other microstimulation studies[7,19]. To reduce the number of the free parameters our logistic fit assumes a similar slope for the psychometric curves in different stimulation conditions. Allowing different slopes did not improve the fit and was not critical to the results.

References:

1. Tanaka K. Inferotemporal cortex and object vision. *Annu Rev Neurosci.* **19**:109-39 (1996) Review.
2. Logothetis NK, Sheinberg DL. Visual object recognition. *Annu Rev Neurosci.* **19**:577-621 (1996) Review.
3. Desimone R, Albright TD, Gross CG, Bruce C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci.* Aug **4**(8):2051-62 (1984).
4. Perrett DI, Rolls ET, Caan W. Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res*. **47**(3):329-42 (1982).
5. Rolls ET. Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces. *Hum Neurobiol.* **3**(4):209-22 (1984).
6. Gross CG. Processing the facial image: a brief history. *Am Psychol.* **60**(8):755-63 (2005).
7. Salzman CD, Murasugi CM, Britten KH, Newsome WT. Microstimulation in visual area MT: effects on direction discrimination performance. *J Neurosci.* **12**(6):2331-55 (1992).
8. Romo R, Hernandez A, Zainos A, Salinas E. Somatosensory discrimination based on cortical microstimulation. *Nature* **392**(6674):387-90 (1998).
9. Romo R, Salinas E. Sensing and deciding in the somatosensory system. *Curr Opin Neurobiol.* Aug;**9**(4):487-93 (1999).
10. Cohen MR, Newsome WT. What electrical microstimulation has revealed about the neural basis of cognition. *Curr Opin Neurobiol.* Apr;**14**(2):169-77. (2004) Review.

11. Stoney SD Jr, Thompson WD, Asanuma H. Excitation of pyramidal tract cells by intracortical microstimulation: effective extent of stimulating current. J *Neurophysiol.* Sep;**31**(5):659-69 (1968).
12. Asanuma H, Arnold A, Zarzecki P. Further study on the excitation of pyramidal tract cells by intracortical microstimulation. *Exp Brain Res.* Dec **26**(5):443-61 (1976).
13. Tolias AS, Sultan F, Augath M, Oeltermann A, Tehovnik EJ, Schiller PH, Logothetis NK. Mapping cortical activity elicited with electrical microstimulation using FMRI in the macaque. *Neuron* Dec **48**(6):901-11 (2005).
14. Nichols MJ, Newsome WT. Middle temporal visual area microstimulation influences veridical judgments of motion direction. *J Neurosci* Nov **22**(21):9530-40 (2002).
15. DeAngelis GC, Newsome WT. Perceptual "read-out" of conjoined direction and disparity maps in extrastriate area MT. *PLoS Biol.* Mar;**2**(3):E77. (2004).
16. Perrett DI, Hietanen JK, Oram MW, Benson PJ. Organization and functions of cells responsive to faces in the temporal cortex. *Philos Trans R Soc Lond B Biol Sci.* Jan **335**(1273):23-30 (1992).
17. Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB. Faces and objects in macaque cerebral cortex. *Nat Neurosci.* Sep;**6**(9):989-95 (2003).
18. Zangenehpour S, Chaudhuri A. Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. *Curr Biol.* Jun **15**(11):993-1005 (2005).

19. Salzman CD, Britten KH, Newsome WT. Cortical microstimulation influences perceptual judgements of motion direction. Nature. 1990 Jul 12;346(6280):174-7. Erratum in: *Nature* Aug **346**(6284):589 (1990).
20. Moore T, Armstrong KM. Selective gating of visual signals by microstimulation of frontal cortex. *Nature* Jan **421**(6921):370-3 (2003).
21. Bruce C, Desimone R, Gross CG. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol.* Aug;**46**(2):369-84 (1981).
22. Sugase Y, Yamane S, Ueno S, Kawano K. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* Aug **400**(6747): 869-73 (1999).
23. Kiani R, Esteky H, Tanaka K. Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *J Neurophysiol* Aug; **94**(2):1587-96 (2005).
24. Horel JA, Pytko-Joiner DE, Voytko ML, Salsbury K. The performance of visual tasks while segments of the inferotemporal cortex are suppressed by cold. *Behav Brain Res.* Jan; **23**(1):29-42 (1987).
25. Buffalo EA, Ramus SJ, Clark RE, Teng E, Squire LR, Zola SM. Dissociation between the effects of damage to perirhinal cortex and area TE. *Learn Mem.* Nov-Dec; **6**(6):572-99 (1999).
26. Keysers C, Xiao DK, Foldiak P, Perrett DI. The speed of sight. *J Cogn Neurosci.* Jan **13**(1):90-101 (2001).

27. Foldiak P, Xiao D, Keysers C, Edwards R, Perrett DI. Rapid serial visual presentation for the determination of neural selectivity in area STSa. *Prog Brain Res.***144**:107-16 (2004).
28. Green DM, A. Swets JA. *Signal Detection Theory and Psychophysics.* (Peninsula Publishing, 1988).
29. Dayan P, Abbott L. F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems.* (MIT Press, 1999).
30. Meeker WQ, Escobar LA. Teaching about approximate confidence regions based on maximum likelihood estimation. *Am Stat*. **49**:48-53 (1995).

Figure Captions:

Figure 1) Each experimental session consisted of two parts: a passive fixation and an active categorization phase. First, the neural stimulus selectivity of several neighboring cortical sites was determined in a fixation task using luminance-matched face and nonface grayscale images. The images were randomly selected from a large set. (a) In the second part of the experiment face and non-face images with varying amounts of noise were used in a categorization task. These noisy face and nonface images could be arranged in a scale where 100 indicates noiseless faces, zero indicates complete noisy images and -100 indicates noiseless non-faces (b). In the categorization task the monkey had to fixate on a fixation point at the center of the screen for 300ms to begin a trial. In each trial a randomly-selected face or non-face noisy image was presented at the center of the screen for 54ms, followed by two targets in the left and right sides of the screen. To indicate whether the image was a face or nonface, the monkey had to make a saccade to the left or the right target, respectively. The monkey was rewarded with a drop of juice for each correct response. Full noise images (100%noise) were rewarded randomly. The two targets disappeared and the trial was aborted after 660ms if no response was made. In this task, there was a random sequence of 4 types of trials with equal proportions. A bipolar current (200Hz, 50uA) was injected for 50ms into cortex in 3 different periods (0-50ms, 50-100ms or 100-150ms from the image onset). In the remaining 25% of trials no current was injected.

Figure 2) Leftward shift in psychometric functions due to microstimulation of two representative face-selective neural clusters in IT cortex of monkey Kh (top) and Fr

(bottom). Data points show the proportion of face choices for different levels of noise in the images for different microstimulation conditions. The curves are logistic regression fits to the data points. The insets show average multiunit responses of the corresponding stimulated site and their neighboring sites. Error bars represent SEM.

Figure 3) Correlation between face selectivity of stimulated sites and the behavioral impacts of microstimulation. The correlation is significant for microstimulation at 50-100ms ($r=0.643$, $p<0.0001$) and 100-150ms ($r=0.539$, $p<0.0001$) after image onset but not for 0-50ms ($r=0.147$, $p=0.18$). Face selectivity was measured by d' index (see Methods). The microstimulation effect was quantified by leftward shift in the psychometric function (positive numbers in the Y-axis). Red data points indicate that microstimulation of the represented cortical site resulted in a statistically significant shift of the psychometric function (a). An average population histogram of multiunit neural responses to face (red line) and non-face (blue line) images (b). Different microstimulation time conditions are depicted by number of stars in (a) and (b) and vertical lines in the average response histogram.

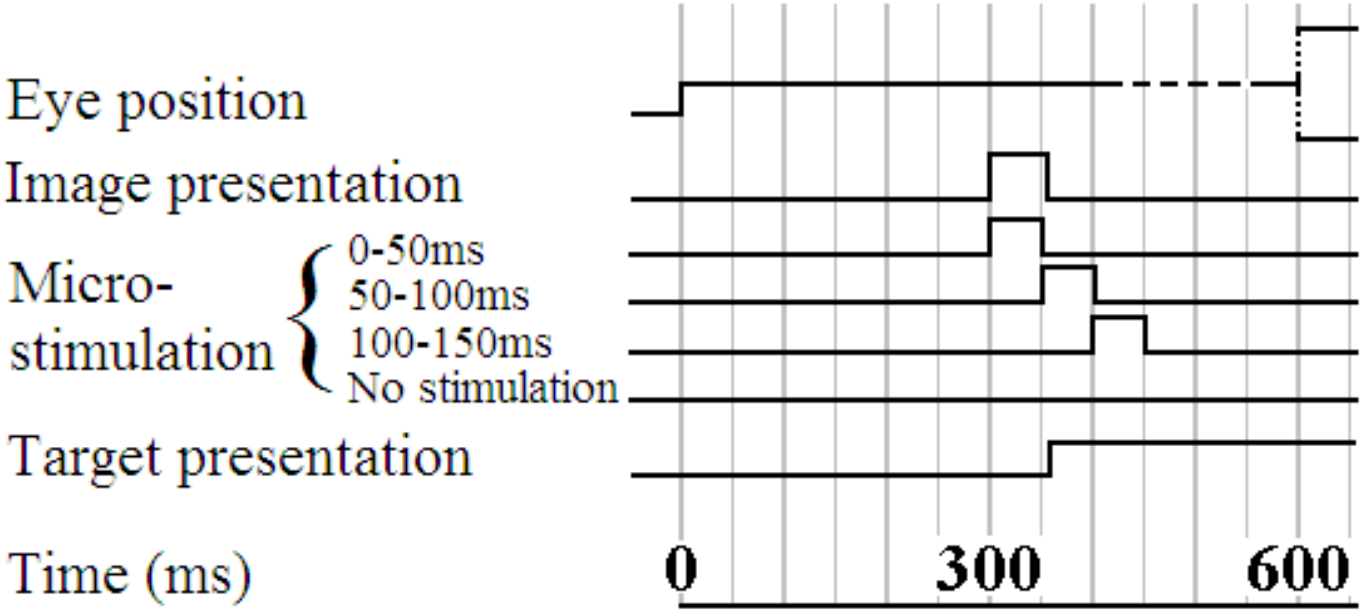
Figure 4) The effect of stimulus selectivity of neighboring cortical sites on microstimulation results. Averaged shift in the psychometric function for the three stimulation conditions is shown for all stimulated face (left plot) and non-selective (right plot) sites. Face and non-selective sites were defined by $d'>1$ and $d'\leq 1$, respectively. Black columns represent sites with face selective neighbors in their 500micron vicinity and gray columns show sites with non-selective neighbor(s). Error bars represent SEM.

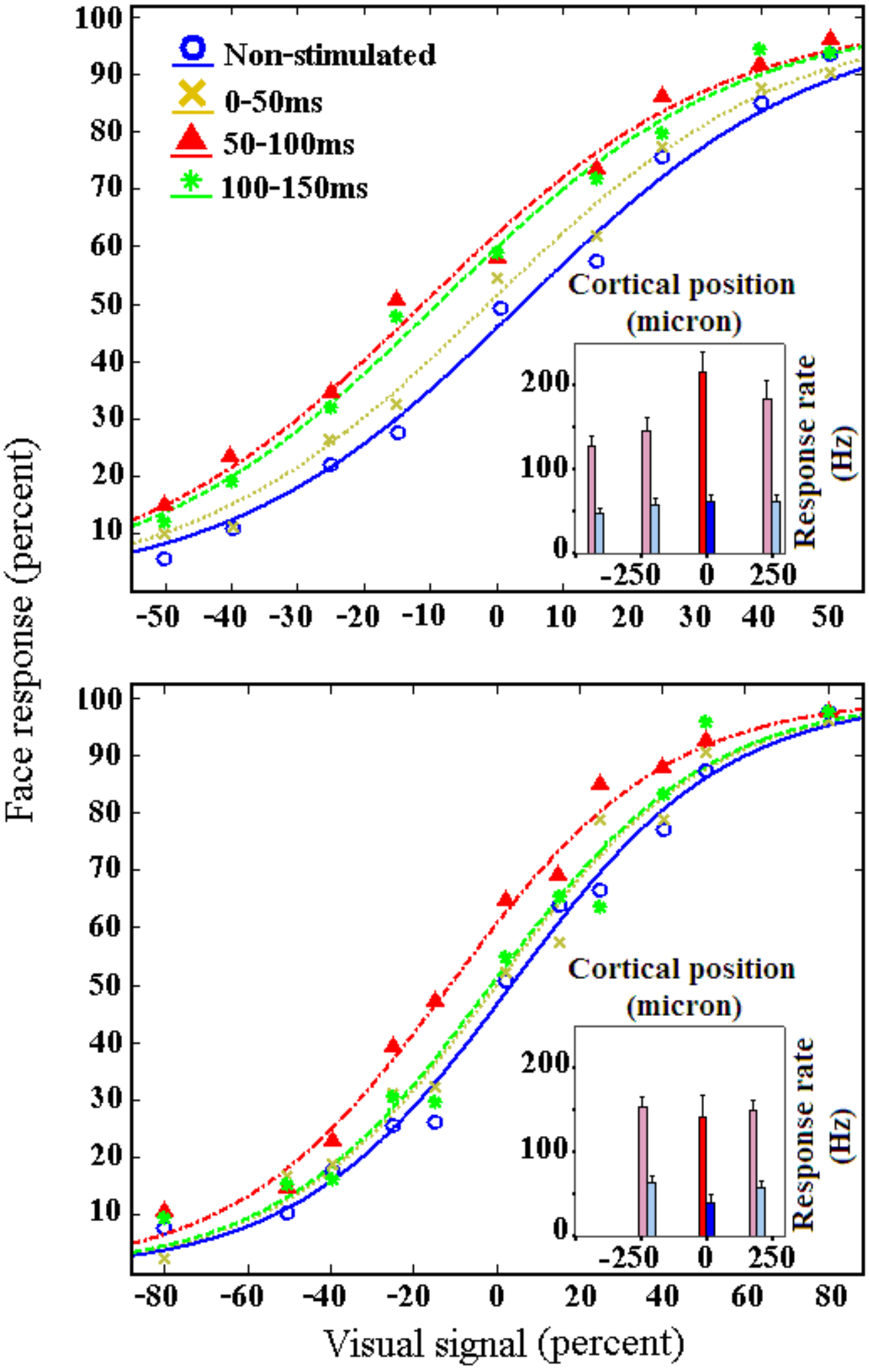
Supplementary Figure 1) Leftward shift in psychometric functions due to microstimulation of 6 cortical sites with a wide range of neuronal stimulus selectivity of the stimulated sites and their neighboring sites. The presented examples cover all of the main conditions of stimulus selectivity profile of neuronal clusters examined in this study. Data points show the proportion of face choices for different levels of noise in the images for different microstimulation conditions. The curves are logistic regression fits to the data points.

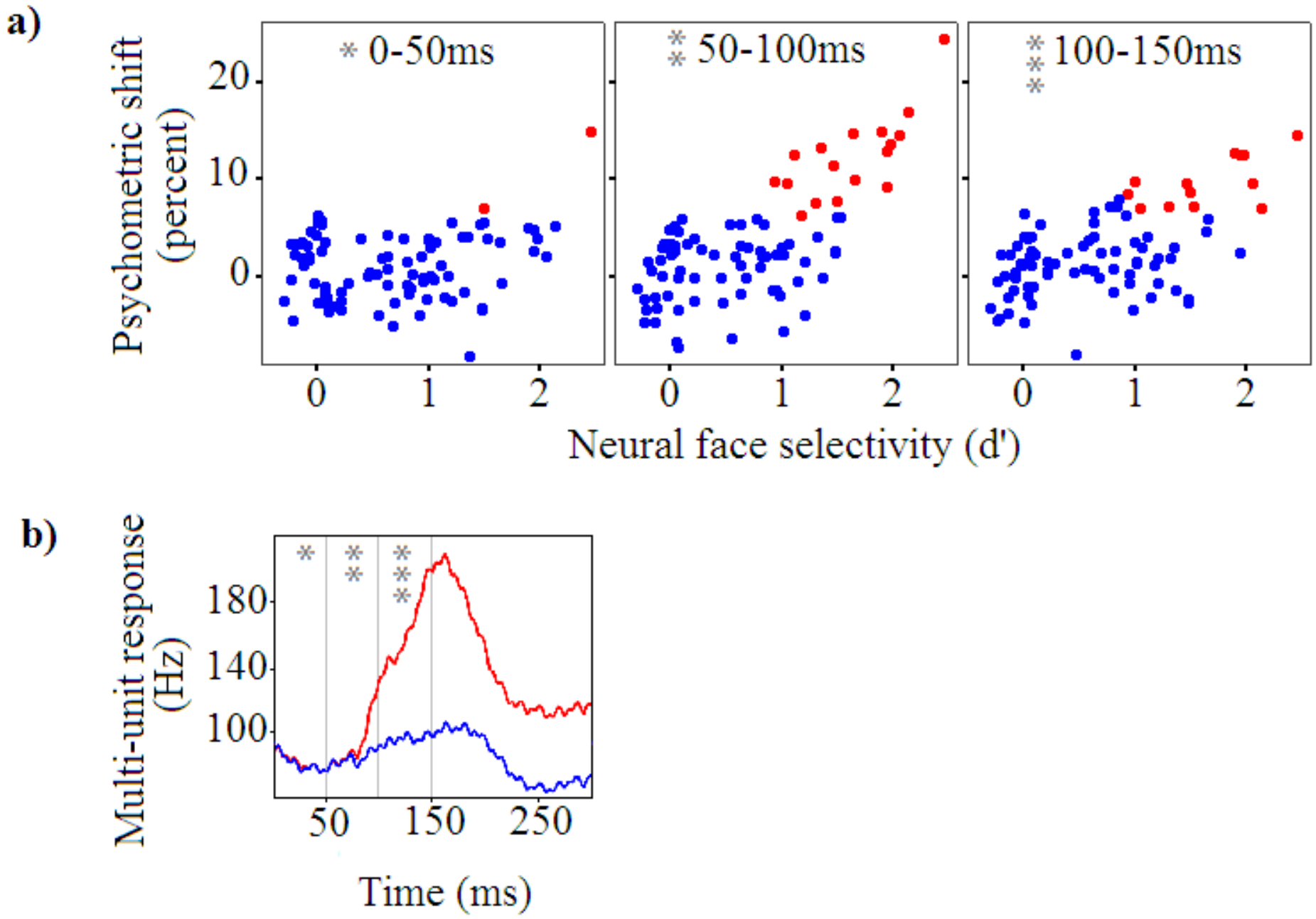
a) Noise level: 20% 75% 100% 75% 20%

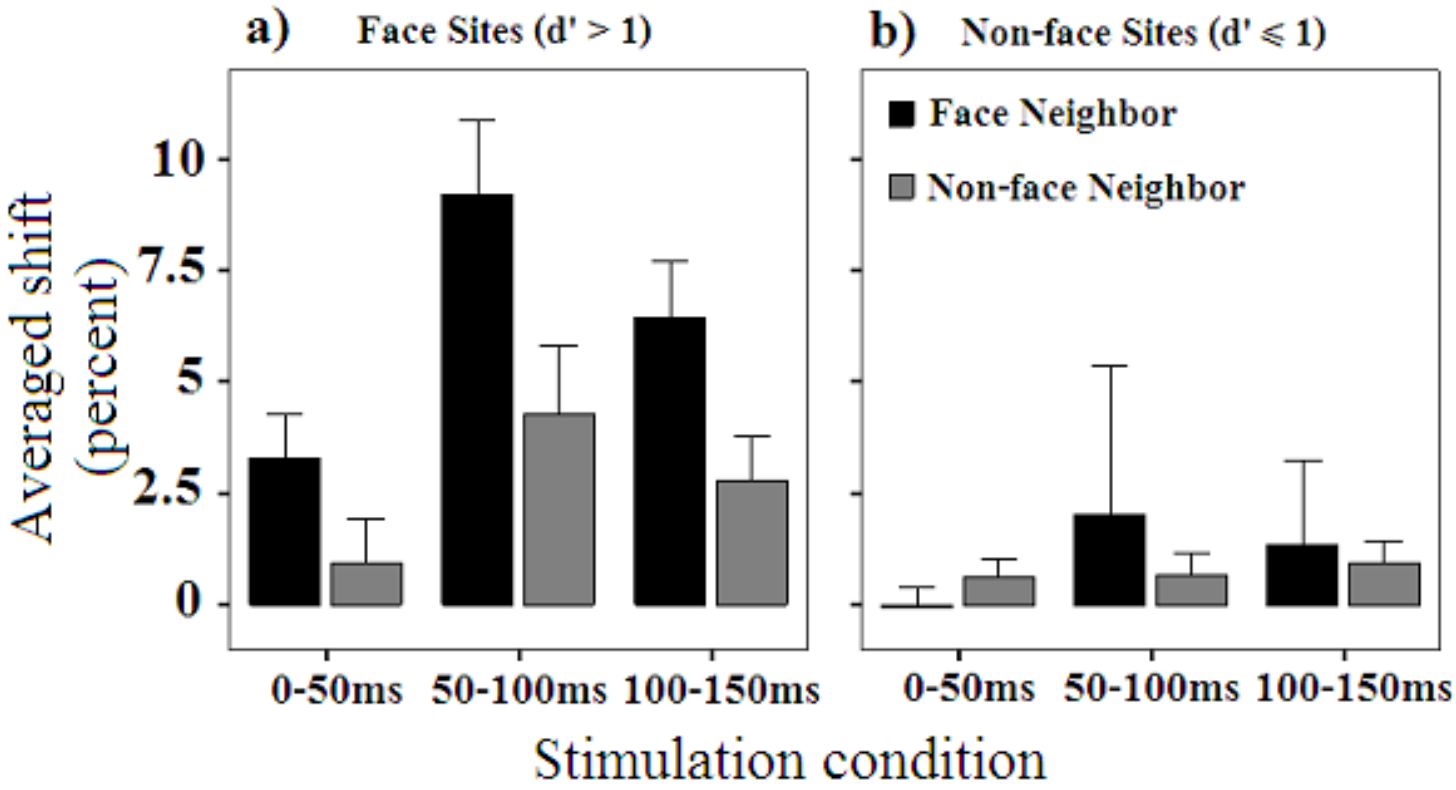
Visual signal: -80 -25 0 25 80

b)




Non-stimulated
0-50ms
50-100ms
100-150ms
Cortical position (micron)
Response rate (Hz)
Face response (percent)
Visual signal (percent)


a)
0-50ms
50-100ms
100-150ms
Psychometric shift (percent)
20
10
0
0
1
2
Neural face selectivity (d')
b)
Multi-unit response (Hz)
180
140
100
50
150
250
Time (ms)


a) Face Sites (d' > 1)
b) Non-face Sites (d' ⩽ 1)
Face Neighbor
Non-face Neighbor
Averaged shift (percent)
10
7.5
5
2.5
0
0-50ms
50-100ms
100-150ms
0-50ms
50-100ms
100-150ms
Stimulation condition