LETTERS

Invariant visual representation by single neurons in the human brain

R. Quian Quiroga[1,2]†, L. Reddy[1], G. Kreiman[3], C. Koch[1] & I. Fried[2,4]

It takes a fraction of a second to recognize a person or an object even when seen under strikingly different conditions. How such a robust, high-level representation is achieved by neurons in the human brain is still unclear[1–6]. In monkeys, neurons in the upper stages of the ventral visual pathway respond to complex images such as faces and objects and show some degree of invariance to metric properties such as the stimulus size, position and viewing angle[2,4,7–12]. We have previously shown that neurons in the human medial temporal lobe (MTL) fire selectively to images of faces, animals, objects or scenes[13,14]. Here we report on a remarkable subset of MTL neurons that are selectively activated by strikingly different pictures of given individuals, landmarks or objects and in some cases even by letter strings with their names. These results suggest an invariant, sparse and explicit code, which might be important in the transformation of complex visual percepts into long-term and more abstract memories.

The subjects were eight patients with pharmacologically intractable epilepsy who had been implanted with depth electrodes to localize the focus of seizure onset. For each patient, the placement of the depth electrodes, in combination with micro-wires, was determined exclusively by clinical criteria[13]. We analysed responses of neurons from the hippocampus, amygdala, entorhinal cortex and parahippocampal gyrus to images shown on a laptop computer in 21 recording sessions. Stimuli were different pictures of individuals, animals, objects and landmark buildings presented for 1 s in pseudo-random order, six times each. An unpublished observation in our previous recordings was the sometimes surprising degree of invariance inherent in the neuron's (that is, unit's) firing behaviour. For example, in one case, a unit responded only to three completely different images of the ex-president Bill Clinton. Another unit (from a different patient) responded only to images of The Beatles, another one to cartoons from *The Simpson's* television series and another one to pictures of the basketball player Michael Jordan. This suggested that neurons might encode an abstract representation of an individual. We here ask whether MTL neurons can represent high-level information in an abstract manner characterized by invariance to the metric characteristics of the images. By invariance we mean that a given unit is activated mainly, albeit not necessarily uniquely, by different pictures of a given individual, landmark or object.

To investigate further this abstract representation, we introduced several modifications to optimize our recording and data processing conditions (see Supplementary Information) and we designed a paradigm to systematically search for and characterize such invariant neurons. In a first recording session, usually done early in the morning (screening session), a large number of images of famous persons, landmark buildings, animals and objects were shown. This set was complemented by images chosen after an interview with the patient. The mean number of images in the screening session was 93.9 (range 71–114). The data were quickly analysed offline to determine the stimuli that elicited responses in at least one unit (see definition of response below). Subsequently, in later sessions (testing sessions) between three and eight variants of all the stimuli that had previously elicited a response were shown. If not enough stimuli elicited significant responses in the screening session, we chose those stimuli with the strongest responses. On average, 88.6 (range 70–110) different images showing distinct views of 14 individuals or objects (range 7–23) were used in the testing sessions. Single views of random stimuli (for example, famous and non-famous faces, houses, animals, etc) were also included. The total number of stimuli was determined by the time available with the patient (about 30 min on average). Because in our clinical set-up the recording conditions can sometimes change within a few hours, we always tried to perform the testing sessions shortly after the screening sessions in order to maximize the probability of recording from the same units. Unless explicitly stated otherwise, all the data reported in this study are from the testing sessions. To hold their attention, patients had to perform a simple task during all sessions (indicating with a key press whether a human face was present in the image). Performance was close to 100%.

We recorded from a total of 993 units (343 single units and 650 multi-units), with an average of 47.3 units per session (16.3 single units and 31.0 multi-units). Of these, 132 (14%; 64 single units and 68 multi-units) showed a statistically significant response to at least one picture. A response was considered significant if it was larger than the mean plus 5 standard deviations (s.d.) of the baseline and had at least two spikes in the post-stimulus time interval considered (300–1,000 ms). All these responses were highly selective: for the responsive units, an average of only 2.8% of the presented pictures (range: 0.9–22.8%) showed significant activations according to this criterion. This high selectivity was also present in the screening sessions, where only 3.1% of the pictures shown elicited responses (range: 0.9–18.0%). There was no significant difference between the relative number of responsive pictures obtained in the screening and testing sessions (*t*-test, $P = 0.40$). Responses started around 300 ms after stimulus onset and had mainly three non-exclusive patterns of activation (with about one-third of the cells having each type of response): the response disappeared with stimulus offset, 1 s after stimulus onset; it consisted of a rapid sequence of about 6 spikes (s.d. = 5) between 300 and 600 ms after stimulus onset; or it was prolonged and continued up to 1 s after stimulus offset. For this study, we calculated the responses in a time window between 300 and 1,000 ms after stimulus onset. In a few cases we also observed cells that responded selectively only after the image was removed from view (that is, after 1 s). These are not further analysed here.

[1]Computation and Neural Systems, California Institute of Technology, Pasadena, California 91125, USA. [2]Division of Neurosurgery and Neuropsychiatric Institute, University of California, Los Angeles (UCLA), California 90095, USA. [3]Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [4]Functional Neurosurgery Unit, Tel-Aviv Medical Center and Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv 69978, Israel. †Present address: Department of Engineering, University of Leicester, LE1 7RH, UK.

Figure 1a shows the responses of a single unit in the left posterior hippocampus to a selection of 30 out of the 87 pictures presented to the patient. None of the other pictures elicited a statistically significant response. This unit fired to all pictures of the actress Jennifer Aniston alone, but not (or only very weakly) to other famous and non-famous faces, landmarks, animals or objects. Interestingly, the unit did not respond to pictures of Jennifer Aniston together with the actor Brad Pitt (but see Supplementary Fig. 2). Pictures of Jennifer Aniston elicited an average of 4.85 spikes (s.d. = 3.59) between 300 and 600 ms after stimulus onset. Notably, this unit was nearly silent during baseline (average of 0.02 spikes in a 700-ms pre-stimulus time window) and during the presentation of most other pictures (Fig. 1b). Figure 1b plots the median number of spikes (across trials) in the 300–1,000-ms post-stimulus interval for all 87 pictures shown to the patient. The histogram shows a marked differential response to pictures of Jennifer Aniston (red bars).

Next, we quantified the degree of invariance using a receiver operating characteristic (ROC) framework[15]. We considered as the hit rate (y axis) the relative number of responses to pictures of a specific individual, object, animal or landmark building, and as

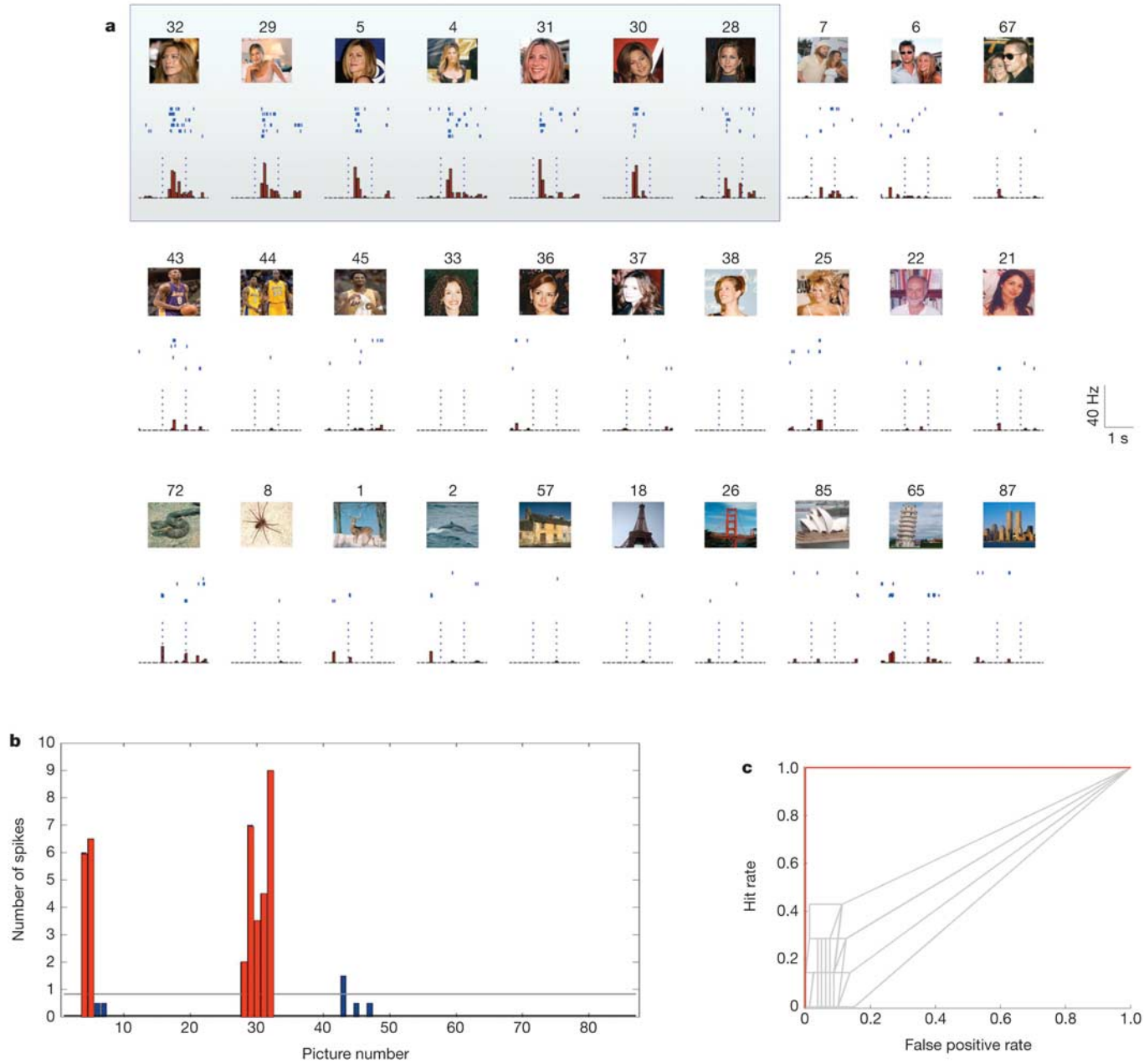


Figure 1 | A single unit in the left posterior hippocampus activated exclusively by different views of the actress Jennifer Aniston. **a**, Responses to 30 of the 87 images are shown. There were no statistically significant responses to the other 57 pictures. For each picture, the corresponding raster plots (the order of trial number is from top to bottom) and post-stimulus time histograms are given. Vertical dashed lines indicate image onset and offset (1 s apart). Note that owing to insurmountable copyright problems, all original images were replaced in this and all subsequent figures by very similar ones (same subject, animal or building, similar pose, similar colour, line drawing, and so on). **b**, The median responses to all pictures. The image numbers correspond to those in **a**. The two horizontal lines show the mean baseline activity (0.02 spikes) and the mean plus 5 s.d. (0.82 spikes). Pictures of Jennifer Aniston are denoted by red bars. **c**, The associated ROC curve (red trace) testing the hypothesis that the cell responded in an invariant manner to all seven photographs of Jennifer Aniston (hits) but not to other images (including photographs of Jennifer Aniston and Brad Pitt together; false positives). The grey lines correspond to the same ROC analysis for 99 surrogate sets of 7 randomly chosen pictures ($P < 0.01$). The area under the red curve is 1.00.

the false positive rate (x axis) the relative number of responses to other pictures. The ROC curve corresponds to the performance of a linear binary classifier for different values of a response threshold. Decreasing the threshold increases the probability of hits but also of false alarms. A cell responding to a large set of pictures of different individuals will have a ROC curve close to the diagonal (with an area under the curve of 0.5), whereas a cell that responds to all pictures of an individual but not to others will have a convex ROC curve far from the diagonal, with an area close to 1. In Fig. 1c we show the ROC curve for all seven pictures of Jennifer Aniston (red trace, with an area equal to 1). The grey lines show 99 ROC surrogate curves, testing invariance to randomly selected groups of pictures (see Methods). As expected, these curves are close to the diagonal, having an area of about 0.5. None of the 99 surrogate curves had an area equal or larger than the original ROC curve, implying that it is unlikely ($P < 0.01$) that the responses to Jennifer Aniston were obtained by chance. A responsive unit was defined to have an invariant representation if the area under the ROC curve was larger than the area of the 99 surrogate curves.

Figure 2 shows another single unit located in the right anterior hippocampus of a different patient. This unit was selectively activated by pictures of the actress Halle Berry as well as by a drawing of her (but not by other drawings; for example, picture no. 87). This unit was also activated by several pictures of Halle Berry dressed as Catwoman, her character in a recent film, but not by other images of Catwoman that were not her (data not shown). Notably, the unit was selectively activated by the letter string 'Halle Berry'. Such an invariant pattern of activation goes beyond common visual features of the different stimuli. As with the previous unit, the responses were mainly localized between 300 and 600 ms after stimulus onset.

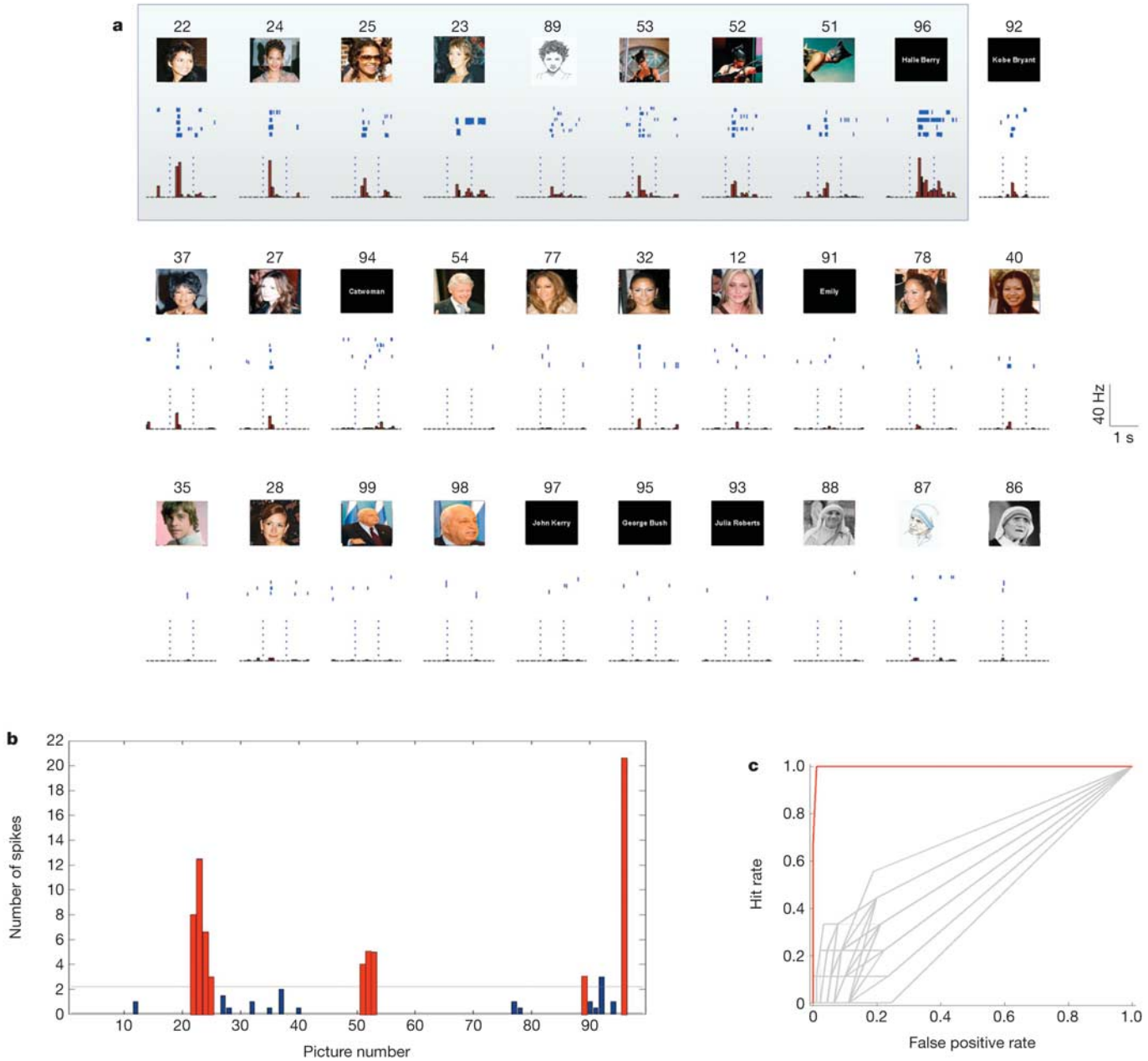


Figure 2 | A single unit in the right anterior hippocampus that responds to pictures of the actress Halle Berry (conventions as in Fig. 1). **a**–**c**, Strikingly, this cell also responds to a drawing of her, to herself dressed as Catwoman (a recent movie in which she played the lead role) and to the letter string 'Halle Berry' (picture no. 96). Such an invariant response cannot be attributed to common visual features of the stimuli. This unit also had a very low baseline firing rate (0.06 spikes). The area under the red curve in **c** is 0.99.

Figure 2c shows the ROC curve for the pictures of Halle Berry (red trace) and for 99 surrogates (grey lines). The area under the ROC curve was 0.99, larger than that of the surrogates.

Figure 3 illustrates a multi-unit in the left anterior hippocampus responding to pictures of the Sydney Opera House and the Baha'i Temple. Because the patient identified both landmark buildings as the Sydney Opera House, all these pictures were considered as a single landmark building for the ROC analysis. This unit also responded to the letter string 'Sydney Opera' (pictures no. 2 and 8) but not to other letter strings, such as 'Eiffel Tower' (picture no. 1). More examples of invariant responses are shown in the Supplementary Figs 2–11.

Out of the 132 responsive units, 51 (38.6%; 30 single units and 21 multi-units) showed invariance to a particular individual (38 units responding to Jennifer Aniston, Halle Berry, Julia Roberts, Kobe Bryant, and so on), landmark building (6 units responding to the Tower of Pisa, the Baha'i Temple and the Sydney Opera House), animal (5 units responding to spiders, seals and horses) or object (2 units responding to specific food items), with $P < 0.01$ as defined above by means of the surrogate tests. A one-way analysis of variance (ANOVA) yielded similar results (see Methods). Eight of these units (two single units and six multi-units) responded to two different individuals (or to an individual and an object). Figure 4 presents the distribution of the areas under the ROC curves for all 51 units that showed an invariant representation to individuals or objects. The areas ranged from 0.76 to 1.00, with a median of 0.94. These units were located in the hippocampus (27 out of 60 responsive units; 45%), parahippocampal gyrus (11 out of 20 responsive units; 55%), amygdala (8 out of 30 responsive units; 27%) and entorhinal cortex

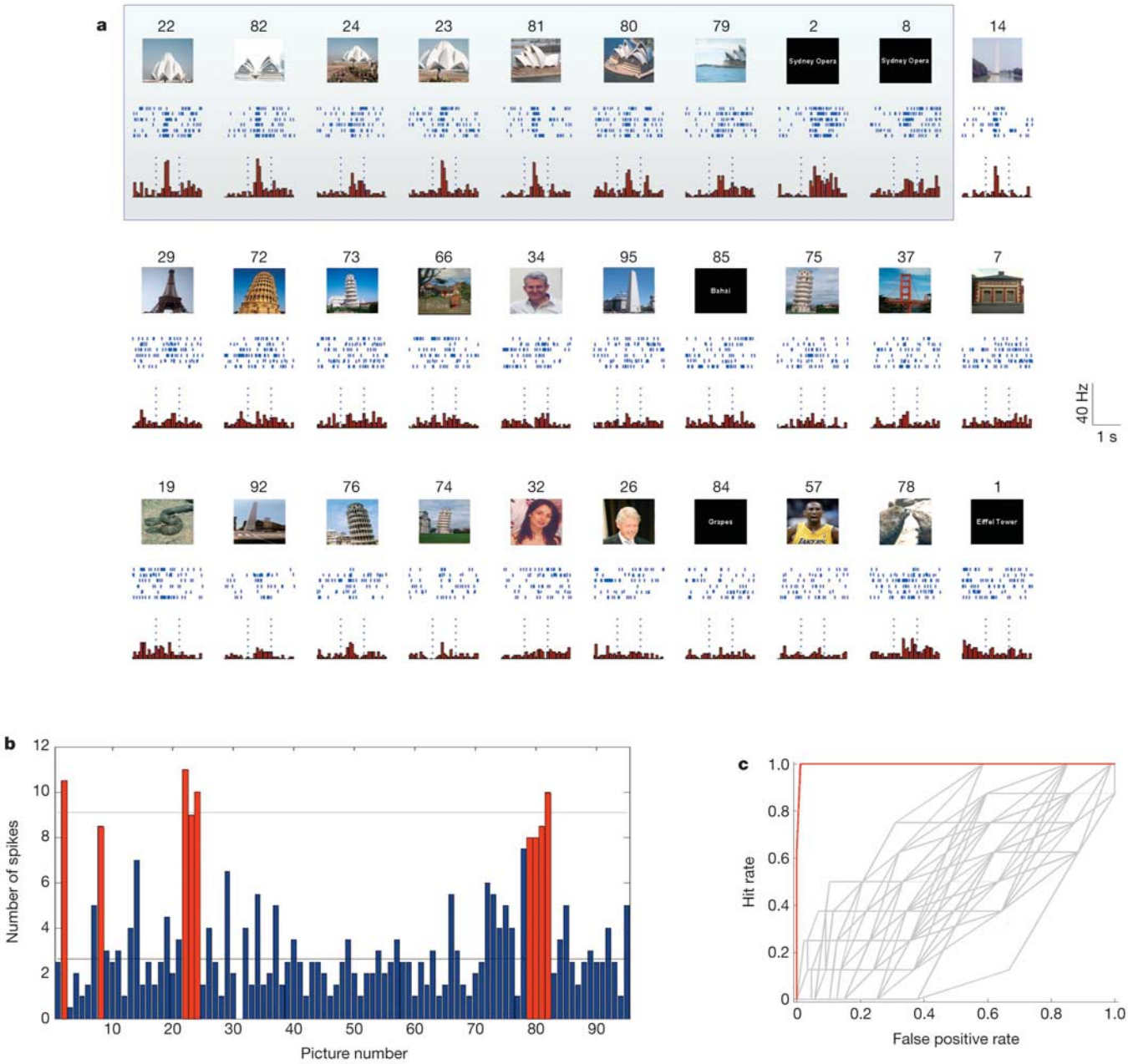


Figure 3 | A multi-unit in the left anterior hippocampus that responds to photographs of the Sydney Opera House and the Baha'i Temple (conventions as in Fig. 1). a–c, The patient identified all pictures of both of these buildings as the Sydney Opera, and we therefore considered them as a single landmark. This unit also responded to the presentation of the letter string 'Sydney Opera' (pictures no. 2 and 8), but not to other strings, such as 'Eiffel Tower' (picture no. 1). In contrast to the previous two figures, this unit had a higher baseline firing rate (2.64 spikes). The area under the red curve in **c** is 0.97.

(5 out of 22 responsive units; 23%). There were no clear differences in the latencies and firing patterns among the different areas. However, more data are needed before making a conclusive claim about systematic differences between the various structures of the MTL.

As shown in Figs 2 and 3, one of the most extreme cases of an abstract representation is the one given by responses to pictures of a particular individual (or object) and to the presentation of the corresponding letter string with its name. In 18 of the 21 testing sessions we also tested responses to letter strings with the names of the individuals and objects. Eight of the 132 responsive units (6.1%) showed a selective response to an individual and its name (with no response to other names). Six of these were in the hippocampus, one was in the entorhinal cortex and one was in the amygdala.

These neuronal responses cannot be attributed to any particular movement artefact, because selective responses started around 300 ms after image onset, whereas key presses occurred at 1 s or later, and neuronal responses were very selective. About one-third of the responsive units had a response localized between 300 and 600 ms. This interval corresponds to the latency of event-related responses correlated with the recognition of 'oddball' stimuli in scalp electroencephalogram, namely, the P300 (ref. 16). Some studies argue for a generation of the P300 in the hippocampal formation and amygdala[17,18], consistent with our findings.

What are the common features that activate these neurons? Given the great diversity of distinct images of a single individual (pencil sketches, caricatures, letter strings, coloured photographs with different backgrounds) that these cells can selectively respond to, it is unlikely that this degree of invariance can be explained by a simple set of metric features common to these images. Indeed, our data are compatible with an abstract representation of the identity of the individual or object shown. The existence of such high-level visual responses in medial temporal lobe structures, usually considered to be involved in long-term memory formation and consolidation, should not be surprising given the following: (1) the known anatomical connections between the higher stages of the visual hierarchy in the ventral pathway and the MTL[19,20]; (2) the well-characterized reactivity of the cortical stages feeding into the MTL to the sight of faces, objects, or spatial scenes (as ascertained using functional magnetic resonance imaging (fMRI) in humans[21,22] and electrophysiology in monkeys[2,4,7–11]); and (3) the observation that any visual percept that will be consciously remembered later on will have to be represented in the hippocampal system[23–25]. This is true even though patients with bilateral loss of parts of the MTL do not, in general, have a deficit in the perception of images[25]. Neurons in the MTL might have a fundamental role in learning associations between abstract representations[26]. Thus, our observed invariant responses probably arise from experiencing very different pictures, words or other visual stimuli in association with a given individual or object.

How neurons encode different percepts is one of the most intriguing questions in neuroscience. Two extreme hypotheses are schemes based on the explicit representations by highly selective (cardinal, gnostic or grandmother) neurons and schemes that rely on an implicit representation over a very broad and distributed population of neurons[1–4,6]. In the latter case, recognition would require the simultaneous activation of a large number of cells and therefore we would expect each cell to respond to many pictures with similar basic features. This is in contrast to the sparse firing we observe, because most MTL cells do not respond to the great majority of images seen by the patient. Furthermore, cells signal a particular individual or object in an explicit manner[27], in the sense that the presence of the individual can, in principle, be reliably decoded from a very small number of neurons. We do not mean to imply the existence of single neurons coding uniquely for discrete percepts for several reasons: first, some of these units responded to pictures of more than one individual or object; second, given the limited duration of our recording sessions, we can only explore a tiny portion of stimulus space; and third, the fact that we can discover in this short time some images—such as photographs of Jennifer Aniston—that drive the cells suggests that each cell might represent more than one class of images. Yet, this subset of MTL cells is selectively activated by different views of individuals, landmarks, animals or objects. This is quite distinct from a completely distributed population code and suggests a sparse, explicit and invariant encoding of visual percepts in MTL. Such an abstract representation, in contrast to the metric representation in the early stages of the visual pathway, might be important in the storage of long-term memories. Other factors, including emotional responses towards some images, could conceivably influence the neuronal activity as well. The responses of these neurons are reminiscent of the behaviour of hippocampal place cells in rodents[28] that only fire if the animal moves through a particular spatial location, with the actual place field defined independently of sensory cues. Notably, place cells have been found recently in the human hippocampus as well[29]. Both classes of neurons—place cells and the cells in the present study—have a very low baseline activity and respond in a highly selective manner. Future research might show that this similarity has functional implications, enabling mammals to encode behaviourally important features of the environment and to transition between them, either in physical space or in a more conceptual space[13].

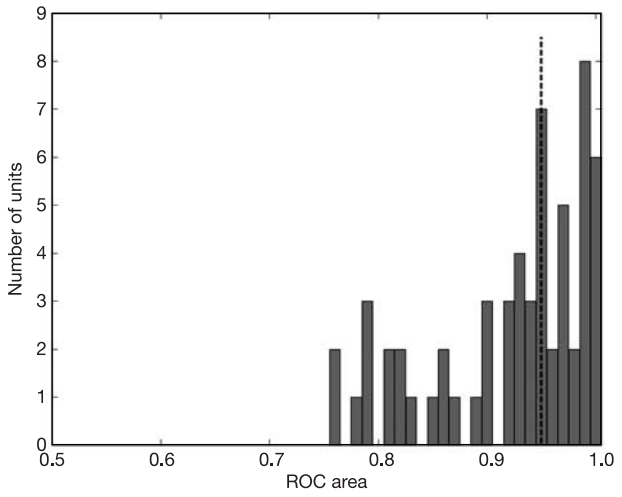


Figure 4 | Distribution of the area under the ROC curves for the 51 units (out of 132 responsive units) showing an invariant representation. Of these, 43 responded to a single individual or object and 8 to two individuals or objects. The dashed vertical line marks the median of the distribution (0.94).

METHODS

The data in the present study come from 21 sessions in 8 patients with pharmacologically intractable epilepsy (eight right handed; 3 male; 17–47 years old). Extensive non-invasive monitoring did not yield concordant data corresponding to a single resectable epileptogenic focus. Therefore, the patients were implanted with chronic depth electrodes for 7–10 days to determine the seizure focus for possible surgical resection[13]. Here we report data from sites in the hippocampus, amygdala, entorhinal cortex and parahippocampal gyrus. All studies conformed to the guidelines of the Medical Institutional Review Board at UCLA. The electrode locations were based exclusively on clinical criteria and were verified by MRI or by computer tomography co-registered to preoperative MRI. Each electrode probe had a total of nine micro-wires at its end[13], eight active recording channels and one reference. The differential signal from the micro-wires was amplified using a 64-channel Neuralynx system, filtered between 1 and 9,000 Hz. We computed the power spectrum for every unit after spike sorting. Units that showed evidence of line noise were excluded from subsequent analysis[14]. Signals were sampled at 28 kHz. Each recording session lasted about 30 min.

Subjects lay in bed, facing a laptop computer. Each image covered about 1.5° and was presented at the centre of the screen six times for 1 s. The order of the pictures was randomized. Subjects had to respond, after image offset, according to whether the picture contained a human face or something else by pressing the

'Y' and 'N' keys, respectively. This simple task, on which performance was virtually flawless, required them to attend to the pictures. After the experiments, patients gave feedback on whether they recognized the images or not. Pictures included famous and unknown individuals, animals, landmarks and objects. We tried to maximize the differences between pictures of the individuals (for example, different clothing, size, point of view, and so on). In 18 of the 21 sessions, we also presented letter strings with names of individuals or objects.

The data from the screening sessions were rapidly processed to identify responsive units and images. All pictures that elicited a response in the screening session were included in the later testing sessions. Three to eight different views of seven to twenty-three different individuals or objects were used in the testing sessions with a mean of 88.6 images per session (range 70–110). Spike detection and sorting was applied to the continuous recordings using a novel clustering algorithm[30] (see Supplementary Information). The response to a picture was defined as the median number of spikes across trials between 300 and 1,000 ms after stimulus onset. Baseline activity was the average spike count for all pictures between 1,000 and 300 ms before stimulus onset. A unit was considered responsive if the activity to at least one picture fulfilled two criteria: (1) the median number of spikes was larger than the average number of spikes for the baseline plus 5 s.d.; and (2) the median number of spikes was at least two.

The classification between single unit and multi-unit was done visually based on the following: (1) the spike shape and its variance; (2) the ratio between the spike peak value and the noise level; (3) the inter-spike interval distribution of each cluster; and (4) the presence of a refractory period for the single units (that is, less than 1% of spikes within less than 3 ms inter-spike interval).

Whenever a unit had a response to a given stimulus, we further analysed the responses to other pictures of the same individual or object by a ROC analysis. This tested whether cells responded selectively to pictures of a given individual. The hit rate (y axis) was defined as the number of responses to the individual divided by the total number of pictures of this individual. The false positive rate (x axis) was defined as the number of responses to the other pictures divided by the total number of other pictures. The ROC curve was obtained by gradually lowering the threshold of the responses (the median number of spikes in Figs 1b, 2b and 3b). Starting with a very high threshold (no hits, no false positives, lower left-hand corner in the ROC diagram), if a unit responds exclusively to an image of a particular individual or object, the ROC curve will show a steep increase when lowering the threshold (a hit rate of 1 and no false positives). If a unit responds to a random selection of pictures, it will have a similar relative number of hits and false positives and the ROC curve will fall along the diagonal. In the first case, for a highly invariant unit, the area under the ROC curve will be close to 1, whereas in the latter case it will be about 0.5. To evaluate the statistical significance, we created 99 surrogate curves for each responsive unit, testing the null hypothesis that the unit responded preferentially to n randomly chosen pictures (with n being the number of pictures of the individual for which invariance was tested). A unit was considered invariant to a certain individual or object if the area under the ROC curve was larger than the area of all of the 99 surrogates (that is, with a confidence of $P < 0.01$). Alternatively, the ROC analysis can be done with the single trial responses instead of the median responses across trials. Here, responses to the trials corresponding to any picture of the individual tested are considered as hits and responses to trials to other pictures as false positives. This trial-by-trial analysis led to very similar results, with 55 units of all 132 responsive units showing an invariant representation. A one-way ANOVA also yielded similar results. In particular, we tested whether the distribution of median firing rates for all responsive units showed a dependence on the factor identity (that is, the individual, landmark or object shown). The different views of each individual were the repeated measures. As with the ROC analysis, an ANOVA test was performed on all responsive units. Overall, the results were very similar to those obtained with the ROC analysis: of 132 responsive units, 49 had a significant effect for factor identity with $P < 0.01$, compared to 51 units showing an invariant representation with the ROC analysis. The ANOVA analysis, however, does not demonstrate that the invariant responses were very selective, whereas the ROC analysis explicitly tests the presence of an invariant as well as sparse representation.

Images were obtained from Corbis and Photorazzi, with licensed rights to reproduce them in this paper and in the Supplementary Information.

Received 1 December 2004; accepted 3 February 2005.

1. Barlow, H. Single units and sensation: a neuron doctrine for perception. *Perception* **1,** 371–394 (1972).
2. Gross, C. G., Bender, D. B. & Rocha-Miranda, C. E. Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science* **166,** 1303–1306 (1969).
3. Konorski, J. *Integrative Activity of the Brain* (Univ. Chicago Press, Chicago, 1967).
4. Logothetis, N. K. & Sheinberg, D. L. Visual object recognition. *Annu. Rev. Neurosci.* **19,** 577–621 (1996).
5. Riesenhuber, M. & Poggio, T. Neural mechanisms of object recognition. *Curr. Opin. Neurobiol.* **12,** 162–168 (2002).
6. Young, M. P. & Yamane, S. Sparse population coding of faces in the inferior temporal cortex. *Science* **256,** 1327–1331 (1992).
7. Logothetis, N. K., Pauls, J. & Poggio, T. Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.* **5,** 552–563 (1995).
8. Logothetis, N. K. & Pauls, J. Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb. Cortex* **3,** 270–288 (1995).
9. Perrett, D., Rolls, E. & Caan, W. Visual neurons responsive to faces in the monkey temporal cortex. *Exp. Brain Res.* **47,** 329–342 (1982).
10. Schwartz, E. L., Desimone, R., Albright, T. D. & Gross, C. G. Shape recognition and inferior temporal neurons. *Proc. Natl Acad. Sci. USA* **80,** 5776–5778 (1983).
11. Tanaka, K. Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* **19,** 109–139 (1996).
12. Miyashita, Y. & Chang, H. S. Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature* **331,** 68–71 (1988).
13. Fried, I., MacDonald, K. A. & Wilson, C. Single neuron activity in human hippocampus and amygdale during recognition of faces and objects. *Neuron* **18,** 753–765 (1997).
14. Kreiman, G., Koch, C. & Fried, I. Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature Neurosci.* **3,** 946–953 (2000).
15. Macmillan, N. A. & Creelman, C. D. *Detection Theory: A User's Guide* (Cambridge Univ. Press, New York, 1991).
16. Picton, T. The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* **9,** 456–479 (1992).
17. Halgren, E., Marinkovic, K. & Chauvel, P. Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalogr. Clin. Neurophysiol.* **106,** 156–164 (1998).
18. McCarthy, G., Wood, C. C., Williamson, P. D. & Spencer, D. D. Task-dependent field potentials in human hippocampal formation. *J. Neurosci.* **9,** 4253–4268 (1989).
19. Saleem, K. S. & Tanaka, K. Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. *J. Neurosci.* **16,** 4757–4775 (1996).
20. Suzuki, W. A. Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: Organization of cortical inputs and interconnections with amygdale and striatum. *Seminar Neurosci.* **8,** 3–12 (1996).
21. Kanwisher, N., McDermott, J. & Chun, M. M. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17,** 4302–4311 (1997).
22. Haxby, J. V. *et al.* Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293,** 2425–2430 (2001).
23. Eichenbaum, H. A cortical-hippocampal system for declarative memory. *Nature Rev. Neurosci.* **1,** 41–50 (2000).
24. Hampson, R. E., Pons, P. P., Stanford, T. R. & Deadwyler, S. A. Categorization in the monkey hippocampus: A possible mechanism for encoding information into memory. *Proc. Natl Acad. Sci. USA* **101,** 3184–3189 (2004).
25. Squire, L. R., Stark, C. E. L. & Clark, R. E. The medial temporal lobe. *Annu. Rev. Neurosci.* **27,** 279–306 (2004).
26. Mishashita, Y. Neuronal correlate of visual associative long-term memory in the primate temporal cortex. *Nature* **335,** 817–820 (1988).
27. Koch, C. *The Quest for Consciousness: A Neurobiological Approach* (Roberts, Englewood, Colorado, 2004).
28. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261,** 1055–1058 (1993).
29. Ekstrom, A. D. *et al.* Cellular networks underlying human spatial navigation. *Nature* **425,** 184–187 (2003).
30. Quian Quiroga, R., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and super-paramagnetic clustering. *Neural Comput.* **16,** 1661–1687 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank all patients for their participation; P. Sinha for drawing some faces; colleagues for providing pictures; I. Wainwright for administrative assistance; and E. Behnke, T. Fields, E. Ho, E. Isham, A. Kraskov, P. Steinmetz, I. Viskontas and C. Wilson for technical assistance. This work was supported by grants from the NINDS, NIMH, NSF, DARPA, the Office of Naval Research, the W.M. Keck Foundation Fund for Discovery in Basic Medical Research, a Whiteman fellowship (to G.K.), the Gordon Moore Foundation, the Sloan Foundation, and the Swartz Foundation for Computational Neuroscience.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and request for materials should be addressed to R.Q.Q. (rodri@vis.caltech.edu).

Supplementary material

Spike detection and sorting

From the continuous wide-band data, spike detection and sorting was carried out by a novel and relatively fast algorithm [30]. Briefly, once spikes are detected via amplitude thresholding, the algorithm uses the wavelet transform to extract features of the spike shapes that are used as inputs to the clustering algorithm. Clustering is done by means of super-paramagnetic clustering, a method from statistical mechanics that does not assume any particular distribution of the clusters. Super-paramagnetic clustering groups the data into clusters as a function of a single parameter, the temperature. In analogy with spin glasses in statistical mechanics, for low temperatures all the data is grouped into a single cluster and for high temperatures, the data is split into many clusters with few members each. Figure S1 shows an example of the clustering outcome for a single channel. A Matlab implementation of the algorithm as well as exemplary data is available from www.vis.caltech.edu/~rodri. After a first unsupervised processing of the data, the temperature is the main parameter that can be changed by the user if the automatic clustering is not satisfactory. At times, we combine results from two different temperatures and/or we assign membership of unclustered spikes to nearby clusters via template matching.

Subsequently, we classify the clusters into single- or multi-units. Multi-unit clusters are those reflecting the activity from several neurons whose spikes can not be further differentiated due to their low signal to noise ratio. The classification between single- and multi-unit was done visually based on: 1) the spike shape and its variance; 2) the ratio between the spike peak value and the noise level; 3) the ISI distribution of each cluster; 4) the presence of a refractory period for the single-units; *i.e.* less than 1% spikes within less than 3ms inter-spike-interval. In Figure S1, the first (blue) cluster corresponds to a multi-unit and the other 3 to single-units.

Difference to our previous work

We reported before the presence of units in the human medial temporal lobe that responded to faces in comparison to objects [13] and to different categories of stimuli [14]. Over the last three years we introduced several modifications to optimize the recording conditions:

(i) we devised a novel spike sorting algorithm that allowed the detection of more units than before [30]; (ii) in previous studies spikes were detected on-line via amplitude thresholding and the very selective units were usually not detected because they remained silent at the beginning of the experiment when the thresholds were set; (iii) we increased the number of electrodes that we could record from 16 to 64; (iv) the access to the continuous data allowed us to estimate the noise level, thus being able to differentiate the multiunit activity from noise (something hard to do with the previous setup where only the spike shapes were stored); finally (v) the screening/testing session design allowed us to present multiple views of specific individuals and objects.

More examples of invariant responses

In order to increase the probability of finding invariant cells, we first explored cellular responses by using screening sessions. In these screening sessions, pictures of different individuals, objects, animals and landmarks were shown. After a fast analysis of the data, we identified the pictures eliciting responses in at least one unit. Then, invariance was tested in later testing sessions by showing between 3 and 8 pictures of these individuals or objects. As discussed above, the neuronal selectivity was statistically the same in both sessions. All the images that elicited a response in the screening session were included in testing sessions.

Figures S2 to S11 correspond to cells recorded during testing sessions following earlier screening sessions. These figures complement the examples shown in the main text.

Figure Legends (Supplementary Material)

Figure S1: The spike sorting algorithm at work. Upper subplot: 60 sec of continuous, bandpass-filtered data and the threshold used for detection (red line). Middle subplots: spike shapes of clustered units and their distribution in a 2-D space of wavelet coefficients (leftmost subplot). Lower subplots: Inter-spike interval distribution of clusters and number of spikes in each cluster as a function of the temperature (leftmost subplot; see text). The optimal temperature chosen by the algorithm is marked by the vertical dotted line. Spike amplitudes and amplitude of the continuous data (y-axes) are in μV.

Figure S2: A multi-unit in the right posterior hippocampus recorded in the same session as the unit shown in Figure 1. Conventions for this and the following figures are the same as in Fig. 1. The order of trial number in the raster plots is from top to bottom. Due to insurmountable copyright problems, in this and subsequent figures, all original images were replaced by very similar ones (same subject, animal or building, similar pose, similar colour, line drawing etc).This unit responded selectively to pictures of Jennifer Aniston together with Brad Pitt. This unit did not respond significantly to pictures of Aniston alone.

Figure S3: A single-unit in the right posterior hippocampus that responded to pictures of the actress Pamela Anderson, including a caricature of her and the letter string “Pamela Anderson”, but not to other letter strings. Note that it is difficult to identify common visual features that would explain this invariant response.

Figure S4: A single-unit in the left posterior hippocampus that responded to pictures of the basketball player Kobe Bryant but also to other images. The area under the ROC curve was 0.78, which is close to the lowest value of the distribution of invariant cells (Fig. 4). Noteworthy, this unit was recorded from the same micro-wire as the one in Fig. 1. This stresses the importance of spike sorting: without good spike sorting, we would have observed just one multi-unit responding both to Bryant and to Aniston. Furthermore, this suggests that nearby units can show different selective responses.

Figure S5: A single-unit in the right anterior hippocampus that responded to pictures of the Bahai Temple and the Sydney Opera. Since the patient did not distinguish these two buildings, we considered them as a single landmark. For this cell, there was also an invariant representation of pictures of Aniston together with Brad Pitt (different from the one shown in Figure S2). Responses are extremely sparse, but consistent across trials. **(C)** ROC curve corresponding to the Sydney Opera/Bahai Temple (green) and Aniston with Pitt (red). Light traces are the surrogates for each case.

Figure S6: A unit in the right posterior hippocampus that responded preferentially to pictures of the Tower of Pisa. There was also a significant response to the Eiffel tower (picture #18).

Figure S7: A single-unit in the left posterior hippocampus that responded to pictures of Jennifer Aniston (different from Figure 1 in the main text). This unit also responded to a similar looking actress (picture #9) that appeared in a famous TV series ("Friends") with Jennifer Aniston.

Figure S8: A single-unit in the right anterior hippocampus responding preferentially to pictures of Mother Teresa. This cell was recorded from the same microwire as the cell shown in Figure 2. This again stresses the importance of spike sorting, since there are no responses to Halle Berry for this cell and no responses to Mother Teresa for the cell shown in Figure 2.

Figure S9: A single-unit in the right anterior hippocampus responding selectively to pictures of Julia Roberts and to the letter string with her name.

Figure S10: A single unit recorded from the same microwire as the one shown in the previous figure. This unit was selectively activated by pictures of Kobe Bryant.

Figure S11: A single-unit in the right anterior hippocampus responding preferentially to pictures of Catwoman. Interestingly, this cell also responded to other animal pictures.

Fig. S1

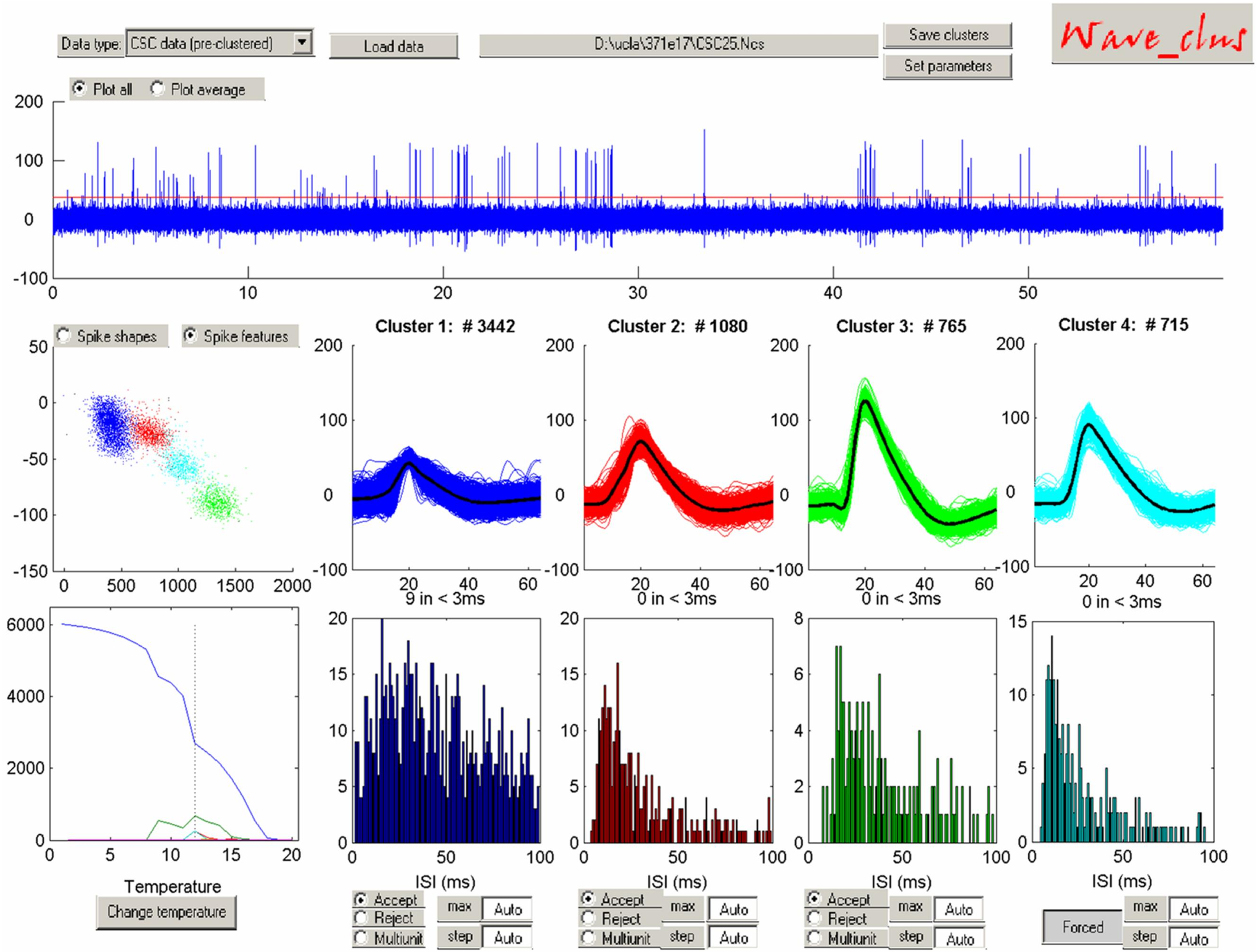
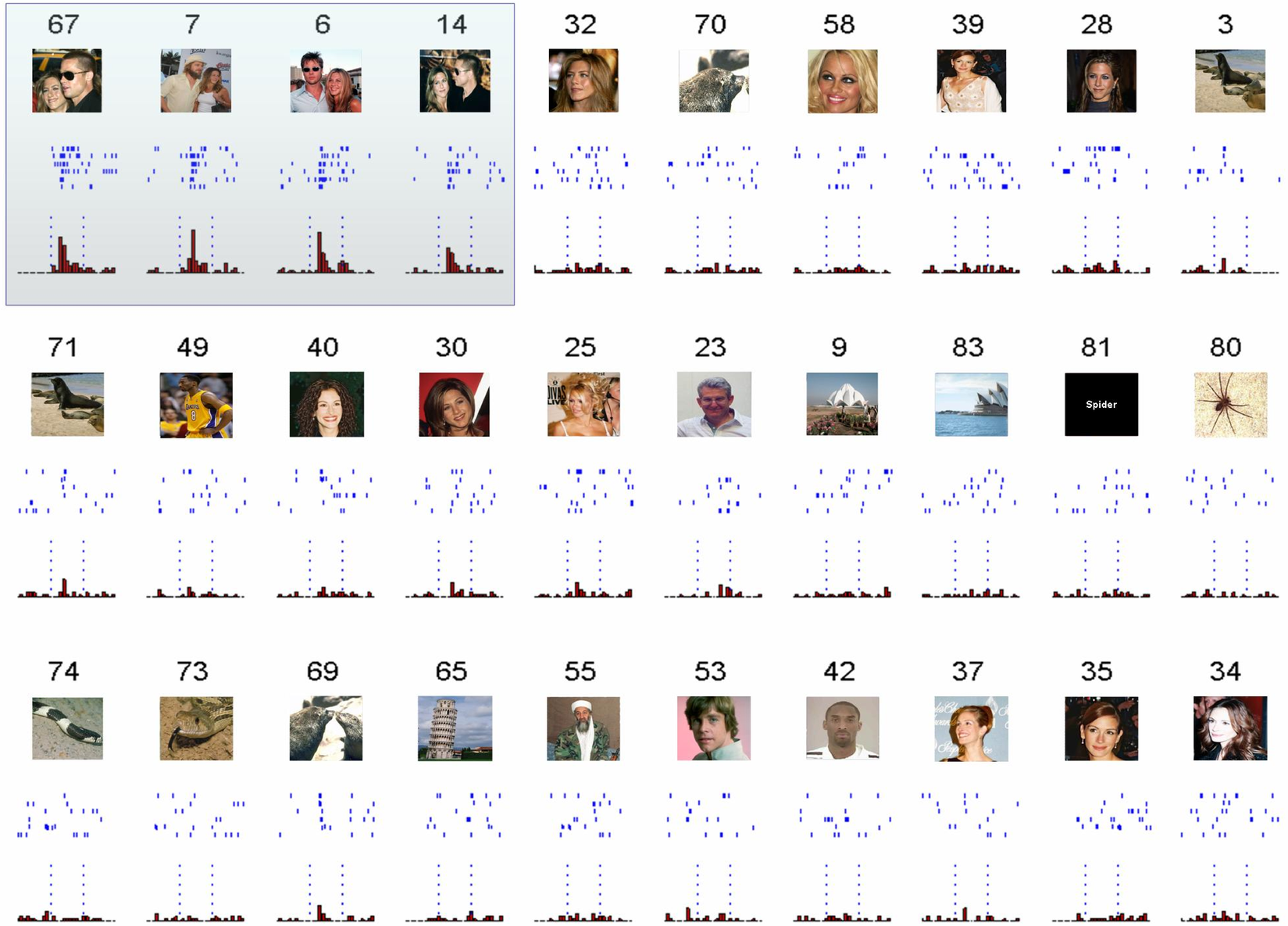
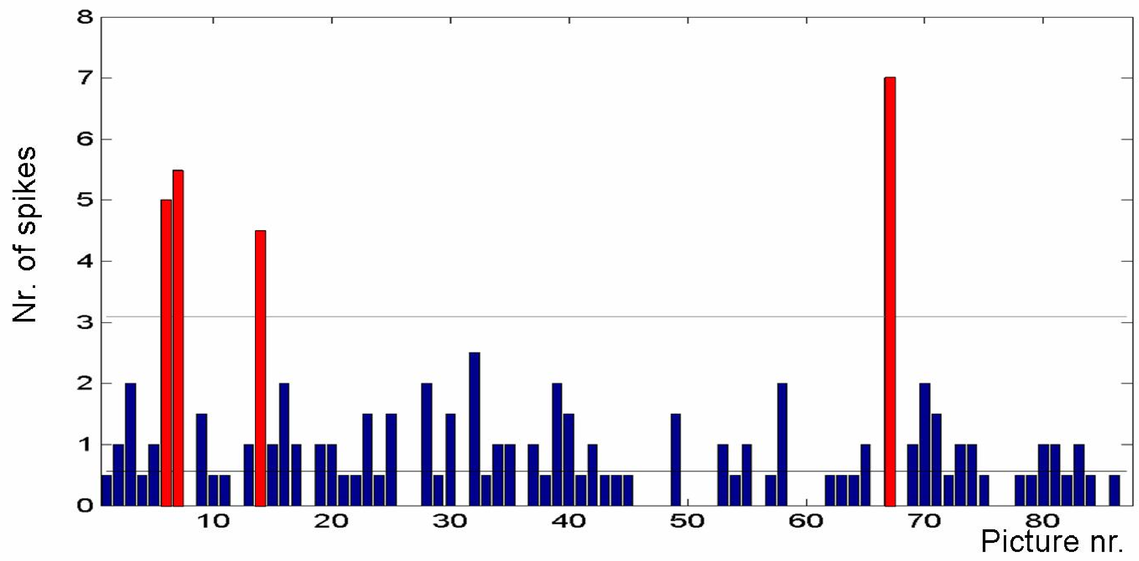

Data type: CSC data (pre-clustered)
Load data
D:\ucla\371e17\CSC25.Ncs
Save clusters
Set parameters
Wave_clus
Plot all
Plot average
Spike shapes
Spike features
Cluster 1: # 3442
Cluster 2: # 1080
Cluster 3: # 765
Cluster 4: # 715
9 in < 3ms
0 in < 3ms
0 in < 3ms
0 in < 3ms
ISI (ms)
Temperature
Change temperature
Accept
Reject
Multiunit
max
Auto
step
Auto
Forced

Fig. S2

A)


67
7
6
14
32
70
58
39
28
3
71
49
40
30
25
23
9
83
81
Spider
80
74
73
69
65
55
53
42
37
35
34
40 Hz
1sec

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
7
8
10
20
30
40
50
60
70
80

C)

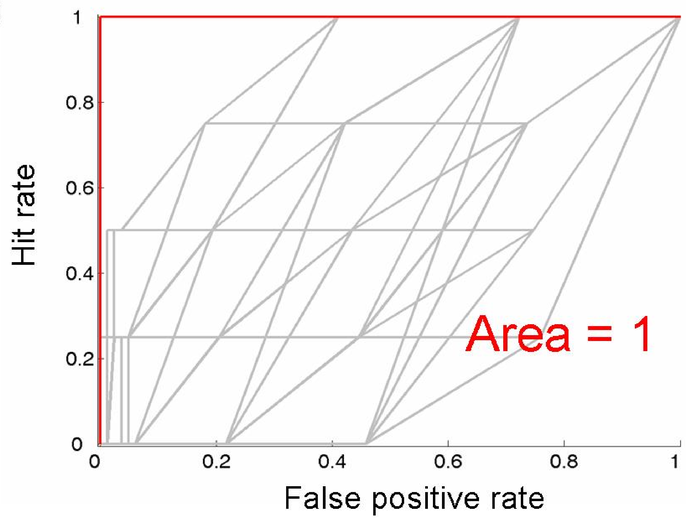
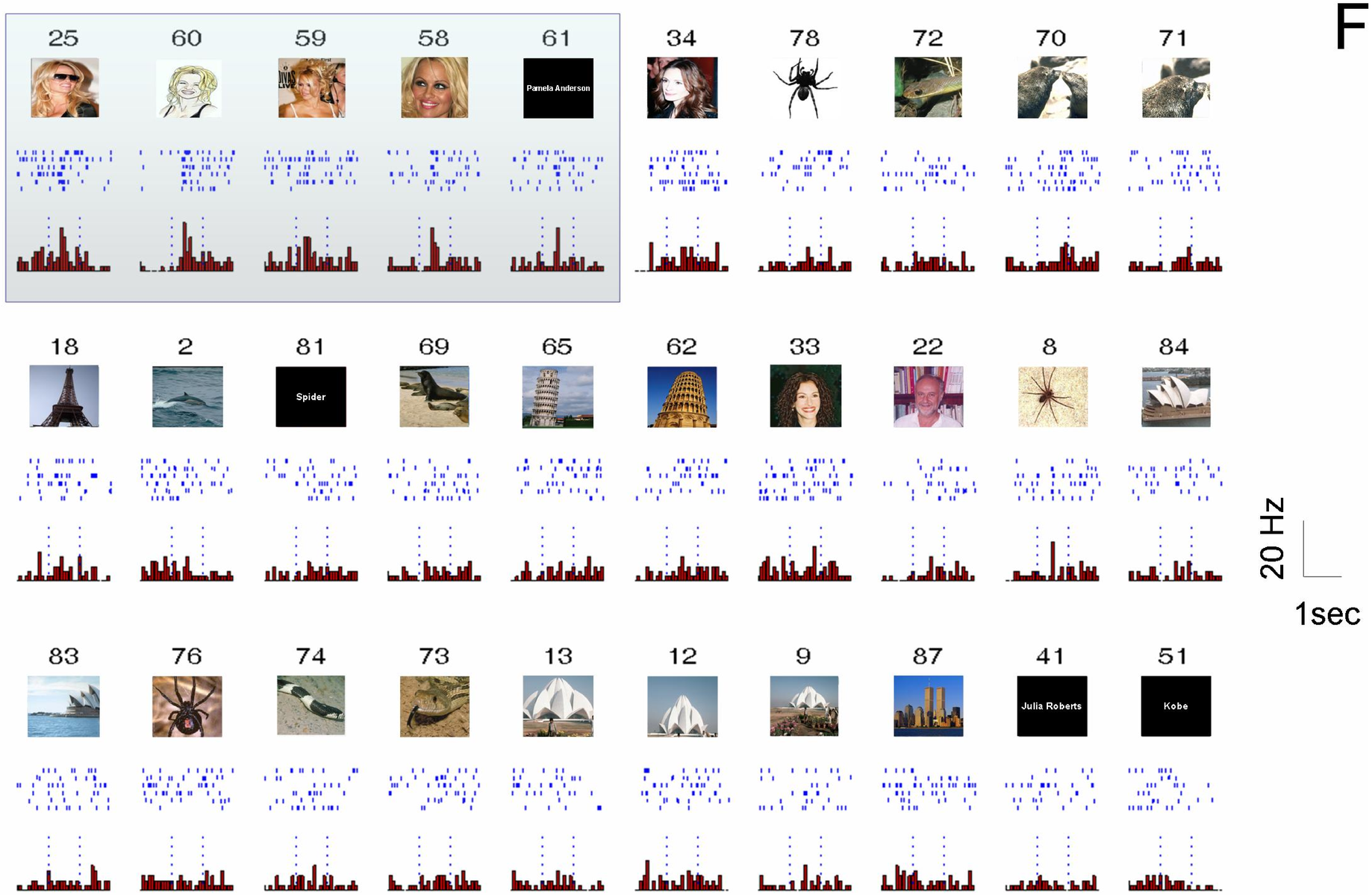
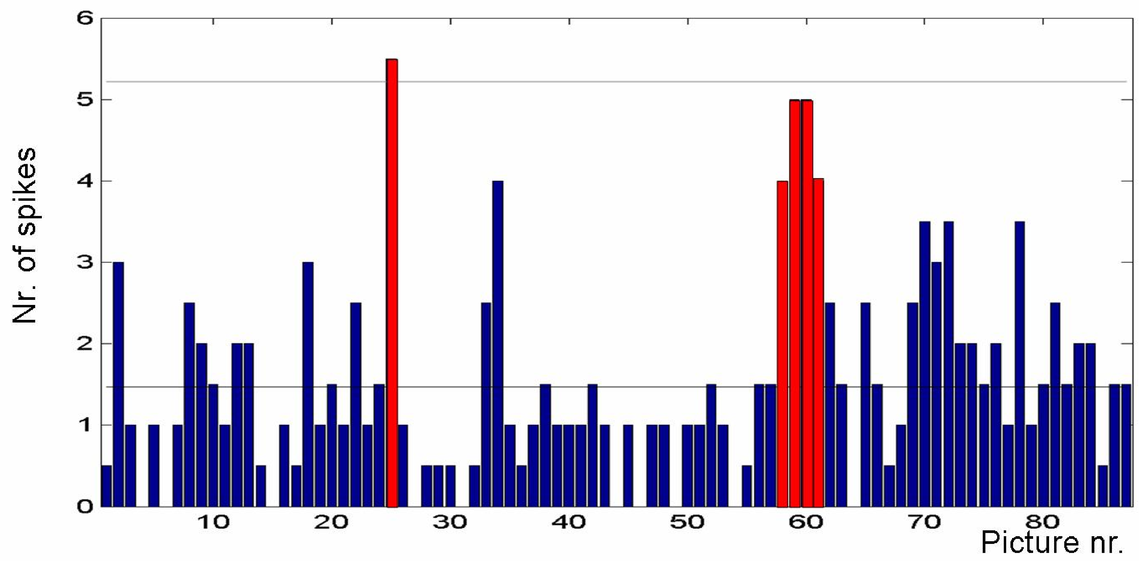

Hit rate
False positive rate
Area = 1
0
0.2
0.4
0.6
0.8
1

Fig. S3

A)


25
60
59
58
61
Pamela Anderson
34
78
72
70
71
18
2
81
Spider
69
65
62
33
22
8
84
20 Hz
1sec
83
76
74
73
13
12
9
87
41
Julia Roberts
51
Kobe

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
10
20
30
40
50
60
70
80

C)

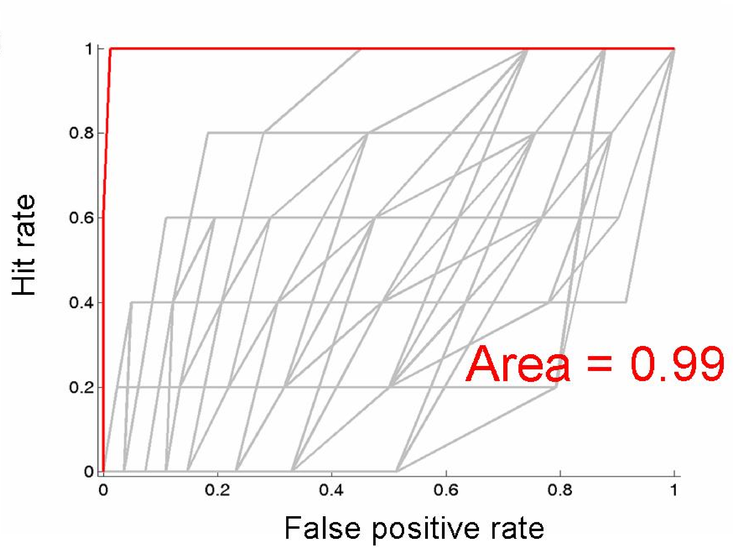
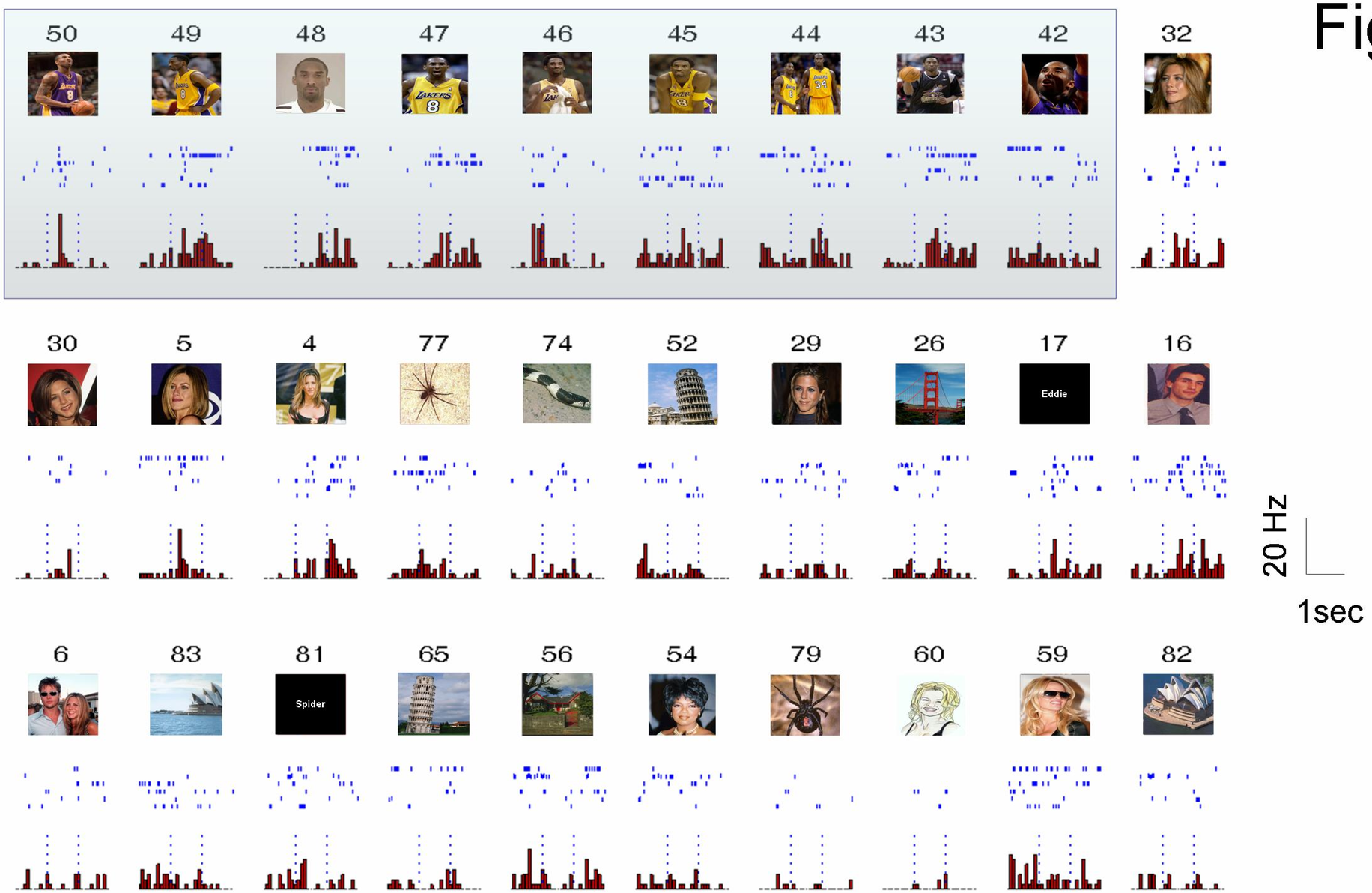
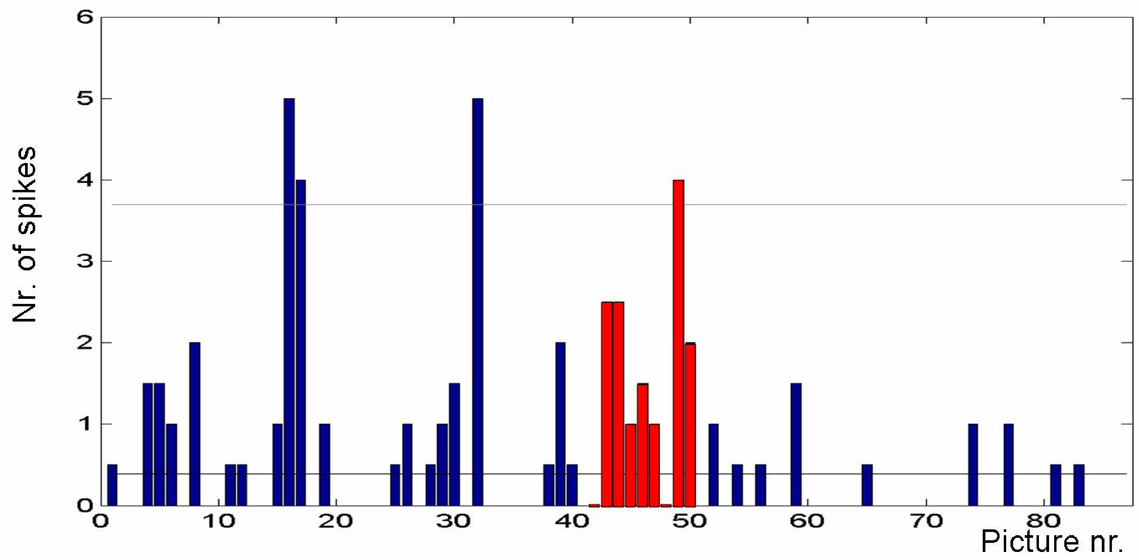

Hit rate
False positive rate
Area = 0.99
0
0.2
0.4
0.6
0.8
1

Fig. S4

A)


50
49
48
47
46
45
44
43
42
32
30
5
4
77
74
52
29
26
17
16
Eddie
20 Hz
1sec
6
83
81
65
56
54
79
60
59
82
Spider

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
0
10
20
30
40
50
60
70
80

C)

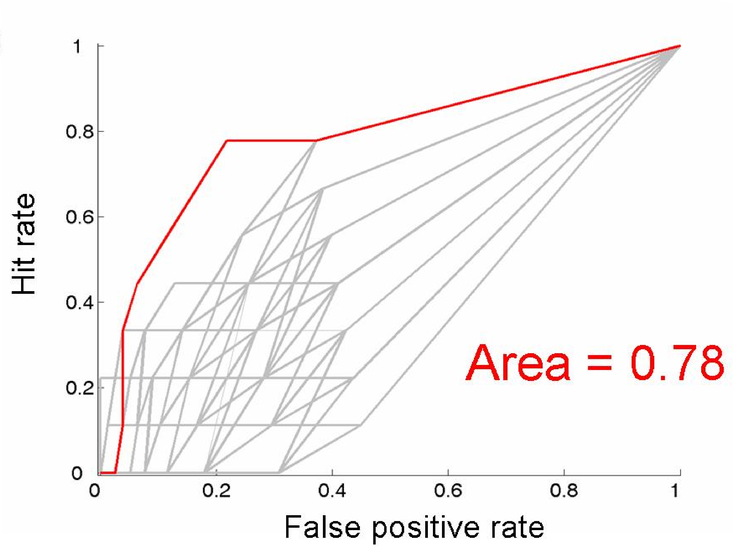
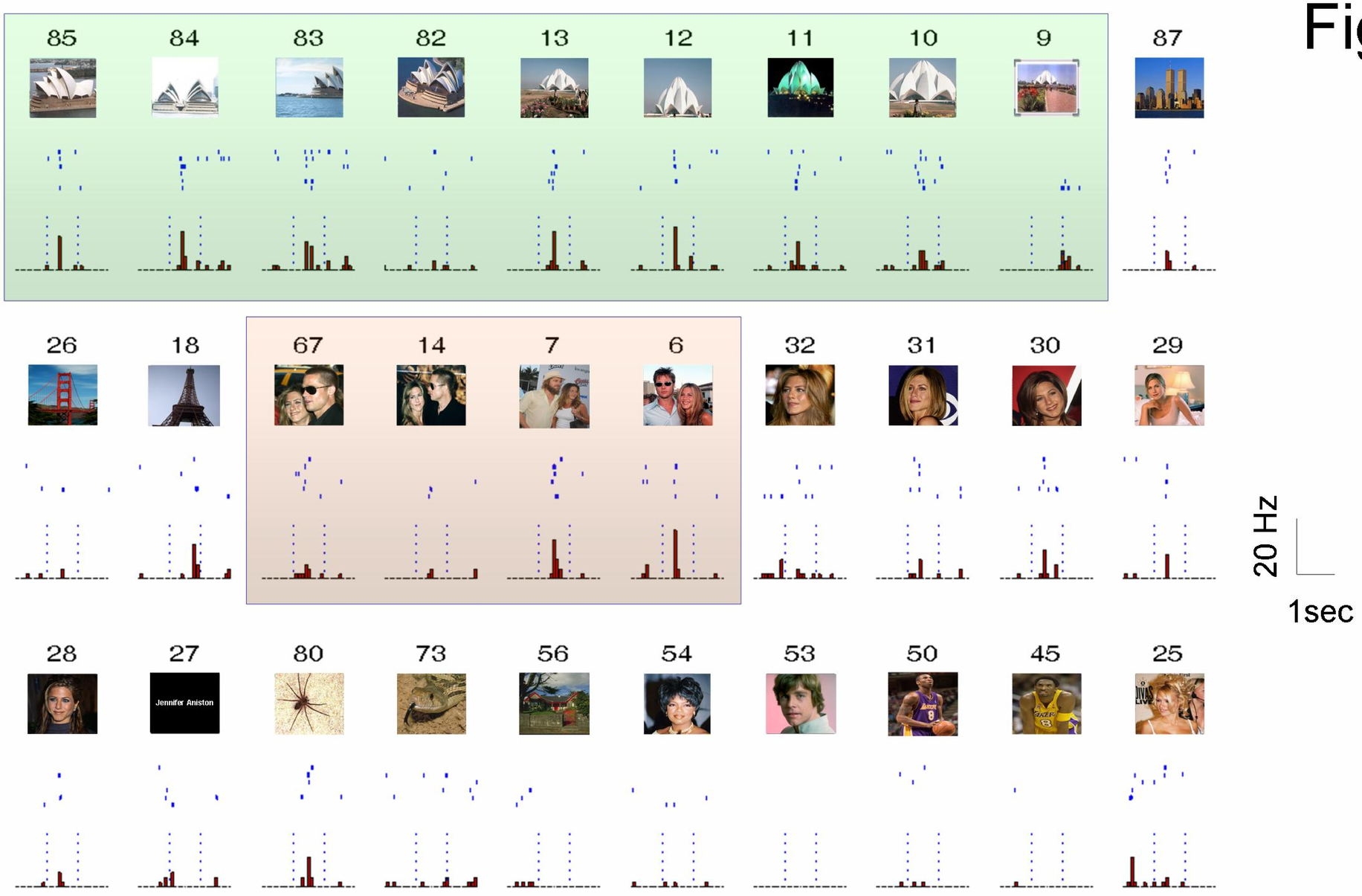
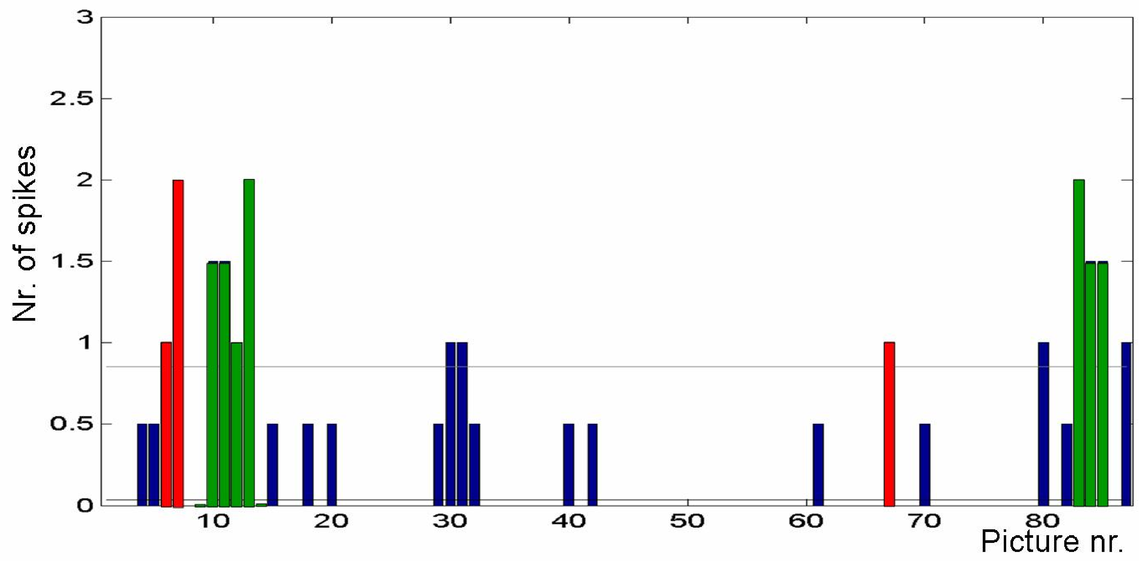

Hit rate
False positive rate
Area = 0.78
0
0.2
0.4
0.6
0.8
1

Fig. S5

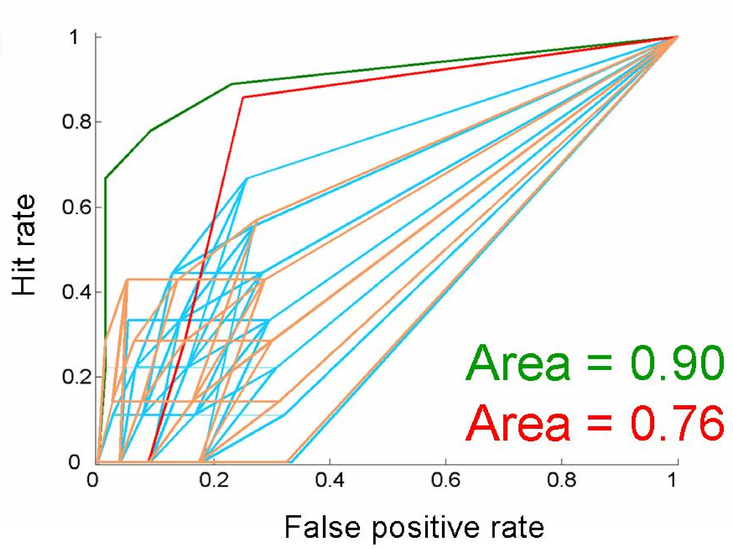
A)


85
84
83
82
13
12
11
10
9
87
26
18
67
14
7
6
32
31
30
29
20 Hz
1sec
28
27
80
73
56
54
53
50
45
25
Jennifer Aniston

B)


Nr. of spikes
Picture nr.
0
0.5
1
1.5
2
2.5
3
10
20
30
40
50
60
70
80

C)


Hit rate
False positive rate
0
0.2
0.4
0.6
0.8
1
Area = 0.90
Area = 0.76

A)

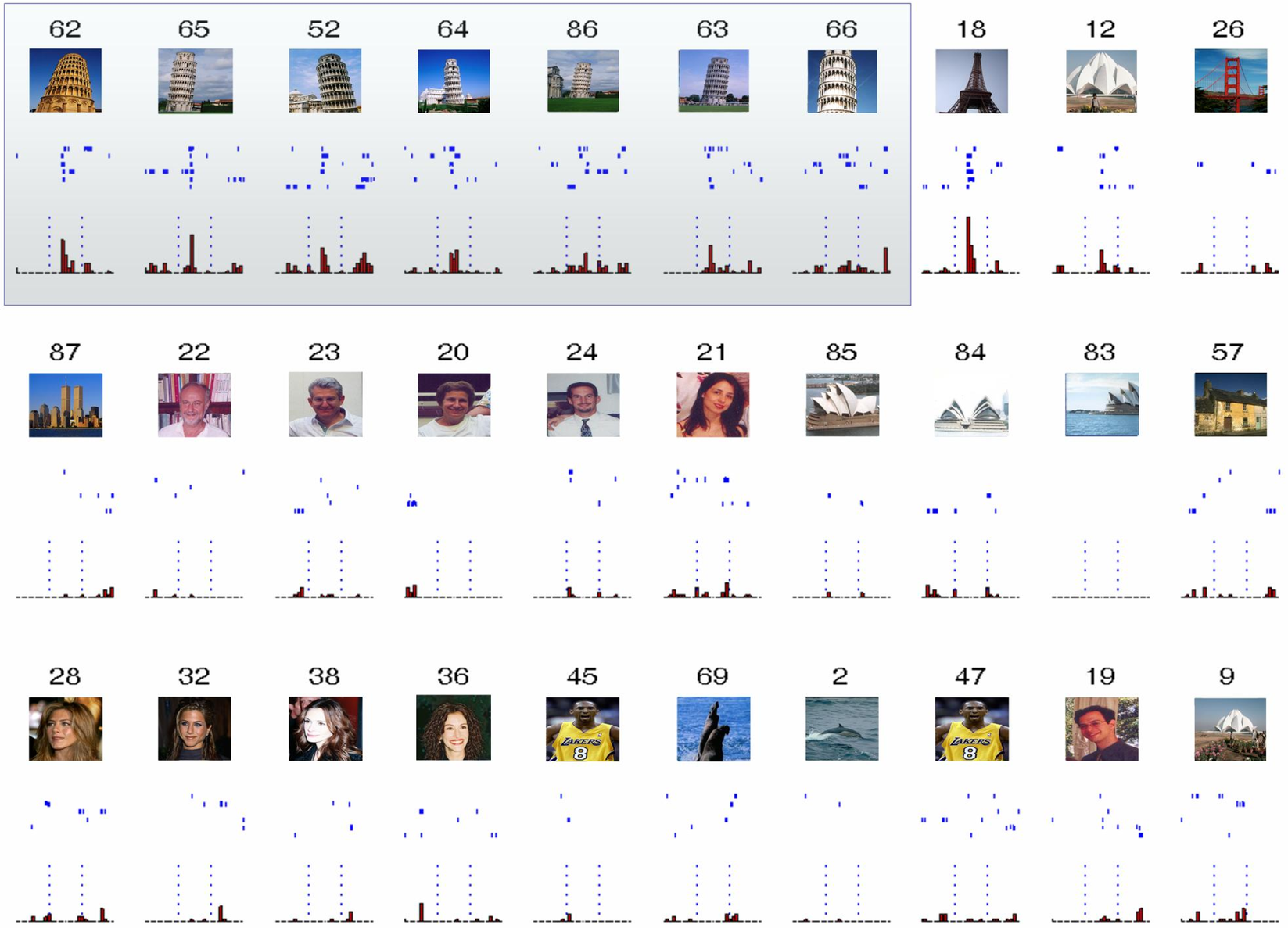
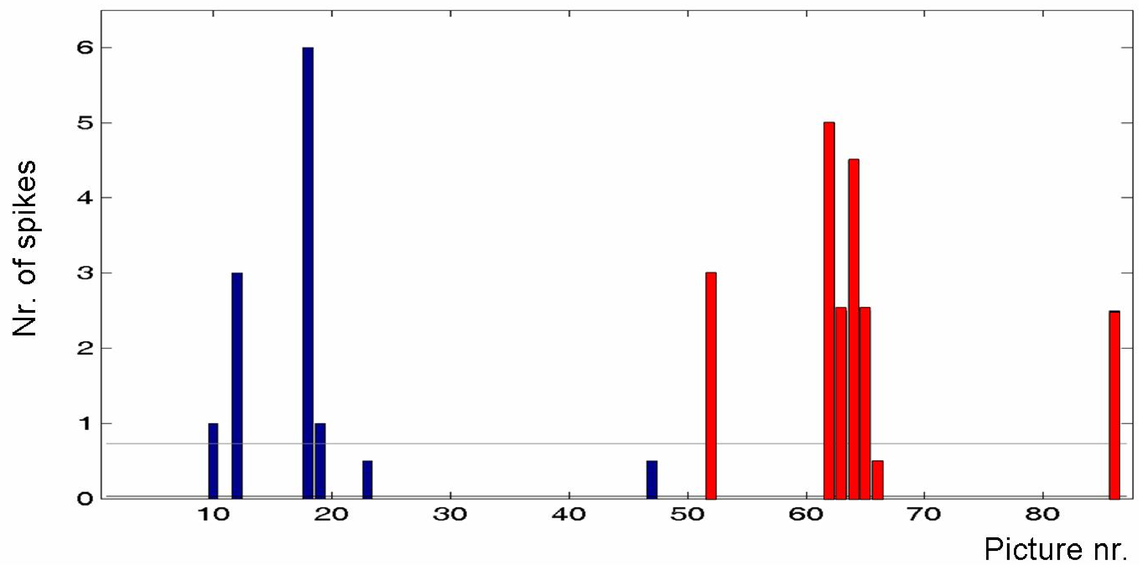

62
65
52
64
86
63
66
18
12
26
87
22
23
20
24
21
85
84
83
57
40 Hz
1sec
28
32
38
36
45
69
2
47
19
9

Fig.S6

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
10
20
30
40
50
60
70
80

C)

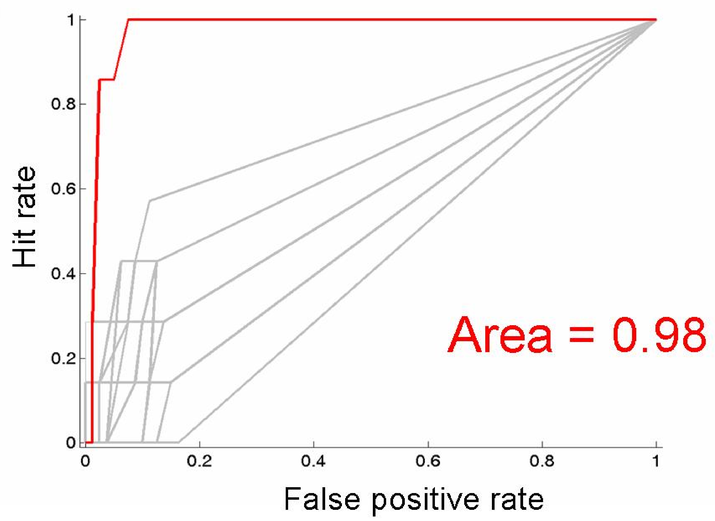
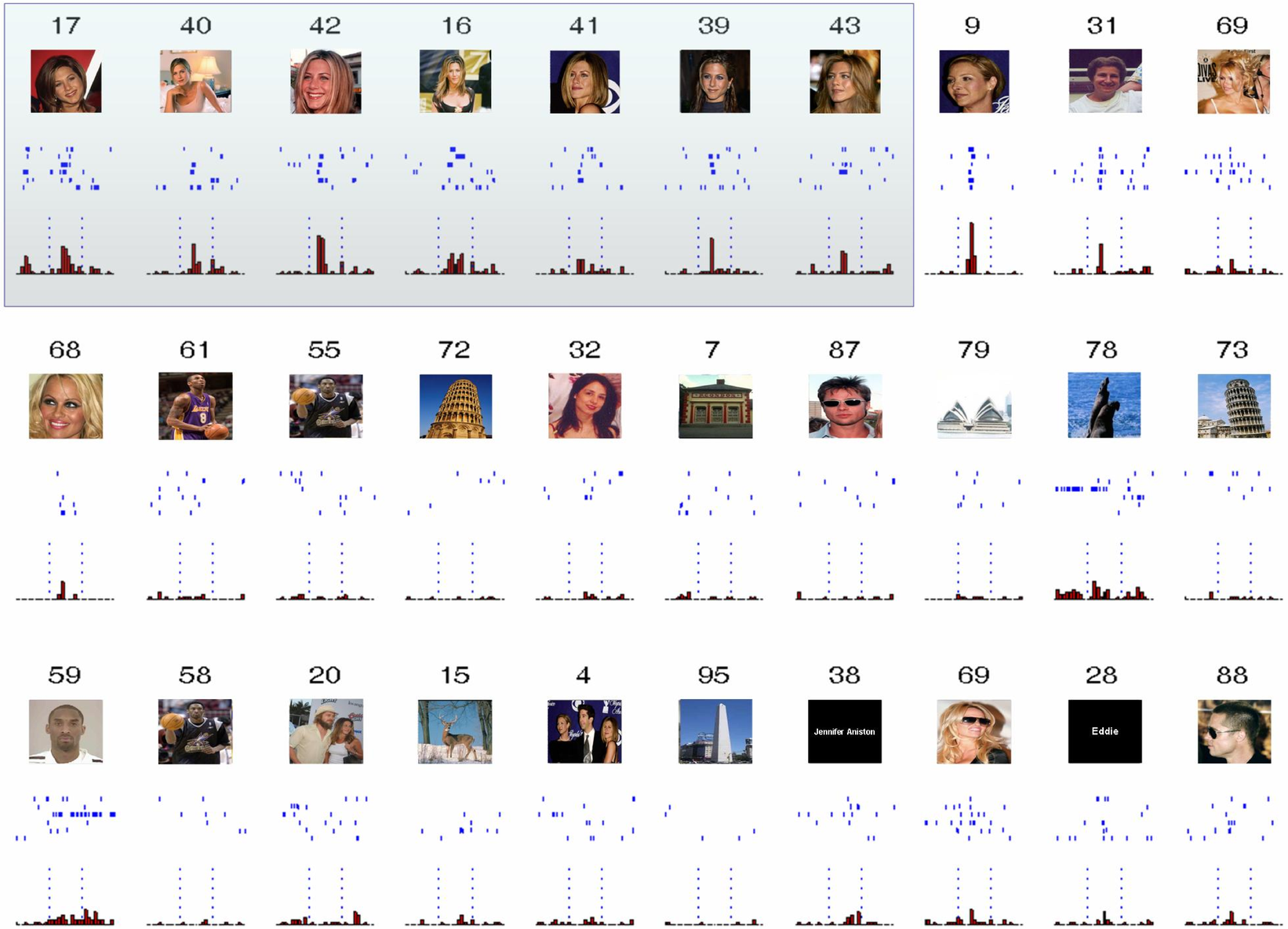
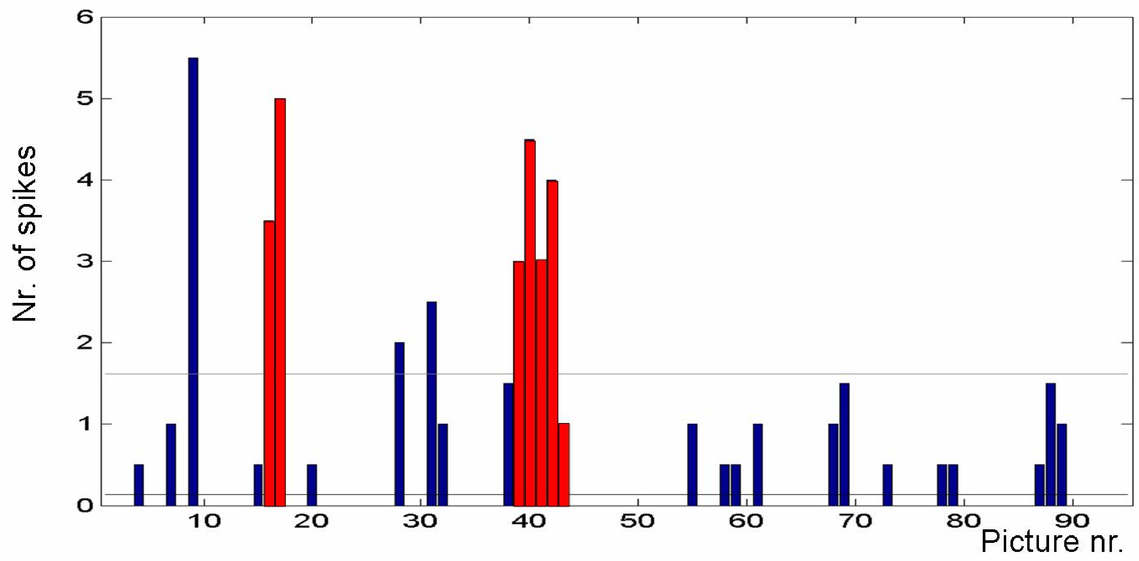

Hit rate
False positive rate
Area = 0.98
0
0.2
0.4
0.6
0.8
1

Fig. S7

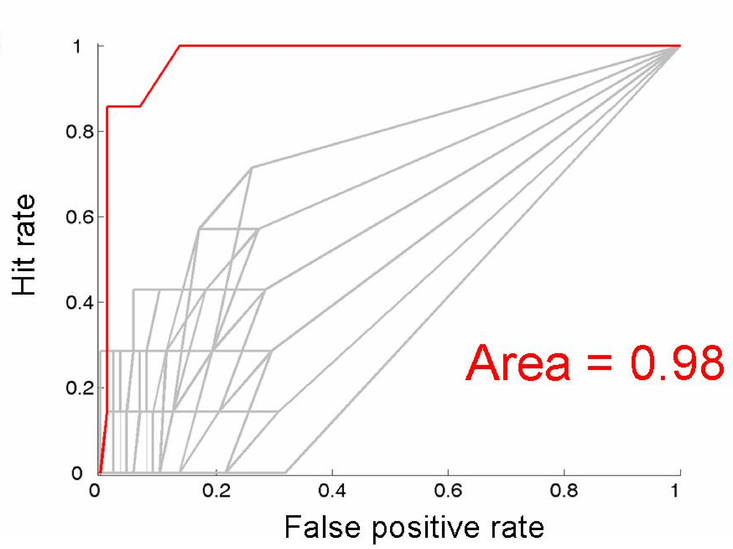
A)


17
40
42
16
41
39
43
9
31
69
68
61
55
72
32
7
87
79
78
73
40 Hz
1sec
59
58
20
15
4
95
38
69
28
88
Jennifer Aniston
Eddie

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
10
20
30
40
50
60
70
80
90

C)

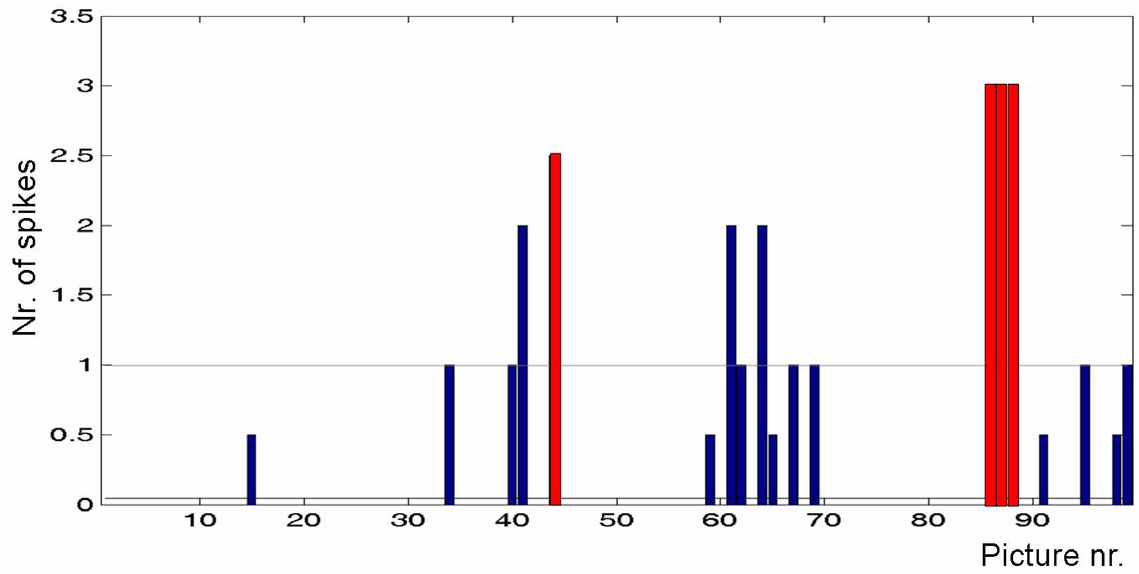

Hit rate
False positive rate
Area = 0.98
0
0.2
0.4
0.6
0.8
1

A)

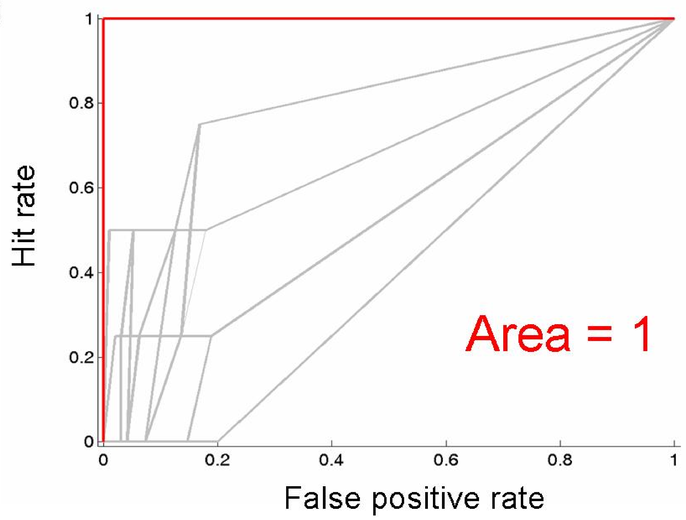

88
87
86
44
64
61
41
62
40
99
95
69
67
34
98
91
65
55
15
97
George Bush
Emily
John Kerry
96
94
22
24
25
23
89
53
52
51
Halle Berry
Catwoman
40 Hz
1sec

Fig. S8

B)


Nr. of spikes
Picture nr.
0
0.5
1
1.5
2
2.5
3
3.5
10
20
30
40
50
60
70
80
90

C)


Hit rate
False positive rate
Area = 1
0
0.2
0.4
0.6
0.8
1

A)

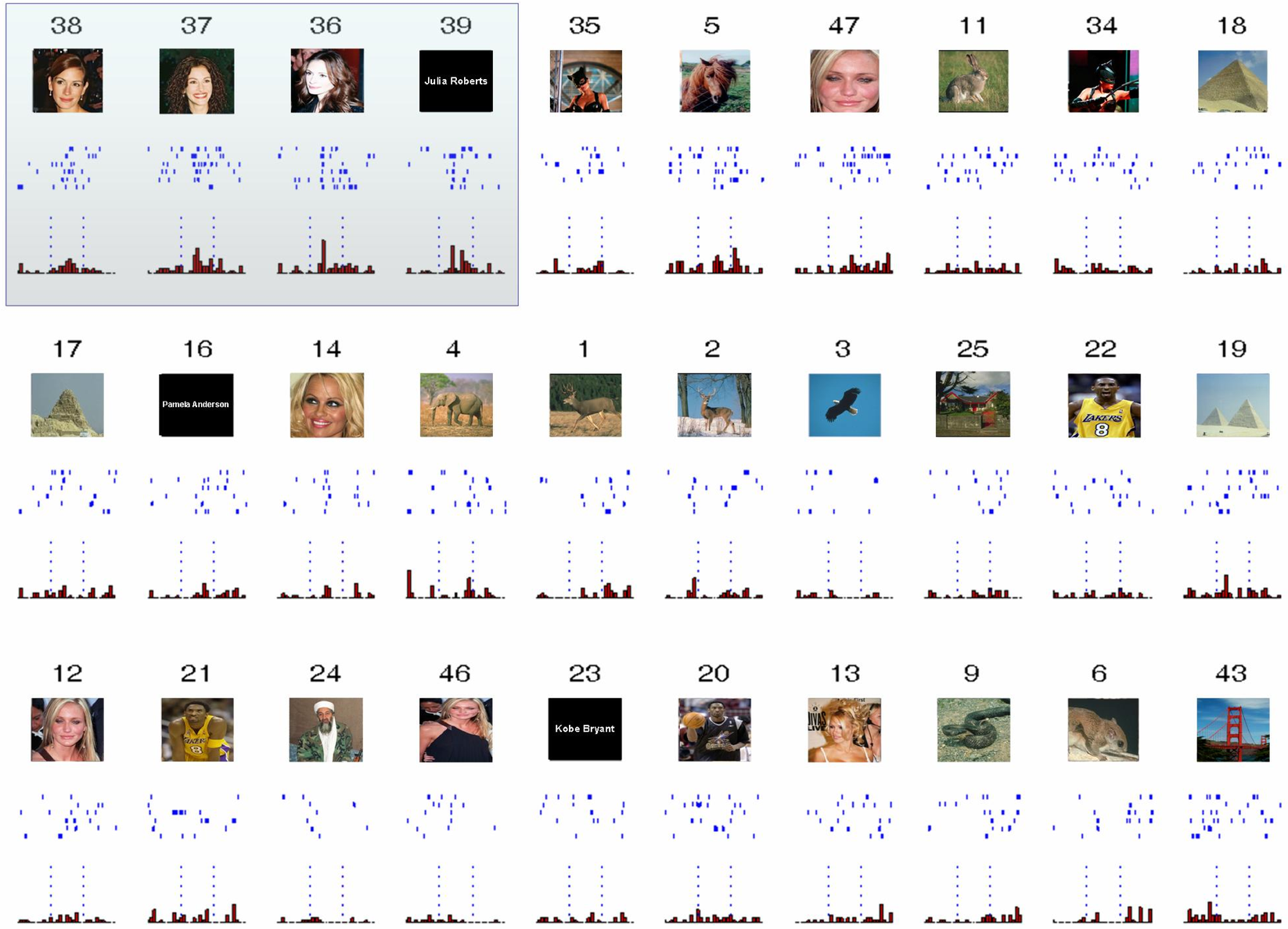
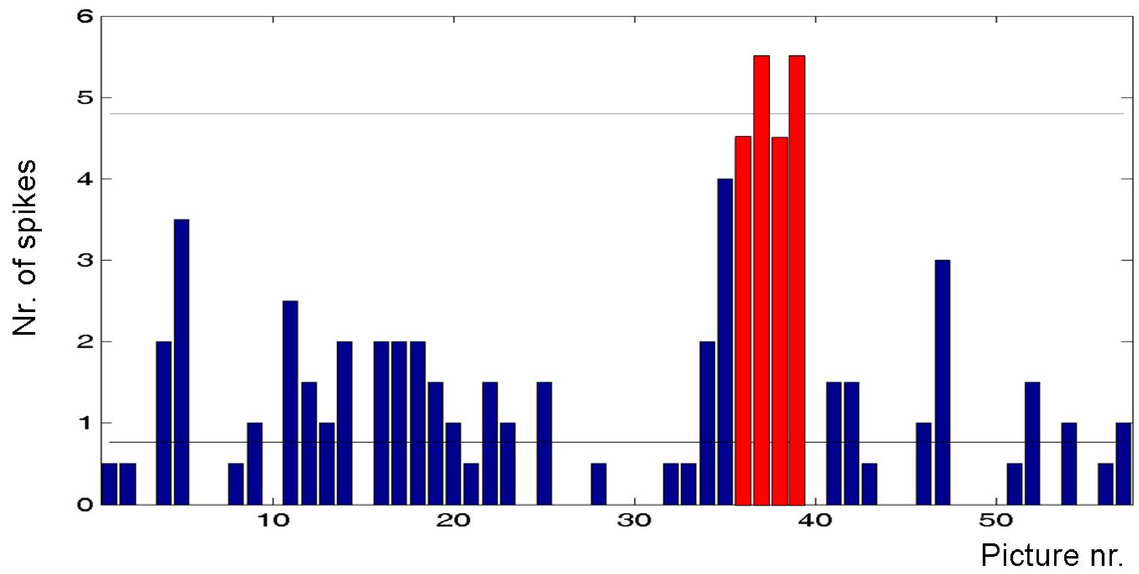
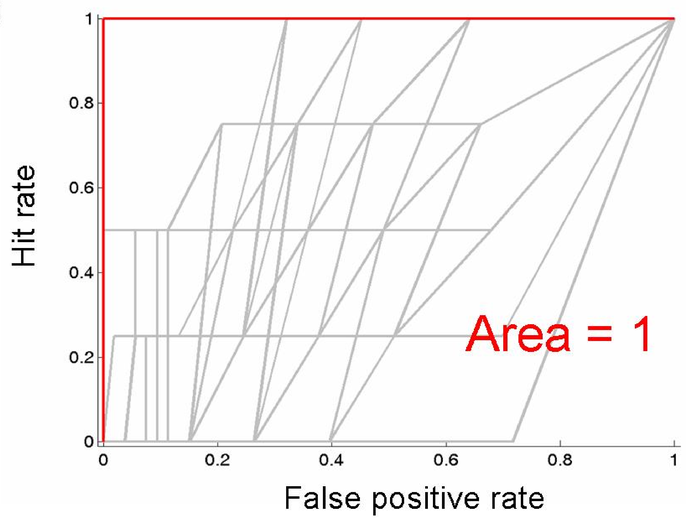

38
37
36
39
Julia Roberts
35
5
47
11
34
18
17
16
14
4
1
2
3
25
22
19
Pamela Anderson
12
21
24
46
23
20
13
9
6
43
Kobe Bryant
40 Hz
1sec

Fig.S9

B)


Nr. of spikes
Picture nr.
0
1
2
3
4
5
6
10
20
30
40
50

C)


Hit rate
False positive rate
Area = 1
0
0.2
0.4
0.6
0.8
1

A)

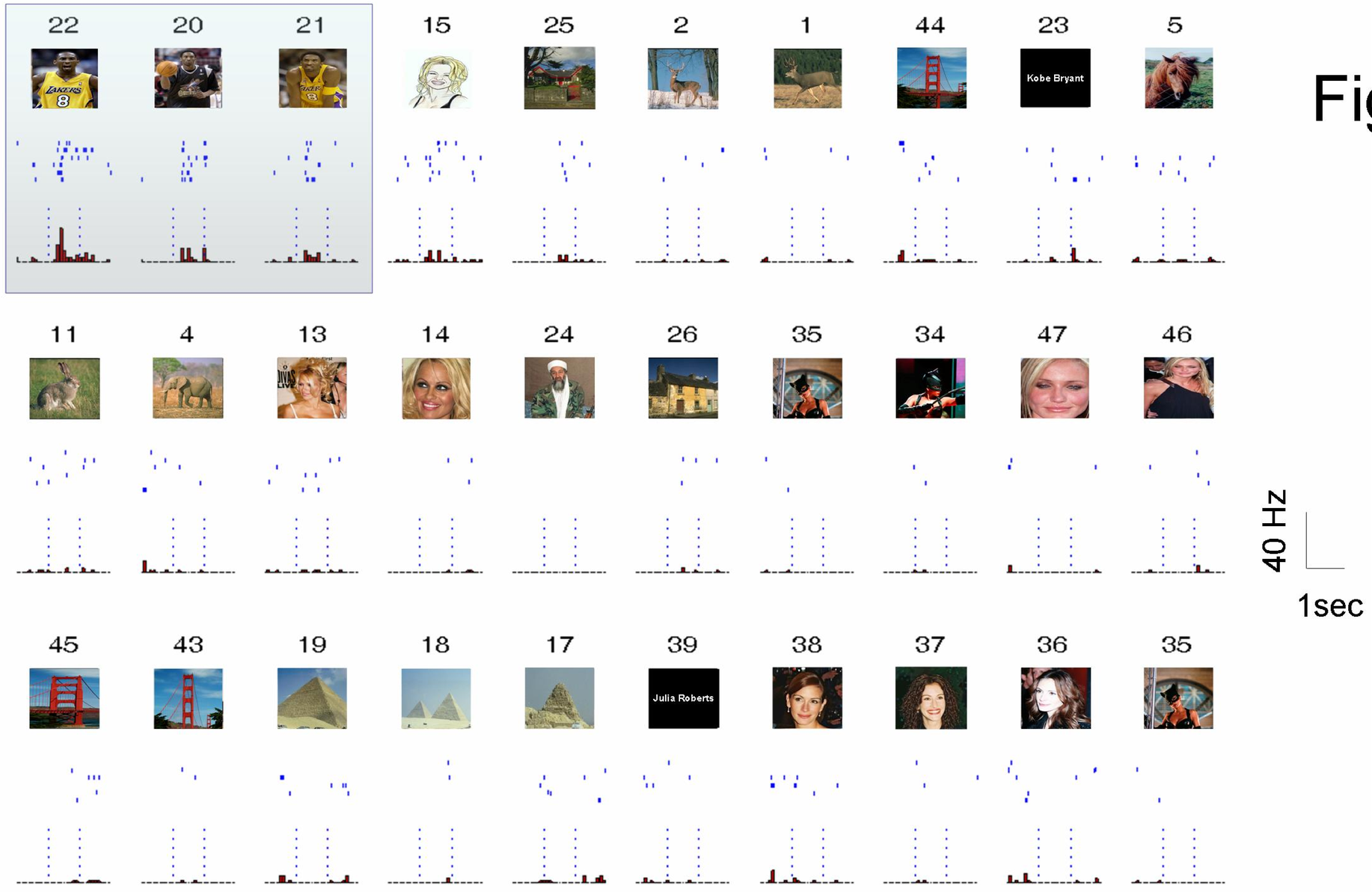
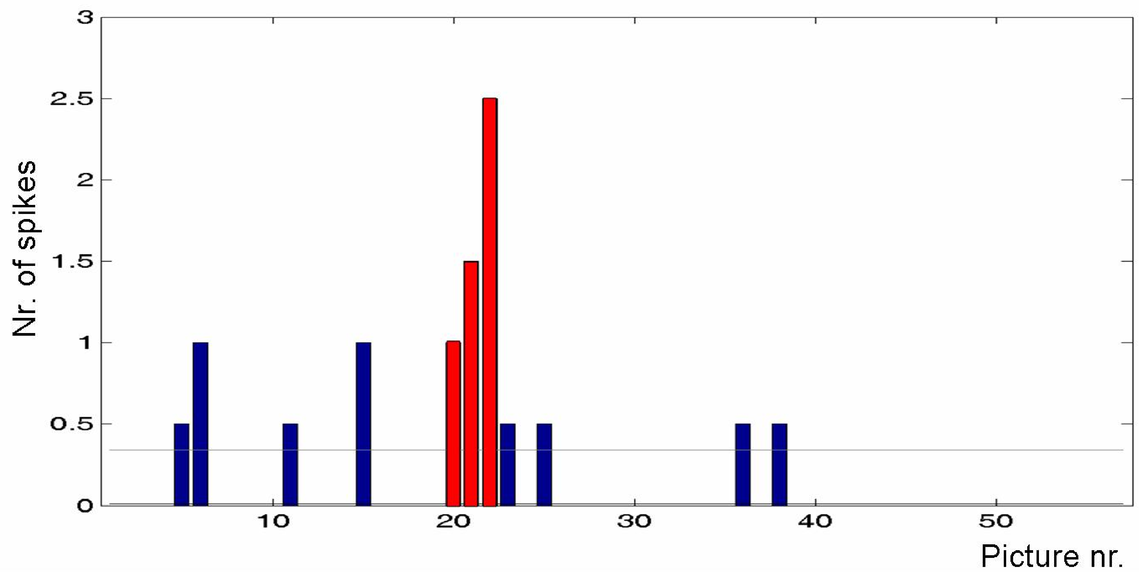
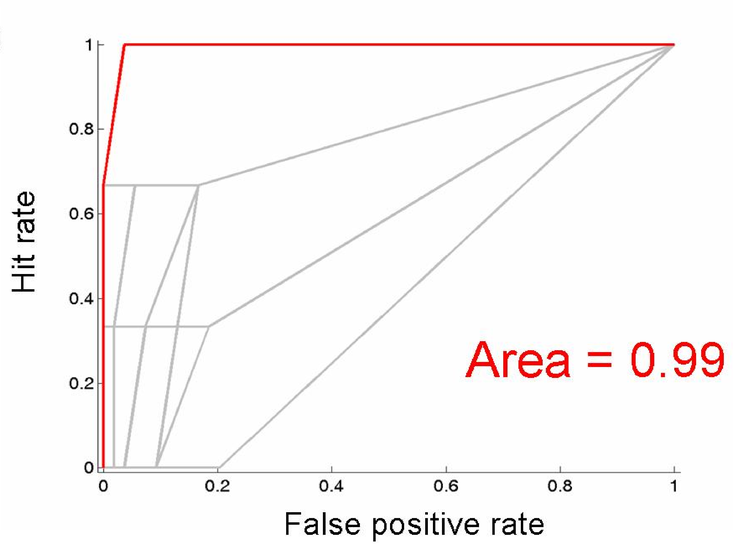

22
20
21
15
25
2
1
44
23
Kobe Bryant
5
11
4
13
14
24
26
35
34
47
46
40 Hz
1sec
45
43
19
18
17
39
Julia Roberts
38
37
36
35

Fig.S10

B)


Nr. of spikes
Picture nr.
0
0.5
1
1.5
2
2.5
3
10
20
30
40
50

C)


Hit rate
False positive rate
Area = 0.99
0
0.2
0.4
0.6
0.8
1

A)

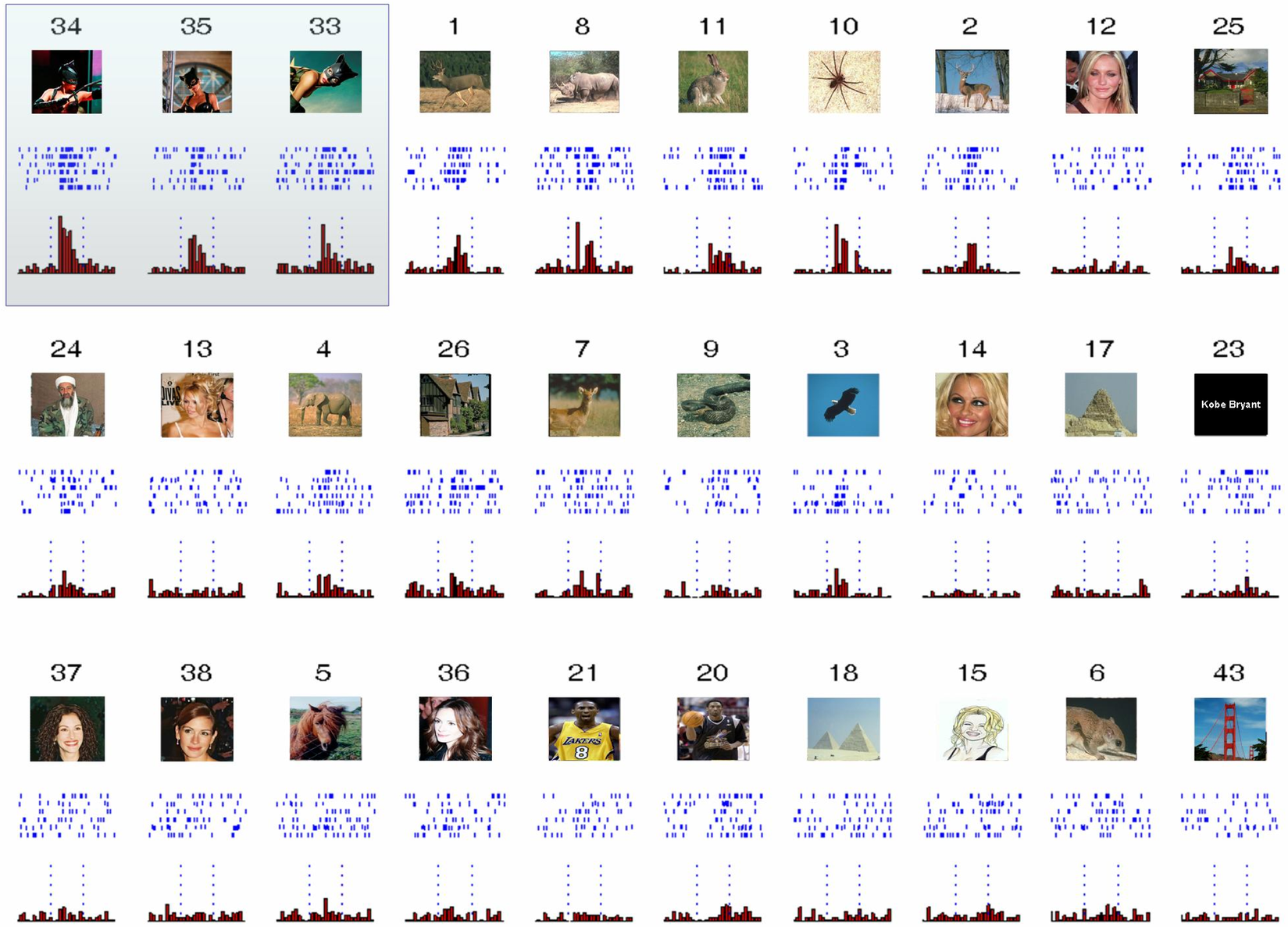
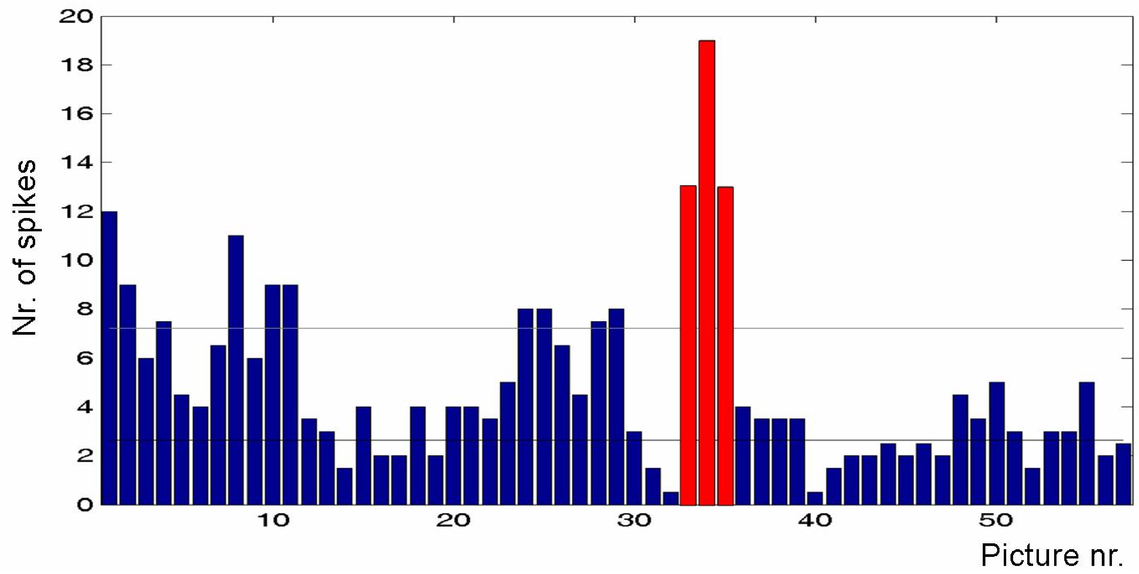
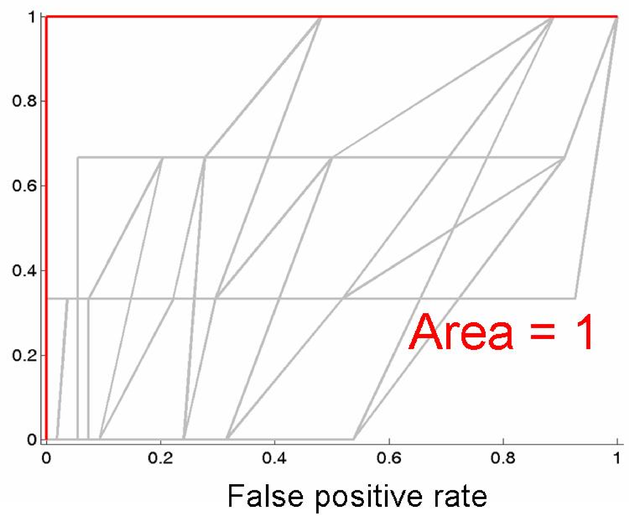

34
35
33
1
8
11
10
2
12
25
24
13
4
26
7
9
3
14
17
23
Kobe Bryant
37
38
5
36
21
20
18
15
6
43
40 Hz
1sec

Fig.S11

B)


Nr. of spikes
Picture nr.
0
2
4
6
8
10
12
14
16
18
20
10
20
30
40
50

C)


Area = 1
False positive rate
0
0.2
0.4
0.6
0.8
1