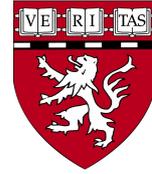
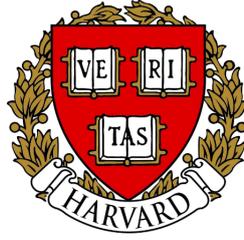


EPFL



MASTER PROJECT

Neuronal correlates of rapid learning in human visual memory task

Author:

Camille GOLLETY

Carried out in the Kreiman lab at Harvard Medical School
Under the supervision of Prof. Gabriel KREIMAN

Under the direction of Prof. Silvestro MICERA

Boston, USA, March 2022

Contents

1	Introduction	3
1.1	Neuroscience of Visual Object Recognition	3
1.2	Responses of single neurons in different human brain regions during object recognition	4
1.3	Motivation	5
2	Experimental procedures	7
2.1	Subjects	7
2.2	Recordings	7
2.3	Stimulus presentation and task	8
2.4	Spiking activity	9
2.5	Further processing	9
3	Neurophysiological Data Analysis	10
3.1	Visual screening	10
3.2	Visual responsiveness	10
3.3	Difference between conditions	10
3.3.1	Neural differences in single units following rapid learning	10
3.3.2	Responses to stimulus features	11
3.4	Selective responses	11
3.5	Multiple-trial classification of intracranial data recordings in a object recognition task using machine learning	12
3.5.1	Classifier input	12
3.5.2	Cross-validation	13
3.5.3	Weights	13
3.6	Neural correlates of Eureka moment	13
3.7	Similarity index	13
3.8	Temporal dynamics underlying object recognition	14
4	Results	15
4.1	Behavioral recognition	15
4.2	Visual responsiveness	15
4.3	Differences between conditions	17
4.3.1	Neural modulation in single units following rapid learning	17
4.3.2	Responses to stimulus features	21
4.4	Selectivity	23
4.5	Inhibitory responses of single neurons during recognition of objects	28
4.6	Classification of visual trials from single neuron recordings	30
4.6.1	Classification of visual object presence	30
4.6.2	Classification of different conditions	31
4.6.3	SVM feature weights	33
4.7	Neural correlates of Eureka moment	33
4.8	Similarity in Neuronal Firing	33
4.9	Timings	34
5	Discussion	38

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Gabriel Kreiman for his continuous support, his patience, motivation, and guidance throughout this project. He offered me the incredible opportunity to work in his laboratory while providing me with all the resources needed to complete my thesis work. It was an honor to join the lab meetings, journal clubs, and multiple additional educational opportunities at the Kreiman's lab and at Harvard in general. I met brilliant and passionate people.

I would like to thank everyone from the Kreiman Lab. To the ones who gave me constructive feedbacks and participated in all the technical discussions about this project. In particular, Yael, Marcelo, Paula, Giorgia, Pranav and Jie.

My very special thanks go to my family and friends who have always given me unconditional love and supports.

Finally, I would like to thank my EPFL advisor, Silvestro Micera, who kindly agreed to monitor and evaluate my project.

Abstract

Our brains do object recognition effortlessly all the time. However, the mechanisms that produce this process remain unclear. There has been notable progress in AI algorithms that can learn to perform different types of pattern recognition tasks. By and large, many of these successes are founded upon supervised training with large number of examples. In contrast, animals including humans excel at unsupervised few-shot learning. We studied a paradigm whereby human subjects were able to rapidly learn to recognize novel images using “Mooney images”. Mooney images consist of black and white blurred objects that are extremely hard to recognize. Recognition rates for these images are close to 0%, yet humans can learn to rapidly recognize them via few-shot learning. We implemented learning by showing subjects the grayscale image counterparts, which are easily recognizable. Even after a single exposure to the grayscale images, subjects dramatically improve their ability to interpret the Mooney images. Here recent experimental findings about single neuron signature of rapid learning are summarized. To evaluate the mechanisms involved in this form of rapid learning, we collaborated with neurosurgeons who implant electrodes in patients with epilepsy. We worked with subjects with pharmacologically intractable epilepsy implanted with electrodes to localize the seizure focus and recorded action potentials from individual neurons along key areas of the occipital and temporal lobes. Visual representations elicited by Mooney images as well as grayscale images could be observed, crucial regions for object recognition in the brain have been identified and category-selective responses could be discriminated. Furthermore, we built a machine learning pipeline to decode whether subjects had learnt to recognize the images from the neuronal population activity.

1 Introduction

One key ability of human brain is object recognition, which refers to the identification of objects based on visual input. This process is rapid, effortless and can impressively happen even after a single exposure. Recognizing a car or a familiar face at a simple glance, seems so easy, so automatic! The brain is capable of comparing visual stimuli with memorized information to accomplish recognition behavior and tasks that assess recognition memory have become increasingly useful tools for investigating the neural basis of object recognition.

Here, we analyzed neuronal responses evoked by Mooney images before and after recognition induced by a learning phase using their grayscale corresponding. Mooney images are two-tone black and white pictures of one object that can be very difficult to recognize. However, after seeing the grayscale image, observers often directly perceive the hidden object, like an eureka effect. From meaningless black and white patches, the image appears as a coherent percept of an object. The cognitive task consisted of several trials during which subjects had to judge by button press whether they recognized the content of the image after brief presentations (500 ms) or not. After several unrecognized presentations of a Mooney image, the Mooney and corresponding grayscale photograph pairs were shown back to back to enforce learning. Finally, the Mooney images were presented again alone. We recorded single unit responses in the human brain, mostly from the medial temporal lobe, from pharmacologically intractable epilepsy patients. They were implanted with electrodes, containing several microwires each, in order to determine the location of the seizure.

How can intracranial field potential recordings help us decoding the process underlying object recognition and understand the mechanisms by which top-down inputs can rapidly orchestrate plastic changes in neuronal circuitry?

The interest of Mooney images lies in the fact that there are stimuli that lead to recognition that is not necessarily obvious or that can even be ambiguous. This makes it possible to report neural changes occurring during differences in perception. Thus, only the variations related to the change in interpretation can be highlighted.

1.1 Neuroscience of Visual Object Recognition

Object recognition is the ability to perceive an object's physical properties (such as shape, colour and texture). It involves a wide variety of functions including the recognition of specific individual objects, their classification into a single category and the understanding of its use and how it relates to others. This recognition requires memory to judge a previously encountered item as familiar and matching with structural descriptions stored in the brain.

Visual recognition processing can be viewed as a bottom-up hierarchy in which visual stimuli are processed sequentially by an intricate system of interconnecting neurons. All the information travels in the form of nerve impulses that are triggered by the photoreceptors in the retina which are neurons specialized on the reception and conduction of visual stimuli [1]. It is then transmitted through the optic nerve in the eye. The two optic nerves travel toward, and meet at the optic chiasm, which lies in the subarachnoid space of the suprasellar cistern [2]. From the optic chiasm, the same axons continue through the optic tract until they synapse with neurons situated in the lateral geniculate nucleus (LGN), a relay system located in the thalamus [3]. From the LGN, axons then carry visual information to the visual cortex through the optic radiations. The pathway can be divided into the upper optic radiation that carries fibers through the parietal lobe to reach the visual cortex and the lower optic radiation where the fibers travel through the temporal lobe, via a pathway known as Meyers' loop, to reach the visual cortex located in the

occipital lobe [4].

Two information processing pathways originate in the occipital cortex, dorsal, which goes to parietal cortex and ventral, which goes to temporal cortex [5]. This two-streams hypothesis was first proposed by Ungerleider and Mishkin (1982) [6] who argued that humans possess two distinct visual systems. The dorsal stream is involved with processing the object's spatial location ("how"), and the ventral stream is involved in the processing of visual object identification and recognition ("what") [7]. The latter, involved in object recognition, is of particular interest in this study. This pathway consists of visual input from primary visual cortex V1 relayed through areas V2 and V4, and finally projected into the inferior temporal cortex.

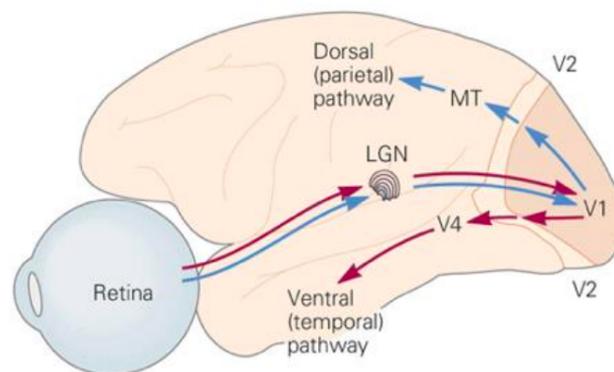


Figure 1: Central Visual Pathway [8]

While areas V1, V2, and V4 are involved in the processing of basic level visual features such as edges, contours, and color, the inferior temporal cortex is suggested to process complex objects [8]. It also makes reciprocal connections with the structures in the limbic system. This subcortical structure plays an important role in processing memory and emotional responses [9], both of which are significant components to visual perception. The two large limbic system structures, the amygdala and hippocampus have important roles in memory. Memory can be categorized as explicit or implicit and involves different brain regions. Explicit memory involves the recall of previously learned information that requires conscious effort to receive, while implicit memory is unconscious and effortless [10]. The amygdala, involved by explicit memory, links visual stimuli with emotions and gives value to items and is associated to implicit basic learning. The hippocampus, on its side, is responsible for learning and memory and helps establish memories of visual stimuli [11]. Both implicit and explicit processes seem to contribute to object recognition and modulate activation that comprise frontal, temporal, parietal and occipital regions [12].

1.2 Responses of single neurons in different human brain regions during object recognition

The neuronal connections originating from the visual cortex are projected to different other cortical areas in the temporal and frontal lobes and to numerous subcortical structures. The medial temporal lobe (MTL), including the hippocampus, entorhinal cortex, amygdala and parahippocampal cortex is known to play an essential role in recognition memory [13]. Within the temporal lobe, the perirhinal cortex is also a major component of recognition memory. It rapidly processes information about the novelty or prior occurrence of individual stimulus items [14]. The hippocampus, due to his critical anatomic position receives converging uni- and multimodal input from neocortical regions and contribute to memory formation [15]. However, entire visual cortical pathways and connecting medial temporal lobe seem important for both perception

and visual memory and there is no functional separation of areas [16]. At the neuronal level, crucial changes could mostly be observed in feedforward neuronal processing during learning in high visual areas. Nevertheless, learning can also be processed through changes in V1 visual cortex based on top-down feedback processing. Thus, early visual areas like the primary visual cortex (V1) have been shown to also be involved in object learning [17]. Indeed, traditional models of object recognition pathway state that processing within the visual system proceeds in a bottom-up, feedforward manner from retina to higher cortical areas [18]. Recent researches have demonstrated strong evidence that recurrent feedback circuits from higher visual areas to lower ones are also engaged during the first hundreds of milliseconds of visual processing and might have contributions to sensory encoding that rely upon top-down [19]. The timing of the responses is a promising tool to distinguish between bottom-up and top-down processing.

Moreover, we know that as we move forward from one layer to a higher one in the visual hierarchy, the neuronal response properties become increasingly complex. In monkeys for example, some V4 neurons respond selectively to particular color, pattern or luminance contrast and IT cells respond only for a specific shape [20]. Additionally, certain neurons respond only to a restricted set of exemplars of a given category [21]. A remarkable degree of category specific firing of individual neurons were previously found, with a selective response to visual stimuli, including faces, natural scenes, famous people and animals [22]. While some incredible progress has been made to describe visual neural plasticity, understanding single-unit responses has proven to be extremely challenging and yet, few studies have reported to examine the neural activity during conscious recognition at the neuronal level and millisecond temporal resolution. Most of our knowledge at the neuronal level about the visual system comes from studies in animal models such as macaque monkeys, cats, rats or mice. Indeed, researchers are limited by ethical consideration and can't do all the measurements they want in a healthy person. Thus, many recordings of neural activity come from the brains of epileptic patients who are required to have brain surgery in order to determine the location of the seizure. Epileptic seizures are defined as jerky or trembling movements in the body due to abnormal neuronal activity and can result in damage to the brain or other parts of the body [23]. We benefit from the fact that these patients need to be implanted for clinical reasons to record physiological data. Here, we take this rare opportunity to highlight the recruitment of neurons that change their firing patterns either in an excitatory or an inhibitory manner related to encoding and retrieval of objects. In addition, we have the chance to have electrodes in several regions of the brain and not to be limited to one location as it is the case in many studies.

1.3 Motivation

Understanding how humans achieve recognition is not only of interest to neuroscientists but also to researchers who are implicated in building computational models that mimic the behavior of real neurons to a high degree of accuracy. Human models are central to the development of these systems and have helped enable the successful emergence of the first generation of BMIs. Those technologies present many potential clinical uses. Recent studies have shown that brain-machine interfaces (BMIs) could offer great potential for applications such as generating images in the mind of a blind person [24]. Indeed, researchers succeeded in identifying patterns of electrical activity in neurons that correspond to a person's attempts to recognize an image, so that the instructions can then be fed to a cortical visual prosthetics.

The research topic of this thesis is fully in line with one of the main goals of the Kreiman lab: elucidating how neural circuits compute and building biologically-inspired artificial intelligence. We now know that bringing together artificial intelligence and neuroscience can be beneficial for both. The architecture of Deep Convolutional Neural Networks (DCNNs) often closely mimics the functioning of the ventral visual stream. These networks consist of layers of nodes that are

analogous to neurons. They have revolutionized machine vision and can now outperform human vision in many object recognition tasks. More recently, Spiking Neural Networks (SNNs), often defined as third generation neural networks have shown promising results and competitive capabilities to deal with numerous cognitive tasks. SNNs are inspired by information processing in biology and operate using spikes, which are discrete events [25]. A clear example of how neural representation in the visual system can be leveraged to build computational models is given by the work of Yamins and his colleagues. They constructed their deep neural network according to the same retinotopic, hierarchical architecture as the brain to predict the brain activity of a monkey [26]. In a more applied way to our project, contours are ambiguous in Mooney images which makes integration of information difficult for both human and computer vision systems. Current computer vision techniques fail to recognize Mooney images and understanding how the visual-processing pathways in the brain are organized could enlighten us on how the eyes take in visual information. Our findings could reveal opportunities to solve issues in computer vision like face recognition in noise, occlusion and darkness, which are often met in video surveillance.

Moreover, neurophysiological recordings in humans can help us understand the computations performed by visual cortex. These discoveries could play a critical role to help the large number of subjects with disorders caused by damage to the parts of the brain that process vision. The study of the responses of single neuron and how this correlates to conscious recognition, whether the activity changed with stimulus, whether this activity depended on the subjects' recognition. The answers of these questions is vital not only for uncovering the neural substrates of these key aspects of learning, but also for understanding the processes of associative visual agnosia and the validity of visual models of this syndrome. Visual agnosia is an impairment in recognizing visually presented objects [27]. This inability is usually caused by lesions on the parietal, temporal, or occipital lobes of the brain. These lobes store semantic information and language [28]. In such cases, messages from the retina of the eye get perfectly transmitted along the optic nerve and visual loss is associated with brain damage.

From another point of view, remembering the content of previously seen Mooney images could be at the origin of a new authentication system. Very promising results have shown that implicit visual memory-based authentication could reduce the password burden [29]. Forgot your password? How many times have you seen this question appear on your screen? Two-tone images can provide an excellent opportunity for realizing implicit authentication and outperform current implicit memory-based methods. The enrollment phase consists in showing both the Mooney images and their corresponding original images to create a link in the user's memory. This phase is similar to what we do in our experiment and reveals that applying the knowledge about how humans store and recall information can be used for real-world applications.

Other possible findings may lead to identify neuronal properties that facilitate the recognition of black and white images. Previous studies showed that 4-5 year-old children are generally unable to recognize Mooney images even after showing them the corresponding grayscale photograph [30]. These results point the ability of the brain to undergo perceptual reorganization over time. This study could allow to know more about the development of top-down mechanisms of perceptual learning and improve young children's interpretation of images.

Understanding object recognition leads us to look at the particular human experience of the Eureka effect or 'Aha!' moment of suddenly understanding a previously insolvable problem. Research on this phenomenon started with Köhler's observations on the problem-solving abilities of chimpanzees [31] and is among the most fascinating mysteries of human cognition. Still, this neuroscientific challenge remains poorly understood and our visual task can be used to assess the perceptual learning component of Eureka moment.

2 Experimental procedures

2.1 Subjects

Subjects were 13 patients (25–50 years old, 11 right-handed, 8 males) with epilepsy, admitted into the hospital to localize the seizure foci for potential surgical resection. Thus, there is a high variability in the number and location of electrodes across subjects. It should be noted that these patients have neurological disorder; however, all the data was collected during periods without any seizure events and most of it was from regions that were found to be nonepileptogenic. Basic information about the 13 subjects are shown in Table 1:

Patient ID	Gender	Age
Subject 1	male	25
Subject 2	female	26
Subject 3	male	50
Subject 4	male	33
Subject 5	female	41
Subject 6	female	40
Subject 7	male	35
Subject 8	female	35
Subject 9	male	48
Subject 10	male	26
Subject 11	male	45
Subject 12	male	46
Subject 13	female	33

Table 1: Subjects Information

2.2 Recordings

Hybrid depth electrodes which contained eight sub-micron-diameter microwires at the tip of each electrode shank were used. The locations of electrode implantation were selected based exclusively on clinical reasons. Targets included the amygdala, entorhinal cortex, hippocampus, occipital lobe, parahippocampal gyrus, anterior temporal lobe, posterior temporal lobe and parietal cortex. The number of recording sites per subject ranged from 14 to 70, for a total of 1424 electrodes. Table 2 reports the number of electrodes in each anatomical region.

Anatomical Region	# of channels	# of clusters
Amygdala	170	337
Entorhinal	38	49
Hippocampus	164	343
Occipital Lobe	73	163
Parahippocampal Gyrus	92	212
Anterior Temporal Lobe	47	90
Posterior Temporal Lobe	105	217
Parietal Cortex	8	13
Total	697	1424

Table 2: Distribution of electrode locations

2.3 Stimulus presentation and task

The cognitive task (Figure 2 and Figure 3) consisted of several trials during which subjects had to judge by button press whether they recognized the content of the image after brief presentations (500 ms) or not. If the subject's response was yes, the experimenter noted the answer as 'Y' and the subject was then asked to verbally provide the answer (for example camel, penguin, Eiffel tower...). If the subject was correct, experimenter input another 'Y'. If incorrect, experimenter input 'N' response. If subject responded "no I did not understand the image", i.e., he/she accidentally pressed the wrong button, experimenter input 'N'. The stimuli were presented in blocks of 60 images. In case of a negative response, indicating no recognition, or no button press, the same image was presented again 1300 ms after the first one. After several unrecognized presentations of a Mooney image (preGS), the grayscale (GS) version of the same images were shown. If the response was correct, the Mooney image version (postGS) was shown back to back with the GS to reinforce learning. Finally, the Mooney images were presented again alone. Trials of the preGS condition, that were recognized by the patient, as well as all trials that were not correctly recognized in the GS and postGS conditions were not considered further. We recorded single unit responses in the human brain, mostly from the medial temporal lobe, from pharmacologically intractable epilepsy patients. They were implanted with electrodes, containing several microwires each, in order to determine the location of the seizure. Between 5 and 10 images were presented per section from several categories, including people, building, sports, clothes... Figure 4 shows an example of a Mooney image and its grayscale corresponding belonging to the animal category.

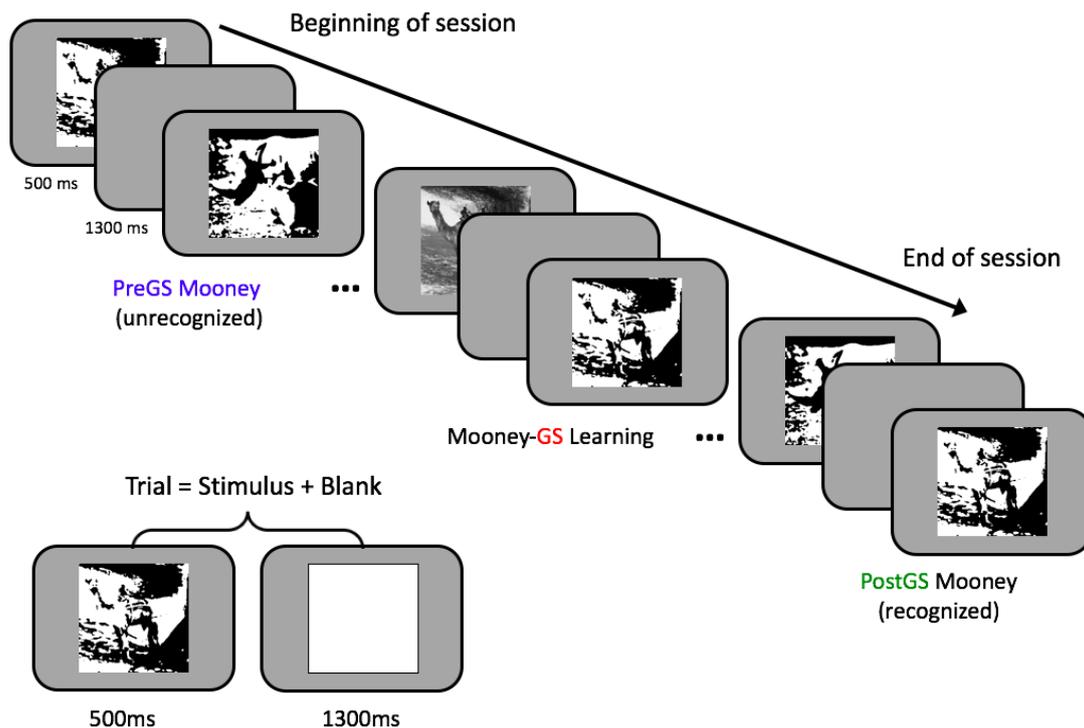


Figure 2: Experimental Design

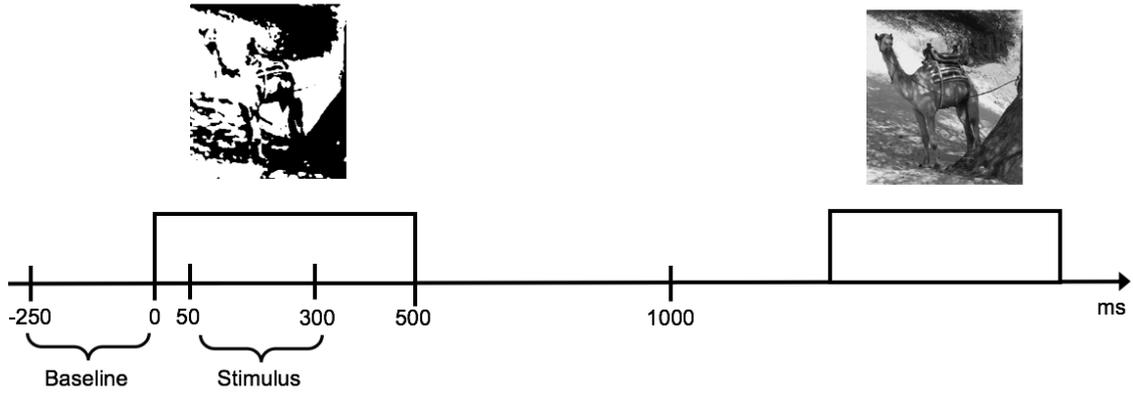


Figure 3: Sequential presentation task and corresponding epochs



Figure 4: Example of a Mooney image and its grayscale corresponding

2.4 Spiking activity

The activity of neurons was recorded from extracellular electrodes capable of recording spiking from multiple nearby neurons. The raw data were band-pass filtered between 300 and 3000 Hz and thresholded for detection of potential spikes. Action potentials were then sorted using a clustering algorithm and manually sorted as spikes or electrical noise [32]. This sorting corresponds to the identification of which spike corresponds to which neuron.

2.5 Further processing

Matlab (MathWorks) was used for data analysis, including statistical analysis as well as data visualization. For classification, we implemented machine learning algorithms in Python (Jupyter lab) using scikit-learn library in addition to standard libraries (NumPy, Pandas, Matplotlib, SciPy, Seaborn).

3 Neurophysiological Data Analysis

Data was collected from epileptic patients willing to provide neurophysiological data while they were implanted with electrodes for clinical purposes. Hence, the choice of the location of these electrodes is dictated by medical reasons and involves recordings from areas of the brain not related to vision.

3.1 Visual screening

Before doing any statistical analyses to find significant units, a visual screening (brief visual examination to look for potential variability) of neuronal responses has been carried out to identify potentially interesting neurons and check that future tests do not miss beautiful responses. Two fundamental methods are used to analyze spike trains of single neurons, which aim to characterize their encoding properties: raster plots and peri-stimulus time histograms. They allow to visually examine spiking activity, trial-by-trial, of a neuron over a period of 900 ms (from -200 ms before stimulus onset to 700 ms after) and show histograms of the times at which neurons fire.

3.2 Visual responsiveness

Several statistical tests have been performed with the aim of identifying visually responsive neurons. A non parametric permutation test was performed to compare the responses during the 50 to 300 ms interval after stimulus onset to the baseline before stimulus presentation (-250 to 0 ms) for every single cluster based on the number of spikes over that period. Non parametric tests come with the advantage that no assumptions are made about the theoretical underlying distribution of test statistics under the H_0 (null hypothesis assuming that the mean is the same between the two sample sets). Instead, the distribution is created from the data we have by shuffling values 1000 times over the two different time windows (baseline and stimulus), computing t-test and thus creating a distribution of t-values. No distinction is made between the conditions for this statistical test and all trials are considered. Statistical significance is then obtained by counting the number t-values that are more extreme (further to the right or left tail of the distribution) than the original t-value and divide it by the total number of iterations to obtain a p-value. A neuron was considered visually responsive when the test showed a significant difference ($p < 0.01$ or $p > 0.99$) between the number of spikes for the baseline and the stimulus epoch.

The t-value is computed as followed for two independent data samples:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (1)$$

where x_1 and x_2 are the sample means, σ_1 and σ_2 are the sample standard deviations, and n_1 and n_2 are the sample sizes.

Excitatory vs inhibitory responses

The response was termed "inhibitory" when it was significantly lower than the prestimulus baseline and "excitatory" when it was higher.

3.3 Difference between conditions

3.3.1 Neural differences in single units following rapid learning

We first performed an analysis to see if there were differences in the firing characteristics between preGS and postGS conditions only, as the main focus is the difference in response between the

unrecognized versus the recognized identical images, rather than differences in stimuli response that might be induced by a different type of stimuli in the GS condition. In order to test the single neuron response to differences in stimuli, we performed t-tests comparing the firing rate in the stimulus window for each neuron separately and for each exemplar. An exemplar is defined as an image type (for example all images 10). The response of a neuron was considered statistically different between two conditions if (i) the response was significantly different between preGS and postGS for the stimulus time period, (ii) the response of the neuron was not different from one condition to the other during baseline and (iii) the response of the neuron was different between stimulus and baseline period. The second condition is necessary in order to eliminate non-task-related effects. All channels that showed a significant difference (t-test, $p \leq 0.01$ or $p \geq 0.99$) for the baseline window were not considered further. For the exemplars, the same conditions were applied at the image level.

3.3.2 Responses to stimulus features

We have also shown interest for modulation of firing rates between stimuli elicited by Mooney and grayscale images. Several studies have demonstrated that contrast and luminance adaptation alter neuronal coding [33] and we found interesting to explore the potential existence of specialized neurons in the perception of either two-tone black and white images or pictures with different shades of gray. To do so, we performed t-tests to see if the responses to the GS condition was significantly different from the preGS and postGS. The response of a neuron was considered statistically different for the GS condition if (i) the response was significantly different between preGS vs GS and postGS vs GS for the stimulus time period, (ii) the response of the neuron was not different from one condition to the other during baseline and (iii) the response of the neuron was different between stimulus and baseline period.

3.4 Selective responses

A neuron was considered selective for a particular stimulus category (i.e. animal, people, building, sports) if (i) the response was significantly different from baseline, (ii) the response of the neuron was different from one category to the others after stimulus onset and (iii) the response of the neuron was not different from one category to the others during baseline. The difference were assessed by doing a one-way analysis of variance (ANOVA) using permutation tests (two tailed, $p < 0.01$ or $p > 0.99$). In this analysis, stimuli were pooled according to the category they belonged to. We simply permuted the number of spikes across categories, calculated an F value (defined as the variation between sample means over the variation within the samples), repeated this 1000 times, and found the percentage of repetitions in which the calculated values of F exceeded the Fs obtained from the original data. This is the p-value under the null hypothesis. Subsequent pairwise t-test comparisons (two tailed, $p < 0.01$ or $p > 0.99$) between the activity during this interval for the specific category and the rest of the categories had to show a statistically significant change. We assessed significance of the response to different categories of visual stimuli by conditions.

For neurons that showed visual selectivity, we further compared the results between the different conditions and illustrated it with a Venm diagram showing the number of neurons with selective responses to one, two or three conditions. Moreover, for each visually selective neuron, we plotted the categorical mean firing rates distribution in order to try to identify category specific responses.

We were also interested in the role of human brain regions in the representation of different categories of visual stimuli and computed the importance of selective neurons in each region.

3.5 Multiple-trial classification of intracranial data recordings in a object recognition task using machine learning

Perception from the same visual stimuli can elicit detectable changes in firing rate recordings and we were interested to see if machine learning classification was able to correctly distinguish this changes with a high accuracy. Before going into deeper classification tasks, we first wanted to determine if an object presentation could be distinguished from the vision of a blank screen. Classification was performed using data from the first 500 ms after each object was shown, from $t = 0$ ms to $t = +500$ ms, relative to the image appearing on the screen and before stimulus onset from $t = -500$ ms to $t = 0$ ms. A 250 ms time window was then used, with data collected from $t = -250$ ms to $t = 0$ ms and from $t = +50$ ms to $t = +300$ ms. Previous studies have shown that neural responses usually occur in the window between 50 ms and 300 ms post stimulus onset [34]. Neurophysiological data collected before and after stimulus onset were respectively labelled as ‘baseline’ or ‘stimulus’. We then trained a support vector machine (SVM) algorithm with a linear kernel to classify the data into one of the two classes according to the presence or absence of specific visual stimuli. This technique has given particularly good results in a wide range of domains, including neural decoding problems in both rats and humans [35][36]. The objective of the SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. To do so, this method finds a plane that has the maximum margin, i.e the maximum distance between data points of both classes. If the data used shows modulation between one state or another (object presentation vs blank screen) when performing a recognition task, it might allow to later classify if the object is recognized and eventually to which category it belongs.

Since we are interested in understanding the ability to rapidly recognize objects and how it is solved in the brain, it is interesting to try to predict if the image is recognized or not by the subject when the same image is shown. Thus, we again trained an SVM classifier with a linear kernel on all the neurons. The trials were respectively labelled as ‘preGS’ or ‘postGS’.

3.5.1 Classifier input

Each session includes several hundreds presentation trials. As this number can be different across sessions, we had to choose a number of trials and remove those who exceeded it. The second minimum number of trials across sessions was chosen as a threshold (552 trials). Thus, the session with the lowest number of trials was taken away. The training data was the number of spikes for each channel either during baseline or stimulus presentation. These training data were labelled as ‘1’ and ‘0’ respectively for baseline and stimulus for each trial when used with our classifier. As a control, the classification was computed with training labels assigned randomly for the stimulus and baseline. Classification was also computed using data from individual sessions separately. On this data, classifiers were trained using input from each individual session in order to potentially identify a small subset of units that consistently gave much higher classification accuracy.

For the image recognition classification, the number of trials chosen was again computed by taking the second minimum across all sessions for preGS and postGS conditions. We then selected the smallest of the two (105 trials). These training data were labelled as ‘1’ and ‘0’ respectively for preGS and postGS images for each trial when used with our classifier. Since we recorded from neurons in several different brain regions, we also trained the classifier using data coming from distinct brain locations. Again, as a control and to assess the performance directly associated with difference of conditions, the classification was computed with training labels assigned randomly for the preGS and postGS condition.

3.5.2 Cross-validation

We used k-fold crossvalidation to evaluate whether the SVM technique achieved significant classification accuracy. It is a standard method for estimating the performance of a machine learning algorithm in which the training set is split into k smaller sets. The model is trained using k-1 of the folds as training data and the resulting model is validated on the remaining part of the data. The performance measure reported by k-fold cross-validation is then the mean of the values computed in the loop. We used the `cross_val_score` function from the sklearn python library to estimate the accuracy of a linear kernel support vector machine, by computing the score 5 consecutive times. Each time, it returns the mean accuracy on the given test data and labels.

3.5.3 Weights

Once the linear SVM is fit to data, the weights can be assessed with `svm.coef_`. They represent the vector coordinates which are orthogonal to the hyperplane and their direction indicates the predicted class. The absolute size of the coefficients in relation to each other can then be used to determine feature importance for the data separation task. Thus, in our analysis, we analyzed the weights assigned to cells by our decoder to determine the importance of input neurons and to make some inference of how the information is decoded.

3.6 Neural correlates of Eureka moment

Initially, Mooney images seem difficult to interpret, but eventually lead to a rich and stable percept of the objects in the image. This process often occurs with a unique phenomenal experience, the ‘Aha!’ moment, referring to the sudden, unexpected recognition of object content, the feeling of finally "seeing" it. Here, we were interested to know if neurons change their activity pattern when that moment happens. For this purpose, we compared the firing rates between the last preGS trials, when the image was not recognized and the first few postGS trials, with correct object identification. Since this part of the study focuses on very few trials, we were not able to perform statistical tests. However, the firing rates of the neurons of the previously mentioned trials were plotted with perhaps the hope of seeing a sudden change in the response.

3.7 Similarity index

To quantify the modulations in firing rates, we computed an index based on the average number of spikes 50-300 ms after the stimulus onset. Next, to compare the change of modulation after recognition, we computed a similarity index to assess whether the firing rates fluctuates towards the GS condition or not.

The **modulation index** is defined as the change in firing rate from one condition to the other as proportion of the variances of the 2 conditions.

The **similarity index** (Figure 5) is the difference between the preGS and postGS modulation index. A positive SI means representation change towards GS.

$$MI_{preGS} = \frac{|GS - preGS|}{\sqrt{\frac{\sigma_{GS}^2 + \sigma_{preGS}^2}{2}}} \quad MI_{postGS} = \frac{|GS - postGS|}{\sqrt{\frac{\sigma_{GS}^2 + \sigma_{postGS}^2}{2}}} \quad (2)$$

$$SI = MI_{preGS} - MI_{postGS} \quad (3)$$

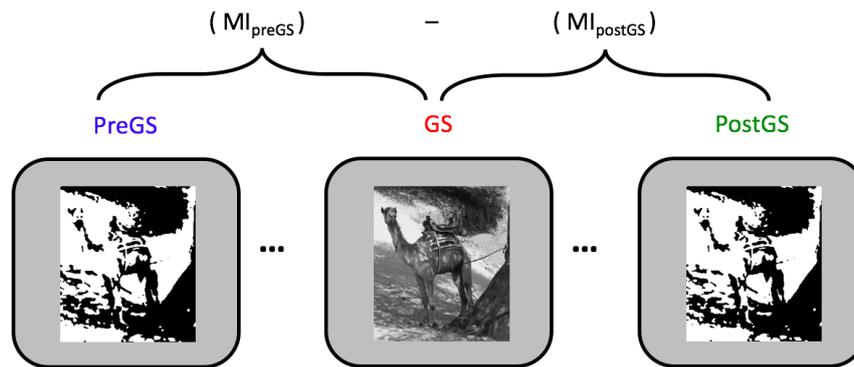


Figure 5: Similarity Index

This index was calculated for all neurons collapsed across exemplars and separately for each exemplar (with at least 10 trials per condition). Then we only kept the indexes of the neurons and exemplars that showed a difference between preGS and postGS conditions according to the previously mentioned criteria. Furthermore, we separated the analysis by region in order to draw conclusions. Moreover, as a control, the similarity indexes were computed for the baseline.

3.8 Temporal dynamics underlying object recognition

In this study, we took advantage of having a temporal resolution of milliseconds to investigate the latency of visually responsive neurons from different regions of the brain. Neuronal response latency is usually defined as the delay between the stimulus onset and the beginning of the response [37]. Quantitative estimates of latency are difficult because they depend on multiple variables, including number of trials, response amplitudes and thresholds. We considered visual response latency (i.e., the time point when a visual response occurred) as a metric to quantify timing. Latency of the visual response was computed at the neuronal level by determining the time, in each unit, when the mean firing rate amplitude from all trials exceeded the mean plus 2.5 standard deviations (s.d.) of the baseline, in the time period after 50 ms post stimulus onset. Only neurons showing a difference between baseline and stimulus were used in the analysis. An additional requirement was added in order to capture truly distinguishable responses. We only took into account neurons showing a mean firing rate after stimulus onset (from 0 to 500 ms) 1.5 times higher than the mean firing rate during baseline. Latency was also computed for neurons that showed inhibition and was defined as the time, in each unit, when the mean firing rate amplitude from all trials was below the mean plus 2.5 standard deviations (s.d.) of the baseline, in the time period after 50 ms post stimulus onset. As before, we have imposed a threshold with a mean firing rate during the baseline 1.25 times higher than the mean firing rate after stimulus onset (from 0 to 500 ms).

4 Results

4.1 Behavioral recognition

All subjects taking part in the neurophysiological experiment were evaluated on their recognition performance at a behavioral level. Figure 6 shows the mean performance of the 13 participants who completed the test. Recognition rates for new Mooney images are close to 0%, yet humans can learn to rapidly recognize them via few-shot learning. We were able to implement learning by showing subjects the grayscale image counterparts, which are easily recognizable. Even after a single exposure to the grayscale images, subjects dramatically improve their ability to interpret the Mooney images. An 80% increase of overall performance was observed in the recognition accuracy between the Mooney images shown before and after the subject has seen the grayscale version.

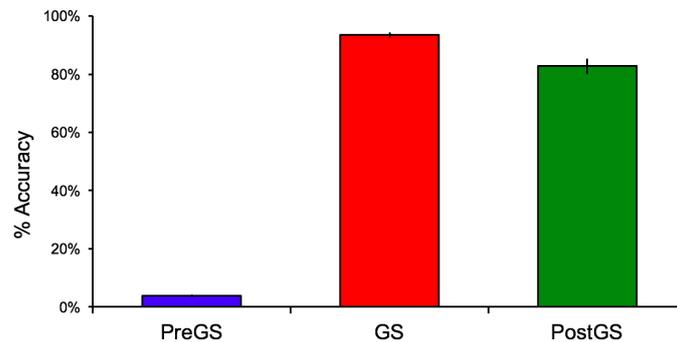


Figure 6: Recognition rate for the 3 image conditions averaged over patients ($n = 13$). The error bars show the standard error of the mean.

4.2 Visual responsiveness

Of the 1424 neurons, 293 (20%) showed changes in firing rate during presentation of the visual stimuli. A comprehensive list of the portion of significant channels for each area is shown in Table 3. No significant response differences were observed between the right and left hemispheres, and therefore the data were pooled. Areas with the highest percentage of modulated channels include the occipital lobe (64.4%) and the parahippocampal gyrus (35.4%).

Anatomical Region	# sig. neurons	% Sig.	Tot. # of neurons
Amygdala	40	11.9	337
Entorhinal	7	14.3	49
Hippocampus	21	6.1	343
Occipital Lobe	105	64.4	163
Parahippocampal Gyrus	75	35.4	212
Anterior Temporal Lobe	8	8.9	90
Posterior Temporal Lobe	37	17	217
Parietal Cortex	0	0	13
Total	293	20.6	1424

Table 3: Units in each region that show a significant difference between the stimulus and baseline

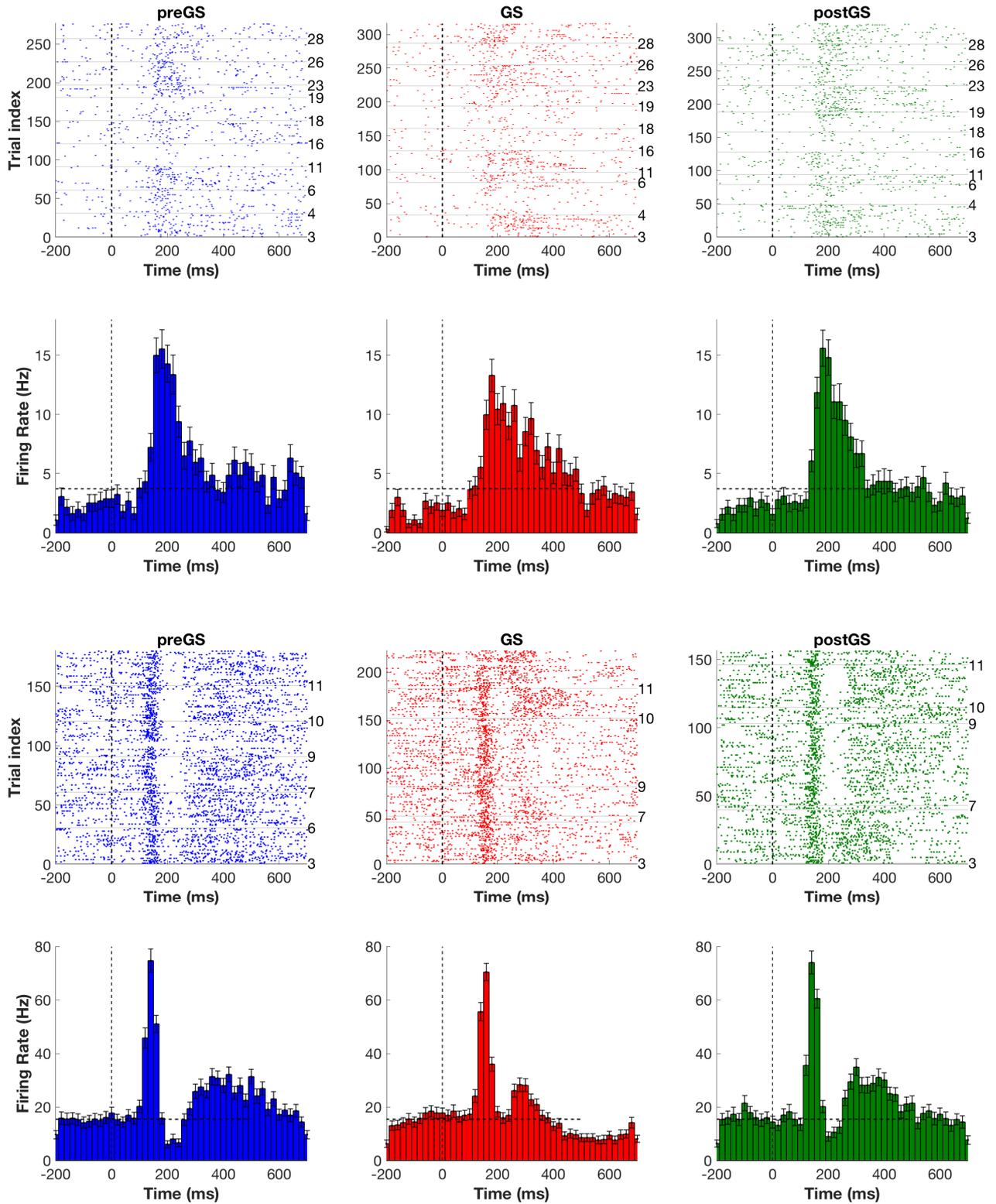


Figure 7: Response of two single units in the occipital lobe of an epileptic patient showing the activity of neurons with a response modulated by visual stimulus. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (bottom) shows the average firing rate of the neuron during each condition.

4.3 Differences between conditions

4.3.1 Neural modulation in single units following rapid learning

24% of the visually responsive neurons showed firing rate modulation dependent on changes in recognition with different responses to preGS and postGS (Table 4). Considering only the responding hippocampal units, 71% had a difference in activity between novel and recognized trials. Other regions of the medial temporal lobe elicited neuronal difference. More than 40% of the neurons in the amygdala and the enthorhinal cortex showed a statistical difference. Figure 8 shows raster plots of evoked responses before and after learning in 2 distinct neurons of the occipital lobe.

Anatomical Region	# Sig. neurons	% Sig.	Tot. # neurons
Amygdala	16	40.0	40
Enthorhinal	3	42.9	7
Hippocampus	15	71.4	21
Occipital Lobe	13	12.4	105
Parahippocampal Gyrus	13	17.3	75
Anterior Temporal Lobe	2	25.0	8
Posterior Temporal Lobe	7	19.0	37
Parietal Cortex	0	0	0
Total	69	23.5	293

Table 4: Units in each region that show a significant difference between preGS and postGS condition

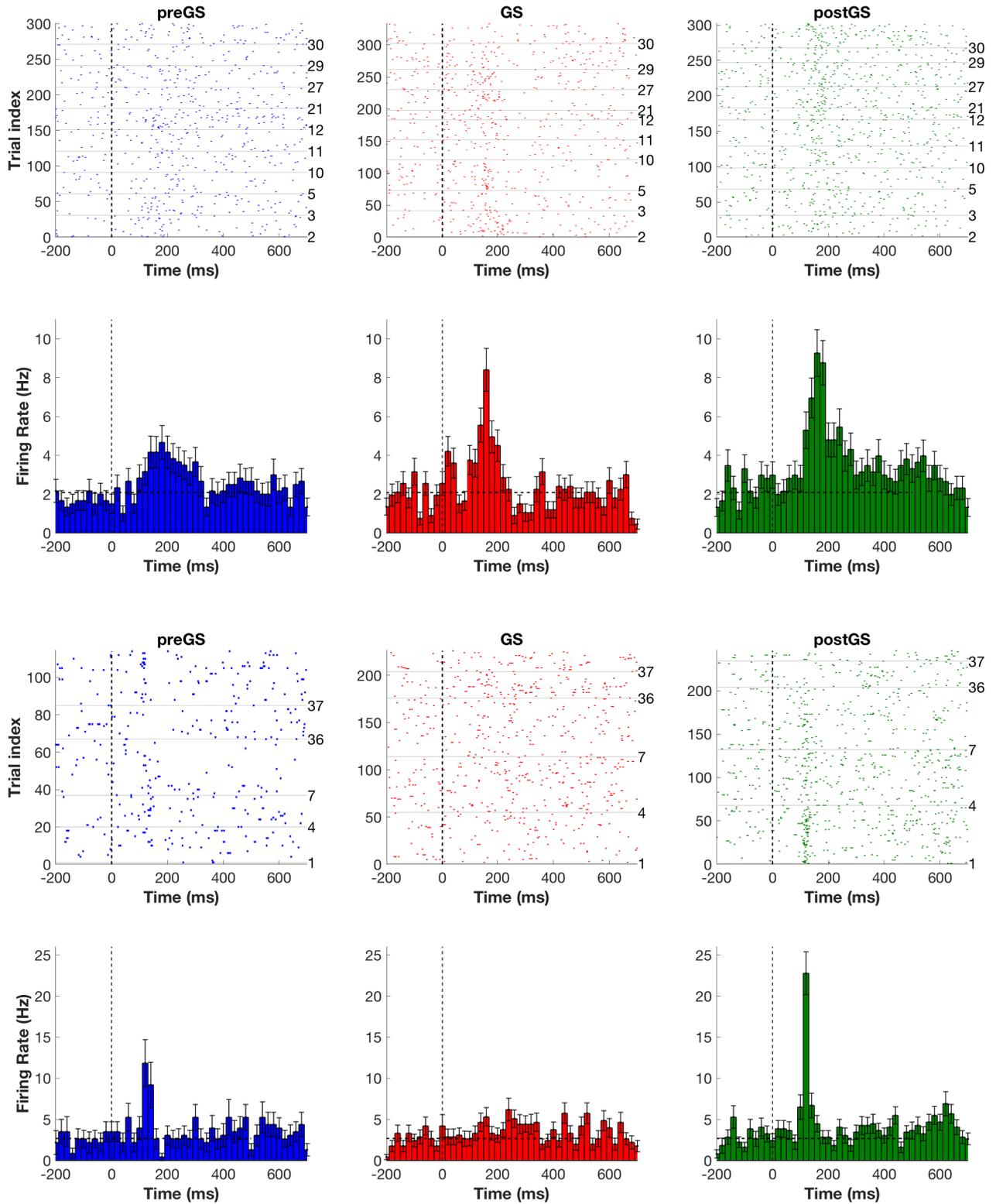


Figure 8: Response of two single units in the occipital lobe of an epileptic patient showing an example of learning during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (bottom) shows the average firing rate of the neuron during each condition.

At the exemplar level, 129 exemplars showed a difference between preGS and postGS (Table 5). This result is well above difference as due to random chance, obtained by shuffling the conditions (on average, 45). The rationale behind the shuffle test is that we wanted to check if the significant exemplars obtained was matter of chance or not. Figure 9 shows raster plots of evoked responses before and after learning in 2 distinct neurons of the occipital lobe and parahippocampal gyrus for a particular exemplar.

Anatomical Region	# Sig. exemplars	% Sig.	Tot. # exemplars
Amygdala	19	6.7	282
Entorhinal	2	3.1	64
Hippocampus	28	8.0	346
Occipital Lobe	34	8.6	393
Parahippocampal Gyrus	22	6.3	350
Anterior Temporal Lobe	4	4.8	83
Posterior Temporal Lobe	19	7.5	254
Parietal Cortex	1	1.0	10
Total	129	7.2	1782

Table 5: Exemplars in each region that show a significant difference between preGS and postGS condition

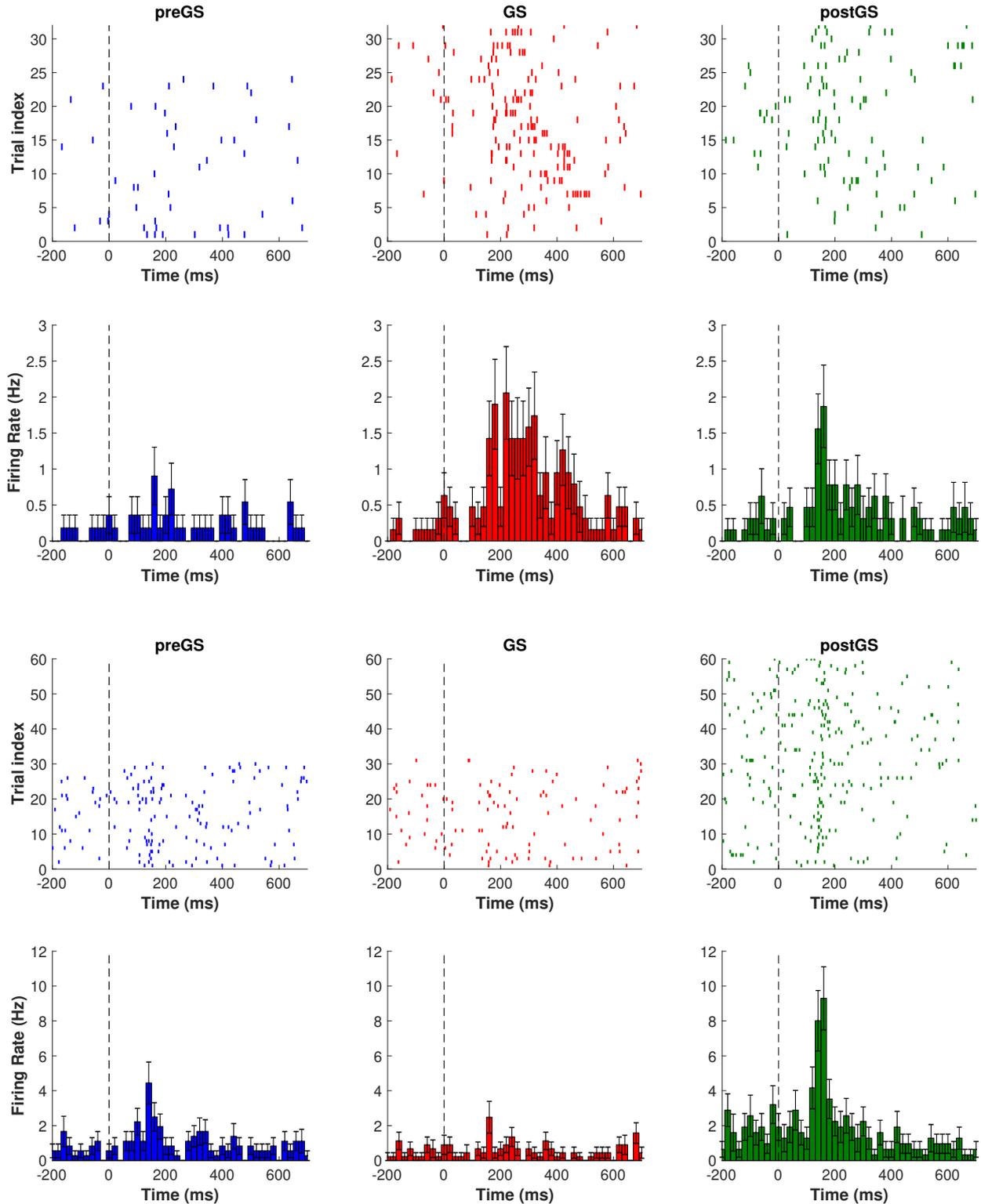


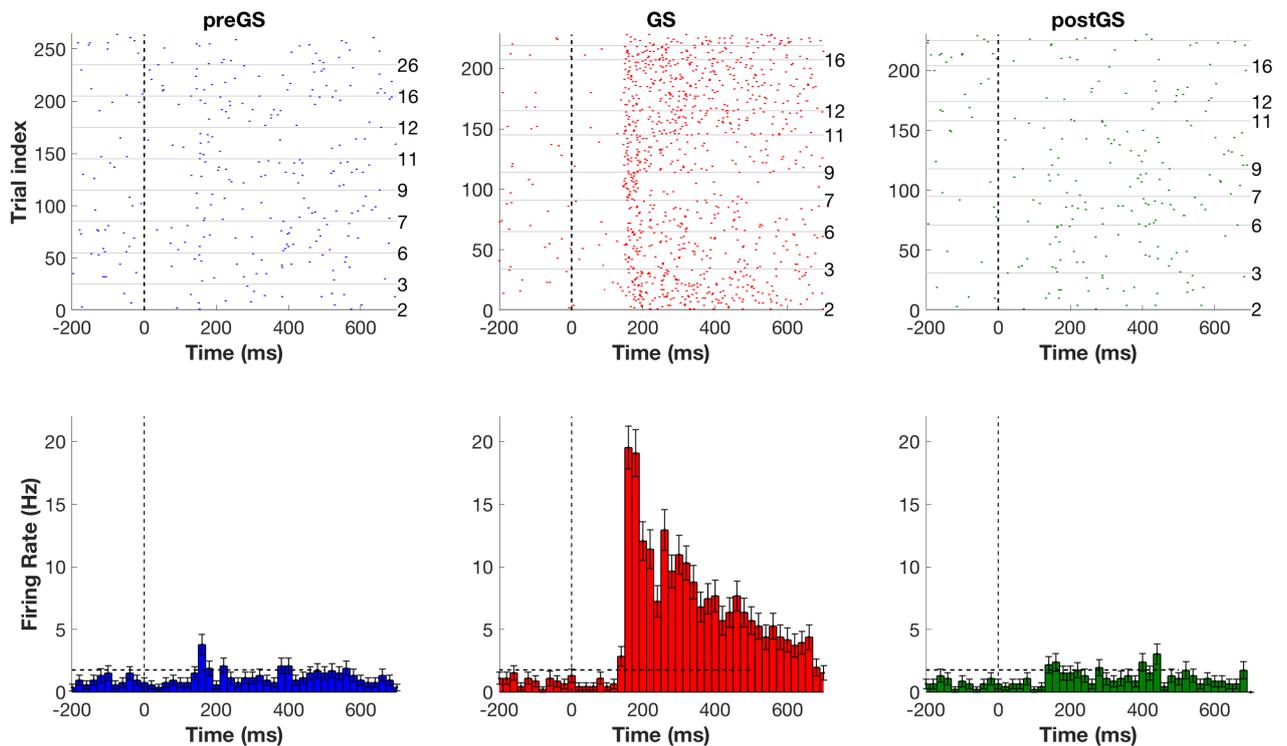
Figure 9: Response of two single units in the occipital lobe and parahippocampal gyrus of an epileptic patient showing an example of learning at the exemplar level during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The post-stimulus time histogram (bottom) shows the average firing rate of the neuron during each condition.

4.3.2 Responses to stimulus features

Out of the 293 responsive units, 70 responded differently to the GS condition than to preGS and postGS (Table 6). One region showed a particularly high number of units, the parahippocampal gyrus with almost half of the responsive units showing a statistically different response. The occipital lobe also showed different responses to Mooney and grayscale images. Figure 10 shows the responses of 3 representative GS-selective single neurons.

Anatomical Region	# Sig. neurons	% Sig.	Tot. # neurons
Amygdala	0	0.0	40
Entorhinal	2	28.6	7
Hippocampus	1	4.8	21
Occipital Lobe	31	29.5	105
Parahippocampal Gyrus	35	46.7	75
Anterior Temporal Lobe	0	0.0	8
Posterior Temporal Lobe	1	2.7	37
Parietal Cortex	0	0.0	0
Total	70	23.9	293

Table 6: Units in each region that show a significant difference between preGS and postGS versus GS condition



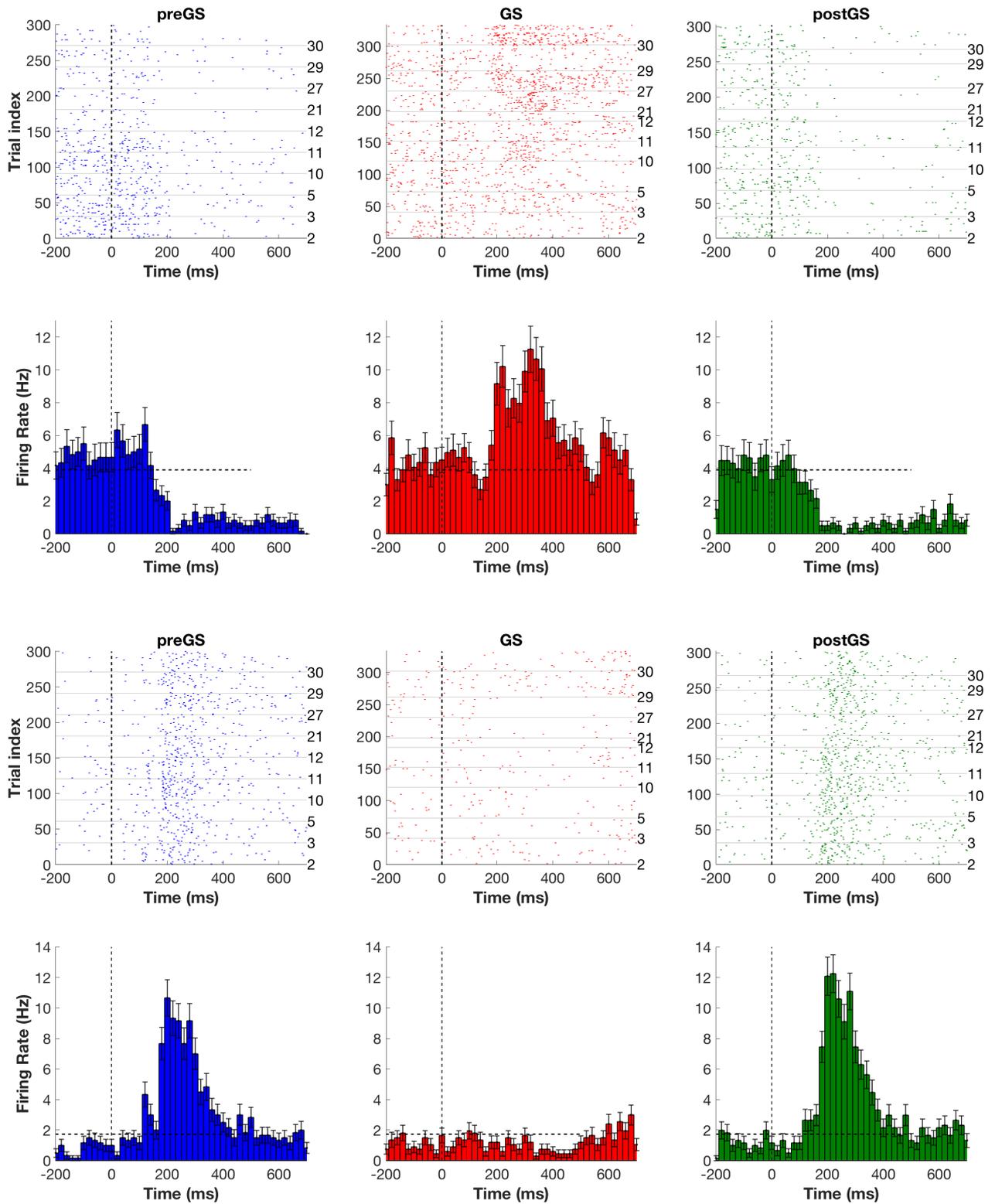


Figure 10: Response of three single units in the occipital lobe and the parahippocampal gyrus of an epileptic patient showing an example of grayscale condition specific response during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (bottom) shows the average firing rate of the neuron during each condition.

4.4 Selectivity

Figure 11 shows the number of neurons responding differently across categories for the different conditions.

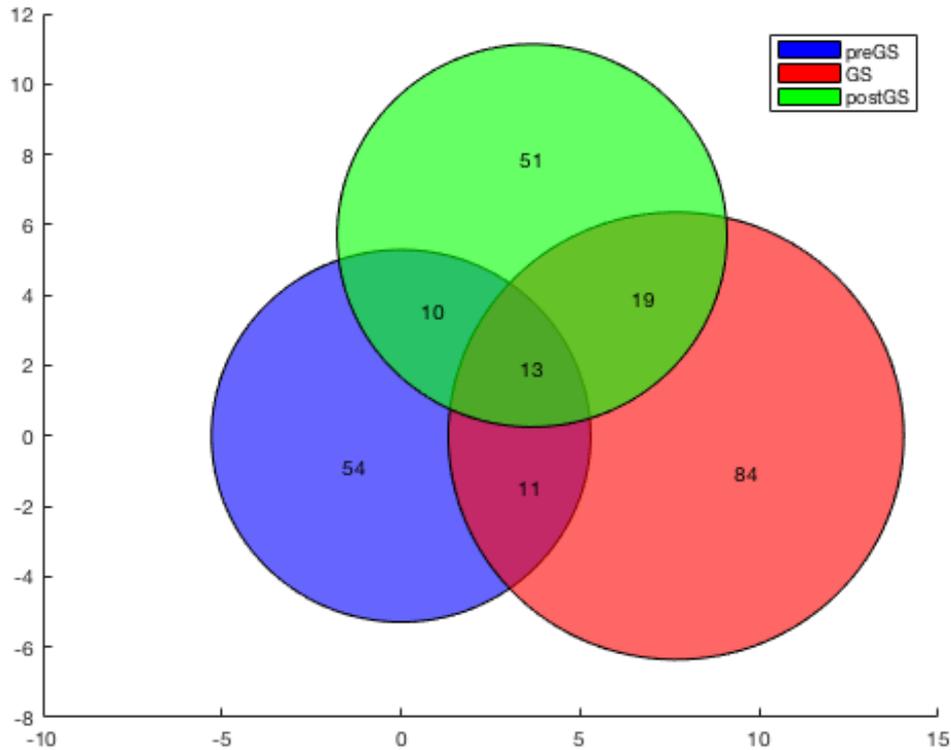


Figure 11: Venn diagram illustrating the number of neurons showing differences in firing rate among the different categories

Thereafter, we only considered the GS and postGS conditions as the images in the preGS condition were not recognized. Figure 12 shows an example of a neuron in the parahippocampal gyrus with a particularly strong response to flowers for the GS condition. A one-way ANOVA yielded $p > 0.99$, and subsequent across-categories pairwise comparisons also showed that the activity during stimulus presentation was significantly higher for this category. One visually selective neuron in the occipital lobe (Figure 13) showed an increased firing rate over baseline in response to animal pictures. Both the ANOVA test and pairwise comparison showed a significant selective response. Some of the neurons also showed changes in firing rate in response to more than one of the categories. Most of these neurons responded to one or two categories.

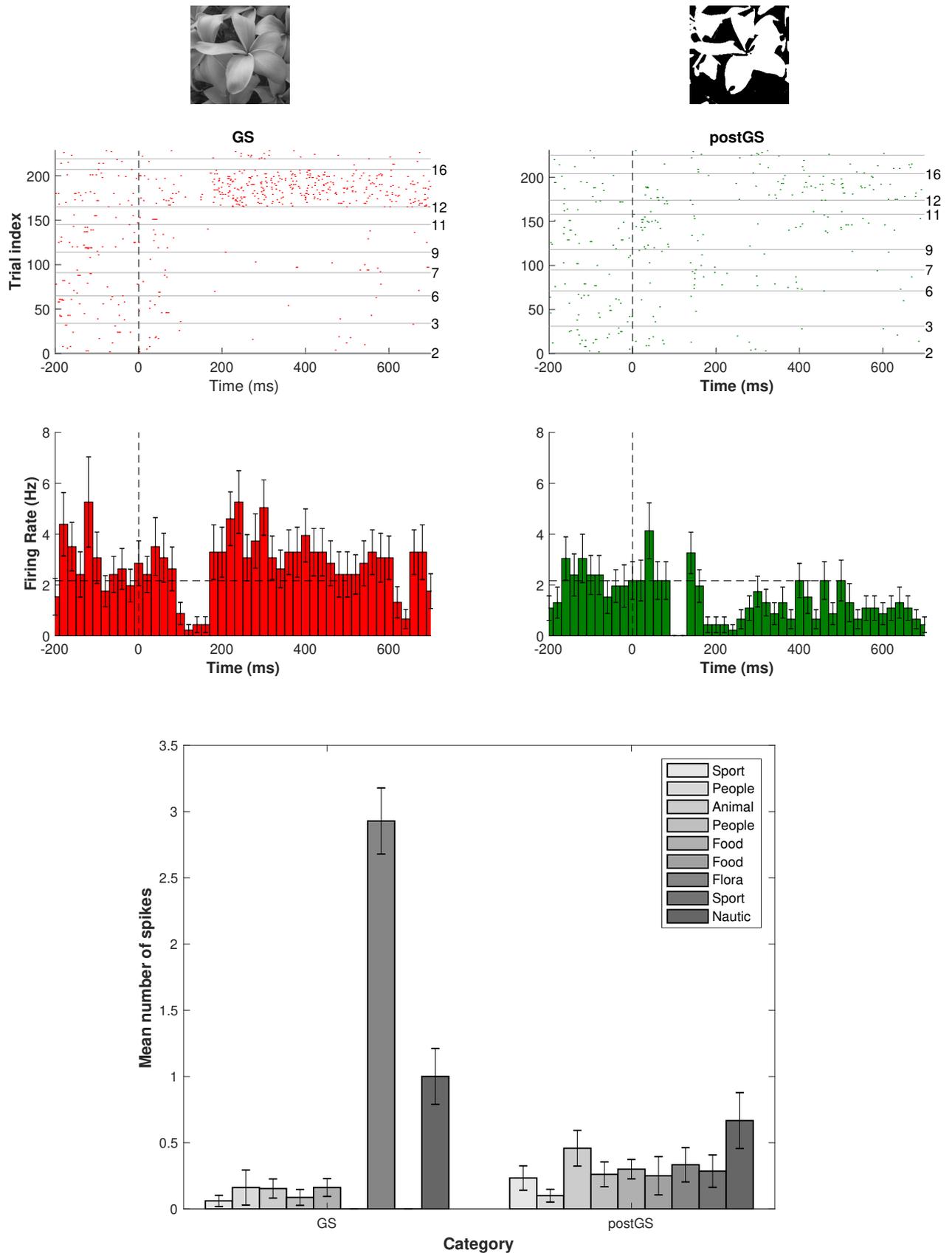


Figure 12: Response of a single unit in the parahippocampal gyrus of an epileptic patient showing an example of selectivity during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (middle) shows the average firing rate of the neuron during each condition. The average number of spikes is shown for each image (bottom).

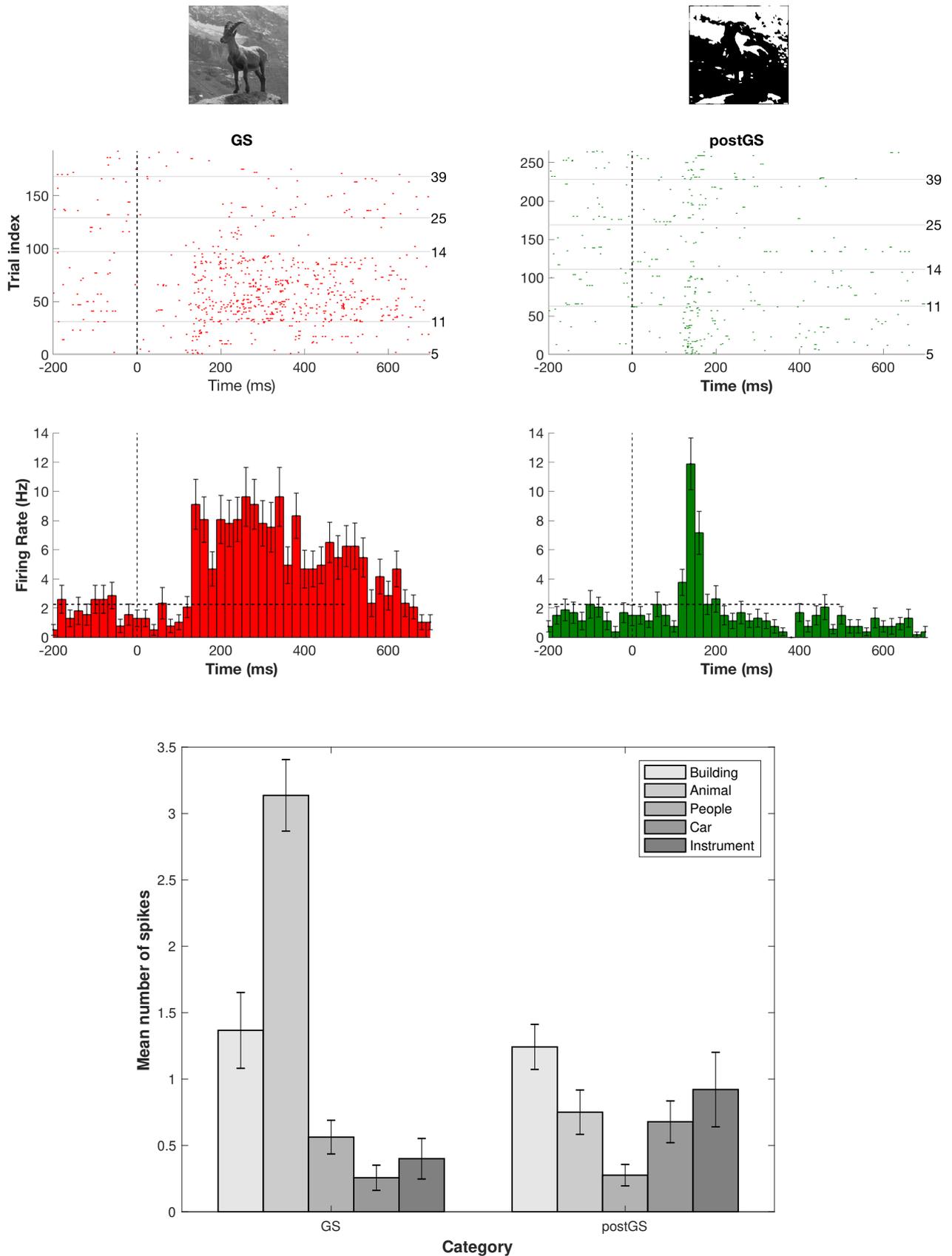


Figure 13: Response of a single unit in the occipital lobe of an epileptic patient showing an example of selectivity during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (middle) shows the average firing rate of the neuron during each condition. The average number of spikes is shown for each image (bottom).

Also, in order to identify neurons responding selectively to a specific category for different stimuli types, we looked at the neurons that showed both a selective response for GS and postGS. Figure [14](#) allows to visualize the stronger activity for animal stimuli than other categories like building, car, people or instrument for both GS and postGS conditions.

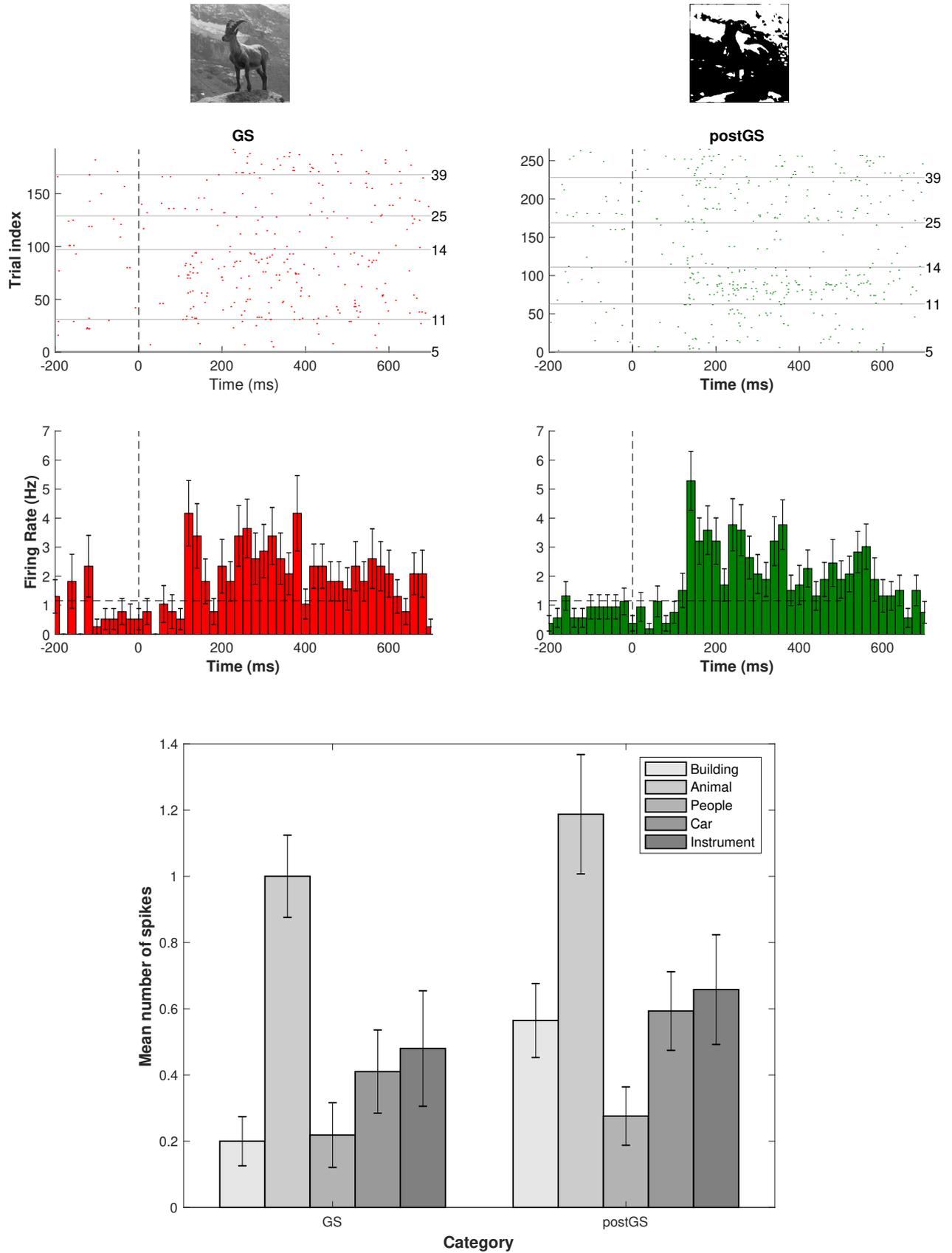


Figure 14: Response of a single unit in the occipital lobe of an epileptic patient showing an example of selectivity during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (middle) shows the average firing rate of the neuron during each condition. The average number of spikes is shown for each image (bottom).

Respectively 27% and 17% of the visually responsive neurons showed selectivity for the GS and postGS conditions (Table 7). Regarding the regions in which we found the most selective responses, the parahippocampal gyrus encompasses respectively 40% and 19% of neurons showing a selective response for GS and postGS conditions. A high number of selective responses were also found in the occipital lobe with respectively 36% and 29% of neurons responding differently to the different categories. For the enthorinal cortex, whereas a significant proportion of neurons are selective for the GS condition, none are listed for the postGS condition. We also observed neurons that showed significant but nonselective changes in firing rate between categories during stimulus presentation, 7% for GS and 9% for postGS condition. However, most neurons for which the ANOVA test indicated difference between categories were also selective.

Anatomical Region	% Selective for GS	% for postGS	# Responsive neurons
Amygdala	7.5 (3)	10.0 (4)	40
Enthorhinal	28.6 (2)	0.0 (0)	7
Hippocampus	0.0 (0)	0.0 (0)	21
Occipital Lobe	36.2 (38)	28.6 (30)	105
Parahippocampal Gyrus	40.0 (30)	18.7 (14)	75
Anterior Temporal Lobe	0.0 (0)	0.0 (0)	8
Posterior Temporal Lobe	13.5 (5)	5.4 (2)	37
Parietal Cortex	0.0 (0)	0.0 (0)	0
Total	26.6 (78)	17.1 (50)	293

Table 7: Units in each region showing a significant selective response according to the criteria mentioned in the text. The numbers in parenthesis indicate the number of visually responsive neurons showing a selective response.

4.5 Inhibitory responses of single neurons during recognition of objects

An important part of the responsive neurons showed an inhibitory response (more than 40%). Recording from single neurons in different regions of the brains during visual recognition, we found that in the anterior temporal lobe and amygdala visual stimulus evoked predominantly suppression of neuronal firing below prestimulus baseline (‘inhibitory responses’). 87.5% of the responsive neurons in the anterior temporal lobe showed an inhibitory response (Table 8).

Anatomical Region	# Sig. neurons	% Sig.	Tot. # neurons
Amygdala	22	55.0	40
Enthorhinal	3	42.9	7
Hippocampus	8	38.0	21
Occipital Lobe	43	40.9	105
Parahippocampal Gyrus	21	28.0	75
Anterior Temporal Lobe	7	87.5	8
Posterior Temporal Lobe	18	48.6	37
Parietal Cortex	0	0.0	0
Total	122	41.6	293

Table 8: Units in each region showing a significant an inhibitory response

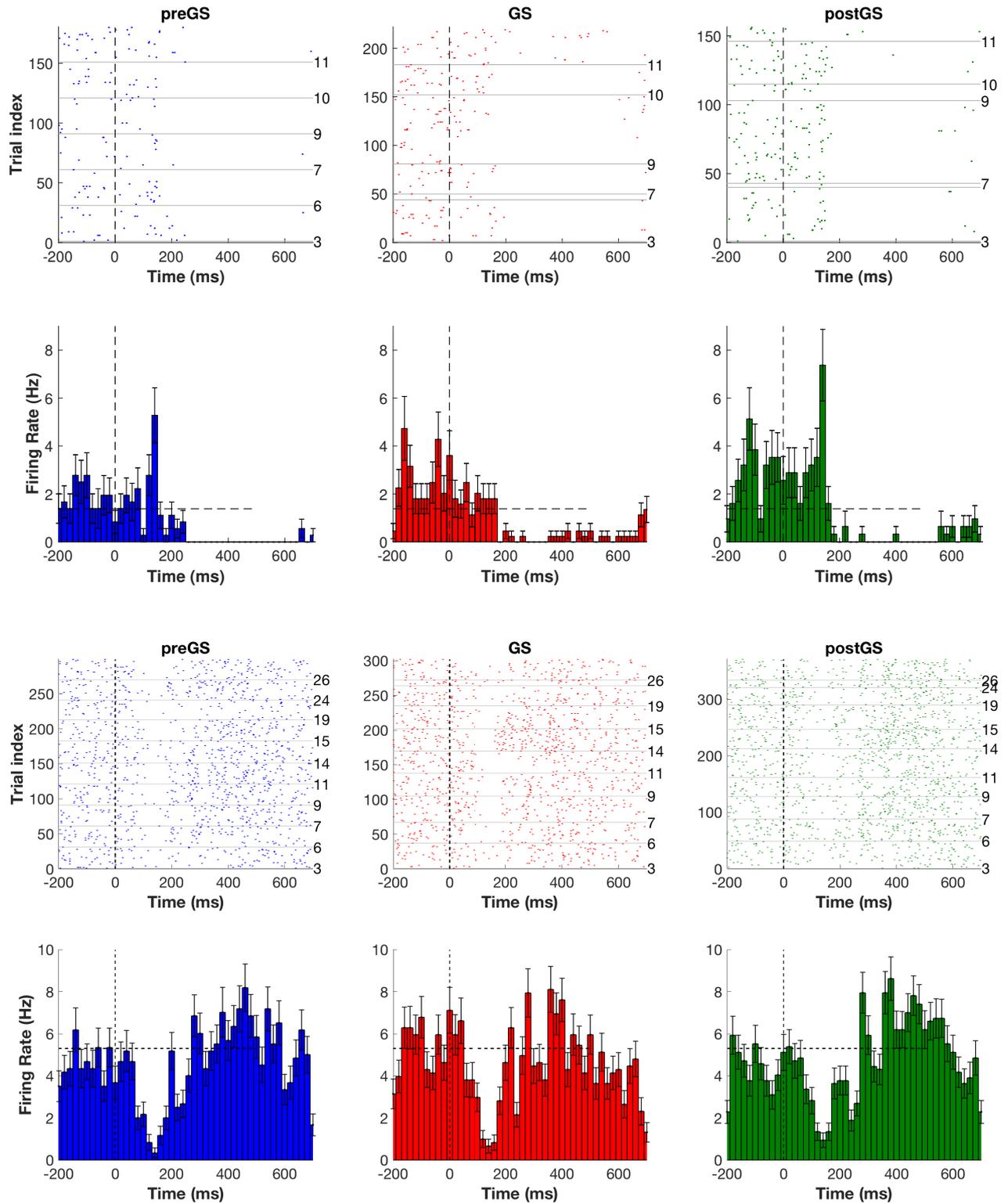


Figure 15: Response of two single units in the parahippocampal gyrus and occipital lobe of an epileptic patient showing an example of inhibition during a visual memory task. The raster plot (top) shows the spikes aligned to stimulus onset (at $t = 0$ ms). The numbers on the right side indicate the label of the image. The post-stimulus time histogram (bottom) shows the average firing rate of the neuron during each condition.

4.6 Classification of visual trials from single neuron recordings

We present results from classifying the activity of neurons as either ‘stimulus’ or ‘baseline’, based on SVM models. Then, the results for the ‘preGS’ versus ‘postGS’ conditions are stated. We report results using spikes from 250 ms or 500 ms time window and from classifiers using data from all neurons, neurons from each of the individual sessions of recordings, and classifiers where only neurons from one region are used.

4.6.1 Classification of visual object presence

We assessed the accuracy of the model in classifying new trial recordings as either ‘baseline’ or ‘stimulus’, based on SVM models built on prior single cell recordings. Training SVM classifiers based on 250 ms or 500 ms time window of all channels from all subjects simultaneously appeared to give a high accuracy of 99.19% and 99.46% respectively (Figure 16). These results are encouraging for the continuation of the study as visual object presentation could be distinguished at all. Data used to train the classifier contain information that is indicative of one state or another as it allowed high classification accuracy.

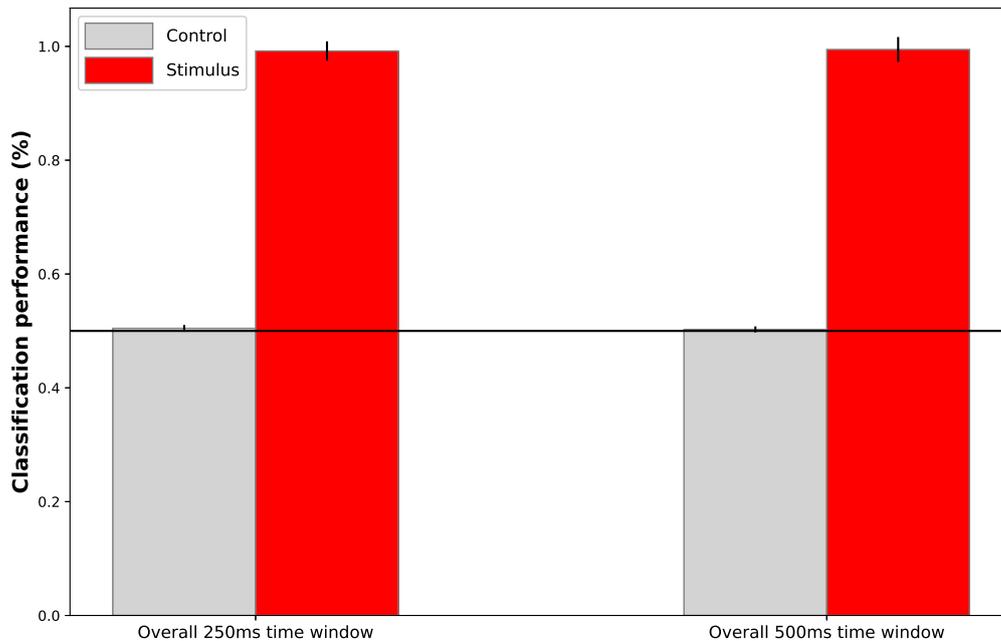


Figure 16: Mean classifier performance across all subjects. The error bars show the standard deviation of the mean. After cross-validation, the model yielded more than 99% classification performance in identifying ‘stimulus’ from ‘baseline’ for the 250 ms and 500 ms time windows based on neural activity

We also show the results of ‘stimulus’ versus ‘baseline’ classification for each subject in Figure 17. Among the 34 sessions, none outperformed the performance for all units combined although some sessions showed a higher performance compared to others.

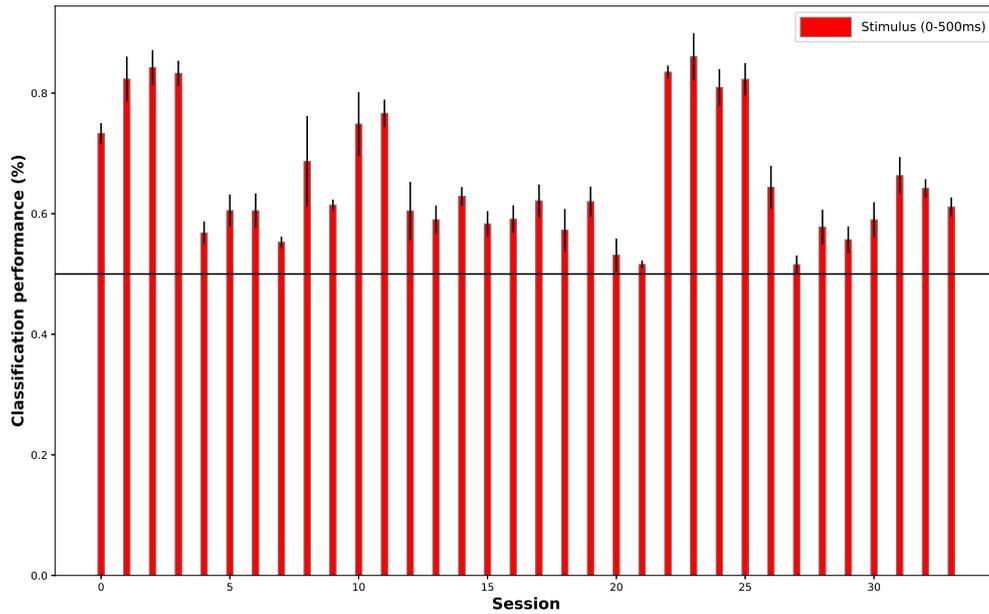


Figure 17: Mean classifier performance for each session. The error bars show the standard deviation of the mean.

4.6.2 Classification of different conditions

We wanted to see if our SVM model could perform well on behavioral state classification. However, the accuracy for the stimulus time period was not different from the baseline (Figure 18). Since we were not able to predict preGS and postGS conditions with all the neurons combined, we focused and separated the analysis by region. The occipital lobe and parahippocampal gyrus neurons clearly yielded higher classifier performance than other regions for the 500 ms time window (Figure 19). For instance, decoding performance of 95% is reached by using the activity of neurons from one of these two regions. They also showed a statistical difference in accuracy from the baseline and post-stimulus which is a good control to ensure the quality of our results. For the 250 ms time window, again the occipital lobe was shown to have a high classification accuracy for the stimulus period and a low one for the baseline, both statistically different (Figure 20). Thus far, we have demonstrated that recognition by a subject can be predicted using the SVM classifier with data from neurons in the occipital lobe. The algorithm achieved a higher decoding by using the 500 ms time window and detected the condition correctly with accuracy greater than 95%.

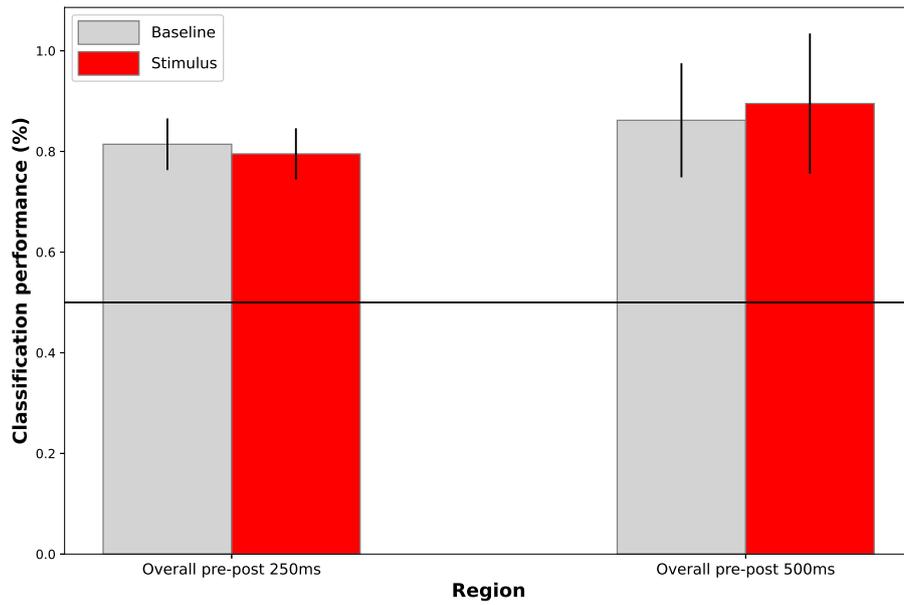


Figure 18: Mean classifier performance for two time windows. The error bars show the standard deviation of the mean. After cross-validation, the model didn't yield a difference in classification performance in identifying preGS or postGS condition for stimulus versus baseline

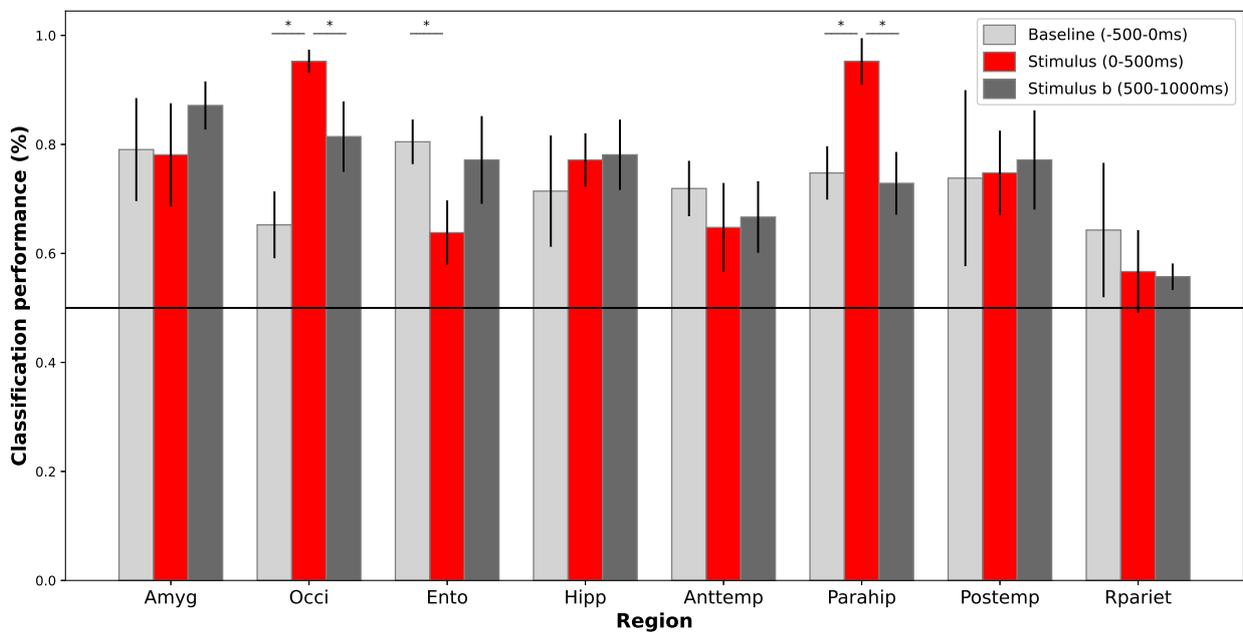


Figure 19: Mean classifier performance for each region. The error bars show the standard deviation of the mean. Only the occipital lobe and the parahippocampal gyrus showed a difference in accuracy from the baseline and post-stimulus. After cross-validation, the model yielded more than 95% classification performance in identifying preGS or postGS condition for both regions.

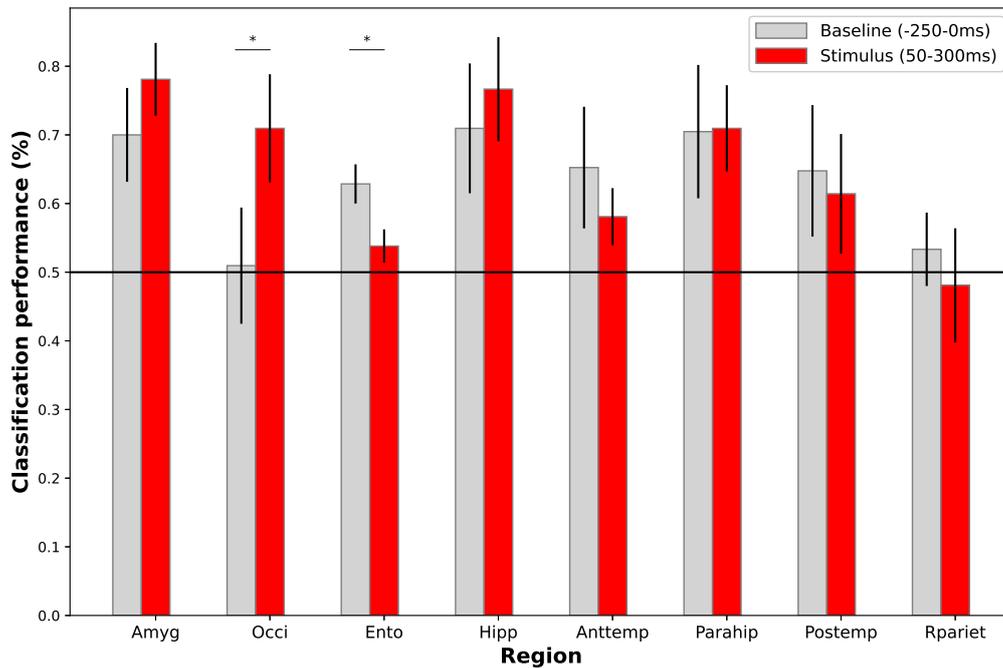


Figure 20: Mean classifier performance for each region. The error bars show the standard deviation of the mean. Only the occipital lobe and the enthorhinal cortex showed a difference in accuracy from the baseline. After cross-validation, the model yielded more nearly 80% classification performance in identifying preGS or postGS condition for the occipital lobe.

4.6.3 SVM feature weights

The occipital lobe being the region showing the strongest accuracy to decode preGS of postGS, we were interested in the weights associated with each input neuron. From the 163 neurons, the 75 (46%) having a weight in absolute value higher than 0.1 represented 70% of the neurons showing a difference between preGS and postGS. Our results from the SVM classifier are thus in line with what was obtained in the permutation tests presented above. Taken together, the results suggest that few neurons in the occipital lobe showed progressive changes in firing rate between the preGS and postGS conditions which could reveal evidence of a modulation in the brain that led to learning.

4.7 Neural correlates of Eureka moment

The visualisation of firing rate of the last preGS trials in parallel with the first postGS trials in order to identify a change in activity didn't show significant difference. The variability in the tasks used and the low number of available trials make it difficult to get a clear picture of the neurobiological underpinnings of the Eureka moment.

4.8 Similarity in Neuronal Firing

As seen in the previous section, the degree of rate modulation can vary considerably across categories within the same neuron. In order to avoid category bias, we decided to analyze the results at the exemplar level instead of doing it per neuron. The control performed by calculating

the indexes over the baseline period gave an average of 54% of positive results. On the other hand, after the stimulus onset, this number was on average 71% across the different regions. This shows a tendency of the firing rate to fluctuate towards the GS condition and possibly a more similar activity between the recognized images than when perception differs. However, not all regions showed an increase in the number of positive indexes. Moreover, one must be aware that other parameters could be taken into account in the calculation of the modulation index like the latency or the magnitude of the peaks. Also, this method is not highly precise and a positive index doesn't necessarily mean learning.

4.9 Timings

A total of 84 units showed a measurable excitatory response according to our previously mentioned criteria for one of the 3 conditions (preGS, GS or postGS) [42 (50%) in the occipital lobe; 32 (38%) in the parahippocampal gyrus]. We will detail the results for these two regions only as we had less than 10 units for the others. Average response latencies for the 3 conditions were earlier (7 ms) in the parahippocampal gyrus than those in the occipital lobe. However, the difference is not significant (Figure 21). Because we used an objective criterion to determine response latencies, we cannot expect our approach to be perfect, and our results may contain a few unreliable latencies. Statistical analysis across all 84 neurons didn't show any difference between preGS and postGS conditions ($p = 0.9737$). Separate regional analysis confirmed no statistically difference between latency for the 2 conditions in the occipital lobe ($p = 0.6395$) and parahippocampal gyrus ($p = 0.2241$).

Regarding inhibitory responses, a total of 37 units showed a measurable response for one of the 3 conditions (preGS, GS or postGS) [13 (35%) in the occipital lobe; 7 (19%) in the parahippocampal gyrus]. Unlike for the excitatory cells, responses latencies for the 3 conditions were higher (20.3 ms) in the parahippocampal gyrus than those in the occipital lobe for the preGS and postGS condition. Each of the 3 conditions showed a later response in the parahippocampal gyrus. However, this difference is not significant. Statistical analysis across all 37 neurons didn't show any difference between preGS and postGS conditions ($p = 0.9984$). Separate regional analysis confirmed no statistically difference between latency for the 2 conditions in the occipital lobe. Not enough neurons showed a measurable response in the parahippocampal gyrus for both conditions to compute a statistical test.

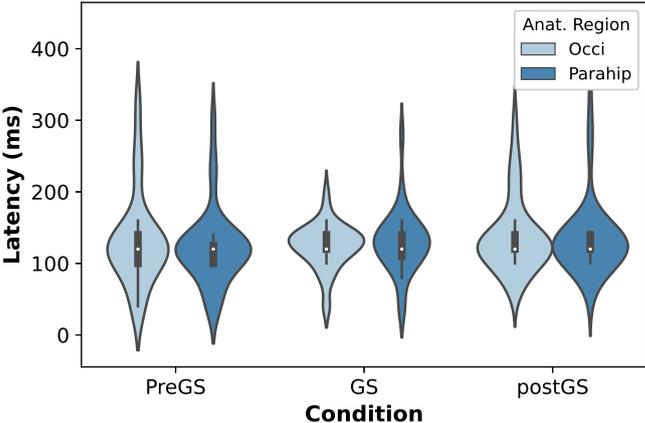


Figure 21: Violin plot showing the distribution of latency for excitatory neurons. For each condition, we separated the analysis by brain regions. The contours of the violins show kernel density estimates of the distributions, truncated at min and max observed values. White circles indicate the medians, thick bars indicate the first and third quartiles, and whiskers indicate $1.5 \times$ interquartile ranges.

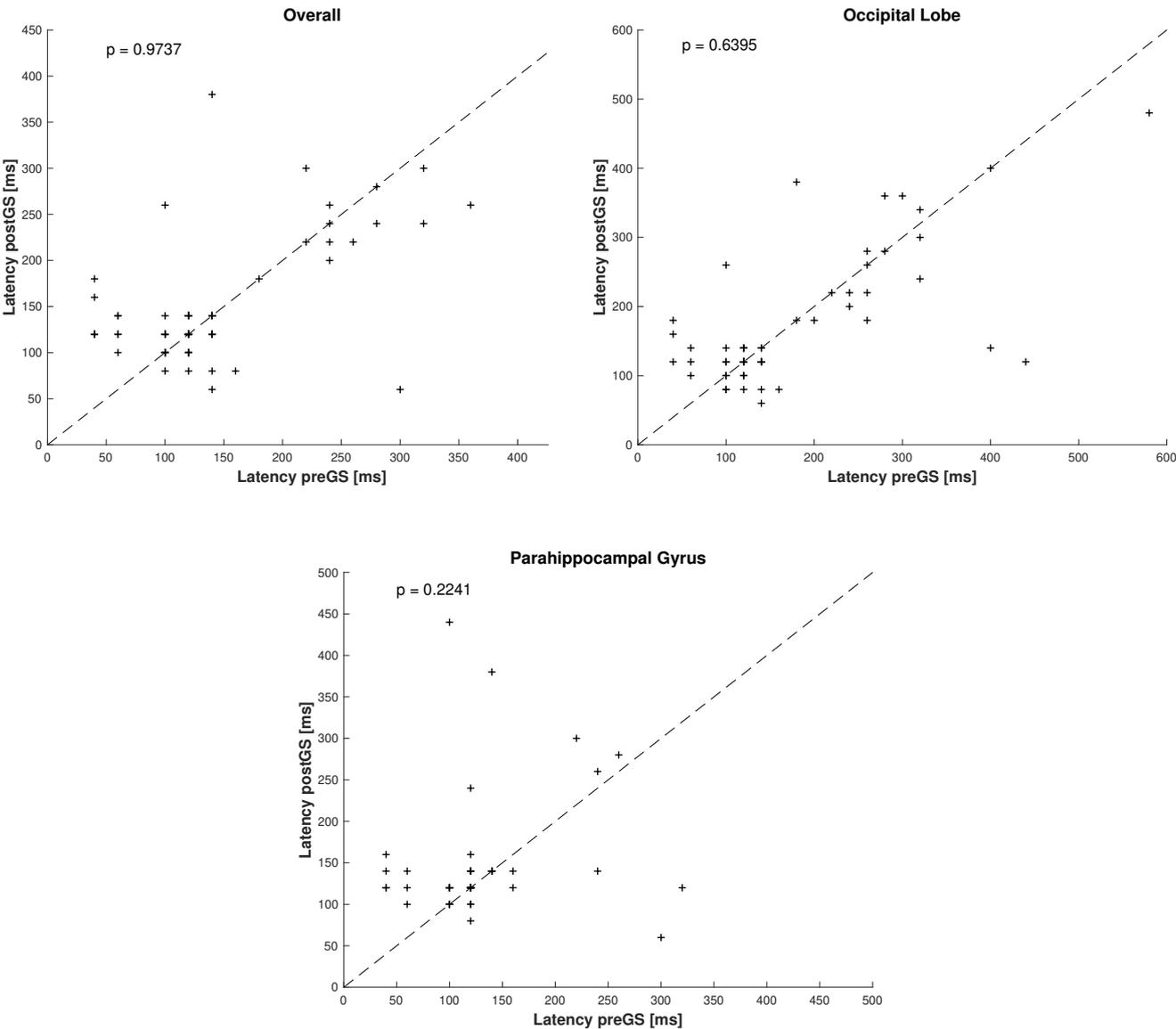


Figure 22: Scatter plots of preGS versus postGS latencies for excitatory neurons. Each point accounts for a neuron showing a responsive and measurable response.

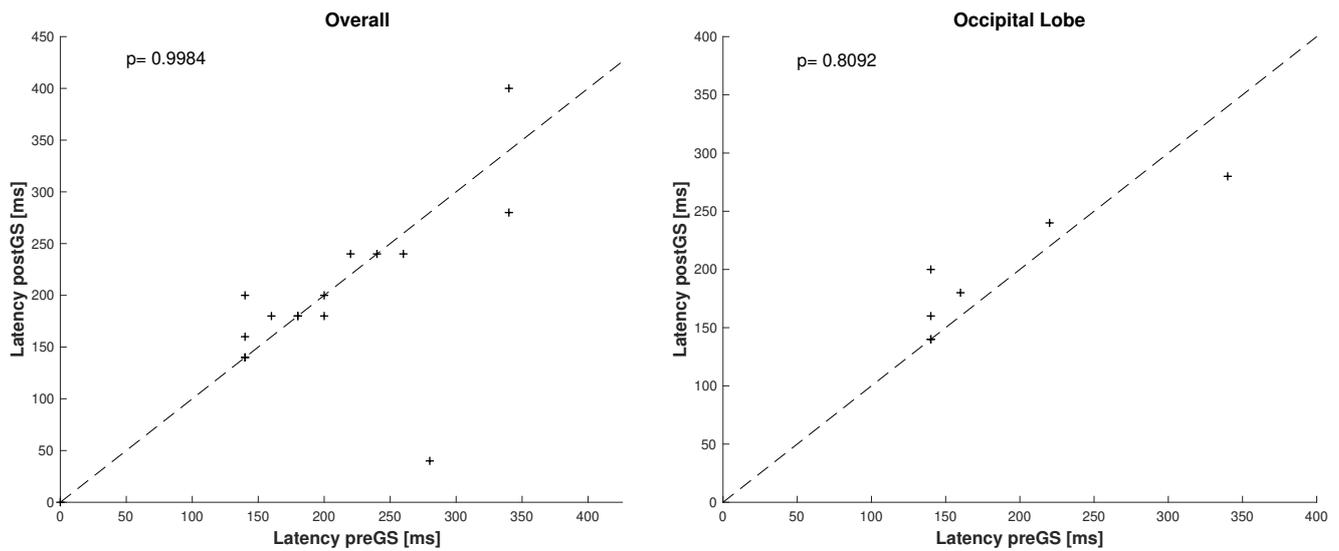


Figure 23: Scatter plots of preGS versus postGS latencies for inhibitory neurons. Each point accounts for a neuron showing a responsive and measurable response.

When looking at the scatter plots (Figure 22 and 23) for both the excitatory and inhibitory latencies, the results hardly show any tendency. Latencies might not characterize the difference in object recognition as well as expected or might reveal the imprecision and the difficulty of its measurement.

5 Discussion

The increase in performance during object recognition between the presentation of novel versus learned Mooney images reveals how this type of stimulus can offer a means of inducing variable perception with constant stimulus and study neuronal activity during learning. The exact same Mooney image presented before and after learning elicited different patterns and these changes in firing rates allowed us to identify relevant regions for learning.

The main goal of this study was to investigate the neural correlates of learning during an object recognition task. Using extracellular single cells recordings in humans, we analyzed neuronal activity evoked by Mooney images before and after they were recognized as meaningful objects. Recognition induced changes, mainly in the hippocampus. Previous studies have pointed at the hippocampus as the central hub for memory formation and highlighted its critical role for long term memory [38]. In line with previous work, we identified significant hippocampal contributions to rapid learning. A large number of visually responsive neurons showed a modulation in firing rate in this region and this study provides further evidence that the hippocampus plays an important role in encoding and recall of memory. This result parallels the increased activation observed in some human studies after learning in visual memory tasks. However, little is known about the features and the time scale of such changes. These are questions that will need to be addressed in future work. At the exemplar level, the results were consistent with those obtained previously with differences mainly in the hippocampus. Importantly, we also found induced modulation in the occipital lobe. Although the occipital lobe is often viewed as a visual processing module, our results might indicate that cortical interactions play a role in learning.

The processing delays in excitatory cells associated with object recognition did not increase along the visual hierarchy as expected in feedforward visual processing. In fact, for an object to be recognized, it has to travel through the ventral visual cortex, from earlier visual areas (V1) to the temporal cortex accumulating delays in processing. Importantly, these delays can increase when additional processing (recurrent or feedback) is required to allow recognition. Our results showed a delayed response in the occipital lobe and revealed conceivable involvement of various feedback processing between the medial temporal lobe and the occipital lobe. The timing of neurophysiological responses gives great insights on the implication of top-down modulation of the visual cortex. However, no difference could be identified in the response time between preGS and postGS condition. Thus, at this point, the analysis of timing cannot show how temporal patterns might be modulated and influenced by perception in our learning task. The location of the electrodes in our study was dictated by clinical criteria. Therefore, the electrode coverage was far from exhaustive. There may well be other areas in the human brain, not interrogated in this study, which might show a direct relationship between latency and recognition. Inhibitory cells showed an early response in the occipital lobe followed by a difference in activity later in the parahippocampal gyrus. Our results suggest that inhibition occurs at all levels of processing of visual information and that it facilitates bottom-up attentional modulation in the primary visual cortex [39].

The images used in this study were both two-tone black and white and corresponding grayscale images, shown in alternation. By using such stimulus set, we were able to test the responses to the same objects with different luminance, contrast and more or less defined contours. While color and contrast selectivities has been studied extensively in macaques, in human, luminance and contrast perception is not well understood. Even if the main focus of this study was not on understanding of the representation of luminance and brightness in the visual cortex, we confirmed that they might have effects on the activity of neurons in the primary visual cortex, which had already been demonstrated in awake monkeys [40]. Moreover, by going back to the fact that young children are generally unable to recognize two-tone images, these results shed

light on where perceptual reorganization might happen in the brain and how important they are to allow latter instant recognition in adults.

Research on the neurophysiological correlates of visual object recognition in a sudden and unexpected Eureka moment is still at its early stages. Here, we wanted to understand more about this mysterious ‘Aha!’ sensation. Learning how to rapidly extract an object identity can occur by an abrupt flash of understanding during which nerve cells in the brain allow sudden insight by changing their activity pattern. However, using the data we had at our disposal, we found that the neuronal firing was not a direct reflection of sudden object recognition behavior, showing a response that changed only marginally between the last preGS trials (unrecognized) and the first postGS trials (recognized).

In this study, we found an important portion of visually responsive neurons showing inhibition, which gave the opportunity to identify the regions in which this phenomenon is most present and plays a role in recognition. Responses of inhibitory cells have attracted less attention, mostly due to their smaller size and larger diversity. Despite the fact that there are fewer of them, inhibitory cells play an essential role in shaping the dynamics and modulation of the neuronal circuits. From a computational point of view, network models increasingly incorporate inhibitory synaptic plasticity [41]. The vast majority of the inhibitory responses to visual stimuli that we found were in the anterior temporal lobe and the amygdala also showed an important part of this type of response. It is possible, however, that the prevalence of inhibitory responses is a function of the epileptic brain. Even if the data were collected during periods without any seizure events, we can’t rule out the possibility of a disturbance leading to an inhibition of certain neurons, particularly in regions with a high prevalence.

Selective neural responses to visual objects is a topic of great interest in neuroscience. This fundamental ability to classify visual inputs into meaningful categories depends on neural computations performed along the visual pathway from the retina to the higher visual areas in ventral temporal cortex. In humans, functional magnetic resonance imaging (fMRI) [42], electrocorticography (ECoG), evoked-potential studies and single-neuron electrophysiology [22] researches have been influential in revealing preferential activation to categories such as faces, animals, building in specialised areas in the brain. They have identified several regions of the occipital and temporal lobe that appear specific for complex stimuli [43]. Studies of patients with agnosia have revealed that the deficit of stimulus recognition does not necessarily impair the recognition of all visual stimuli, but can selectively affect certain categories (faces, spatial layouts, letters, body parts), leaving others unaffected [44]. We showed that in the occipital lobe and the parahippocampus gyrus, there is a relatively high degree of category-specific neurons, showing a firing patterns representing certain objects. Surprisingly, we did not identify as many selective neurons as expected in some regions of the medial temporal lobe while its role in the representation of different categories of visual stimuli has been previously demonstrated [45] [46]. No selective neurons were found in the hippocampus, less than 10% of the responses in the amygdala are selective and a significant proportion of neurons are selective in the enthorinal cortex only for the GS condition. We established that single neurons in humans explicitly respond to specific categories of stimuli. However, another definition of the categories given here may show other results, with neurons changing their activity only to specific examples. Moreover, it could be important to examine recognition for an entire population of neurons by pooling together responsive units.

We described a method for classifying visual perception from neuronal activity patterns using a machine learning approach. We demonstrated that our classifier can distinguish responses to visual object presentation on a trial-by-trial basis from data when a blank screen was pre-

sented to the subject at a 95% accuracy. The question of whether the decoder was subsequently able to decode learning in single trials based on the activity of a neuronal pseudopopulation led to no significant performance accuracy when using activity from units in all regions. However, when classifying preGS versus postGS condition using classifier inputs of the occipital lobe gave an accuracy well above chance. This presents evidence that activity from 0 ms to 500 ms after stimulus onset in the occipital lobe can be used to predict recognition. On the other hand, the weights associated with the input neurons are consistent with the results obtained from the statistical tests. The same units showed both absolute weights with large values and a significant difference between the two conditions. Taken together, these results show an evidence of rapid plastic changes in the occipital lobe associated with object recognition.

There are rare cases where Mooney images were not recognized even after seeing the grayscale corresponding image. By considering those cases, future analyses of the neuronal responses associated could be conducted to try to identify neurons that fired differently depending on the perception and understand what happened at the neuronal level that made the patient unable to memorize the image. However, one must be aware that these cases are much less numerous and it is therefore more difficult to conduct analysis based on a few trials.

References

- [1] M. Gupta, A. C. Ireland, and B. Bordoni, “Neuroanatomy, Visual Pathway,” in *StatPearls*. Treasure Island (FL): StatPearls Publishing, 2022. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK553189/>
- [2] “Anatomy of the Visual Pathways : Journal of Glaucoma.” [Online]. Available: https://journals.lww.com/glaucomajournal/fulltext/2013/06001/anatomy_of_the_visual_pathways.4.aspx
- [3] S. Prasad and S. L. Galetta, “Anatomy and physiology of the afferent visual system,” in *Handbook of Clinical Neurology*. Elsevier, 2011, vol. 102, pp. 3–19. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/B9780444529039000078>
- [4] “The Optic Nerve - Visual Pathway - Chiasm - Tract - TeachMeAnatomy.” [Online]. Available: <https://teachmeanatomy.info/head/cranial-nerves/optic-cnii/>
- [5] L. G. Ungerleider and L. Pessoa, “What and where pathways,” *Scholarpedia*, vol. 3, no. 11, p. 5342, Nov. 2008. [Online]. Available: http://www.scholarpedia.org/article/What_and_where_pathways
- [6] M. Mishkin, L. G. Ungerleider, and K. A. Macko, “Object vision and spatial vision: two cortical pathways,” *Trends in Neurosciences*, vol. 6, pp. 414–417, Jan. 1983. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/016622368390190X>
- [7] “Cognitive neuroscience of visual object recognition.” [Online]. Available: https://psychology.fandom.com/wiki/Cognitive_neuroscience_of_visual_object_recognition
- [8] L. G. Ungerleider and J. V. Haxby, “‘What’ and ‘where’ in the human brain,” *Current Opinion in Neurobiology*, vol. 4, no. 2, pp. 157–165, Jan. 1994. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0959438894900663>
- [9] V. Rajmohan and E. Mohandas, “The limbic system,” *Indian Journal of Psychiatry*, vol. 49, no. 2, pp. 132–139, 2007. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917081/>
- [10] “Implicit and Explicit Memory | Simply Psychology.” [Online]. Available: <https://www.simplypsychology.org/implicit-versus-explicit-memory.html>
- [11] C. Henley, “Vision: Central Processing,” Jan. 2021, book Title: Foundations of Neuroscience Publisher: Michigan State University Libraries. [Online]. Available: <https://openbooks.lib.msu.edu/neuroscience/chapter/vision-central-processing/>
- [12] U. Martens, P. Wahl, U. Hassler, U. Frieese, and T. Gruber, “Implicit and Explicit Contributions to Object Recognition: Evidence from Rapid Perceptual Learning,” *PLoS ONE*, vol. 7, no. 10, p. e47009, Oct. 2012. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3466249/>
- [13] V. Cutsuridis and M. Yoshida, “Editorial: Memory Processes in Medial Temporal Lobe: Experimental, Theoretical and Computational Approaches,” *Frontiers in Systems Neuroscience*, vol. 11, 2017. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnsys.2017.00019>
- [14] “Recognition memory: What are the roles of the perirhinal cortex and hippocampus? | Nature Reviews Neuroscience.” [Online]. Available: <https://www.nature.com/articles/35049064>

-
- [15] I. Fried, K. A. MacDonald, and C. L. Wilson, “Single Neuron Activity in Human Hippocampus and Amygdala during Recognition of Faces and Objects,” *Neuron*, vol. 18, no. 5, pp. 753–765, 1997. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0896627300803153>
- [16] Z. U. Khan, E. Martín-Montañez, and M. G. Baxter, “Visual perception and memory systems: from cortex to medial temporal lobe,” *Cellular and Molecular Life Sciences*, vol. 68, no. 10, pp. 1737–1754, May 2011. [Online]. Available: <https://doi.org/10.1007/s00018-011-0641-6>
- [17] “Neural correlates of perceptual learning: A functional MRI study of visual texture discrimination | PNAS.” [Online]. Available: <https://www.pnas.org/content/99/26/17137>
- [18] T. Ro, B. Breitmeyer, P. Burton, N. S. Singhal, and D. Lane, “Feedback contributions to visual awareness in human occipital cortex,” *Current Biology*, vol. 13, no. 12, pp. 1038–1041, Jun. 2003. [Online]. Available: <http://www.scopus.com/inward/record.url?scp=0038613429&partnerID=8YFLogxK>
- [19] A. Sillito, J. Cudeiro, and H. Jones, “Always returning: Feedback and sensory processing in visual cortex and thalamus,” *Trends in neurosciences*, vol. 29, pp. 307–16, 2006.
- [20] N. K. Logothetis, J. Pauls, and T. Poggio, “Shape representation in the inferior temporal cortex of monkeys,” *Current Biology*, vol. 5, no. 5, pp. 552–563, 1995. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0960982295001084>
- [21] E. Thomas, M. M. Van Hulle, and R. Vogels, “Encoding of categories by noncategory-specific neurons in the inferior temporal cortex,” *Journal of Cognitive Neuroscience*, vol. 13, no. 2, pp. 190–200, Feb. 2001.
- [22] G. Kreiman, C. Koch, and I. Fried, “Category-specific visual responses of single neurons in the human medial temporal lobe,” *Nature Neuroscience*, vol. 3, no. 9, pp. 946–953, Sep. 2000, number: 9 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/nm0900_946
- [23] H. Anwar, Q. U. Khan, N. Nadeem, I. Pervaiz, M. Ali, and F. F. Cheema, “Epileptic seizures,” *Discoveries*, vol. 8, no. 2, p. e110. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7305811/>
- [24] M. Ptito, M. Bleau, I. Djerourou, S. Paré, F. C. Schneider, and D.-R. Chebat, “Brain-Machine Interfaces to Assist the Blind,” *Frontiers in Human Neuroscience*, vol. 15, 2021. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnhum.2021.638887>
- [25] M. Pfeiffer and T. Pfeil, “Deep Learning With Spiking Neurons: Opportunities and Challenges,” *Frontiers in Neuroscience*, vol. 12, 2018. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnins.2018.00774>
- [26] N. Savage, “How AI and neuroscience drive each other forwards,” *Nature*, vol. 571, no. 7766, pp. S15–S17, Jul. 2019, bandiera_abtest: a Cg_type: Outlook Number: 7766 Publisher: Nature Publishing Group Subject_term: Brain, Computer science, Neuroscience, Information technology. [Online]. Available: <https://www.nature.com/articles/d41586-019-02212-4>
- [27] “Visual Agnosia - an overview | ScienceDirect Topics.” [Online]. Available: <https://www.sciencedirect.com/topics/neuroscience/visual-agnosia>
- [28] “What Is Agnosia?” Sep. 2013. [Online]. Available: <https://www.healthline.com/health/agnosia>
-

- [29] C. Castelluccia, M. Duermuth, M. Golla, and F. Deniz, “Towards Implicit Visual Memory-Based Authentication,” in *Proceedings 2017 Network and Distributed System Security Symposium*. San Diego, CA: Internet Society, 2017. [Online]. Available: <https://www.ndss-symposium.org/ndss2017/ndss-2017-programme/towards-implicit-visual-memory-based-authentication/>
- [30] J. Yoon, J. Winawer, N. Witthoft, and E. Markman, “Striking deficiency in top-down perceptual reorganization of two-tone images in preschool children,” in *2007 IEEE 6th International Conference on Development and Learning*. London, UK: IEEE, Jul. 2007, pp. 181–186. [Online]. Available: <http://ieeexplore.ieee.org/document/4354071/>
- [31] “animal learning - Insight and reasoning | Britannica.” [Online]. Available: <https://www.britannica.com/science/animal-learning/Insight-and-reasoning>
- [32] R. Quian, L. Reddy, G. Kreiman, C. Koch, and I. Fried, “Invariant Visual Representation by Single Neurons in the Human Brain,” *Nature*, vol. 435, pp. 1102–7, 2005.
- [33] “Contrast and luminance adaptation alter neuronal coding and perception of stimulus orientation | Nature Communications.” [Online]. Available: <https://www.nature.com/articles/s41467-019-08894-8>
- [34] H. Liu, Y. Agam, J. R. Madsen, and G. Kreiman, “Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex,” *Neuron*, vol. 62, no. 2, pp. 281–290, Apr. 2009.
- [35] A. A. Gajadhar, R. C. Moiola, B. K. A. S. de Melo, A. C. B. Kunicki, A. S. C. Peres, and T. G. d. Rêgo, “Neural decoding with SVM and feature selection in a rat active tactile discrimination task,” in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–6, iSSN: 2161-4407.
- [36] C. Humber, K. Ito, and C. Bouton, “Nonsmooth Formulation of the Support Vector Machine for a Neural Decoding Problem,” 2010.
- [37] M. Levakova, M. Tamborrino, S. Ditlevsen, and P. Lansky, “A review of the methods for neuronal response latency estimation,” *Biosystems*, vol. 136, pp. 23–34, Oct. 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0303264715000660>
- [38] J. E. Kragel, S. Schuele, S. VanHaerents, J. M. Rosenow, and J. L. Voss, “Rapid coordination of effective learning by the human hippocampus,” *Science Advances*, vol. 7, no. 25, p. eabf7144, publisher: American Association for the Advancement of Science. [Online]. Available: <https://www.science.org/doi/10.1126/sciadv.abf7144>
- [39] E. Gajewska-Dendek, A. Wróbel, and P. Suffczynski, “Lateral Inhibition as the organizer of the bottom-up attentional modulation in the primary visual cortex,” *BMC Neuroscience*, vol. 16, no. Suppl 1, p. P243, Dec. 2015. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4699165/>
- [40] M. Kinoshita and H. Komatsu, “Neural Representation of the Luminance and Brightness of a Uniform Surface in the Macaque Primary Visual Cortex,” *Journal of Neurophysiology*, vol. 86, no. 5, pp. 2559–2570, Nov. 2001, publisher: American Physiological Society. [Online]. Available: <https://journals.physiology.org/doi/full/10.1152/jn.2001.86.5.2559>
- [41] H. Sprekeler, “Functional consequences of inhibitory plasticity: homeostasis, the excitation-inhibition balance and beyond,” *Current Opinion in Neurobiology*, vol. 43, pp. 198–203, 2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959438817300909>

-
- [42] F. Pulvermüller and O. Hauk, “Category-specific conceptual processing of color and form in left fronto-temporal cortex,” *Cerebral Cortex (New York, N.Y.: 1991)*, vol. 16, no. 8, pp. 1193–1201, Aug. 2006.
- [43] J. Vidal, T. Ossandón, K. Jerbi, S. Dalal, L. Minotti, P. Ryvlin, P. Kahane, and J.-P. Lachaux, “Category-Specific Visual Responses: An Intracranial Study Comparing Gamma, Beta, Alpha, and ERP Response Selectivity,” *Frontiers in Human Neuroscience*, vol. 4, 2010. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnhum.2010.00195>
- [44] A. Berti and M. Neppi-Modona, “Agnosia (including Prosopagnosia and Anosognosia),” in *Encyclopedia of Human Behavior (Second Edition)*, V. S. Ramachandran, Ed. San Diego: Academic Press, Jan. 2012, pp. 60–67. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780123750006000124>
- [45] G. Kreiman, C. Koch, and I. Fried, “Category-specific visual responses of single neurons in the human medial temporal lobe,” *Nature Neuroscience*, vol. 3, no. 9, pp. 946–953, Sep. 2000, number: 9 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/nn0900_946
- [46] R. Q. Quiroga, R. Mukamel, E. A. Isham, R. Malach, and I. Fried, “Human single-neuron responses at the threshold of conscious recognition,” *Proceedings of the National Academy of Sciences*, vol. 105, no. 9, pp. 3599–3604, Mar. 2008, publisher: National Academy of Sciences Section: Biological Sciences. [Online]. Available: <https://www.pnas.org/content/105/9/3599>





Figure 24: Mooney images presented







Figure 25: Corresponding grayscale images