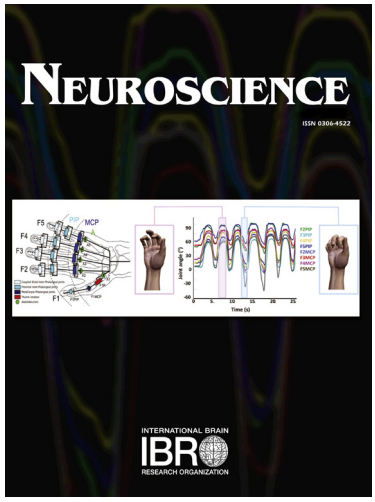
Accepted Manuscript



Average activity, but not variability, is the dominant factor in the representation of object categories in the brain

Hamid Karimi-Rouzbahani, Nasour Bagheri, Reza Ebrahimpour

PII: S0306-4522(17)30008-8
DOI: http://dx.doi.org/10.1016/j.neuroscience.2017.01.002
Reference: NSC 17538

To appear in: *Neuroscience*

Received Date: 16 September 2016
Revised Date: 18 December 2016
Accepted Date: 2 January 2017

Please cite this article as: H. Karimi-Rouzbahani, N. Bagheri, R. Ebrahimpour, Average activity, but not variability, is the dominant factor in the representation of object categories in the brain, *Neuroscience* (2017), doi: http://dx.doi.org/10.1016/j.neuroscience.2017.01.002

Average activity, but not variability, is the dominant factor in the representation of object categories in the brain

Hamid Karimi-Rouzbahani[1,3], Nasour Bagheri[1,3], Reza Ebrahimpour[2,3*]

[1]Department of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

[2]Department of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

[3]School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

*To whom correspondence should be addressed. Email: rebrahimpour@srttu.edu

Telephone number: +982122970060 (ext. 2501)

Abstract- To categorize the perceived objects, brain utilizes a broad set of its resources and encoding strategies. Yet, it remains elusive how the category information is encoded in the brain. While many classical studies have sought the category information in the across-trials-averaged activity of neurons/neural populations, several recent studies have observed category information also in the within-trial correlated variability of activities between neural populations (i.e. dependent variability). Moreover, other studies have observed that independent variability of activity, which is the variability of the measured neural activity without any influence from correlated variability with other neurons/populations, could also be modulated for improved categorization. However, it was unknown how important each of the three factors (i.e. average activity, dependent and independent variability of activities) was in category encoding. Therefore, we designed an EEG experiment in which human subjects viewed a set of object exemplars from four categories. Using a computational model, we evaluated the contribution of each factor separately in category encoding. Results showed that the average activity played a significant role while the independent variability, although effective, contributed moderately to the category encoding. The inter-channel dependent variability showed an ignorable effect on the encoding. We also investigated the role of those factors in the encoding of variations which showed similar effects. These results imply that the brain, rather than variability, seems to use the average activity to convey information on the category of the perceived objects.

Keywords: Object category encoding, Signal variability, EEG, Computational model, Representational analysis

Introduction

When humans perceive objects, neural representations form highly distributed and dynamical patterns of activity across several regions of the brain (Cox and Savoy, 2003; Ishai et al., 1999). In addition to the large body of literature which reported dominant roles for the lateral occipital and ventral temporal cortices in category encoding (Grill-Spector and Weiner, 2014; Riesenhuber and Poggio, 2002; Grill-Spector et al., 2001), there are evidence suggesting contributions from the frontal and prefrontal cortices to the encoding of categories (Jiang et al., 2007; Goddard et al., 2016; Freedman et al., 2001). While several studies have found clusters of category-specific neurons (such as for faces, body parts, etc.) on the temporal cortex (Freiwald and Tsao, 2010; Connolly et al., 2012; Orlovet al., 2010), many others argued for distributed and overlapping category maps in the same area (Haxby et al., 2001;

Edelman et al., 1998; Grill-Spector and Weiner, 2014; Grill-Spector and Kanwisher, 2001). What has remained mostly unanswered by the mentioned studies is whether the separate neural populations (i.e. in the same or different areas) encode category information independently or in cooperation with each other. Indeed, despite much improvements in the methods of population-level representational analysis (Kriegskorte et al., 2008a; Kaneshiro et al., 2015), it has not yet been investigated whether the distinct category-related regions of the brain provide category information in independent activities (which are finally stacked up to form population codes) or whether extra category-related information would appear by considering the correlation of activities between those regions.

A large set of studies have shown a significant role for the inter-neuron and inter-area correlated trial-to-trial variability of activities (i.e. known as noise correlation) in visual perception (Ruff and Cohen, 2016; Zenon and Krauzlis, 2012; Cohen and Maunsell, 2009) and decision-making (Neiborg et al., 2012; Shadlen et al., 1996). Other findings have shown that the trial-to-trial variability of activities (i.e. known as Fano factor) was modulated to improve the visual perception (Herrero et al., 2013; Churchland et al., 2010). It was revealed by Cohen and Maunsell, (2009), that dependent (or correlated) trial-to-trial variability of activities between neurons contributed more dominantly to stimulus perception in an attention-deployed task compared to independent variability of activities (i.e. which is the variability measured by Fano factor from which the effect of correlated variability is removed) for each neuron. Thus, these studies revealed a significant role for dependent and independent variability in the encoding of categories in the brain. However, as these experiments investigated the trial-to-trial variability of activities, they remain silent on the role of within-trial variability of activities on the neural encoding.

Generally, activities of individual neurons, voxels and EEG channels are averaged within trials and treated independently from other sources (i.e. neurons, voxels or EEG channels) in the current population representational analyses to obtain information on the perceived stimuli (Kriegskorte et al., 2008b; Kiani et al., 2007; Kaneshiro et al., 2015). This averaging overlooks several time-dependent dimensions of neural processing which have been shown to play roles in the encoding of information. Theoretical studies have shown that gamma-band synchronization (Engel et al., 1991), millisecond precision spike trains (Abeles, 1991; Oram, 1999) and the order of response latencies (Thorpe et al., 2001; Van Rullen et al., 1998) can also provide suitable representational brain spaces for information. Moreover, experimental studies have shown that the temporal patterns of activity convey information regarding the edge co-occurrences (Eckhorn et al., 1988), orientations in primary visual cortex (Celebrini et al., 1993) and light intensity in the retina (Gollisch, 2008). For the object category representation, an

EEG study has shown the involvement of temporal phase encoding at work (Behroozi et al., 2015). Another study has shown that the within-trial correlation of activities between regions of the temporal cortex played a prominent role in object category representation (Majima et al., 2014).

Motivated by these latter studies on the importance of within-trial variability, we asked which component of the recorded activity carried the highest amount of information regarding categories within the time course of trials: is it the average activity of the region, each region's independent variability of activity or the variability of activities which are correlated between distinct regions? It should be noted that, by variability we mean the changes of signal amplitude within the time course of trials which is totally different from the classical definitions of noise correlation and Fano factor which measure the trial-to-trial variability of neural activities (Churchland et al., 2010; Abbot and Dayan, 1999).

To answer this question, we designed a whole-brain EEG recording paradigm to measure the object-evoked activity from human subjects. EEG allowed us to record high-resolution temporal signals from different brain regions simultaneously, which was generally difficult to obtain from fMRI or single cell recording. We provided a method to quantitatively compare the contribution of the within-trial average activity, independent variability as well as the dependent variability of different brain regions in conveying category-related information. Accordingly, a modified version of a previously developed computational model (Shadlen et al., 1996) was used to generate artificial trials whose statistics closely matched those measured in the experiment. The model received three input parameters, namely the mean of activity, the variance and the correlation of activities, which were extracted from category representations. The mean activity represented the within-trial sample mean of the activities recorded in individual electrodes evoked by each stimulus. While the within-trial variance of the activity showed the accumulation of independent variability and dependent variability of individual electrodes, the within-trial correlations between (all possible) pairs of electrodes were measures of in-phase synchronization between every pair of electrodes (Majima et al., 2014). By neutralizing an individual parameter at a time, the model allowed us to investigate the contribution of each parameter independently from the other two parameters in the representational enhancement. To measure the category separability in the brain and the model representational spaces, we defined a separability index which provided several key advantages to the common decoding schemes which were previously used to extract information regarding the object categories (Kaneshiro et al., 2015; Majima et al., 2014; Goddard et al., 2016), object variations (Isik et al., 2014; Hong et al., 2016) in both invasive (Hong et al.,

2016; Majima et al., 2014) and non-invasive (Kaneshiro et al., 2015; Goddard et al., 2016) recording experiments.

Results showed that the mean activity of individual brain areas contributed more significantly to category representation compared to the independent variability of individual brain areas and the correlated variability between different brain areas.

Methods

Stimulus set

Since our goal was to investigate the representations of objects in the brain, we needed realistic forms of objects under everyday variations, rather than the effects of background, color or other category-unrelated parameters which could interfere with the brain representations of the target objects (Martinovic et al., 2008). Therefore, an object image set was generated in which 16 object exemplars were put in four object categories, namely animal, car, face and plane (Fig 1A).

To generate the image set we downloaded free 3D object models from (http://tf3dm.com/) and rendered them under different variation conditions (Fig 1B) using Blender software (https://www.blender.org/). Since the variations were integral parts of object encoding which accompany the object representations even at the highest levels of the visual streams (Hong et al., 2016), we had to include them into the generation of the image set to fully activate the object-related mechanisms of the brain for an inclusive representational analysis. Therefore, the variations were applied in four dimensions including the light source direction, pose, size and position each of which in three conditions. The variation conditions were chosen so as to expose the dynamics of object representations in different visual areas such as the occipital, occipito-temporal and parietal cortices as they have shown the highest contributions to object representations (Kaneshiro et al., 2015). To generate the lighting conditions we put the light source at different directions pointing to the objects. For the pose conditions, the objects were rotated simultaneously around the X, Y and Z Cartesian axes in steps of 45 degrees and sampled at 0, 135 and 270 degrees of orientation. For the size conditions, the objects were resized from the very small to the very large scales spanning respectively from 2.5 to 13.5 degrees of visual angle when presented on the screen. As for the position, the objects were put randomly at three different eccentricities from the image center to generate distances between 0.8 to 7.7 degrees of visual angle in the experiment. A unique 512 by 512 pixel image was generated from the

objects in different conditions adding up to a total of 192 images in the image set (i.e. 16 exemplars in 12 conditions).

Experimental paradigm

In order to reach the object representations on the human scalp with EEG, we designed a behavioral task for the human subjects. 10 human subjects (mean age 22, 3 females) volunteered in the experiment which was held in a darkened room and lasted for about half an hour. Subjects had normal or corrected-to-normal vision and were seated in front of a monitor which was 60 cm away from them. Each subject was presented with the 192 unique images in the image set in random order each of which was presented three times to increase the signal to noise ratio in the analyses (i.e. totally 576 images were presented). The stimuli were presented in four blocks of 144 trials with a 5-minute rest time between each block. We used Matlab PsychoToolbox (Brainard, 1997) for stimulus presentation and response recording.

Each trial began with the presentation of a black fixation point in the center of the screen which remained the same for 200 ms, after which the first stimulus was presented for 50 ms being accompanied with the fixation point in either red or green. After the disappearance of the first stimulus, and 1200 ms of inter-stimulus interval, the second stimulus appeared with either a green or red fixation point in the center of the screen and remained there for 50 ms. After the second stimulus, the subject was supposed to give a response indicating whether the stimulus-accompanying fixation points were the same color or not by pressing one of predefined keys on the keyboard. Although the subjects had to give a response to proceed to the next trial, there was no time limit for the subjects' responses. The colors of the stimulus-accompanying fixation points were different in 50% of the trials. The subjects were acquainted with the task in a training session immediately before the main task on a different set of stimuli. The important aspect of the EEG paradigm was that it was performed in a passive format in which the subjects' task was irrelevant to object categorization which was to be studied in the representational analysis (Isik et al., 2014). The reason for that choice was that, in case of categorization tasks which have been frequently used in recent studies (Taghizadeh-Sarabi et al., 2014), the top-down categorical (Thorpe et al., 1996) and attentional effects (Chikkerur et al., 2010; Milner 1974) could have a huge impact on the representational space of categories with influences on the correlation (Cohen and Maunsell, 2009), variance or the mean of activities of recorded signals (McAdams and Maunsell, 1999).

For the same reason, we chose a short stimulus presentation time which has been shown to cause the domination of the feed-forward mechanisms of the brain in visual processing (Hong et al., 2016).

EEG recording and preprocessing

We used a 32-channel amplifier for the EEG signal recording (eWave, produced by ScienceBeam, http://www.sciencebeam.com/) which provided 1K sample/second of time resolution. The recorded data was taken to Matlab (http://www.mathworks.com/) and preprocessed as follows. All the recorded data went through the preprocessing phase and were then used in the representational analyses and in the extraction of the parameters to feed the computational model with. The recorded signals were first digitally band-passed filter in the range 0.1-100 Hz for DC and high-frequency noise removal, and notch-filtered at 50 Hz to avoid mains noise. To do this, we used FIR linear filters as implemented in Matlab. To investigate the effect of high pass filtering on the time course of the signals (Widmann and Schröger, 2012), another set of analyses were performed with 1-100 Hz filters which showed no significant changes to the signals. Muscle and eye-blink artifacts were removed from the signals using the ICA analysis as implemented in EEGLAB (Delorme and Makeig, 2004). To reach the disturbing components in the ICA analysis, we used ADJUST plugin (Mognon et al., 2011), which calculated several statistical measures from the filtered signals and suggested the ICA components which passed a set of predefined thresholds for removal. Averagely, 98.62% of the trials (5681 out of 5760 trials, s.d. = 1.3%) remained for the following analyses after the artifact removal from all subjects. After the signal filtering and artifact removal, we epoched the stimulus-aligned channel activities from 200 ms pre-stimulus to 800 ms post-stimulus. In order to decrease the unwanted fluctuations in the signal and reach a reasonable time resolution we applied a 5-ms non-overlapping moving average window on the signals and decreased the sampling rate to 200 samples per second. The chosen window was fine enough not to lose time precision and coarse enough to avoid the non-informative parts of the signals.

The epoching and smoothing of the signals resulted in a 3-dimensional data matrix ‘X’ for each subject with 31 rows (i.e. number of channels), 200 columns (i.e. number of samples in the -200 to +800 ms window relative to the stimulus) and 576 layers (i.e. number of trials for a subject with no removed trials). The values in the matrix were the measured activities on the EEG electrodes (i.e. values in microvolts) which were either positive or negative relative to the reference electrode. To reach the

separability results for each time point in the raw analyses (i.e. results of Fig 4 and 5), we used the measured values from the mentioned data matrix on every time sample (i.e. every matrix column).

For the model, however, we calculated three parameters from the values in the data matrix on 10 consecutive samples (i.e. corresponding to a 50-ms window): the mean (i.e. sample mean) and the variance of individual channels' activity and the inter-channel correlation of time sample activities between every possible pair of channels (465 pairs for the 31 channels) using Pearson linear correlation. To obtain the results for the whole peri-stimulus span (i.e. -200 to 800 ms span), we then moved the window across the time samples, obtained new parameters, fed them to the model and calculated the resultant separability at the model's output. More information on how the separability values were calculated is given below.

Representational analysis

Despite the domination of the decoding methods in the extraction of information from the signals in electrophysiological studies (Isik et al., 2014; Hong et al., 2016), we used a separability index, known as sensitivity, for representational analysis, since it provided several key advantages to decoding. Among them is the avoidance of any need for defining hyper-parameters in the analysis of representations such as the train/test sampling strategy, the cross validation method, and the classifier, each of which can bias the analysis. Another advantage is that the sensitivity will not saturate in case of totally separable classes as it is the case in decoding. As previously mentioned, we had four categories of objects: animals, faces, cars and planes, each of which had four exemplars (Fig 1A). On every time sample (i.e. each column of the data matrix), to measure the separability between a sample pair of categories (e.g. animal vs. car, each of which had 144 trials), we dedicated one point for each trial in the 31-dimensional space (Fig 3). For instance, to find the corresponding point for the 8^{th} trial at the +210 ms post stimulus, we used the 31 values in all rows of the 82^{th} column (from left to right) of the 8^{th} layer of the matrix. The 31 dimensions referred to the recording channels (electrodes) of the EEG amplifier. After finding all the 288 points corresponding to the trials for of the first (i.e. animal) and second (i.e. car) categories, we calculated the average response for each of those sample categories and projected all points on the imaginary line of discrimination which passed through the two averages of the sample categories' data clusters (Fig 3, black crosses). This projection led to a one-dimensional distribution for the two categories each of which had a mean and a standard deviation value which were used in the calculation of separability as in (1):

$$d' = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{1}{2}(\sigma_1^2 + \sigma_2^2)}} \qquad (1)$$

where μ_1 and μ_2 are the means and the σ_1^2 and σ_2^2 are the variances of the two sample categories, respectively. d' is the sensitivity and reflects the separability of the two sample categories. As the sensitivity measured the separability between a single pair of categories, we calculated the sensitivity between every possible pairs of the four mentioned categories (i.e. six combinations) in Fig 4B, and reported the average of the six combinations in Fig 4A. As the separability was directly proportional to the difference between the means of two categories and inversely proportional to the sum of the root mean square of the standard deviations of the categories, it could measure how separately the categories positioned relative to one another in the representational space. For the representational analysis through time, the same procedure was repeated to calculate the separability index at every time point. We repeated this procedure for every subject individually and plotted the across-subject average and standard error (with shaded regions) in Fig 4. For the computational model, the same procedure was followed, but with the data matrix X obtained from equation (2).

Computational model

We implemented a computational model to separate the contributions of the brain regions' mean and variance of activities and the inter-region correlations of activities in object category representation. The model was first developed to simulate neural activities at population level (Shadlen et al., 1996). For that purpose, in one study, the model was provided with the individual neuron's firing rates, trial-to-trial variance and inter-neuron correlated variability which had been measured between simultaneously recorded pairs of neurons (Cohen and Newsome, 2008). The model then provided simulated activities for the population which, to an acceptable extent, replicated the recorded activities. This allowed a recent study to evaluate the contribution of each of the inputs in the representational enhancement in a covert attention task on monkeys (Cohen and Maunsell, 2009). Here, however, as we measured the brain activities using EEG, we adapted the model to the EEG by replacing the neurons with the EEG channels. Moreover, as we aimed to evaluate the short-term contributions of the parameters, we fed the model with the within-trial measures as opposed to the previous studies which used inter-trial measures (Cohen and Newsome, 2008; Cohen and Maunsell, 2009).

The model is formulated as in (2):

$$X = M + Y * V \qquad (2)$$

where M and V are respectively the 31×1 vectors of the mean and variance of activities from the 31 EEG channels, and Y is calculated as in (3):

$$Y = QZ \qquad (3)$$

where Z is a 31×1 normally distributed zero-mean random vector with unit variance, and Q is the square root of the 31×31 correlation matrix R (i.e. $R = Q'Q$). The correlation matrix R incorporates the inter-channel correlation values calculated from the activities of every possible combination of channel pairs (i.e. 465 combinations of channel pairs) as summarized in corresponding rows and columns of R as shown in (4):

$$R = \begin{bmatrix} 1 & r_{12} & \dots \\ r_{21} & 1 & r_{23} \\ \vdots & r_{32} & \ddots \end{bmatrix} \qquad (4)$$

where r_{nm} (which has an equal value to r_{mn}) indicates the correlation value calculated between channels n and m, respectively. As the correlation matrix R can have several square roots depending on the correlation values, we used the principal square root of the matrix for Q, as implemented by the Matlab function 'sqrtm' which used the Blocked Schur method (Deadman et al., 2013), among other possible choices (Denman and Beavers, 1976). Finally, X provided the simulated activities for the 31 channels which resembled the experimentally measured EEG channels' activities. For a detailed description of the model and the differences of the original model with the current version read (Shadlen et al., 1996).

In order to obtain the three parameters as the inputs to the model, as explained earlier, we used the preprocessed data matrix named 'X' from the preprocessing section. These parameters were fed to the model to provide a simulated data vector X which estimated the activity of the 31 channels on the same time window as in which the parameters were calculated.

Results and Discussions

This study was aimed at clarifying the contributions of different signal attributes to object category representations. For that purpose, ten human subjects participated in the passive EEG recording experiment (Fig 2), and their brain representations to different object categories were analyzed. The subjects performed the color matching task with high accuracy (mean accuracy = 94.68%, s.d. = 3.72%),

which was significantly above chance ($p < 0.001$, Wilcoxon signed-rank test). They took an average response time of 734 ms (s.d. = 95 ms) to provide their responses. These results indicate that the subjects were attentive and alert during the experiment.

Dynamics of category separability in the human brain

In order to ensure that the category-related information could be extracted from the brain signals by the above-explained representational analysis method, we calculated the separability index as a function of time for the pool of category pairs (Fig 4A) and each pair (Fig 4B) separately. For the pooled result, after a moderate value prior to the stimulus onset, the across-pair averaged separability rose to significance at 85 ms and showed three peaks at 110 ms, 190 ms and 265 ms post-stimulus time points. The separability remained significant until 800 ms post-stimulus. The significance of separability was evaluated by comparing the average separability value in the last 200 ms pre-stimulus with the separability value at every post-stimulus time point, using Wilcoxon signed-rank test. The timing of the appearance of category information was consistent with a large body of studies reporting the temporal dynamics of object representations in the brain (Liu et al., 2009; Freiwald and Tsao, 2010).

We also investigated whether the calculated separability values replicated previous findings on the temporal dynamics of different levels of object representations in the brain (Dehaqani et al., 2016) (Fig 4B). The first significant separability indices appeared for the car-face, face-plane, animal-car, animal-plane, car-plane and animal-face pairs respectively at 95, 95, 100, 100, 105 and 115 ms post-stimulus time points. These results, while confirming the previous finding on the precedence of mid-level category representations (e.g. car-face and face-plane) to super-ordinate (e.g. animal-car and animal-plane) and subordinate representations (e.g. car-plane) (Dehaqani et al., 2016), seems to be in line with the suggestion that shape similarity, rather than the semantics, determines the separability of category representations in the brain (Sofer et al., 2016).

These results show that the stimulus set could drive the expected category-related information in the brain for the computational analysis. In addition, they confirmed that the separability index was capable of extracting category information from the EEG signals which adds confidence to its suitability for the representational analyses.

Contribution of different brain areas to category representation

In order to investigate the relative contribution of different brain regions in the generation of category information, we measured the separability indices for individual electrodes one at a time, and

summarized their superposition results on scalp maps (Fig 5). Several representative time points (i.e. three of which were the peak separability time points observed in Fig 4A) were chosen to reveal the temporal flow of information on the scalp. The separability indices have declined as they have been computed on a single dimension rather than 31 dimensions, which was the case in Fig 4.

The occipital and parietal regions (O1, O2, Oz, P3, P4 and Poz) showed higher category information in the 95 and 120 ms time points compared to other brain areas along with some information around the Fz region. The contributions of the occipital areas, as the first stages in visual processing, were previously observed by a line of studies (Gerlach, 2007). At 185 ms, however, the whole posterior, central (Cz, Fc1 and Fc2) and frontal areas (Afz, F3, F4, Fz and Fpz) contributed to the category separability. The category information which was observed at the parietal as well as frontal brain areas have been recently well supported by the role of these areas in object representation (Rishel et al., 2013; Swaminathan and Freedman, 2012) and the encoding of object variations such as position and scale (Hoogmoed et al., 2012). The early category-related information observed at the frontal areas raised the possibility of a preparatory object category encoding in those areas as previously suggested by Bar et al. (2005). Interestingly, the posterior information then moved to more occipito-temporal regions (P7 and P8) in the 220 ms window, which is consistent with many studies showing the role of ventral visual stream in object category encoding (Hong et al., 2016; Isik et al., 2014). At 260 ms, which corresponded to the last peak of separability (Fig 4A), the information was mainly concentrated on the occpito-temporal, central and frontal areas of the brain which raised the potential cooperative object and variation encoding scheme between those areas (Hoogmoed et al., 2012; Underleider and Mishkin, 1982). The category information started to decline in the following time points with lingered activities probably for more difficult conditions (Hagen et al., 2006). The observed temporal dynamics of object category encoding was in agreement with a large body of literature on the timing of visual processing in the brain (Simanova et al., 2010; Daliri et al., 2012). In the following sections we will compare the contribution of three most-studied parameters of the EEG signals in the encoding of category information.

Peri-stimulus modulations of the contributing parameters

To contribute to the category representations, the three supposedly influential parameters of mean, variance and correlation needed to change dynamically in response to the presented stimuli. To investigate this, we calculated the changes of the three parameters across time in the peri-stimulus span (Fig 6). For variance, however, we reported Fano factor or the mean-normalized variance, which is

defined as the ratio of the variance of activity to the absolute mean, since it has been reported to directly indicate the modulation of the representational space (Churchland et al., 2010). Yet, the results for the variance did not reflect a different trend from that of Fano factor (data not shown).

The pooled channels' mean activities showed a typical ERP waveform with consecutive P1, N1, P2, N2, P3 peaks which repeated many classical ERP studies (Daliri et al., 2012; Hoogmoed et al., 2012) (Fig 6A). A significantly ($p < 0.05$, Wilcoxon signed-rank test) lower mean activity for the animal category at 110 ms and a higher mean activity for the face category at 180 ms post-stimulus were among the most notable observations. The latter replicated the P1/P170 effect which has been often associated with face processing in the brain (Dering et al., 2011).

As for the Fano factor (Fig 6B), after a slow negative trend which lasted until 240 ms post-stimulus, it started to rise again and approached its pre-stimulus value in the following time windows until 800 ms post-stimulus. This means that the appearance of the stimuli decreased the non-informative fluctuations of the recorded signals which lead to a lower noise to signal ratio (as it is measured by Fano factor (Churchland et al., 2010)). It has been previously reported that the reduction in neural variability (i.e. Fano factor) could enhance the representational space for stimuli, as this have been the case in many attention-deployed cognitive processes (Cohen and Maunsell, 2009). This parameter did not reveal many category-specific time points in its time course. The pattern of the reported Fano factor, which most probably has never been reported on the EEG data before, resembled those reported in the single and multi-unit neural data (Cohen and Maunsell, 2009; Churchland et al., 2010). This is interesting since all these previous studies, as opposed to the current study which reported within-trial measurements of variability, used the trial-to-trial variability as an aspect of neuronal representation (Cohen and Maunsell, 2009) and as a signature of learning in EEG (McIntosh et al., 2008).

Although more fluctuating compared to the two previous parameters, the inter-channel correlation of activities also showed a very stimulus-locked behavior with a rising trend between 105 to 265 ms which totally damped at around 800 ms post-stimulus. Moderate increased correlations were previously reported between neurons on different brain regions as a result of attention employment in monkeys (Ruff and Cohen, 2016). The correlation parameter did not show any significant category-specific signature.

In fact, none of the parameters showed significant ($p < 0.05$) time points of differences between all 6 category pairs. It was interesting that none of the parameters alone could repeat the separability of

categories (see the time points indicated by diamonds in Fig 6) which was observed in Fig 4A (as indicated by the circles in a continuous fashion). This implied that the parameters must have interacted to provide the observed separability of categories. We then asked how much each parameter contributed to the observed category separability, which is answered by the computational analyses of this study.

To gain a deeper insight into the temporal dynamics of the three parameters on the brain, we measured how the pre-stimulus values of the parameters were affected by the stimulus appearance, using a modulation index (Fig 7). The modulation index was defined as the difference between the post-stimulus and the pre-stimulus values of the corresponding parameter. The pre-stimulus window in which the parameters were calculated was the last pre-stimulus window and had the same width as the post-stimulus windows which were 50 ms wide. For this analysis, we chose three post-stimulus time points in which the modulation indices were calculated. These time points were 95 ms, 185 ms and 260 ms post-stimulus, in which the separability of categories peaked (Fig 4A), as we expected to see considerable amounts of category-related information on the scalp at those time points.

We observed a large amount of negative mean modulation at 95 ms at the central (Cz) and frontal (F3, F4, Fz, Fc1, Fc2 and Afz) brain areas with an opposite effect at the occipital (O1, O2 and Oz) and occipito-temporal (P7 and P8) regions (Fig 7A). The positive mean modulation became more concentrated on the parietal regions reaching the P3, P4 and Pz electrodes as the negative mean modulation vanished from the frontal areas at 185 ms. The positive and negative mean modulations reappeared simultaneously at 260 ms with the dominance of the positive effect at the occipital and occipito-temporal regions. These results showed that, exactly at the times which showed the highest category-related information, the mean amplitude of the EEG channels revealed a highly bipolar behavior for the frontal and occipito-parietal regions of the brain each of which has been supported to be involved in the representation of objects (Goddard et al., 2016). However, neither the positive nor the negative modulations of the mean could alone explain the observed distribution of the category information on the scalp (Fig 5), as the explanation may lie in the brain bipolarity.

The mean-normalized variance (Fano factor) showed a negative modulation with concentration on the central and frontal (C3, C4, Fz and Afz) areas at 95 ms, an occipito-parietal (Pz, Poz, O1, O2, and Oz) domination in 185 ms, and an anterior to central and central to posterior distribution in 260 ms (Fig 7B). Therefore, the Fano factor alone seems to fail to explain the observed separability shown in Fig 5.

The correlated variability between the EEG channels could be determined by the inter-channel distance in the post-stimulus span (Fig 7C). More specifically, the correlation values were inversely proportional to the distances between the pair of channels ($r = -0.32$, $p < 10^{-13}$, Pearson linear correlation). This could already be predicted, as in the EEG, the recording channels influence one another by the skin and the volume conductivity with decreasing effects at longer distances resulting in moderate spatial resolution for the signals. However, as the modulation of correlations showed highly dynamical behavior in time (Fig 6C), we predicted that these distance-correlation relationships could have been violated in many time instances as a result of the dynamics of category-related processing in the brain. To examine this, we gathered all the inter-channel correlation modulations in matrix format (Fig 7E). This was done for the three representative post-stimulus time points as in Fig 7A and B, which provided interesting insights into the brain's inter-area cooperation. The rows and columns of the three matrices are indexed with electrode numbers as shown in Fig 7D. Here we only highlighted the most amplified (indicated by red boxes in Fig 7E) and declined (indicated by black boxes) correlations between channels.

In the first window (95 ms), the parieto-temporal areas at both hemispheres (electrodes 1-3 and 17-19) showed the most notably enhanced correlations with each other, with their contra-lateral temporal (respectively electrodes 20 and 4) and centro-frontal areas (respectively electrodes 21-24 and 5-8) as well as with the centro-frontal areas (electrodes 25-30). The highest decline in the correlated variability was observed between the occipital electrodes (9-12) and the parietal (13-15), temporal (4 and 20) as well as the parieto-temporal electrodes (1-3 and 17-19). Altogether, it seems that at this time, the occipital brain regions were processing differently from higher levels of the ventral visual stream while the latter areas were cooperating with the frontal cortex on the same processing matters. The timing of the different processing behavior at the occipital areas seemed to be reflecting the early processing of visual information at the early visual cortex. Please note that the plotted results were averaged in the 50 ms window around the indicated times. Interestingly, the cooperative activities of the temporal as well as the parieto-temporal cortex with the frontal areas have previously been reported as around the same time for category prediction (Bar et al., 2005; Goddard et al., 2016). However, in the within-trial context, the current study seems to be the first to report an increase of correlated variability between occipito-temporal and frontal brain areas (Fig 7E).

While becoming even more decorrelated with the parieto-temporal areas, the parieto-occipital areas (electrodes 9-13) showed increased correlation with the whole frontal area (electrodes 6-8 and 22-30) in the 185 ms window. Interestingly, the central occipital electrode (10) which had a negative trend of

correlation with almost all brain areas has become correlated with the frontal regions of the brain showing collaborations between the anterior and posterior brain regions. This was also reflected in the separability maps of the two windows with increased separability on the two regions (Fig 5). The results in the last window (260 ms) resembled those in the 185 ms window, but with less correlation between the frontal channels (electrodes 24-31) as they become involved in the task (i.e. this was reflected in Fig 5).

These results showed that, the correlated variability between the EEG channels, although affected by the inter-channel distance, was to a greater extent determined by the cognitive processes involved in the task. Having observed the dynamics of the changes in the three parameters, we used the computational model to measure the contribution of the three parameters in object category encoding.

Model evaluation

To use the computational model for separating the contribution of the three parameters, we had to make sure that the model could simulate the actually recorded activities with sufficient accuracy. To investigate this and to understand whether the parameters were informative enough to reconstruct the channels' activities recorded in the experimental procedure, we measured the differences between the actual and the simulated channel activities for all subjects. The Root-Mean-Square Error (RMSE) was calculated for the model estimates of the actually recorded activities and showed an average of 1.3484 (s.d. = 0.2589). The estimated activities for several representative channels of a sample subject are shown in Fig 8.

Although not perfectly at all peaks and valleys of the signals, the model was able to estimate the activities at all presented channels with an acceptable error rate. These results showed that the model could predict the actually recorded activities only based on the three parameters. This means that the three parameters, as well as how they were incorporated into the model, were informative enough to replicate the trend of the recorded signals. Yet, the observed error might be the result of other overlooked information existing in the signals and not necessarily the inaccuracy of the model. Therefore, the next step was to evaluate the contribution of each parameter in object representation.

Role of signal parameters in category separability as clarified by the computational model

In order to assess the contribution of each of the mentioned parameters in the representations of objects, we used the computational model and calculated the separability index as a result of each parameter. To do so, we considered the last pre-stimulus window as a null window in which the three

parameters were calculated and used as substitute parameters. For instance, to evaluate the contribution of the mean, we replaced the variance and the correlation of channels calculated in the post-stimulus window by their corresponding values calculated in the last pre-stimulus window, while leaving the mean parameter intact (i.e. using its post-stimulus value). The same procedure was repeated to calculate the effect of variance and correlation by replacing the other two parameters with their pre-stimulus values each time. We also measured the contribution of every possible pairs of parameters in the representational space by keeping the pair of parameters and replacing the other parameter by its pre-stimulus value. Altogether, we had eight conditions for the two time windows (pre- and post-stimulus windows) for all combinations of the three parameters (Fig 9). The pre-stimulus window had the same time width as the post-stimulus windows (i.e. here 50 ms) and was always the last time window before the stimulus onset (i.e. here -50 ms to 0).

We calculated the separability indices for all post-stimulus windows in all 8 conditions (Fig 9A). Results of Fig 9A showed significantly ($p < 0.01$) higher contribution in the object category representation for the mean of activities (i.e. compare conditions 1 and 2 with 3 and 4, and 5 and 6 with 7 and 8), significant ($p < 0.05$) but moderate contribution for the variance (i.e. compare the even and odd conditions) and insignificant ($p > 0.1$) contribution for the correlation of activities (i.e. compare conditions 1, 3 and 4 with 5, 7 and 8). However, for the sake of brevity, we showed the results of a few critical time points in which either the three parameters changed their trends or the categories showed highly separable clusters based on one of the parameters (e.g. these points were indicated with diamonds in Fig 6). The results of these critical post-stimulus time points which included 95, 120, 185, 260, 340, 400, 510 and 570 ms are shown in Fig 9B. The pooled results were also reflected in the representative critical time points with the dominance of mean activity, moderate contribution for the variance and insignificant negative contribution for the correlation of activities (Fig 9B).

The percentages of contribution to category-encoding were finally calculated for the mean of activity, as well as independent variability and correlated (as measured by correlation) variability (Fig 10A). It should be noted that the variance measured the variations which were observed on a single channel. However, this variability could be a result of variability in the channel's activity, the shared variability between the channel and other channels or a combination of both. As the correlation measured the shared variability (i.e. the inter-channel variability), we determined the contribution of the independent variability by removing the contribution of shared variability from the contribution of the variance. This resulted in contributions of 88.09%, 10.40% and 0.59% respectively for the mean of activities,

independent and correlated variability (all of which were significantly different from one another, $p < 0.01$, Wilcoxon signed-rank test). In order to calculate the contribution of a parameter, the difference of the separability indices between the conditions in which the parameter was present and the conditions in which it was absent was calculated and then normalized between the first and the eighth conditions which were respectively considered the 100% and 0% contribution boundaries. These results indicated that the mean of activity at different brain areas played a significantly more important role in the representation of objects compared to the independent or dependent variability.

These results, while reappraised the role of mean activity as a substrate for neural encoding, as it is used in the classical ERP analyses, cast doubt on the previously-suggested usefulness of time-correlations in the representation of object categories (Majima et al., 2014). This discrepancy can be explained in light of several methodological differences between this previous study and ours. First, as we aimed to evaluate the rapid object processing, we presented the stimuli for only 50 ms while the presentation time was 300 ms in the previous study (Majima et al., 2014). This could have lowered the impact of correlations on the category encoding, as it is a time-dependent parameter. Yet, we re-ran the whole computational analyses, with correlations (and the other parameters) calculated on 300-ms windows (also on 100, 150, 200 and 250 ms), which showed no significant changes in the results (data not shown). Second, as the signals used in the previous study were obtained from an expectedly higher spatial resolution setup (ECoG) and with concentration on the IT cortex (i.e. which is a limited area with the highest category selectivity on the brain), they could have presented more category-specific correlations. On the other hand, that setup missed a wider range of cross-area correlations which were observed in the current study (Fig 7E) which replicated the increase previously reported by Ruff and Cohen, (2016). Third, as opposed to the power and phase features, we used the independent channel variability and activities each of which could have outperformed the correlation feature used by Majima et al. (2014), if used in the same context. Finally, the computational methodology we used here was basically different from the decoding procedure used in (Majima et al., 2014). This methodology allowed us to reconstruct the channel activities of the actual experiment with high accuracy and evaluate the resultant representational space using a simple representational analysis rather than applying a rigorous classifier (SVM) on a set of extracted correlation values.

Observing these results, we next asked whether the contributions of the three parameters could alter if they had been modulated to different extents. As it was shown in Fig 6, upon the appearance of the stimulus, the three parameters experienced different levels of modulation, with the modulation being as

high as 421% for the mean activity and as low as 6.32% for the correlation and -24.76% for the variance of activities in the parameters' peaks. Therefore, it raised the possibility that the two less important parameters (variance and correlation) could contribute more to the category encoding if they had been modulated as much as the mean of activity. To test this hypothesis, we multiplied the modulations of the three parameters by different scaling factors before feeding them to the model and recalculated the separability indices at the output of the model for conditions 4, 6 and 7 (Fig 10B). The resultant category separability showed dependence on the level of the three parameters. However, this dependence was significantly higher for the mean of activities compared to the other parameters. The correlation moderately reduced the separability in inverse proportion to the scaling factors as opposed to the other parameters which improved the separability almost an order of magnitude at the highest scaling factors. This showed that the level of modulation of the parameters could determine their contribution to the representational space.

In order to compare the impact of the three parameters in case of equal modulations, in each of the three runs of the model (for conditions 4, 6 and 7), we rescaled the other two parameters (by scaling factors) to reach the experimentally measured modulation of the other parameter and finally averaged the results of the three runs (Fig 10C). Results showed that, in case of equal modulations for the three parameters, the mean, independent and the correlated variability would contribute to the category representations as much as 18.18%, 60.25% and -18.53%, respectively. This first implies that, potentially all of the measured parameters could have positive (e.g. mean and independent variability) or negative (e.g. correlated variability) impacts on the representational space of categories which is captured by the computational model. Second, these results suggested that independent variability, could have taken the role of mean of activity if it had been equally modulated. Moreover, it showed that any increase in correlated variability (i.e. the correlation was increased in the two other cases to reach the modulation level equal to the other two parameters) could significantly deteriorate its impact in representational enhancement. Therefore, there seems to be a reason for the observed levels of modulation of the three parameters. As it was shown in Fig 10B, lower levels of variance (as indicated by increased scaling factor) could result in higher separability between categories. However, as the variance of activities is mostly suppressed in large population of neurons, the brain must have dedicated many more neurons to the task (Cohen and Maunsell, 2009). Moreover, to increase as much separability that can be achieved by very small modulations of the mean, the correlation needed to increase many orders of magnitude. Therefore, the brain seems to reflect the category information in an optimum dimension of the signals.

Next we asked whether the three parameters contributed differently to the encoding of category-orthogonal variations (Hong et al., 2016). In other words, we hypothized that the variability of activities in neural populations might have carried information regarding other aspects of object recognition such as variations. To evaluate this, instead of category information, for each variation in the dataset, we calculated the separability of the conditions in that variation. This included three conditions for each variation as shown in Fig 1B, between which the separability index was measured (3 combinations of pairs) disregarding the categorical classifications of the constituent stimuli. The average contributions for the mean, independent variability and correlated variability were 90.28%, 9.42% and 0.19%, respectively (Fig 10D), which showed no significant differences between variations. More importantly, the variation-encoding results did not show any different roles for the three parameters from those obtained for categories. This implies that, as for the category, the mean activity plays a much dominant role in the representation of variations, compared to the independent variability and correlated variability. Yet, a neural-level recording experiment seems to be needed to address this problem more conclusively.

Conclusion

The object decoding methodology along with the computational model used in this study revealed significant contribution (88.09%) from the mean signal activity, moderate contribution (10.40%) from independent signal variability and insignificant contribution (0.59%) from dependent signal variability in object category encoding in the brain. These results suggest that, in invariant object encoding (i.e. object encoding under variations), the average activity of neurons/neural populations provides one of the most probable dimensions of neural codes from which the category information can be extracted, while other dimensions (such as independent/dependent variability) may play complementary roles in the enhancement of object representations in situations such as in attention-deployed recognition tasks.

References

Abbott LF, Dayan P (1999), The effect of correlated variability on the accuracy of a population code. Neural Comput 11:91–101.

Abeles M (1991), Corticonics: Neural circuits of the cerebral cortex. Cambridge UP, New York.

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hamalainen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E (2005), Top-down facilitation of visual recognition, P NATL ACAD SCI USA 103:449-454.

Behroozi M, Daliri MR, Shekarchi B (2015), EEG phase patterns reflect the representation of semantic categories of objects. Med Biol Eng Comput 54:205-221.

Brainard DH (1997), The psychophysics toolbox. Spatial Vision 10:433–436.

Celebrini S, Thorpe S, Trotter Y, Imbert M (1993), Dynamics of orientation coding in area V1 of the awake primate. Vis Neurosci 10:811–825.

Chikkerur S, Serre T, Tan C, Poggio T (2010), What and where: a bayesian inference theory of attention. Vision Res 50:2233–2247.

Churchland MM, Yu BM, Cunningham JP, et al. (2010), Stimulus onset quenches neural variability: a widespread cortical phenomenon, Nat Neurosci 13:369-378.

Cohen M, Maunsell JHR (2009), Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12:1594-1601.

Cohen MR, Newsome WT (2008), Context-dependent changes in functional circuitry in visual area MT. Neuron 60:162–173.

Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu YC, Abdi H, Haxby JV (2012), The representation of biological classes in the human brain. J Neurosci 32:2608–2618.

Cox DD, Savoy RL (2003), Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. NeuroImage 19:261–270.

Daliri MR, Taghizadeh M, Niksirat KS (2012), EEG Signature of Object Categorization from Event‐related Potentials. J Med Sign Sens 3:37-44.

Deadman E, Higham NJ, Ralha, R (2013), Blocked Schur algorithms for computing the matrix square root. Lect Not Comput Sci 7782:171–182.

Dehaqani MRA, Vahabie AH, Kiani R, Ahmadabadi MN, Araabi BN, Esteky H (2016), Temporal dynamics of visual category representation in the macaque inferior temporal cortex. J Neurophysiology, In press.

Delorme A, Makeig S (2004), EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neuroscience Meth 134:9–21.

Denman ED, Beavers AN (1976), The matrix sign function and computations in systems. Appl Maths Comput 2:63–94.

Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988), Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. Biol Cybern 60:121–130.

Edelman S, Grill-Spector K, Kusnir T, Malach R (1998), Towards direct visualization of the internal shape space by fMRI. Psychobiology 26:309–321.

Engel AK, König P, Kreiter AK, Singer W, (1991), Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. Science 252:1177–1179.

Freedman D, Riesenhuber M, Poggio T, Miller EK (2001), Comparison of primate prefrontal and anterior temporal cortex activity during visual categorization. Soc Neurosci Abs, 852.14.

Freiwald WA, Tsao DY (2010), Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science 330:845–851.

Gerlach C (2007), A review of functional imaging studies on category specificity. J Cogn Neurosci 19: 296–314.

Goddard E, Carlson TA, Dermody N, Woolgar A (2016) Representational dynamics of object recognition: Feedforward and feedback information flows. NeuroImage 128:385–397.

Gollisch T, Meister M (2008), Rapid neural coding in the retina with relative spike latencies. Science 319, 1108–1111.

Grill-Spector K, Kanwisher NG (2001), The functional organization of human ventral temporal cortex is based on stimulus selectivity, not recognition task. Soc Neurosci Abs 122.10.

Grill-Spector K, Kourtzi Z, Kanwisher N (2001), The lateral occipital complex and its role in object recognition. Vis Res 41:1409-1422.

Grill-Spector K, Weiner KS (2014), The functional architecture of the ventral temporal cortex and its role in categorization. Nat Rev Neurosci 15:536-548.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001), Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430.

Hong H, Yamins DKL, Majaj NJ, Dicarlo JJ (2016), Explicit information for category-orthogonal object properties increases along the ventral stream. Nat Neurosci 19:613-622.

Hoogmoid AHV, Brink DVD, Janzen G (2012), Electrophysiological correlates of object location and object identity processing in spatial scenes. PLoS One 7.

Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999), Distributed representation of objects in the human ventral visual pathway. Proc Natl Acad Sci USA 96:9379–9384.

Isik L, Meyers EM, Leibo JZ, Poggio T (2014) The dynamics of invariant object categorization in the human visual system. J Neurophysiology 111:91-102.

Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M (2007), Categorization training results in shape and category-selective human neural plasticity. Neuron 53:891–903.

Kaneshiro B, Perreau Guimaraes M, Kim H-S, Norcia AM, Suppes P (2015) A Representational Similarity Analysis of the Dynamics of Object Processing Using Single-Trial EEG Classification. PLoS One 10.

Kiani R, Esteky H, Mirpour K, Tanaka K (2007), Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296–4309.

Kriegeskorte N, Mur M, Bandettini PA (2008a), Representational similarity analysis – connecting the branches of systems neuroscience. Front Syst Neurosci 2. 10.3389/ neuro.06.004.2008.

Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA (2008b), Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60:1126–1141.

Liu H, Agam Y, Madsen J R, Kreiman G (2009), Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. Neuron 62:281–290.

Majima K, Matsuo T, Kawasaki K, Kawai K, Saito N, Hasegawa I, Kamitani Y (2014), Decoding visual object categories from temporal correlations of ECoG signals. NeuroImage 90:74-83.

Martinovic J, Gruber T, Müller MM (2008), Coding of visual object features and feature conjunctions in the human brain. PLoS One.

McAdams CJ, Maunsell JH (1999), Effects of attention on the reliability of individual neurons in monkey visual cortex. Neuron 23, 765–773.

McIntosh AR, Kovacevic N, Itier RJ (2012), Increased brain signal variability accompanies lower behavioral variability in development. PLoS Comput Biol 7.

Milner PM (1974), A model for visual shape recognition. Psychol Rev 81:521–535.

Mognon A, Jovicich J, Bruzzone L, Buiatti M (2011), ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48:229-240.

Nienborg H, Cohen MR, Cumming BG (2012), Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu Rev Neurosci 35:463– 483.

Oram MW, Wiener MC, Lestienne R, Richmond BJ (1999), Stochastic nature of precisely timed spike patterns in visual system neuronal responses. J Neurophysiol 81:3021–3033.

Orlov T, Makin TR, Zohary E (2010), Topographic representation of the human body in the occipitotemporal cortex. Neuron 68:586–600.

Reisenhuber M, Poggio T (2002), Neural mechanisms of object recognition. Curr Opin Neurobiol 12:162:168.

Rishel CA, Huang G, Freedman DJ (2013), Independent category and spatial encoding in parietal cortex. Neuron 77:969–979.

Ruff DA, Cohen MR (2016), Attention increases spike count correlations between visual cortical areas. J Neurosci 36:7523-7534.

Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996), A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neuroscience 16:1486–1510.

Simanova I, Gerven MV, Oostenveld R, Hagoort P (2010), Identifying object categories from event-related eeg: toward decoding of conceptual representations. PLoS One 5.

Sofer I, Crouzet SM, Serre T (2015), Explaining the Timing of Natural Scene Understanding with a Computational Model of Perceptual Categorization. PLoS Comput Biol 11.

Swaminathan SK, Freedman DJ (2012), Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. Nat Neurosci 15:315–320.

Taghizadeh-Sarabi M, Daliri MR, Niksirat KS (2014), Decoding objects of basic categories from electroencephalographic signals using wavelet transform and support vector machines. Brain Topogr 28:33-46.

Thorpe S, Delorme A, Van Rullen R (2001), Spike-based strategies for rapid processing. Neural Netw 14:715–725.

Thorpe S, Fize D, Marlot C (1996), Speed of processing in the human visual system. Nature 381:520-522.

Underleider LG, Mishkin M (1982), Two cortical visual systems. In: Ingle MA, Analysis of Visual Behavior, pp 549–586. Cambridge, MA: MIT Press.

Van Rullen R, Gautrais J, Delorme A, Thorpe S (1998), Face processing using one spike per neuron. Biosystems 48:229–239.

Widmann A, Schröger E (2012), Filter effects and filter artifacts in the analysis of electrophysiological data. Front Psych 3.

Ze´non A, Krauzlis RJ (2012), Attention deficits without cortical neuronal deficits. Nature 489:434–437.

Fig 1. The image set. (A), Rows show animal, car, face and plane categories and their constituent exemplars. The exemplar images are chosen from the frontal light source condition. (B), The conditions that each category exemplar in (A) underwent (here only the conditions for the first animal exemplar are shown). Exemplar images, in lighting and pose variations, are shown larger than used in experiments for better illustration. Rows from top show lighting, pose, size and position conditions, respectively. Information about each condition is provided below it.

Fig 2. Experimental paradigm. After a 200-ms fixation point, two consecutive visual stimuli were presented for 50 ms with 1200 ms of inter-stimulus interval. After the second stimulus, the subject was supposed to indicate if the stimulus-accompanying fixation points had different colors or not.

Fig 3. Procedure for calculating the separability of categories. A point was dedicated to the responses of the 31 EEG channels as evoked by each stimulus (i.e. here only responses of two channels are shown). Then the points are projected on the line connecting the mean of the two categories data clusters (as shown by crosses). The separability of two sample categories is then measured in units of d' using equation (1).

Fig 4. Category separability as a function of time. (A), the averaged separability for all pairs of categories. (B), the separability for different pairs of categories shown in different colors. The shaded error regions show the standard error across the 10 subjects. The circles indicate the time points at which the corresponding separability value was significantly ($p < 0.5$, Wilcoxon signed-rank test) above the average separability value in the last 200 ms pre-stimulus window (as shown by the horizontal dashed lines). The vertical dashed lines indicate the time of the stimulus onset.

Fig 5. The dynamics of category separability on the scalp. In order to reach these plots, the separability index was calculated on the signals from individual channels and the final maps were constructed by the superposition of the individual electrodes at several representative time points.

Fig 6. The dynamical behavior of the three parameters under study. The mean (A) and Fano factor (B) of single channel activities as well as the inter-channel correlations (C) are plotted as a function of time for each category. To reach these plots, the parameters were calculated in 50 ms moving windows.

Therefore, the result at each time point shows the parameters calculated in the -25 to +25 ms relative to that time point. Results of (A) and (B) are averaged across the channels and the results of (C) are averaged across all channel combinations (i.e. 465 combinations). The shaded areas show the standard error across subjects. The circles indicate the time points in which the corresponding parameter was significantly ($p < 0.5$, Wilcoxon signed-rank test) above the average parameter value in the last 200 ms pre-stimulus window (as shown with the horizontal dashed line). The diamonds indicate the time points at which at least one category showed significantly different parameters from the other three categories.

Fig 7. Modulation of parameters relative to their pre-stimulus values. (A), The scalp maps showing the modulation of the mean of activities at three peak separability time points. (B), The same as in (A), but for the mean-normalized variance (Fano factor) of activities. (C), The relationship between inter-channel distance (465 pairs) and their time correlations which were averaged in the whole post-stimulus span. (D), The channel numbers and their positions used in the correlation matrices in (E). (E), The inter-channel matrices of correlation modulations for the peak time points of the separability indicated shown in Fig 4A.

Fig 8. Model's accuracy at estimating the experimentally recorded activities at several candidate channels for a sample subject. The plots show the channel activities which are averaged across categories and trials. The Root-Mean-Square Error (RMSE) is shown for each channel. The channels were selected so since they provided a sub-sample of different areas of the scalp.

Fig 10. Contribution of the mean, independent and dependent variability to the representational space. (A), The percentage of contribution for each parameter. (B), The changes of separability as a function of different levels of modulation of the three parameters. To reach these plots, we multiplied the levels of modulations which were obtained from the experiment by different scaling factors. (C), The same as (A), but for when the modulation levels were rescaled to match between the three parameters. The stars indicate significant differences ($p < 0.01$, as evaluated by Wilcoxon signed-rank). (D), The contribution of the three parameters to the representation of different variations.

Highlights of the research "***Average activity, but not variability, is the dominant factor in the representation of object categories***" by Hamid Karimi-Rouzbahani et al.

- Within-trial variability of neural activities provide information regarding object categories
- The mean of activities plays a significant role in category encoding in the brain
- The independent variability, although effective, contribute only moderately to the category representation
- The correlated variability between the EEG channels has an ignorable effect on the representations