bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Temporal codes provide additional category-related information in object category decoding: a systematic comparison of informative EEG features

Hamid Karimi-Rouzbahani[1,2,3]*, Mozhgan Shahmohammadi[4], Ehsan Vahab[5], Saeed Setayeshi[6], Thomas Carlson[7,2]

[1]Medical Research Council Cognition and Brain Sciences Unit, University of Cambridge, UK

[2]Perception in Action Research Centre and Department of Cognitive Science Macquarie University, Australia

[3]Department of Computing, Macquarie University, Australia

[4]Department of Computer Engineering, Central Tehran Branch, Islamic Azad University, Iran

[5]Department of Computer and Information and Technology Engineering, Qazvin Branch, Islamic Azad University, Iran

[6]Department of Medical Radiation Engineering, Amirkabir University of Technology, Iran

[7]School of Psychology, University of Sydney, Australia

* to whom correspondence should be addressed.

Abstract

Humans are remarkably efficent at recognizing objects. Understanding how the brain performs object recognition has been challenging. Our understanding has been advanced substantially in recent years with the development of multivariate decoding methods. Most start-of-the-art decoding procedures, make use of the 'mean' neural activation to extract object category information, which overlooks temporal variability in the signals. Here, we studied category-related information in 30 mathematically distinct features from electroencephalography (EEG) across three independent and highly-varied datasets using multivariate decoding. While the event-related potential (ERP) components of N1 and P2a were among the most informative features, the informative original signal samples and Wavelet coefficients, selected through principal component analysis, outperformed them. The four mentioned informative features showed more pronounced decoding in the Theta frequency band, which has been suggested to support feed-forward processing of visual information in the brain. Correlational analyses showed that the features, which were most informative about object categories, could predict participants' behavioral performance (reaction time) more accurately than the less informative features. These results suggest a new approach for studying how the human brain encodes object category information and how we can read them out more optimally to investigate the temporal dynamics of the neural code. The codes are available online at https://osf.io/wbvpn/.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Keywords

object category processing; multivariate pattern analysis; multivariate decoding; electroencephalography (EEG); feature extraction

Introduction

How does the brain encode information about visual object categories? This question has been studied for many years using different neural recording techniques including invasive neurophysiology (Hung et al., 2005) and electrocorticography (ECoG; Majima et al., 2014; Watrous et al., 2015; Rupp et al., 2017; Lie et al., 2009; Miyakawa et al., 2018; Liu et al., 2009), as well as non-invasive methods such as functional Magnetic Resonance Imaging (fMRI; Haxby et al., 2001), magnetoencephalography (MEG; Contini et al., 2017; Carlson et al., 2013) and electroencephalography (EEG; Kaneshiro et al., 2016; Simanova et al., 2010) or a combination of them (Cichy et al., 2014). Despite the recent successes in neuroimaging in "reading-out" or "decoding" neural representations of semantic object categories, it is unclear whether conventional decoding analyses leverage the main feature in which the recorded neural activity reflects object category information. Majority of these studies, rely on the signal's 'mean' amplitude (i.e. average voltage across EEG electrodes), which although informative, but might be a sub-optimal feature for decoding the object category information from neural activations as it ignores many subtle fluctuations that can be informative. The use of this potentially sub-optimal feature might thus hide the true temporal dynamics of object category encoding in the brain, which is still debated in cognitive neuroscience (Grootswagers et al., 2019; Majima et al., 2014; Karimi-Rouzbahani et al., 2017a; Behroozi et al. 2016; Isik et al., 2013; Cichy et al., 2014). Here, we quantitatively compare the information content of a large set of relevant features extracted from EEG activity, which have successfully provided object category information in previous studies, and evaluate their neural relevance by measuring how well each feature explains behavioral object recognition performance.

Multivariate pattern analysis (MVPA), especially multivariate decoding, has become a central method for the analysis of neuroimaging data (i.e. fMRI, M/EEG), especially in studying neural coding of object categories (Norman et al., 2006; Tong and Pratte, 2012; Haynes et al., 2015; Grootswagers et al., 2017; Hebart and Baker, 2018). MVPA incorporates activations across multiple recording sites (i.e. sensors/electrodes in M/EEG or brain voxels in fMRI) to detect subtle but widespread differences between patterns of activity across conditions (e.g. object categories). These differences might not be detectable when comparing univariate brain activations (Norman et al., 2006; Grootswagers et al., 2017), such as when comparing conventional single-electrode event-related potentials (ERPs) in EEG (Ambrus et al., 2019). While fMRI provides millimeter-scale spatial resolution allowing us to localize the brain areas involved in object category processing (Haxby et al., 2001), the recorded activations are usually modelled using generalized linear regression techniques for both univariate and multivariate analyses. This statistical modelling approach, which is a prerequisite step when analyzing evoked activations in fMRI, removes many critical/relevant temporal variabilities (i.e. potential codes) which have been shown to provide information about visual stimuli when adopting methods with high temporal resolution such as ECoG, EEG or single cell recording (Eckhorn et al., 1988; Celebrini et al., 1993; Gollisch and Meister, 2008; Majima et al., 2014; Karimi-Rouzbahani et al., 2017a).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The temporal resolution of M/EEG allows us to analyze the temporal neural variabilities with millisecond resolution. Utilizing this property along with the sensitivity of multivariate decoding, the communities of neuroscience and brain-computer interface (BCI) have gained deeper insights into how we can decode information about object categories from recorded neural activity. Earlier decoding studies have shown that individual mean-based ERP components such as N1, P1, P2a and P2b, which refer to arbitrary time windows in neural time series extracted from 100 to 300 ms post-stimulus onset, could reliably differentiate visual object categories (Chan et al., 2011). Later studies used Linear Discriminant Analysis (LDA) classifiers to discriminate up to four object categories utilizing the information content of those ERP components (Wang et al., 2012), which were later fused (combined) to improve previous decoding accuracies (Qin et al., 2016). However, these studies and others (Taghizadeh-Sarabi et al., 2015; Torabi et al., 2017; Wang et al., 2018) overlooked the temporal dynamics of object category encoding which is determined by within-trial dynamics of neural category processing. To address this issue, researchers repeated the decoding procedure in short (e.g. 4~50 ms) sliding time windows within trials and revealed the dynamical profile of category information within trials, which significantly varied by factors such as task, image presentation time, etc. (Cichy et al., 2014; Kaneshiro et al., 2015; Karimi-Rouzbahani et al., 2017a; Karimi-Rouzbahani et al., 2017b; Karimi-Rouzbahani et al., 2018; Karimi-Rouzbahani et al., 2019; Grootswagers et al., 2017; Grootswagers et al., 2019). While providing a temporal account for the neural object category processing, these time-resolved decoding studies overlooked other possible features of neural activations that could provide additive object category information. More importantly, these potentially different temporal profiles might help explaining the behavioral performance more accurately (Williams et al., 2007; Grootswagers et al., 2018; Woolgar et al., 2019).

To address this issue, studies have investigated other features of brain activations such as the phase of the signal (Behroozi et al., 2016; Watrous et al., 2015; Torabi et al., 2017; Wang et al., 2018; Voloh et al., 2020), signal power across frequency bands (Rupp et al., 2017; Miyakawa et al., 2018; Behroozi et al., 2016; Majima et al., 2014; Miyakawa et al., 2018, with band-specific contents), time-frequency features such as Wavelet coefficients (Hatamimajoumerd and Talebpour, 2019; Taghizadeh-Sarabi et al., 2015), inter-electrode correlations (Majima et al., 2014; Karimi-Rouzbahani et al., 2017a), nonlinear statistical features (Joshi et al., 2018; Torabi et al., 2017; Stam, 2005). Behroozi et al., (2016) also decoded object categories using phase patterns in the Delta frequency band (i.e. 1-4 Hz). This finding was later investigated by another group using Hilbert transform, but found the information in the Theta frequency band (4-8 Hz; Wang et al., 2018). Other studies found that signal power contained significant category-related information (Rupp et al., 2017; Majima et al., 2014; Miyakawa et al., 2018). An ECoG investigation found the information in the synchronization of low- (~2.5 Hz) and high-frequency (~84 Hz) oscillations (Watrous et al., 2015). Other research showed that, nonlinear statistical features, such as fractal dimensions, Hjorth complexity, and entropy, which measure the nonlinear structures (complexity) of signals as an indicator of its information richness, could discriminate object categories (Torabi et al., 2017; Joshi et al., 2018). These feature-based studies, however, were done on whole-trial time windows, thus they provided no insight into temporal dynamics.

Together, these extant literature leaves three unanswered questions about the neural decoding of object categories. First, which features of the recorded signals are most informative about object

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

categories? Specifically, while several of the above studies have compared multiple features (Chan et al., 2011; Taghizadeh-Sarabi et al., 2015; Torabi et al., 2016), they focused on specific classes of features (e.g. ERPs, nonlinear features, etc.), limiting our understanding of how different feature classes compare to one another. Moreover, some features (e.g. signal power (Behroozi et al., 2016 vs. Wang et al., 2018), or inter-electrode correlation (Majima et al., 2014 vs. Karimi-Rouzbahani et al., 2017a) have shown discrepancy across datasets. This suggest that their results might have been driven by category-irrelevant task features such as attentional load or task demands, which can modulate the neural activity to a greater level than that evoked by a stimulus (Karimi-Rouzbahani et al., 2019; Karimi-Rouzbahani et al., 2020b). Therefore, a wider set of features and datasets should be evaluated to provide more generalizable conclusions. Our prediction is that, as the processing of object categories in the brain is both spatially and temporally specific to regions and time windows (Cichy et al., 2014), the information should be detected by features which are spatially and/or temporally specific (e.g. ERP components and multi-valued features). Our second prediction is that, as the visual object category processing is dominantly supported by the feed-forward visual streams (Dicarlo et al., 2012; Vaziri-Pashkam and Xu, 2017), the information should be mainly reflected in the Theta frequency band which has been previously suggested to support feed-forward flow of visual information in the brain (Bastos et al., 2015).

Second, what is the temporal dynamics of object category decoding when using distinct features of brain activations? Very few studies have investigated the role of features, other than the mean of activations, in time-resolved decoding procedures (Majima et al., 2014; Stewart et al., 2014; Karimi-Rouzbahani et al., 2017a). An ECoG study by Majima et al., 2014 showed that inter-electrode correlation of time samples carried more categorical information than signal power and phase features in 100 and 300-ms wide sliding time windows. We have recently found that inter-electrode correlation of time samples carried categorical object information, but it was much weaker than the signal variance within 50 ms sliding time windows (Karimi-Rouzbahani et al., 2017a). In time-resolved decoding, we predict that we can obtain more information using the features which detect informative samples or extract information from those samples compared to when time-averaging samples as in the conventional mean-based time-resolved decoding procedures (Grootswagers et al., 2017). The reason for this prediction is that, although time-averaging of consecutive samples in (short or long) time series increases the signal-to-noise ratio at the expense of temporal specificity, it will treat all the samples contained in the window similarly; informative or not. Therefore, informative signal samples and features can average out (when averaging) or get ignored (when down-sampling) if combined with no-informative samples. Potential improvement in decoding can allow us to observe and interpret the sub-threshold neural information, which were ignored because of failing to reach the threshold of significance for inference.

Third, which features of brain activations explain behavioral recognition performance? A major open question in neuroimaging is whether the information that is extracted from neural activity is relevant or is just epiphenomenal to the target condition. To answer this question, recent efforts have tried to explain behavioral performance using multivariate decoding accuracies (Williams et al., 2007; Grootswagers et al., 2018; Woolgar et al., 2019). These studies found that the decoding accuracy

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

obtained by analyzing mean signal activations can predict the behavioral performance in object recognition (Ritchie, et al., 2015). However, it has remained unknown how (if at all) the decoding accuracies obtained from other features can explain more variance of the behavioral performance. Our prediction is that as the more informative features access more of the subtle and overlooked aspects of neural activation, as reflected in their improved decoding, they should explain the behavioral performance more accurately.

To address these questions, we re-implemented a set of 30 features from the literature and quantitatively evaluated their information about object categories from neural activity on the whole trial data as well as sliding time windows. We evaluated the features across three independent datasets, which varied in many parameters including the image set, task, and paradigm. This allowed us to obtain more generalizable results about the role of each feature in category decoding and explaining behavioral object recognition performance.

Methods

Overview of datasets

We chose three previously published EEG datasets for this study, which differed across a wide range of parameters including the recording set-up (e.g. amplifier, number of electrodes, preprocessing steps, etc.), characteristics of the image-set (e.g. number of categories and exemplars within each category, colorfulness of images, etc.), paradigm and task (e.g. presentation length, order and the participants' task; Table 1). All three datasets previously successfully provided object category information from electrical brain activity using multivariate decoding methods.

Dataset 1. We have previously collected Dataset 1 while participants were briefly (i.e. 50 ms) presented with gray-scale images from four synthetically-generated 3D object categories (Karimi-Rouzbahani et al., 2017a). The objects underwent systematic variations in scale, positional periphery, in-depth rotation and lighting conditions, which made their perception difficult, especially in extreme variation conditions. Randomly ordered stimuli were presented in consecutive pairs (Figure 1, top row). The participants' task was unrelated to object categorization- they pressed one of two pre-determined buttons to indicate if the fixation dots, superimposed on the first and second stimuli, were the same/different color (2-alternative forced-choice).

Dataset 2. We have collected Dataset 2 in an active experiment, in which participants pressed a button if the presented object image was from a target category (go/no-go), which was cued at the beginning of each block of 12 stimuli (Karimi-Rouzbahani et al., 2019; Figure 1, middle row). The object images, which were cropped from real photographs, were part of the well-stablished benchmark image set for object recognition developed by Kiani et al., (2007), which has been previously used to extract object category information from both human and monkey brain using MEG (Cichy et al., 2014), fMRI

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

197 (Cichy et al., 2014; Kriegeskorte et al., 2008) and single-cell electrophysiology (Kriegeskorte et al., 2008;
198 Kiani et al., 2007).

199

200 **Dataset 3.** We also adopted another dataset (Dataset 3) which was not collected in our lab. This
201 dataset was collected by Kaneshiro et al., (2015) on 6 sessions for each participant, from which we used
202 the first session only, as it could represent the whole dataset- other sessions were repetitions of the first
203 session and aimed at increasing the signal to noise ratio by repeating the presentation of the same stimuli.
204 The EEG data was collected during passive viewing (participants had no task; Figure 1, bottom row) of 6
205 categories of objects with stimuli chosen from Kiani et al. (2007) as explained above. We used a pre-
206 processed (i.e. band-pass-filtered in the range 0.03 to 50 Hz) version of the dataset which was available
207 online[1].

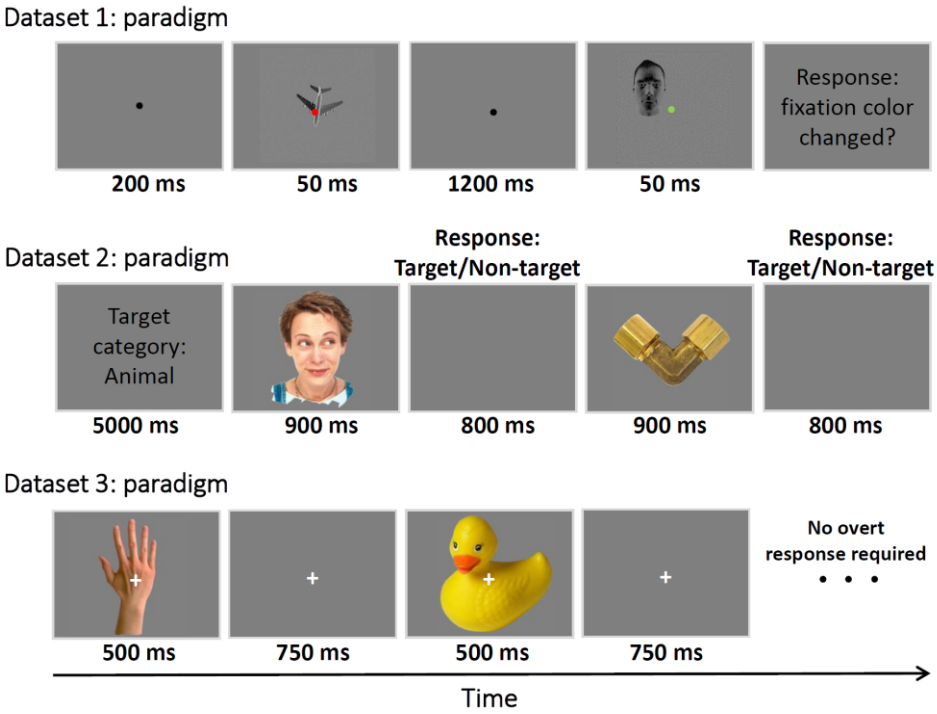


Figure 1. Paradigms of the datasets used in this study. Dataset 1 (top row) presented two consecutive object images with a fixation dot. Participants' task was to indicate if the fixation dot was the same or different colors across the object images (passive task). Dataset 2 (middle row) presented objects from the target and non-target categories in sequences of 12 images. Participant's task was to indicate, for each image, if it was from the target/non-target category (active task). Dataset 3 (bottom row), presented sequences of object images from 6 different categories. Participants did not have any specific task, except for looking at the center of object images (no overt task). Please see the details of the datasets in the relevant references cited in Table 1.

208

209 All three datasets were collected at a sampling rate of 1000 Hz. For Datasets 1 and 2, only the trials which
210 led to correct responses by participants, were used in the analyses. Each dataset consisted of data from

[1] https://purl.stanford.edu/tc919dd5388

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

10 participants. Each object category in each dataset included 12 exemplars. To make the three datasets as consistent as possible, we pre-processed them differently from their original papers. Specifically, the band-pass filtering range of Dataset 3 was 0.03 to 50 Hz, and we did not have access to the raw data to increase the upper bound to 200 Hz. Datasets 1 and 2 were band-pass-filtered in the range from 0.03 to 200 Hz before the data was split into trials. We used finite-impulse-response filters with 12 dB roll-off per octave for band-pass filtering of Datasets 1 and 2 and when evaluating the sub-bands of the three datasets. We did not remove artifacts (e.g. eye-related and movement-related) from the signals, as we and others have shown that sporadic artifacts have minimal effect in multivariate decoding (Grootswagers et al., 2017). To increase signal to noise ratios in the analyses, each unique stimulus had been presented to the participants 3, 6 and 12 times in Datasets 1, 2, 3, respectively. Trials were defined in the time window from 200 ms before to 1000 ms after the stimulus onset to cover most of the range of event-related neural activations. The average pre-stimulus (-200 to 0 ms relative to the stimulus onset) signal amplitude was removed from each trial of the data. For more information about each dataset see Table 1 and the references to their original publications.

Table 1. Details of the three datasets used in the study.

	Dataset	# electrodes	Band-pass filtering	Notch filtering	# object categories	# stimulus repetition	Stimulus presentation time	Stimulus size (periphery)	Task	Participants' accuracy
1	Karimi-Rouzbahani et al., 2017a	31	0.03-200 Hz	50 Hz	4	3	50 ms	2~13.5° (0.7~8.8°)	Color matching (passive)	%94.68
2	Karimi-Rouzbahani et al., 2019	31	0.03-200 Hz	50 Hz	4	6	900 ms	8° × 8° (0)	Object category detection (active)	%94.65
3	Kaneshiro et al., 2017	124	0.03-50 Hz	No	6	12	500 ms	7.0° × 6.5° (0)	No task (fixation)	N/A

Features

EEG signals are generated by inhibitory and excitatory post-synaptic potentials of the cortical neurons. These potentials extend to the scalp surface and are recorded through electrodes as amplitudes of voltage in units of microvolts. Researchers have been using different aspects of these voltage recordings to obtain meaningful information about human brain processes. Below we explain the mathematical formulation of each individual feature used in this study. We also provide brief information about underlying neural mechanisms which lead to the information content provided by those EEG features. We classified the features into five arbitrary classes based on their mathematical similarity to simplify the presentation of the results and their interpretations. The five classes consist of Moment, Complexity, ERP, Frequency-domain and Multi-valued features. However, the classification of the features is not strict and the features might be classified based on other criteria and definitions. For example, complexity itself has different definitions (Tononi et al., 1998), such as degree of randomness, or degrees of freedom in a large system of interacting elements. Therefore, each definition may exclude or include some of our features in the class. It is of note that, we only used the features which were previously used to decode categories of evoked potentials from EEG signals mainly through multivariate

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

decoding methods. Nonetheless, there are definitely other features, especially, those extracted from EEG time series collected during long-term monitoring of human neural representations in health and disorder. In presenting the features' formulations, we avoided repeating the terms repeatedly from the first feature to the last one. Therefore, the reader might need to go back a few steps to find the definitions of terms.

Moment features

These features are the most straightforward and intuitive features from which we might be able to extract information about neural processes. Mean, Variance, Skewness and Kurtosis are the 1st to the 4th moments of EEG time series and can provide information about the shape of the signals and their deviation from stationarity which is the case in evoked potentials (Rasoulzadeh et al., 2016; Wong et al., 2006). These moments have been shown to be able to differentiate visually evoked responses (Pouryzdian and Erfaninan, 2010; Alimardani et al., 2018).

Mean

Mean amplitude of an EEG signal changes in proportion to the neural activation of the brain. It is by far the most common feature of the recorded neural activations used in analyzing brain states and cognitive processes either in univariate and multivariate analysis (Vidal et al., 2010; Hebart and Baker, 2017; Grootswagers et al., 2017; Karimi-Rouzbahani et al., 2019). In EEG, the brain activation is reflected as the amplitude of the recorded voltage across each electrode and the reference electrode at specific time points. To calculate the Mean feature, which is the first moment in statistics, the sample mean is calculated for recorded EEG time series as:

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t \quad (1)$$

where $\bar{x}$ is the mean of the N time samples contained in the time series and x_t refers to the amplitude of the recorded sample at time point t. N can be as small as unity as in the case of time-resolved EEG analysis (Grootswagers et al., 2017) or as large as required by the analysis. In this study, we set $N = 50$ and $N = 1000$ for the time-resolved and whole-trial decoding analyses, respectively.

Median

Compared to the Mean feature, Median is less susceptible to outliers (e.g. spikes) in the time series, which might not come from neural activations but rather from artifacts caused by the recording hardware, preprocessing, eye-blinks, etc. Provided that the signal probability distribution is P, signal median m, is calculated so that it meets the following conditions:

$$P(x > m) \geq \frac{1}{2} \text{ and } P(x \leq m) \geq \frac{1}{2} \quad (2)$$

Variance

Variance of an EEG signal is one of the best indicators showing how much the signal is deviated from stationarity i.e. deviated from its original baseline statistical properties (Wong et al., 2006). It is a measure of signal variability, has been shown to squeeze upon the stimulus onset as a result of neural co-activation

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Churchland et al., 2010) and has provided information about object categories in a recent EEG decoding study (Karimi-Rouzbahani et al., 2017a). Variance is calculated as:

$$\sigma^2 = \frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2 \quad (3)$$

Skewness

While Variance is silent about the direction of the deviation, Skewness, which is the third signal moment, measures the degree of asymmetry in the signal's probability distribution. In symmetric distribution (i.e. when samples are symmetrically around the mean) skewness is zero. Positive and negative skewness indicates right- and left-ward tailed distribution, respectively. As the visually evoked ERP responses usually tend to be asymmetrically deviated in either positive or negative direction, even after baseline correction (Mazaheri and Jensen, 2008), we assume that Skewness should provide information about the visual stimulus if each category modulates the deviation of the samples differentially. Skewness is calculated as:

$$\gamma_1 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{x_t - \bar{x}}{\sigma}\right)^3 \quad (4)$$

Kurtosis

Kurtosis reflects the degree of 'tailedness' or 'flattedness' of the signal's probability distribution. Accordingly, the more heaviness in the tails, the less value of the Kurtosis and vice versa. Based on previous studies, Kurtosis has provided distinct representations corresponding to different classes of visually evoked potentials (Alimardani et al., 2018; Pouryzdian and Erfaninan, 2010). We test to see if Kurtosis plays a more generalized role in information coding e.g. coding of semantic aspects of visual information as well. It is the fourth standardized moment of the signal defined as:

$$Kurt = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{x_t - \bar{x}}{\sigma}\right)^4 \quad (5)$$

Complexity features

There can potentially be many cases in which simple moment statistics such as Mean, Median, Variance, Skewness and Kurtosis provide equal values for distinct time series (e.g. A: 10, 20, 10, 20, 10, 20, 10, 20 vs. B: 20, 20, 20, 10, 20, 10, 10, 10) for both of which the five above-mentioned features provide equal results. Therefore, we need more complex and possibly nonlinear measures which can capture subtle but significant differential patterns across distinct time series. The analysis of nonlinear signal features has recently been growing, following the findings showing that EEG reflects weak but significant nonlinear structures (Stam, 2005; Stepien, 2002). Importantly, many studies have shown that the complexity of EEG time series can significantly alter during cognitive tasks such as visual (Bizas et al., 1999) and working memory tasks (Sammer et al., 1999; Stam, 2000). Therefore, it was necessary to evaluate the information content of nonlinear features for our multivariate decoding of object categories. As mentioned above, the grouping of these nonlinear features as "complexity" here is not strict and the features included in this class are those which capture complex and nonlinear patterns across time series. Although the accurate detection of complex and nonlinear patterns generally need more time samples compared to linear patterns (Procaccia, 1988), it has been shown that nonlinear

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

structures can be detected from short EEG time series (i.e. through fractal dimensions; Preißl et al., 1997). Moreover, to ensure that we do not miss the detection of nonlinear structures as a result of the short time series, we extracted the nonlinear features from the time-resolved (50 samples) and the whole-trial data (1000 samples).

Lempel-Ziv complexity (LZ Cmplx)

Lempel-Ziv complexity measures the complexity of a time series (Lempel et al., 1976). Basically, the algorithm counts the number of unique sub-sequences within a larger binary sequence. Accordingly, a sequence of samples with a certain regularity does not lead to a large LZ complexity. However, the complexity generally grows with the length of the sequence and its irregularity. In other words, it measures the generation rate of new patterns along a digital sequence. In a comparative work, it has been shown that, compared to many other frequency metrics of time series (e.g. noise power, stochastic variability, etc.), LZ complexity has the unique feature of providing a scalar estimate of the bandwidth of a time series and the harmonic variability in quasi-periodic signals (Aboy et al., 2006). It is widely used in biomedical signal processing and has provided successful results in the decoding of visual stimuli from neural responses in primary visual cortices (Szczepański et al., 2003). We used the code by Quang Thai[2] implemented based on "exhaustive complexity" which is considered to provide the lower limit of the complexity as explained by Lempel et al. (1976). We compared the results obtained from this implementation to two other implementations, which provided identical results. We used the signal median as a threshold to convert the signals into binary sequences for the calculation of LZ complexity. The LZ complexity provided a single value for each signal time series.

Fractal dimension

In signal processing, fractal is an indexing technique which provides statistical information determining the complexity of the time series. A higher fractal value indicates more complexity for a sequence as reflected in more nesting of repetitive sub-sequences at all scales. Fractal dimensions are widely used to measure two important attributes: self-similarity and the shape of irregularity. A growing set of studies have been using fractal analyses for the extraction of information about semantic object categories (such as living and non-living categories of visual objects; Ahmadi-Pajouh et al., 2018; Torabi et al., 2017) as well as simple checkerboard patterns (Namazi et al., 2018) from visually evoked potentials. These results support the coding of visual information in EEG signal patterns through the modulation of their nonlinear structure. In this study, we implemented two of the common methods for the calculation of fractal dimensions of EEG time series, which have been previously used to extract information about object categories as explained below. We used the implementations by Jesús Monge Álvarez[3] after verifying it against other implementations.

- **Higuchi's fractal dimension (Higuchi FD)**

[2] https://www.mathworks.com/matlabcentral/fileexchange/38211-calc_lz_complexity

[3] https://ww2.mathworks.cn/matlabcentral/fileexchange/50290-higuchi-and-katz-fractal-dimension-measures

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In this method (Higuchi et al., 1988), a set of sub-sequences x_k^m is generated in which k and m refer to the step size and initial value, respectively. Then, the length of this fractal dimension is calculated as:

$$L_k^m = \frac{\left\{\left[\sum_{i=1}^{\left[\frac{N-m}{k}\right]} |x_{(m+ik)} - x_{(m+(i-1).k)}|\right] \frac{N-1}{\left[\frac{N-m}{k}\right].k}\right\}}{k} \quad (6)$$

where $\frac{N-1}{\left[\frac{N-m}{k}\right].k}$ is the normalization factor. The length of the fractal curve at step size of k is calculated by averaging k sets of L_k^m. Finally, the resultant average will be proportional to k^{-D} where D is the fractal dimension. We set the free parameter of k equal to half the length of signal time series in the current study.

- **Katz's fractal dimension (Katz FD)**

We also calculated fractal dimension using the Katz's method (Katz, 1988) as it showed a significant amount of information about object categories in a previous study (Torabi et al., 2017). The fractal dimension (D) is calculated as:

$$D = \frac{log_{10}(\frac{L}{a})}{log_{10}(\frac{d}{a})} = \frac{log_{10} r}{log_{10}(\frac{d}{L}) + log_{10} r} \quad (7)$$

where L and a refer to the sum and average of the consecutive signal samples, respectively. Also d refers to the maximum distance between first sample of the signal and i^{th} sample of the signal which has the maximum distance from first sample as:

$$L = \sum_{i=2}^{N} |x_i - x_{i-1}| \quad (8)$$

$$d = max(distance(1, i)) \quad (9)$$

$$r = L/a \quad (10)$$

Hurst exponent (Hurst Exp)

Hurst exponent is widely used to measure the long-term memory in a time-dependent random variable such as biological time series (Racine, 2011). In other words, it measures the degree of interdependence across samples in the time series and operates like an autocorrelation function over time. Hurst values between 0.5 and 1 suggest consecutive appearance of high signal values on large time scales while values between 0 and 0.5 suggest frequent switching between high and low signal values. Values around 0.5 suggest no specific patterns among samples of a time series. It is defined as an asymptotic behavior of a rescaled range as a function of the time span of the time series defined as:

$$E\left[\frac{\max(z_1, z_2, \ldots, z_N) - \min(z_1, z_2, \ldots, z_N)}{\sqrt{\frac{1}{N}\sum_{t=1}^{N}(x_t - \bar{x})^2}}\right] = C.N^H \; as \; N \to \infty \quad (11)$$

$$z_t = \sum_{i=1}^{t} y_i \; ; t = 1, \ldots, N \quad (12)$$

$$y_t = x_t - \bar{x} \quad (13)$$

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where E is the expected value, C is a constant and H is the Hurst exponent (Racine, 2011). We used the open-source implementation of the algorithm[4], which has also been used previously for the decoding of object category information in EEG (Torabi et al., 2017). We compared two more implementations all of which provided identical results.

Entropy

Entropy can measure the perturbation in time series. A higher value for entropy suggests a higher irregularity in the given time series. Precise calculation of entropy usually requires considerable number of samples and is also sensitive to noise. Here we used two methods for the calculation of entropy, each of which has its advantages over the other.

- **Approximate entropy (Apprx Ent)**

Approximate entropy was initially developed to be used for medical data analysis (Pincus and Huang, 1992), such as heart rate, and then was extended to other areas such as brain data analysis. It has the advantage of requiring a low computational power which makes it perfect for real-time applications on low sample sizes (N<50). However, the quality of this entropy method is impaired on lower length of the data. This metric detects changes in episodic behavior which are not represented by peak occurrences or amplitudes (Pincus and Huang, 1992). We used an open-source code[5] for calculating approximate entropy. We compared the results obtained from this implementation to one other implementation, which provided identical results. We set the embedded dimension and the tolerance parameters to 2 and 20% of the standard deviation of the data respectively, to roughly follow a previous study (Shourie et al., 2014) which compared approximate entropy in visually evoked potentials and found differential effects across artist vs. non-artist participants when looking at paintings.

- **Sample entropy (Sample Ent)**

Sample entropy, which is a refinement of the approximate entropy, is frequently used to calculate regularity of biological signals (Richman et al., 2000). Basically, it is the negative natural logarithm of the conditional probability that two sequences (subset of samples), which are similar for m points remain similar at the next point. A lower sample entropy also reflects a higher self-similarity in the time series. It has two main advantages to the approximate entropy: it is less sensitive to the length of the data and is simpler to implement. However, it does not focus on self-similar patterns in the data. We used the Matlab 'entropy' function for the implementation of this feature, which has already provided category information in a previous study (Torabi et al., 2017). See (Richman et al., 2000; Subha et al., 2010) for the details of the algorithm.

Autocorrelation (Autocorr)

Autocorrelation determines the degree of similarity between the samples of a given time series and a time-lagged version of the same series. It detect periodic patterns in signal time series, which is an integral

[4] https://www.mathworks.com/matlabcentral/fileexchange/9842-hurst-exponent

[5] https://www.mathworks.com/matlabcentral/fileexchange/32427-fast-approximate-entropy

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

part of EEG time series. Therefore, following recent successful attempts in decoding neural information using the autocorrelation function from EEG signals (Wairagkar et al., 2016), we evaluated the information content of the autocorrelation function in decoding visual object categories. As neural activations reflect many repetitive patterns across time, the autocorrelation function can quantify the informationcontents of those repetitive patterns. Autocorrelation is calculated as:

$$R(\tau) = \frac{1}{(N-\tau)\sigma^2}\sum_{t=1}^{N-\tau}(x_t - \bar{x})(x_{t+\tau} - \bar{x}) \quad (14)$$

where τ indicates the number of lags in samples of the shifted signal. A positive value for autocorrelation indicates a strong relationship between the original time series and its shifted version, whereas a negative autocorrelation refers to an opposite pattern between them. Zero autocorrelation indicates no relationship between the original time series and its shifted version. In this study, we extracted autocorrelations for 30 consecutive lags ([τ=1, 2, . . ., 30]) used their average in classification. Please note that each lag refers to 1 ms as the data was sampled at 1000 Hz.

Hjorth parameters

Hjorth parameters are descriptors of statistical properties of signals introduced by Hjorth (1970). These parameters are widely used in EEG signal analysis for feature extraction across a wide set of applications including visual recognition (Joshi et al., 2018; Torabi et al., 2017). These features consist of Activity, Mobility and Complexity as defined below. As the Activity parameter is equivalent to signal Variance, which we already had in the analyses, we did not repeat it.

- **Hjorth complexity (Hjorth Cmp)**

It determines the variation in time series' frequency by quantifying the similarity between the signal and a pure sine wave leading to a value of 1 in case of perfect match. In other words, values around 1 suggest lower complexity for a signal. It is calculated as:

$$Complexity = \frac{Mobility\ (\frac{dx_t}{dt})}{Mobility\ (x_t)} \quad (15)$$

- **Hjorth mobility (Hjorth Mob)**

It determines the proportion of standard deviation of the power spectrum as is calculated below, where var refers to the signal variance.

$$Mobility = \sqrt{\frac{var(\frac{dx_t}{dt})}{var(x_t)}} \quad (16)$$

ERP components (N1, P1, P2a and P2b)

An ERP is a measured brain response to a specific cognitive, sensory or motor event that provides an approach to study the correlation between the event and neural processing. According to the latency and amplitude, ERP is split into specific sub-windows called components. Here, we extracted ERP components

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

by calculating mean of signals in specific time windows to obtain the P1 (80 to 120 ms), N1 (120 to 200 ms), P2a (150 to 220 ms) and P2b (200 to 275 ms) components, which were shown previously to provide significant amounts of information about visual object and face processing in univariate (Rossion et al., 2000; Rousselett et al., 2007) and multivariate analyses (Chan et al., 2011; Jadidi et al., 2016; Wang et al., 2012).

Frequency-domain features

Frequency-domain analysis, which is the conventional and yet one of the most powerful approaches in EEG data analysis, generally inform us of the distribution of signal power over frequency bands through the use of variety of frequency-domain features such as Fourier coefficients, etc. Therefore, motivated by previous studies showing signatures of object categories in the frequency domain (Behroozi et al., 2016; Rupp et al., 2017; Iranmanesh and Rodriguez-Villegas, 2017; Joshi et al., 2018; Jadidi et al., 2016) and the reflection of temporal coding of visual information in the frequency domain (Eckhorn et al., 1988), we also extracted frequency-domain features to see if they provide additional category-related information to time-domain features. There are limitations to our frequency analysis as follow. While the whole-trial analysis can provide results that we can compare with previous studies, the evoked EEG potentials are likely nonstationary (i.e. their statistical properties change across time), potentially biasing the frequency features towards the most dominant frequency components and hiding the subtle (i.e. high-frequency) fluctuations of the signal. On the other hand, while the time-resolved analysis, which is done in 50ms sliding time windows, enable us to detect the time-varying characteristics in the frequency domain, it will only incorporate frequencies lower than 25 Hz according to the Nyquist theorem. Despite these limitations, we still extracted and analyzed the frequency-domain features as below.

Signal power (Signal Pw)

Power spectrum density (PSD) represents the intensity or the distribution of the signal power into its constituent frequency components. This feature was motivated by previous studies showing associations between aspects of visual perception and certain frequency bands (Rupp et al., 2017; Behroozi et al., 2016; Majima et al., 2014). According to the Fourier analysis, signals can be broken into its constituent frequency components or a spectrum of frequencies in a specific frequency range. Here, we calculated signal power using the PSD as (17).

$$\tilde{S}_{xx}(w) = \frac{(\Delta t)^2}{T}\left|\sum_{n=1}^{N} x_n e^{-iwn\Delta t}\right|^2 \quad (17)$$

where $x_n = x_{n\Delta t}$ is signal sampled at a rate of $T = \frac{1}{\Delta t}$ and w is the frequency at which the signal power is calculated. As signal power is a relatively broad term, including the whole power spectrum of the signal, we also extracted a few more parameters from the signal frequency representation, to see what specific features in the frequency domain (if any) can provide information about object categories.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

493 *Mean frequency (Mean Freq)*

494 Motivated by the successful application of mean and median frequencies in the analysis of EEG signals
495 and their relationship to signal components in the time domain (Intrilligator and Polich, 1995;
496 Abootalebi et al., 2009), we extracted these two features from the signal power spectrum to obtain a
497 more detailed insight into the neural dynamics of category representations. Mean frequency is the
498 average of the frequency components available in a signal. Assume a signal consisting of two frequency
499 components of f_1 and f_2. The mean frequency of this signal is $f_{mean} = \frac{f_1+f_2}{2}$. Generally, the mean
500 normalized (by the intensity) frequency is calculated using the following formula:

501 $$f_{mean} = \frac{\sum_{i=0}^{n} l_i f_i}{\sum_{i=0}^{n} l_i} \quad (18)$$

502 where n is the number of splits of the PSD, f_i and l_i are the frequency and the intensity of the PSD in its
503 i^{th} slot, respectively. It was calculated using Matlab 'meanfreq' function.

504

505 *Median frequency (Med Freq)*

506 Median frequency is the median normalized frequency of the power spectrum of a time-domain signal.
507 It is calculated similarly to the signal median in the time domain, however, here the values are the
508 power intensity in different frequency slots of the PSD. This feature was calculated using Matlab
509 'medfreq' function.

510

511 *Power and Phase at median frequency (Pw MdFrq and Phs MdFrq)*

512 Interestingly, apart from the median frequency itself, which reflects the frequency aspect of the power
513 spectrum, the power and phase of the signal at the median frequency have been shown to be
514 informative about aspects of human perception (Joshi et al., 2018; Jadidi et al., 2016). Therefore, we
515 also calculated the power and phase of the frequency-domain signals at the median frequency as
516 features.

517

518 *Average frequency (Avg Freq)*

519 As the evoked potentials show a few number of positive and negative peaks after the stimulus onset,
520 and the observation that they might show deviation in the positive or negative direction depending on
521 the information content (Mazaheri and Jensen, 2008), we also evaluated the zero-crossing frequency of
522 the ERPs. Basically, for measuring the Average Frequency, we measured the number of times the signal
523 swapped signs during the trial. Note that each trial is baselined according to the average of the same
524 trial in the last 200 ms immediately before the stimulus onset. We calculated the zero crossing rate on
525 the post-stimulus time span from the time point of 0 to 1000 ms.

526

527 *Spectral edge frequency (SEF 95%)*

528 SEF is a common feature used in monitoring the depth of anesthesia and stages of sleep using EEG
529 (Iranmanesh and Rodriguez-Villegas, 2017). It measures the frequency which covers X percent of the PSD.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

X is usually set between 75% to 95%. Here we set X to 95%. Therefore, this reflects the maximum frequency observed in a signal which covers 95% of a signal power spectrum.

Multi-valued features

The main hypothesis of the present study is that, we can potentially obtain more information about object categories as well as behavior if we take into account the temporal variability of signal samples within the analysis window rather than just averaging them as in the conventional decoding studies. Therefore, we extracted other features, which provide more than one value per analysis window (i.e. 50 ms for the time-resolved analysis or 1000 ms for the whole-trial analysis), so that we can select the most informative one for decoding. The reason for selecting only one feature per time window is to be able to directly compare the results with those obtained with single-valued features explained above. We also included the original signal samples as the last feature, so that we know how much (if at all) our feature extraction and selection helps.

Inter-electrode correlation (Cros Cor)

Following up on recent studies, which have successfully used this feature in decoding object category information from brain activations (Majima et al., 2014; Karimi-Rouzbahani et al., 2017a), we extracted inter-electrode correlation to measure the similarity between pairs of signals, here, from coming from pairs of electrodes. Through the concept of correlation, this feature can detect any subtle co-activation or co-deactivation of neural populations from district pairs of electrodes. Although closer electrodes tend to provide more similar (and therefore correlated) activation, compared to further electrodes, the inter-electrode correlation can capture correlations which are functionally relevant and are not explained by the distance (Karimi-Rouzbahani et al., 2017a). Please note that as correlation is an amplitude-independent variable, this feature detects the similarities across pairs of signals, which cannot be determined by the mean-based features from individual signals. It is calculated as:

$$R_{xy} = \frac{1}{N\sigma_x\sigma_y}\sum_{t=1}^{N}(x_t - \bar{x})(y_t - \bar{y}) \quad (19)$$

where x and y refer to the signals obtained from electrodes x and y, respectively. We calculated the cross-correlation between every given electrode and all the other electrodes before finally averaging them to obtain a single value per electrode. Therefore, for the datasets with 31 (i.e. Datasets 1 and 2) and 128 (i.e. Dataset 3) electrodes, we obtained 30 and 127 inter-electrode correlations per electrode.

Wavelet transform (Wavelet)

Recent studies have shown remarkable success in decoding of object categories using the Wavelet transformation of the EEG time series (Taghizadeh-Sarabi et al., 2015; Torabi et al., 2017). Considering the time-dependent nature of ERP signals, Wavelet transform seems to be a very reasonable choice as it provides a time-frequency representation of signal components determining the primary frequency components of a specific signal and their temporal location in the time series. To do that, the transformation passes the signal time series through digital filters (Guo et al., 2009; equation (20)), using

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the convolution operator, each of which are adjusted to extract a specific frequency (scale) at a specific time as in (20):

$$y_n = (x * g) = \sum_{k=-\infty}^{+\infty} x_k g_{n-k} \quad (20)$$

where g is the digital filter and $*$ is the convolution operator. This filtering procedure is repeated for several rounds (levels) filtering low- (approximations) and high-frequency (details) components of the signal to provide more fine-grained information into the constituent components of the signal. This can lead to coefficients which can potentially discriminate signals evoked by different conditions. As in a previous study (Taghizadeh-Sarabi et al., 2015), and to make the number of Wavelet features comparable in number to the signal samples (which were 1000 after the stimulus onset), we used detail coefficients at five levels $D1, \dots, D5$ as well as the approximate coefficients at level 5, $A5$. This led to 1015 features in the whole-trial and 57 in the 50 ms sliding time windows, respectively. We used the '*Symlet2*' basis function for our Wavelet transformations as implemented in Matlab.

Hilbert transform (Hilb Amp and Hilb Phs)

There were two motivations for using Hilbert transform in the current study. First, this transformation technique, which provides amplitude and phase information of the signal upon the transformation, has recently shown success in decoding visual letter information from ERP signals (Wang et al., 2018). Second, although previous systematic comparison of Hilbert and Wavelet transforms has shown very minor differences in evaluating neuronal synchrony (Le Van Quyen et al., 2001), it is still unclear which method can detect category-relevant information from the nonstationary ERP components more effectively. Specifically, the phase component of the Hilbert transform can qualitatively provide the spatial information which you obtain from the Wavelet transformation. In signal processing, Hilbert transform is described as a mapping function that takes a function $x(t)$ of a real variable, and using convolution with the function, $\frac{1}{\pi t}$, produces another function of a real variable $H(u)(t)$ as:

$$H(u)(t) = \frac{1}{n} \int_{-\infty}^{+\infty} \frac{u(\tau)}{t-\tau} d\tau \quad (21)$$

where $H(u)(t)$ is a frequency-domain representation of the signal, which has simply shifted all the components of the input signal by $\frac{\pi}{2}$. In the current study, Hilbert transform was applied on every trial (1000 samples) or 50 ms sliding time windows, which produced one amplitude and one phase component per sample, so 1000 and 50 for the whole-trial and 50 ms sliding time windows, respectively. We used the amplitude and phase components separately to discriminate object categories in the decoding analyses.

Signal samples (Samples)

We also used the post-stimulus signal samples (i.e. 1000 or 50 samples for the whole-trial and sliding time windows, respectively) to decode object category information without any feature extraction. This allowed us to compare the information content of the extracted features with the original signal samples to see if the former provided any extra information.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Multivariate decoding

We used multivariate decoding to extract information about object categories from the EEG datasets. Basically, multivariate decoding, which has been dominating neuroimaging studies recently (Haynes et al., 2015; Grootswagers et al., 2017; Hebart and Baker, 2018), utilizes within- and cross-condition similarity/distance to determine the amount of neural information when contrasting those conditions. We used linear discriminant analysis (LDA) classifiers to measure the information content across all possible pairs of object categories in each dataset throughout our multivariate decoding. Specifically, we trained and tested the classifier on e.g. animal vs. car, animal vs. face, animal vs. plane, car vs. plane, face vs. car and plane vs. face categories, then averaged the 6 decoding results and reported them for each participant. The decoding for each of the categories, is reported in the original references of the datasets. The LDA classifier has been shown to be robust when decoding object categories from M/EEG (Grootswagers et al., 2017; Grootswagers et al., 2019), has provided higher decoding accuracies than Euclidean distance and Correlation based decoding methods (Carlson et al., 2013) and was around 30 times faster to train in our initial analyses compared to the more complex classifier of support-vector machine (SVM). We ran our initial analysis and found similar results for the LDA and SVM, so used LDA to save the time. We used a 10-fold cross-validation procedure in which we trained the classifier on 90% of the data and tested it on the left-out 10% of the data, repeating the procedure 10 times until all trials from the pair of categories participate once in the training and once in the testing of the classifiers. We repeated the decoding across all possible pairs of categories within each dataset, which were 6, 6 and 15 pairs for Datasets 1, 2 and 3, which consisted of 4, 4 and 6 object categories, respectively. Finally, we averaged the results across all combinations and reported them as the average decoding for each participant.

In the whole-trial analyses, we extracted the above-mentioned features from the 1000 data samples after the stimulus onset (i.e. from 1 to 1000 ms). In the time-resolved analyses, we extracted the features from 50 ms sliding time windows in steps of 5 ms across the time course of the trial (-200 to 1000 ms relative to the stimulus onset time). Therefore, in time-resolved analyses, the decoding results at each time point reflect the data for the 50 ms window around the time point, from -25 to +24 ms relative to the time point. Time-resolved analyses allowed us to evaluate the evolution of object category information across time as captured by different features.

Dimensionality reduction

The multi-valued features explained above resulted in more than a single feature value per trial per sliding time window (e.g. inter-electrode correlation, wavelet, Hilbert amplitude and phase and signal samples), which could provide higher decoding values compared to the decoding values obtained from single-valued features merely because of including a higher number of features. Moreover, when the features outnumber the observations (i.e. trials here), the classification algorithm can over-fit to the data (Duda et al., 2012). Therefore, to obtain comparable decoding accuracies across single-valued and multi-valued features and to avoid potential over-fitting of classifier to the data we used principle component analysis (PCA) to reduce the dimension of the data for multi-valued features. Accordingly,

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

we reduced the number of the values in the multi-valued features to ***one*** per electrode per time window per trial, which equaled the number of values for the single-valued features. Specifically, the data matrix before dimension reduction, had a dimension of n rows by $e \times f$ columns where n, e and f were the number of trials (from all categories), the number of electrodes and the number of values obtained from any given feature (concatenated in columns), respectively. As $f = 1$ for the single-valued features, for the multi-valued features, we only retained the e most informative columns that corresponded to the e eigen values with highest variance and removed the other columns using PCA. Therefore, we reduced the dimension of the data matrix to $n \times e$ which was equal between single- and multi-valued features and used the resulting data matrix for multivariate decoding. For example, before the dimension reduction, the data matrix for Dataset 3, and the inter-electrode correlation (Cros Cor) feature had a dimension of 864 rows (corresponding to all correct trials) by 16256 columns (i.e. 128 electrodes by 127 inter-electrode correlation values). However, after the dimension reduction procedure, it had a dimension of 864 rows (corresponding to all correct trials) by 128 columns (i.e. a combination of electrodes and correlation values with maximum variance across the four categories). This means that, for the multi-valued features, we only retained the most informative value of the extracted values from each feature and electrode. In other words, we sub-sampled in both the space (across electrodes) and time (across time window). To avoid potential leakage of information from testing to training (Pulini et al., 2019), we applied the PCA algorithm on the training data (folds) only and used the training PCA parameters (i.e. eigen vectors and means) for both training and testing sets for dimension reduction in each cross-validation run separately. We only applied the dimension-reduction procedure on the multi-valued features. Note that, we did not reduce the dimension of the neural space (columns in the dimension-reduced data matrix) to below the number of electrodes "e" (as was done in Hatamimajoumerd et al., 2019) as we were interested in qualitatively comparing our results with the vast literature currently using multivariate decoding with all sensors (Grootswagers et al., 2017; Karimi-Rouzbahani et al., 2017a; Hebart and Baker 2017). Also, we did not aim at finding more than one feature per trial, as we wanted to compare the results of multi-valued features with those of single-valued features, which only had a single value per trial.

Statistical analyses

Bayes factor analysis

As in our previous studies (Grootswagers et al., 2019; Robinson et al., 2019), to determine the evidence for the null and the alternative hypotheses, we used Bayes analyses as implemented by Bart Krekelberg based on Rouder et al. (2012). We used standard rules of thumb for interpreting levels of evidence (Lee and Wagenmakers, 2014; Dienes, 2014): Bayes factors of >10 and <1/10 were interpreted as strong evidence for the alternative and null hypotheses, respectively, and >3 and <1/3 were interpreted as moderate evidence for the alternative and null hypotheses, respectively. We considered the Bayes factors which fell between 3 and 1/3 as suggesting insufficient evidence either way.

In the whole-trial decoding analyses, we asked whether there was a difference between the decoding values obtained from all possible pairs of features and also across frequency bands within every feature. Accordingly, we performed the Bayes factor analysis and calculated the Bayes factors as the probability of the data under alternative (i.e. difference) relative to the null (i.e. no difference) hypothesis between

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

all possible pairs of features and also across frequency bands within every feature and dataset separately. The same procedure was used to evaluate evidence for difference (i.e. alternative hypothesis) or no difference (i.e. null hypothesis) in the maximum and average decoding accuracies, the time of maximum and above-chance decoding accuracies across features for each dataset separately.

We also evaluated evidence for the alternative of above-chance decoding accuracy vs. the null hypothesis of no difference from chance. For that purpose, we performed Bayes factor analysis between the distribution of actual accuracies obtained and a set of 1000 accuracies obtained from random permutation of class labels across the same pair of conditions (null distribution) on every time point (or only once for the whole-trial analysis), for each feature and dataset separately. No correction for multiple comparisons have been performed when using Bayes factors as they are much more conservative than frequentist analysis in providing false claims with confidence (Gelman and Tuerlinckx, 2000; Gelman et al., 2012).

The priors for all Bayes factor analyses were determined based on Jeffrey-Zellner-Siow priors (Jeffreys, 1961; Zellner and Siow, 1980) which are from the Cauchy distribution based on the effect size that is initially calculated in the algorithm (Rouder et al., 2012). The priors are data-driven and have been shown to be invariant with respect to linear transformations of measurement units (Rouder et al., 2012), which reduces the chance of being biased towards the null or alternative hypotheses.

Random permutation testing

To evaluate the significance of correlations between decoding accuracies and behavioral reaction times, we calculated the percentage of the actual correlations that were higher (if positive) or lower (if negative) than a set of 1000 randomly generated correlations. These random correlations were obtained by randomizing the order of participants' data in the behavioral reaction time vector (null distribution) on every time point, for each feature separately. The correlation was considered significant if surpassed 95% of the randomly generated correlations in the null distribution in either positive or negative directions ($p < 0.05$) and the p-values were corrected for multiple comparisons across time using Matlab mafdr function which works based on fix rejection region (Storey, 2002).

Results

Which features of the recorded signals are most informative about object categories?

To answer the first question of this study, we compared decoding accuracies in the whole-trial time span across all features and for each dataset separately (Figure 2, black bars). This is a more conventional pass on the data, which incorporates the whole trial time span, and gives insight about the information content of the recorded signals for decoding object category information (Kaneshiro et al., 2015). For direct comparison of the decoding accuracies across features and frequency bands, see the bar plots in Figure 2 and check their corresponding evidence for difference (the alternative hypothesis) and/or no difference (the null hypothesis) in Supplementary Figure 1A and Supplementary Figure 1B.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Not all features could provide strong (BF>10) evidence for above-chance decoding (Figure 2). There was strong evidence (BF>10) that the four ERP components (i.e. P1, N1, P2a and P2b), and a few multi-valued features consisting of Wavelet, Hilb Phs and Samples provided above-chance decoding. In addition, consistently across the three datasets, there was moderate or strong evidence (BF>3; Supplementary Figure 1A; black boxes) that most of the ERP components (i.e. N1, P2a and P2b) and multi-valued features (i.e. Wavelet, Hilb Phs and Samples) provided higher decoding accuracies compared to the rest of the features. This result is consistent with our prediction that features which are spatially and/or temporally specific (and targeted) can potentially detect more information about object categories as category processing is a spatially and temporally specific neural process generally observed in the window from 50 to 300 ms post-stimulus onset (Karimi-Rouzbahani et al., 2017b; which overlap with the ERP features) and mainly in around the occipito-temporal, occipito-parietal and frontal areas (Vaziri-Pashkam and Xu, 2017; Karimi-Rouzbahani 2018; these electrodes were probably selected through the PCA procedure for the multi-valued features). As explained in the methods, for keeping the dimension of the data identical to the single-valued features, the multi-valued features underwent a PCA-based sample selection process, which selected the most informative samples/features from across electrodes (space) and samples (time) in the analysis window. Accordingly, the reason for comparable decoding accuracies obtained for the Samples feature and the ERP components across the three datasets, might be that, the signal samples used in the Samples may have come roughly from nearby time windows of the trial as the ERP components, through the PCA procedure. The higher decoding values for Dataset 2 compared to the other datasets can be potentially explained by the active object detection task and longer image presentation time. In summary, the spatial and/or temporally specific ERP components (i.e. N1, P2a and P2b) and the multi-valued features (i.e. Wavelet, Hilb Phs and Samples) were most informative about object categories as predicted.

Following evidence from previous studies reporting pronounced information in specific frequency sub-bands such as Theta (Behroozi et al., 2016; Bastos et al., 2015), we also compared the decoding accuracies across different frequency sub-bands to see what frequency band(s) provided the most information. Specifically, we evaluated the information contents of features in the well-known EEG frequency bands of Delta (0.5-4 Hz), Theta (4-8 Hz), Alpha (8-12 Hz), Beta (12-16 Hz) and Gamma (>25 Hz) against the broad-band signals, which covered the whole available spectrum after pre-processing (>0.03Hz, Figure 2). We did not perform the decoding for the frequency-domain features (except Signal pw), as they would be meaningless in limited frequency bands by definition. For Signal Pw, however, as in previous studies (Rupp et al., 2017; Miyakawa et al., 2018; Behroozi et al., 2016) we calculated it in the time domain and performed the decoding. Showing comparable patterns of decoding to the broad-band results, the decoding across features especially in the mid-frequency bands of Theta, Alpha and Beta (Figure 2) showed moderate to strong evidence (BF>3; Supplementary Figure 1B; red boxes) that ERP components and multi-valued features were the most informative about object categories.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

766

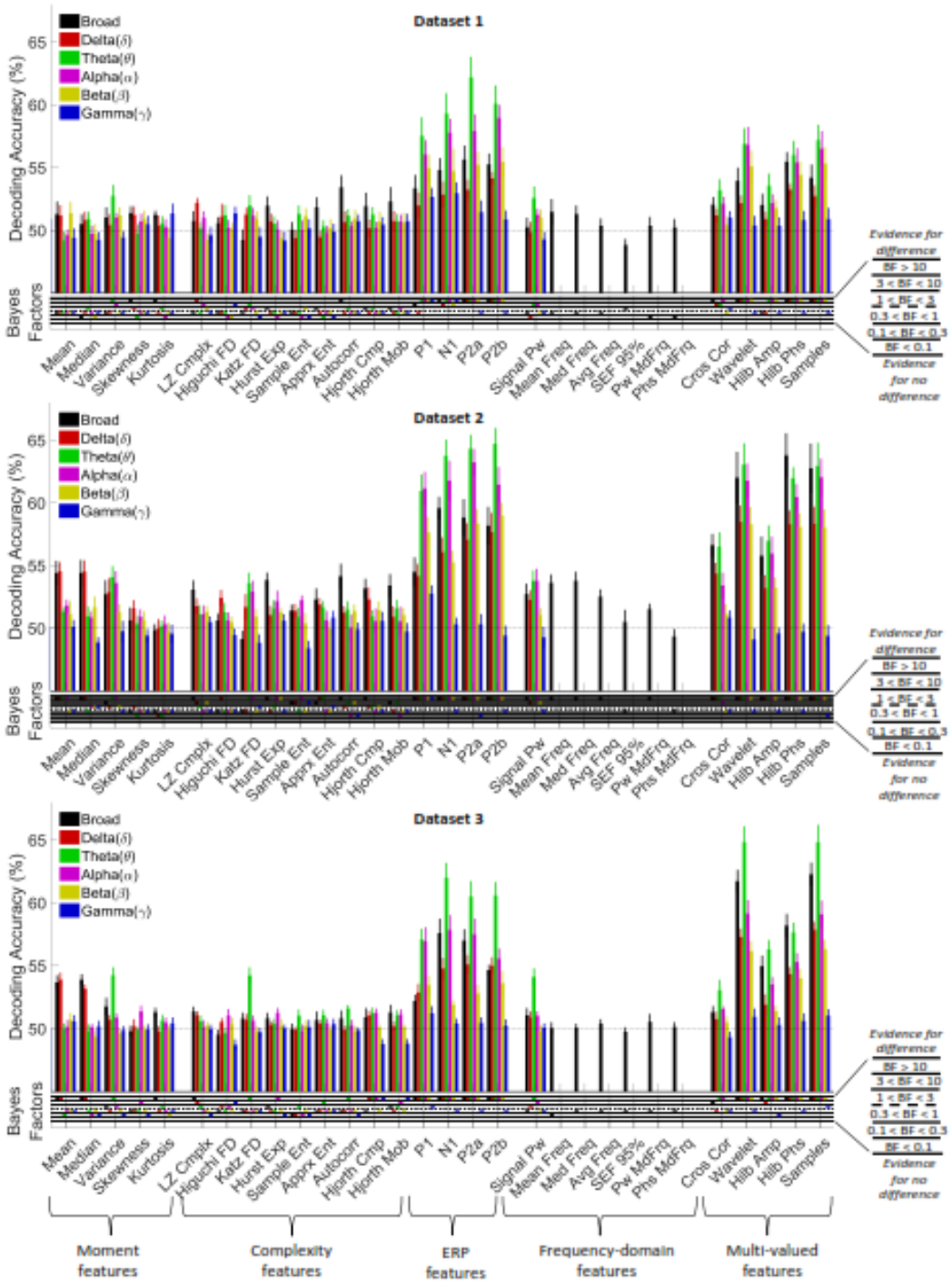


Figure 2. Whole-trial decoding of object categories from the three datasets using 30 features on different frequency bands (for Bayesian evidence analyses see Supplementary Figure 1). Decoding of category information using the 30 features in the 6 frequency bands. The black horizontal dashed lines on the top panel refer to chance-level decoding. Thick bars show the average decoding across participants (error bars Standard Error across participants). Bayes Factors are shown in the bottom panel of each graph: Filled circles show moderate/strong evidence for either hypothesis and empty circles indicate insufficient evidence. They show the results of Bayes factor analysis when evaluating the difference from chance-level decoding.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In Datasets 1 and 2, where the upper cut-off frequency of filtering was 200 Hz rather than 50 Hz as in Dataset 3, the broad-band decoding accuracies were on average higher for many of complexity features (consisting of Hurst Exp, Apprx Ent, Autocorr, Hjorth Cmp and Mob). The reason is likely that the inclusion of more frequency components (i.e. higher frequency fluctuations) leads to the survival of meaningful repetitive (e.g. higher frequency harmonics) and complex patterns in the time series, which are detected by the features which are sensitive to repetitive sinusoidal (i.e. Hurst Exp, Autocorr, Hjorth Cmp and Mob) and complex (i.e. Apprx Ent) patterns.

Interestingly, however, the most informative features in the datasets (i.e. ERP and multi-valued features) showed pronounced information in the Theta band over other sub-bands, with strong evidence BF>10 for the advantage of Theta over broad-band decoding for P1 and P2b in both datasets. There was also moderate (3<BF<10) or strong (BF>10) evidence that the Gamma band accuracies were lowest among other frequency bands for many features including ERP components and multi-valued features (Supplementary Figure 1B; red boxes).

Together, frequency-resolved results showed that broad-band signals can provide more information than its sub-bands when using features sensitive to repetitive and complex patterns. Importantly, however, ERP components and multi-valued features, which showed the highest information about object categories in the feature set, showed the greatest information in the Theta band, even more than which could be achieved using broad-band signals. The coding of category information in the Theta band for the informative features suggests that this information is dominantly processed by the feed-forward visual mechanisms of the brain as suggested previously to reflect in the Theta band (Bastos et al., 2015).

What is the temporal dynamics of object category decoding when using distinct features of brain activations?

To answer the second question, we adopted a time-resolved decoding procedure for each feature and dataset separately (Grootswagers et al., 2017; see *Methods*). In this method, we repeated the decoding of category information in 50 ms sliding time windows, using a 5 ms step size (45 ms overlap between consecutive time windows) across the time course of the trials (Figure 3). For a justification of choosing the 50ms time window see Supplementary Text 2 and Supplementary Figure 2A. This analysis provided the temporal profile of information encoding as revealed by different features. Note that, by definition, we do not have the time-resolved decoding results of the time-specific features of ERP components (i.e. P1, N2, P2a and P2b).

The advantage of the Theta-band to broad-band frequency range, observed for the whole-trial time window (Figure 2), was also observed for the time-resolved decoding when using the Wavelet feature (especially for Dataset 2 which showed the effect across many time points; moderate to strong (BF>3) evidence starting to appear after 50 ms post-stimulus onset), but not the Mean feature (Supplementary Figure 1A). However, we used broad-band signals for the time-resolved decoding analyses, because we not only aimed at comparing the information content of the features, but we were also interested in comparing our results with the previous category decoding studies which used the broad-band

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

frequency range. Note that we do not aim to maximize the decoding values in this study, but rather aim to compare the decoding dynamics and their correlations to behavior across a wide range of relevant features.

While no pairs of features provided identical patterns of decoding in any of the three datasets, for all features there was moderate (3<BF<10) or strong (BF>10) evidence for difference from chance-level decoding at some time points/windows in the three datasets (Figure 3). This means that, all features, including the complexity and frequency-domain features, which have been respectively suggested to need longer time windows (Procaccia, 2000) and stationary signals, could be successfully used to decode object category information from evoked ERP signals. We did not plot the variance of decoding accuracies across participants, as it would make the figures cluttered (see Supplementary Figure 2 for decoding results using Mean and Wavelet features along with their variance across participants to get a feeling of the variance for other features). Interestingly, while the Mean and Median features generally showed lower decoding accuracies in the whole-trial analyses (Figure 2), they provided comparable or even higher decoding accuracies than several of the multi-valued features in the time-resolved analysis (Figure 3). This is because Mean and Median of the signals lost most of their information (as a result of baseline removal in preprocessing) when averaged across a whole-trial time span (i.e. 1000 samples; Figure 2). While the decoding curves in Dataset 1 showed two comparable early peaks at around 180 and 300 ms for several features such as Mean and Median, most other highly informative features (e.g. Variance, Wavelet and Samples) showed only one peak at around 180 ms. There was only one peak for Mean, Median and frequency-domain and multi-valued features in Datasets 2 and 3 and two dominant peaks for the other features including the complexity and moment features. This discrepancy across datasets can be explained by many parameters which differ across them. For example, image presentation onset and offset can lead to bumps in decoding patterns (Carlson et al., 2013) which was the case for Dataset 1 (50 ms) happening in close temporal succession, but further apart for the other datasets. Consistently across the three datasets, for all features, there was moderate (3<BF<10) or strong (BF>10) evidence for above-chance decoding starting to appear from around 80 ms. The decoding curves returned back to the chance-level around than 500 ms, particularly for the most informative features such as Mean, Median and multi-valued features in Dataset 1. The same features remained above chance (BF>3) up until 550 ms (Dataset 2) or even later than 800 ms (Dataset 3). This difference can be potentially because of the longer stimulus presentation time in Datasets 2 and 3, which provided stronger sensory input for processing of category information (Grootswagers et al., 2019).

To quantitatively compare the decoding patterns across features, we calculated several time and amplitude-related parameters from the decoding curves (Figure 4). These parameters consist of maximum and average decoding accuracies, the time to the first above-chance (BF>3) and maximum decoding. All parameters were calculated in the post-stimulus window (0 to 1000 ms), and have previously provided important implications for studying the dynamics of object recognition in the human brain (Isik et al., 2014). There was strong evidence that all features from all three datasets had above-chance maximum decoding accuracy (Figure 4A; colored dots in the bottom Bayes Factors' panels). The features of Mean, Median, Wavelet and Samples obtained the highest maximum and average decoding accuracies among other features (Figure 4A; black boxes).

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

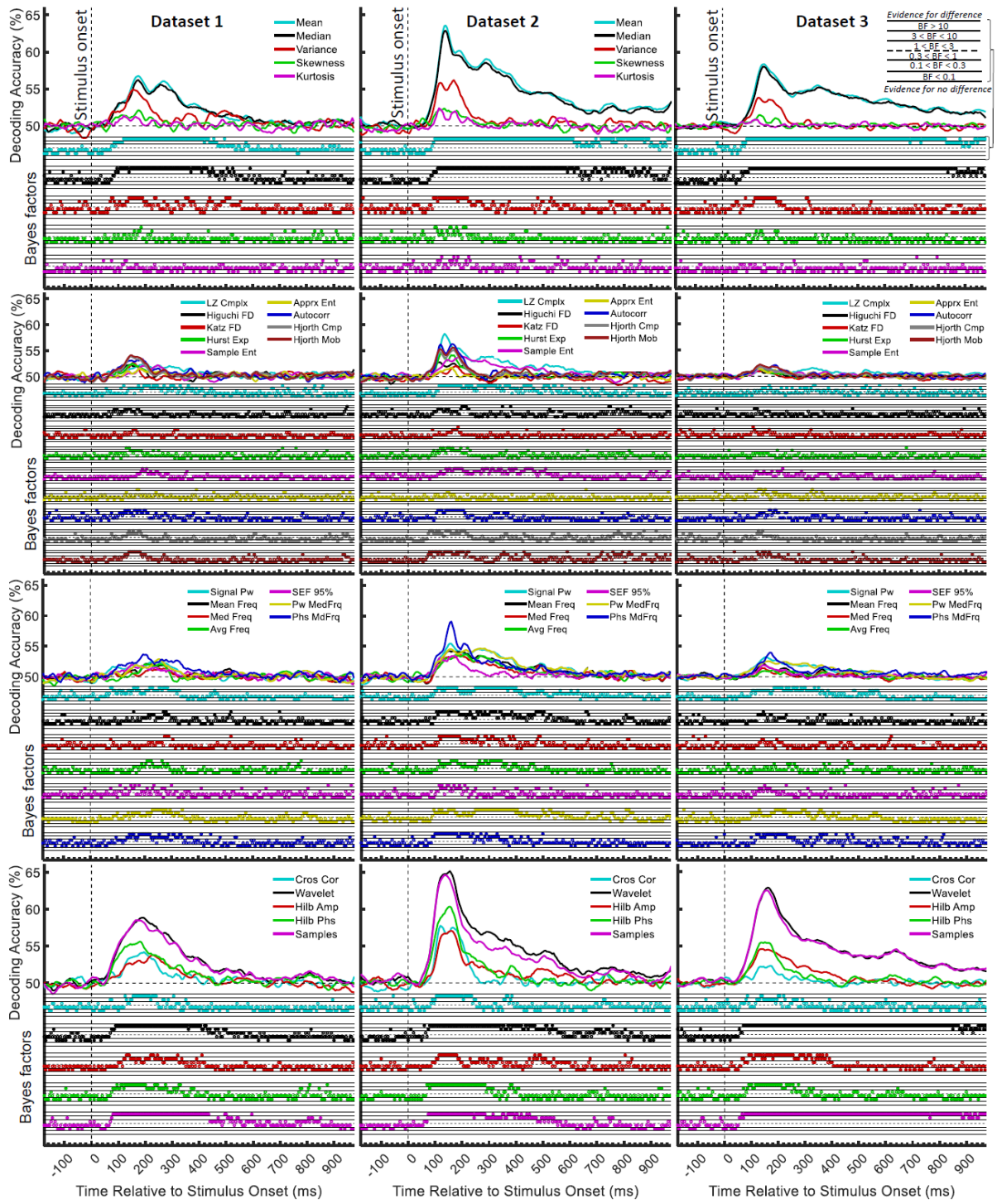


Figure 3. Time-resolved decoding of object categories from the three datasets using 26 features and Bayesian evidence analyses. Each row shows the results of one type of feature (i.e. moment, complexity, frequency-domain and multi-valued features from top to bottom, respectively). Curves show the average decoding across participants. Each column shows the results for one dataset. Top section in each panel shows the decoding accuracies across time and the bottom section shows the Bayes factor evidence for the difference of the decoding accuracy compared to chance-level decoding. The horizontal dashed lines on the top panel refer to chance-level decoding. Filled circles in the Bayes Factors show moderate/strong evidence for either difference or no difference from chance-level decoding and empty circles indicate insufficient evidence for either hypotheses.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

There was strong (BF>10) evidence that Wavelet had the highest maximum decoding accuracy compared to all other features (except Samples in all three datasets and Mean and Median in Dataset 2). Although there was strong (BF>10) evidence that Mean, Median, Wavelet and Samples provided above-chance average accuracy across the three datasets, there was insufficient (0.3<BF<1) evidence for above-chance average accuracy for several other features (e.g. Katz FD).

The temporal dynamics of different features seem to reflect a similar decoding pattern in the sense that the most informative features can lead to both a higher maximum decoding and a more sustained decoding pattern along the trial and vice versa. This can suggest a general advantage for the more vs. less informative features which is reflected both in their maxima as well as their sustained decoding patterns. Alternatively, it can be the case that features with higher maxima have lower average decoding across the trial or vice versa. This, on the other hand, suggests that different features detect an equal amount of information but represent it either in their maximum or average decoding accuracy. A second alternative can be that there is no relationship between the peaks and the average decoding across features, reflecting potentially different pieces of neural code that each feature is sensitive to. To test this question, we calculated the correlation between the average and maximum decoding values for all features, which showed highly correlated results ($r > 0.9$; $p < 0.01$; Supplementary Text 3 and Supplementary Figure 3A). This suggests that, all features followed a generally similar pattern of decoding with more informative features providing higher decoding maxima and a more sustained level of information decoding.

In terms of the temporal pattern, the time to the maximum decoding was quite constrained as many of the features showed a maximum decoding between 150 to 220 ms post-stimulus onset, and there was no clear trend towards any classes of features (Figure 4C). This is consistent with many decoding studies showing the temporal dynamics of category processing in the brain (Isik et al., 2013; Cichy et al., 2014). There was moderate (3<BF<10) or strong (BF>10) evidence that Wavelet, Hilb Phs and Samples were among the earliest features to reach their peaks in Datasets 2 and 3. The time of first above-chance (BF>3) decoding did not show the priority of any specific class of features relative to others (Figure 4D). There was moderate (3<BF<10) or strong (BF>10) evidence that Mean, Median, Wavelet and Samples showed an earlier appearance of above-chance decoding compared to Sample Ent, Avg Freq, Med Freq, SEF 95%, Cros Cor and Katz FD in Dataset 1. There was moderate (3<BF<10) evidence that Hjorth Cmp and Mob, were among the earliest features to show decoding in Dataset 2.

There has been no consensus yet about whether the time of the maximum or the first above-chance decoding reflects the speed of category processing in the brain (Grootswagers et al., 2017; Ritchie et al., 2015). Hence, we calculated the correlation of these temporal parameters across features to see if they both possibly reflect the dynamics of the same processing mechanism in the brain. The time of first above-chance and maximum decoding correlated in Datasets 2 and 3 but not Dataset 1 ($r=0.67$, $r=0.51$ and $r=0.02$ respectively for Datasets 1, 2 and 3; Supplementary Text 3 and Supplementary Figure 3A). Lack of significant correlation for Dataset 1 can be explained by the lower decoding values in Dataset 1 compared to the other datasets making the correlations noisier.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

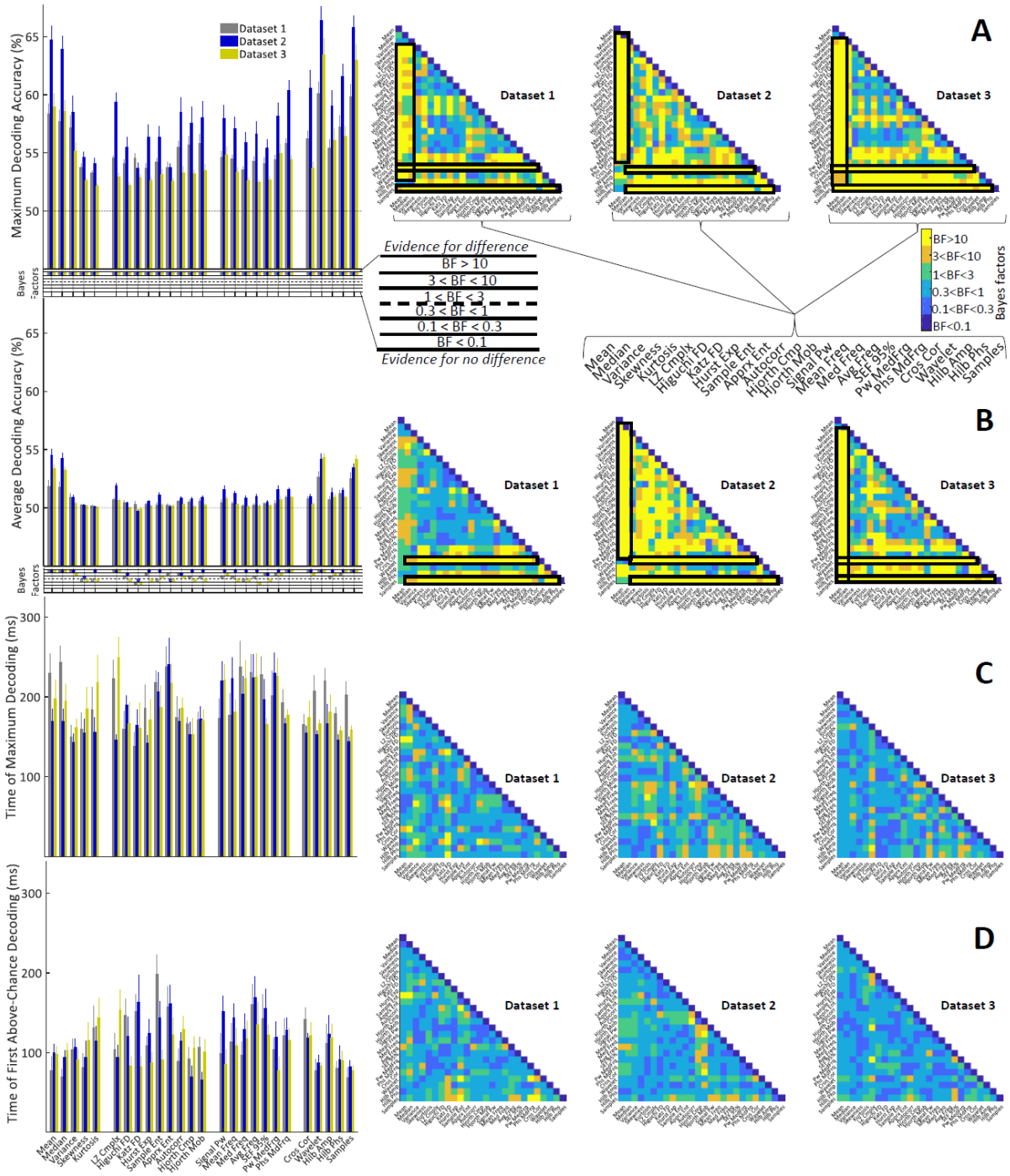


Figure 4. Timing and amplitude parameters extracted from the time-resolved accuracies of each feature and each dataset and their Bayesian evidence analyses. (A-D) Left: the maximum and average decoding accuracies, the time of maximum and the first above-chance decoding. Thick bars show the average across participants (error bars Standard Error across participants). Bottom section on A and B show the Bayes factor evidence for the difference of the decoding accuracy compared to chance-level decoding; Right: matrices compare the right parameters obtained from different features. Different levels of evidence for existing difference (moderate 3<BF<10, Orange; strong BF>10, Yellow), no difference (moderate 0.1<BF<0.3, light blue; strong BF<0.1, dark blue) or insufficient evidence (1<BF<3 green; 0.3<BF<1 Cyan) for either hypotheses. Black and red boxes show moderate or strong evidence for higher decoding values for specific features compared other sets of features as explained in the text. The horizontal dashed lines on the left panels of (A) and (B) refer to chance-level decoding. Filled circles in the Bayes Factors show moderate/strong evidence for either hypothesis and empty circles indicate insufficient evidence.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Therefore, features that reached their above-chance decoding earlier also reached their maximum decoding earlier leading to the suggestion that they both reflect the temporal dynamics of the same cognitive processes with some delay.

In conclusion, the features of Mean, Wavelet and Samples, not only provided the maximum and average amount of category-related information in the brain, but they also were among the earliest features that provided signatures of object category information in the brain at around 100 ms post stimulus onset. Results suggest that our prediction that reliance on temporally and spatially specific features improves the accuracy in reading out the dynamics of object category processing in the brain. Moreover, it shows that object category information peaked at around 180 ms after the stimulus onset, irrespective of the feature used for decoding of information, which further constrains the temporal window of category processing in the object processing literature to the span of 100 to 200 ms.

Which features of brain activations explain behavioral recognition performance?

Although the results above showed the advantage of specific features to others in providing the highest and the earliest signatures of object category processing, this can all be a by-product of the actual neural processing that underlies human object recognition behavior (Vidaurre et al., 2019). Therefore, to see if the decoding patterns provided by these features can explain behavioral performance, we calculated the *correlation* between the decoding accuracies obtained from each feature and the reaction times of participants at every time point around the stimulus onset (Ritchie et al., 2015; Karimi-Rouzbahani et al., 2020a). Participants' reaction time in object recognition has been previously shown to be predictable from decoding accuracy (Ritchie et al., 2015). We expected to observe negative correlation values between decoding accuracies and participants' reaction time in the post-stimulus span, meaning that greater separability across neural representations of object categories would lead to/correlate with faster recognition of the corresponding categories. For this analysis, we only used Dataset 2 as it was the only dataset with an active object detection task; therefore relevant reaction times were available. To calculate the correlations, we generated a 10-dimensional vector of neural decoding accuracies at every time point and a 10-dimensional vector for behavioral reaction times obtained from the group of 10 participants and correlated the two vectors across the time course of trials using Spearman's rank-order correlation (Cichy et al., 2014). This resulted in a single correlation value for each time point for the group of 10 participants.

All features, except Katz FD, showed negative trends after the stimulus onset (Figure 5A). The negative correlations did not remain significant for long windows of the time, except for multi-valued features, which showed more pronounced negative correlations that remained significant for larger spans of time. Correlations also showed larger negative peaks (generally < -0.5) for multi-valued features especially Wavelet, compared to other features (generally > -0.5). Specifically, while higher-order moment features (i.e. Variance, Skewness and Kurtosis) as well as many complexity features showed earlier peaks of negative correlations at around 150 ms, Mean, Median, frequency-domain features and multi-valued features showed later negative peaks after 300 ms.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

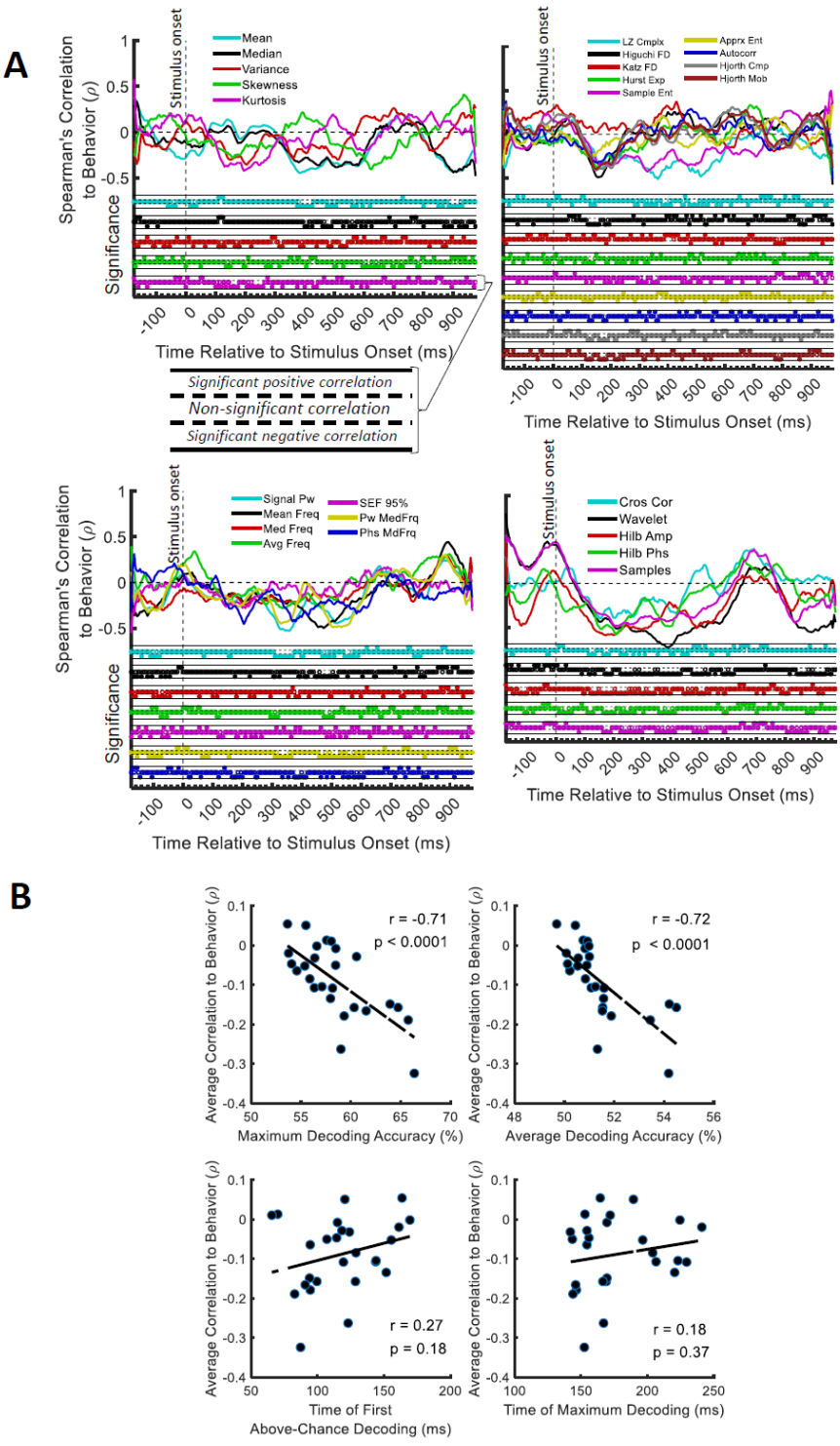


Figure 5. Correlation between the decoding accuracies and behavioral reaction time for Dataset 2 (other datasets did not have an object recognition/detection task). (A) Top section in each panel shows the (Spearman's) correlation coefficient obtained from correlating the decoding values and the reaction times for each feature separately. Correlation curves were obtained from the data of all participants. Bottom section shows positively or negatively significant (P<0.05; filled circles) or non-significant (p>0.05; empty circles) correlations as evaluated by random permutation of the variables in correlation. (B) Correlation between each of the amplitude and timing parameters of time-resolved decoding (i.e. maximum and average decoding accuracy and time of first and maximum decoding) with the average time-resolved correlations calculated from (A) for the set of N=26 features. The slant line shows the best linear fit to the distribution of the data.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The multi-valued features (especially Wavelet), Mean, Median, which dominated other features in terms of decoding accuracy (Figures 2-4), showed also the largest window of negative correlations to behavior.

Together, it seems that features with the highest decoding accuracy were also better at explaining behavioral performance. This was reflected in their longer time windows and larger negative peaks of correlations. To quantitatively assess this hypothesis, we calculated the correlation between different parameters of the decoding accuracy curves (Figure 4) and the average correlation-to-behavior obtained from the same features (Figure 5A). For the decoding accuracy parameters, we used the average and maximum decoding accuracies which are relevant for our hypothesis; and the time to the first above-chance and maximum decoding accuracies which were irrelevant to our hypothesis, and were used as control for comparison. To obtain the average correlation-to-behavior, we simply averaged the 'correlation to behavior' in the post-stimulus time span (from Figure 5A). Results showed that (Figure 5B), while the temporal parameters of 'time of first above-chance' and 'maximum' decoding (which were our control irrelevant parameters) failed to predict the level of correlation to behavior ($r=0.27$, $p=0.27$, and $r=0.18$, $p=0.37$, respectively), the amplitude parameters of 'maximum' and 'average' decoding accuracies of features significantly ($r=-0.71$ and $r=-0.72$ respectively, with $p<0.0001$; Pearson's correlation) predicted the average correlation between the features' decoding accuracies and the behavioral performance.

This result showed that features which provided more information about object categories (i.e. provided higher decoding accuracy), also predicted behavioral performance better. Therefore, the extraction of temporally and spatially specific information from brain activations, which has been achieved through features such as PCA-based Wavelet transform, can lead to more accurate prediction of behavioral performance in object recognition. This suggests that, finding the category-related informative samples and features can not only provide us with a more accurate view of the dynamics of neural processing in the brain, but it will also improve our ability in predicting the object recognition performance. This is not a trivial result, which might incorrectly be expected to be obtained by a higher decoding value for the more informative features leading to a higher correlation to behavior. This is because, the mathematical calculation of the 'correlation coefficient' standardizes the variables, so that the changes in scale/amplitude will not affect its value.

Discussion

None of the methods or algorithms discussed in this paper are new; the novel contribution of this work is an empirical and quantitative comparison of a large set of statistical and mathematical EEG features which have been suggested to provide object category information in multivariate decoding. In the whole-trial analyses, we showed that the features, which were temporally and spatially specific (ERP components and multi-valued features), could provide more information about object categories than non-specific features. Results also showed that, the Theta frequency band provided more information about object categories than the conventionally used broad-band activity, whether decoding the whole trial or the time-resolved data. We showed that multi-valued features such as Samples and Wavelet

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

coefficients can provided information up and above the generally-used feature of Mean brain activity, which overlooks the temporal codes within the sliding time windows across the signal time series. We also showed that the Wavelet feature, not only provided the highest decoding accuracy, but it also explained the behavioral object recognition performance better than all the other features. As the mentioned results were generally consistent across three datasets, which had been collected across a wide range of variations, the generalizability of the results are far more than previous studies. Our results suggest that improving the performance in decoding could be a path for improving the behavioral explanatory power in multivariate decoding, which can help filling the gap between neuroimaging and behavior.

This study provides new insights for the fields of cognitive neuroscience and BCI at the same time. In the past two decades, many neuroimaging studies in cognitive neuroscience have tried to provide insights into the spatiotemporal dynamics of object category processing in the human brain (Haxby et al., 2001; Contini et al., 2017; Carlson et al., 2013). This study does not provide any information about the spatial location of object category information processing in the brain. From the temporal viewpoint, however, it suggests that to access more variance in the neural code, we need to take the temporal and spatial information of the neural activations into account when running multivariate decoding. This aligns with the recent shift towards taking into account the temporal variability of trials when decoding visual information from high-temporal resolution methods such as MEG (Vidaurre et al., 2019). Importantly, our results showed that, multi-valued features such as original signal Samples and Wavelet coefficients, could provide information up and above which could be achieved using the conventional Mean-based decoding analyses (Grootswagers et al., 2017). This is because, with (i.e. Wavelet) or without any transformation (i.e. Samples), these features extracted the most informative samples across electrodes within each sliding time window. Therefore, through the use of PCA, the selection of samples from different electrodes were directed towards the most informative samples/features (i.e. wavelet coefficients which even provided more information). This supports previous studies which show that the temporal patterns of activity could provide information regarding the co-occurrences of visual edges (Eckhorn et al., 1988) and orientation in primary visual cortex (Celebrini et al., 1993) as well as light intensity in the retina (Gollisch and Meister, 2008). This is also consistent with more recent findings in object recognition suggesting a role for the temporal phase (Behroozi et al., 2016) or within-trial correlation in the temporal cortex (Majima et al., 2014). However, none of the mentioned studies have validated their results across multiple datasets to provide a more generalizable view on object category encoding in the human brain. Moreover, the advantage of the Theta band, to all the other frequency bands evaluated in this study, is consistent with the observation of Theta band being involved in the processing of feed-forward visual information (Bastos et al., 2015), which is dominant in the evaluated object recognition datasets used here.

Importantly, even the complexity features which have been suggested to suffer when analyzing short sequences of the data (50 samples here; Procaccia, 1988) provided information about object categories showing above-chance decoding in the 100 ms to 300 ms time window at some point, with more pronounced results for Dataset 2. This supports previous suggestions (Preißl et al., 1997; Ahmadi-Pajouh et al., 2018; Torabi et al., 2017; Namazi et al., 2018) that even short EEG time series can show highly

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

complex and nonlinear but meaningful structures, which if read appropriately (e.g. through LZ complexity) can provide significant amounts of information about sensory processes. Another interesting observation was the information content provided by the frequency-domain features (e.g. phase at median frequency). These features have been suggested to be more suitable for the extraction of information from stationary time series and not EEG evoked potentials. Comparing the whole-trial and time-resolved results suggest that splitting the signal into its sub-sequences can lead to a more stationary time series allowing frequency features to become more informative.

Another important question in cognitive neuroscience has been whether (if at all) neuroimaging data can explain behavior (Williams et al., 2007; Ritchie et al., 2015; Woolgar et al., 2019). Although many recent studies have found correlations between the neural decoding and behavioral performance in object and face recognition (Karimi-Rouzbahani et al., 2019; Karimi-Rouzbahani et al., 2020a; Dobs et al., 2019), as we also did in the current study, one question that had remained unanswered was whether a more optimal decoding of object category processing, which is searched for here using feature extraction, could explain the behavioral performance more accurately. Here, we showed that, this can be the case. The reason for this observation seems to be that, if there is any explanatory power in the conventional Mean-based decoding analyses, it should improve if we can detect and utilize the additional neural codes which have been ignored as a result of down-sampling/temporal averaging. Interestingly, here we observed that there seems to be a linear relationship between the decoding accuracy that we can obtain and the explanatory power of the features suggesting that in order to bring the neuroimaging closer to behavior, we might need to work on how we can read out the neural code more optimally.

As for the field of visual-representation-based BCI, in which the aim is to improve the decoding performance for improved brain-computer interaction (Wang et al., 2012; Van Gerven et al., 2009), this study provides new suggestions. Specifically, we showed that, when considering the whole-trail data, which is often the case in BCI, the ERP components and the multi-valued features provided the highest amount of information about object categories, and it is most pronounced in the Theta band. None of the previous studies, which used the ERP components (Wang et al., 2012; Qin et al., 2016; Jadidi et al., 2016) or the Wavelet transformation (Taghizadeh-Sarabi et al., 2015; Torabi et al., 2017), limited their frequency band to the Theta band, which here showed higher decoding accuracy in both the whole-trial ERP components (Figure 2) as well as sliding-window Wavelet component (Supplementary Supplementary Figure 1A). Therefore, the suggestion that this work can have for BCI might be to concentrate at specific frequency sub-bands relevant to the cognitive or sensory processing undergoing in the brain; i.e. looking at the Theta band which has been suggested to reflect feed-forward visual information processing in the brain (Bastos et al., 2015) when doing visual-representation-based BCI.

While many studies have used supervised computational algorithms such as Common Spatial Patterns (Murphy et al., 2011), Voltage Topographies (Tzovara et al., 2011), Independent Component Analysis (Stewart et al., 2014) and Convolutional Neural Networks (Seeliger et al., 2017), the focus of the current study was to compare the inherent codes already available in brain representations through statistical

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and computational feature extraction. In other words, we made no supervised adjustments to these features to improve category-separable representations and extracted them directly from the time series data. This is because, rather than trying to maximize the separability/decodeability of representations across object categories, we mainly aimed at gaining insight into how the brain encodes this information by determining which of the available statistical features could capture/detect these codes. Accordingly, there might be supervised algorithms, particularly in the area of BCI, which can provide higher decoding accuracies than those obtained from individual features of this study. Moreover, if we had not reduced the dimensionality of the data for the multi-valued features for the sake of comparison with single-valued features, the decoding performance would likely increase as it has been the case in previous studies (Taghizadeh-Sarabi et al., 2015; Torabi et al., 2017).

There are several future directions for this research. One main question is how the results of this study generalize to other cognitive processes such as attention, memory, decision making, etc. In other words, it is interesting to know what the most informative features would be when decoding different conditions in memory and attention tasks. Another interesting observation that we had in the current study was that shorter time windows of signals (i.e. 5 ms) tended to provide larger initial (0 ms<t<200 ms) peaks in decoding compared to longer ones (i.e. 100 ms), while the latter provided higher decoding values in later stages of the trial (t > 200; Supplementary Figure 2A). This was more pronounced for Dataset 2, which had an active task and a longer presentation time. It suggests that, it could be the case that initial stages of object category processing (e.g. extraction of visual features) takes a shorter time spans, while later stages (e.g. association of visual information to categories or recurrence/feedback processing) take longer time spans. Therefore, we might need a time-variable sliding window for different stages of object recognition or any other cognitive processes to truly capture the temporal dynamics of cognition. Another interesting extension to this work can be studying how (if at all) a combination of the features used in this study could provide added information about object categories and/or behavior. In other words, although all of the individual features evaluated here covered some variance of category object information, to obtain the full variance of the actual neural code, we might need to combine multiple features. To that end, we can combine the extracted features using a variety of supervised and un-supervised methods as have previously provided additional information (Karimi Rouzbahani et al., 2011; Qin et al., 2016). Finally, although explored to some extent (Hatamimajoumerd et al., 2019), it can be interesting to look at specific time points and electrodes selected in our PCA-based dimension reduction to see when and where in the brain the optimal neural codes were extracted from.

The cross-dataset, large-scale analysis methods implemented in this study aligns with the growing trend towards meta-analyses in cognitive neuroscience. Recent studies have also adopted and compared several datasets to facilitate forming more rigorous conclusions about how the brain performs different cognitive processes such as sustained attention (Langner et al., 2013) or working memory (Adam et al., 2020). Our exploratory analysis presented here, has also the advantage that it was not biased towards any of the findings. We simply compared sets of features from the EEG signals to see which one provides more information about object categories and which one best explain behavior. The findings of this

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

study provide generalizable insights into the most informative features of EEG signals in object category processing.

Acknowledgements

This research was funded by UK Royal Society's Newton International Fellowship SUAI/059/G101116 to H.K.R.

References

Abootalebi, V., Moradi, M.H. and Khalilzadeh, M.A., 2009. A new approach for EEG feature extraction in P300-based lie detection. Computer methods and programs in biomedicine, 94(1), pp.48-57.

Aboy, M., Hornero, R., Abásolo, D. and Álvarez, D., 2006. Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. IEEE transactions on biomedical engineering, 53(11), pp.2282-2288.

Adam, K.C., Vogel, E.K. and Awh, E., 2020. Multivariate analysis reveals a generalizable human electrophysiological signature of working memory load. *bioRxiv*.

Ahmadi-Pajouh, M.A., Ala, T.S., Zamanian, F., Namazi, H. and Jafari, S., 2018. Fractal-based classification of human brain response to living and non-living visual stimuli. Fractals, 26(05), p.1850069.

Alimardani, F., Cho, J.H., Boostani, R. and Hwang, H.J., 2018. Classification of bipolar disorder and schizophrenia using steady-state visual evoked potential based features. IEEE Access, 6, pp.40379-40388.

Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H. and Fries, P., 2015. Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), pp.390-401.

Behroozi, M., Daliri, M.R. and Shekarchi, B., 2016. EEG phase patterns reflect the representation of semantic categories of objects. *Medical & biological engineering & computing*, *54*(1), pp.205-221.

Bizas, E., Simos, P.G., Stam, C.J., Arvanitis, S., Terzakis, D. and Micheloyannis, S., 1999. EEG correlates of cerebral engagement in reading tasks. Brain Topography, 12(2), pp.99-105.

Carlson, T., Tovar, D.A., Alink, A. and Kriegeskorte, N., 2013. Representational dynamics of object vision: the first 1000 ms. *Journal of vision*, *13*(10), pp.1-1.

Celebrini, S., Thorpe, S., Trotter, Y. and Imbert, M., 1993. Dynamics of orientation coding in area V1 of the awake primate. *Visual neuroscience*, *10*(5), pp.811-825.

Chan, A.M., Halgren, E., Marinkovic, K. and Cash, S.S., 2011. Decoding word and category-specific spatiotemporal representations from MEG and EEG. *Neuroimage*, *54*(4), pp.3028-3039.

Churchland, M.M., Byron, M.Y., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B. and Bradley, D.C., 2010. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nature neuroscience, 13(3), pp.369-378.

Cichy, R.M., Pantazis, D. and Oliva, A., 2014. Resolving human object recognition in space and time. *Nature neuroscience*, *17*(3), p.455.

Contini, E.W., Wardle, S.G. and Carlson, T.A., 2017. Decoding the time-course of object recognition in the human brain: From visual features to categorical decisions. *Neuropsychologia*, *105*, pp.165-176.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1136 DiCarlo, J.J., Zoccolan, D. and Rust, N.C., 2012. How does the brain solve visual object recognition?.
1137 Neuron, 73(3), pp.415-434.

1138 Dienes, Z., 2014. Using Bayes to get the most out of non-significant results. *Frontiers in psychology*, *5*,
1139 p.781.

1140 Dobs, K., Isik, L., Pantazis, D. and Kanwisher, N., 2019. How face perception unfolds over time. *Nature*
1141 *communications*, *10*(1), pp.1-10.

1142 Duda, R.O., Hart, P.E. and Stork, D.G., 2012. Pattern classification. John Wiley & Sons.

1143 Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. and Reitboeck, H.J., 1988. Coherent
1144 oscillations: A mechanism of feature linking in the visual cortex?. *Biological cybernetics*, *60*(2), pp.121-
1145 130.

1146 Gelman, A. and Tuerlinckx, F., 2000. Type S error rates for classical and Bayesian single and multiple
1147 comparison procedures. Computational Statistics, 15(3), pp.373-390.

1148 Gelman, A., Hill, J. and Yajima, M., 2012. Why we (usually) don't have to worry about multiple
1149 comparisons. Journal of Research on Educational Effectiveness, 5(2), pp.189-211.

1150 Gollisch, T. and Meister, M., 2008. Rapid neural coding in the retina with relative spike
1151 latencies. *Science*, *319*(5866), pp.1108-1111.

1152 Grootswagers, T., Cichy, R.M. and Carlson, T.A., 2018. Finding decodable information that can be read
1153 out in behaviour. NeuroImage, 179, pp.252-262.

1154 Grootswagers, T., Robinson, A.K. and Carlson, T.A., 2019. The representational dynamics of visual
1155 objects in rapid serial visual processing streams. NeuroImage, 188, pp.668-679.

1156 Grootswagers, T., Wardle, S.G. and Carlson, T.A., 2017. Decoding dynamic brain patterns from evoked
1157 responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. Journal of
1158 cognitive neuroscience, 29(4), pp.677-697.

1159 Guo, L., Rivero, D., Seoane, J.A. and Pazos, A., 2009. Classification of EEG signals using relative
1160 wavelet energy and artificial neural networks. In Proceedings of the first ACM/SIGEVO Summit on
1161 Genetic and Evolutionary Computation (pp. 177-184).

1162 Hatamimajoumerd, E. and Talebpour, A., 2019. A Temporal neural trace of wavelet coefficients in human
1163 object vision: an MEG study. F*rontiers in neural circuits*, *13*, p.20.

1164 Hatamimajoumerd, E., Talebpour, A. and Mohsenzadeh, Y., 2019. Enhancing multivariate pattern
1165 analysis for magnetoencephalography through relevant sensor selection. *International Journal of Imaging*
1166 *Systems and Technology*.

1167 Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L. and Pietrini, P., 2001. Distributed and
1168 overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), pp.2425-
1169 2430.

1170 Haynes, J.D. and Rees, G., 2006. Decoding mental states from brain activity in humans. Nature Reviews
1171 Neuroscience, 7(7), pp.523-534.

1172 Hebart, M.N. and Baker, C.I., 2018. Deconstructing multivariate decoding for the study of brain function.
1173 Neuroimage, 180, pp.4-18.

1174 Higuchi, T., 1988. Approach to an irregular time series on the basis of the fractal theory. *Physica D:*
1175 *Nonlinear Phenomena*, *31*(2), pp.277-283.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1176 Hjorth, B., 1970. EEG analysis based on time domain properties. *Electroencephalography and clinical*
1177 *neurophysiology*, *29*(3), pp.306-310.

1178 Hung, C.P., Kreiman, G., Poggio, T. and DiCarlo, J.J., 2005. Fast readout of object identity from macaque
1179 inferior temporal cortex. Science, 310(5749), pp.863-866.

1180 Intriligator, J. and Polich, J., 1995. On the relationship between EEG and ERP variability. International
1181 Journal of Psychophysiology, 20(1), pp.59-74.

1182 Iranmanesh, S. and Rodriguez-Villegas, E., 2017. An ultralow-power sleep spindle detection system on
1183 chip. IEEE transactions on biomedical circuits and systems, 11(4), pp.858-866.

1184 Isik, L., Meyers, E.M., Leibo, J.Z. and Poggio, T., 2014. The dynamics of invariant object recognition in
1185 the human visual system. *Journal of neurophysiology*, *111*(1), pp.91-102.

1186 Jadidi, A.F., Zargar, B.S. and Moradi, M.H., 2016, November. Categorizing visual objects; using ERP
1187 components. In *2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International*
1188 *Iranian Conference on Biomedical Engineering (ICBME)* (pp. 159-164). IEEE.

1189 Jeffreys, H., 1998. *The theory of probability*. OUP Oxford.

1190 Joshi, D., Panigrahi, B.K., Anand, S. and Santhosh, J., 2018. Classification of Targets and Distractors
1191 Present in Visual Hemifields Using Time-Frequency Domain EEG Features. *Journal of healthcare*
1192 *engineering*, *2018*.

1193 Kaneshiro, B., Guimaraes, M.P., Kim, H.S., Norcia, A.M. and Suppes, P., 2015. A representational
1194 similarity analysis of the dynamics of object processing using single-trial EEG classification. *Plos*
1195 *one*, *10*(8).

1196 Karimi Rouzbahani, H. and Daliri, M.R., 2011. Diagnosis of Parkinson's disease in human using voice
1197 signals. *Basic and Clinical Neuroscience*, *2*(3), pp.12-20.

1198 Karimi-Rouzbahani, H., 2018. Three-stage processing of category and variation information by entangled
1199 interactive mechanisms of peri-occipital and peri-frontal cortices. *Scientific reports*, *8*(1), pp.1-22.

1200 Karimi-Rouzbahani, H., Bagheri, N. and Ebrahimpour, R., 2017a. Average activity, but not variability, is
1201 the dominant factor in the representation of object categories in the brain. *Neuroscience*, *346*, pp.14-28.

1202 Karimi-Rouzbahani, H., Bagheri, N. and Ebrahimpour, R., 2017b. Hard-wired feed-forward visual
1203 mechanisms of the brain compensate for affine variations in object recognition. *Neuroscience*, *349*,
1204 pp.48-63.

1205 Karimi-Rouzbahani, H., Bagheri, N. and Ebrahimpour, R., 2017c. Invariant object recognition is a
1206 personalized selection of invariant features in humans, not simply explained by hierarchical feed-forward
1207 vision models. Scientific reports, 7(1), pp.1-24.

1208 Karimi-Rouzbahani, H., Ramezani, F., Woolgar, A., Rich, A.N., Ghodrati, M., 2020a. Perceptual difficulty
1209 modulates the direction of information flow in familiar face recognition.

1210 Karimi-Rouzbahani, H., Vahab, E., Ebrahimpour, R. and Menhaj, M.B., 2019. Spatiotemporal analysis of
1211 category and target-related information processing in the brain during object detection. *Behavioural brain*
1212 *research*, *362*, pp.224-239.

1213 Karimi-Rouzbahani, H., Woolgar, A. and Rich, A.N., 2020b. Neural signatures of vigilance decrements
1214 predict behavioural errors before they occur. *bioRxiv*.

1215 Katz, M.J., 1988. Fractals and the analysis of waveforms. *Computers in biology and medicine*, *18*(3),
1216 pp.145-156.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1217 Kiani, R., Esteky, H., Mirpour, K. and Tanaka, K., 2007. Object category structure in response patterns of
1218 neuronal population in monkey inferior temporal cortex. *Journal of neurophysiology*, *97*(6), pp.4296-4309.

1219 Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K. and Bandettini, P.A.,
1220 2008. Matching categorical object representations in inferior temporal cortex of man and
1221 monkey. *Neuron*, *60*(6), pp.1126-1141.

1222 Langner, R. and Eickhoff, S.B., 2013. Sustaining attention to simple tasks: a meta-analytic review of the
1223 neural mechanisms of vigilant attention. *Psychological bulletin*, *139*(4), p.870.

1224 Le Van Quyen, M., Foucher, J., Lachaux, J.P., Rodriguez, E., Lutz, A., Martinerie, J. and Varela, F.J.,
1225 2001. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony.
1226 Journal of neuroscience methods, 111(2), pp.83-98.

1227 Lee, M.D. and Wagenmakers, E.J., 2005. Bayesian statistical inference in psychology: Comment on
1228 Trafimow (2003).

1229 Lempel, A. and Ziv, J., 1976. On the complexity of finite sequences. *IEEE Transactions on information*
1230 *theory*, *22*(1), pp.75-81.

1231 Liu, H., Agam, Y., Madsen, J.R. and Kreiman, G., 2009. Timing, timing, timing: fast decoding of object
1232 information from intracranial field potentials in human visual cortex. *Neuron*, *62*(2), pp.281-290.

1233 Majima, K., Matsuo, T., Kawasaki, K., Kawai, K., Saito, N., Hasegawa, I. and Kamitani, Y., 2014.
1234 Decoding visual object categories from temporal correlations of ECoG signals. *Neuroimage*, *90*, pp.74-83.

1235 Mazaheri, A. and Jensen, O., 2008. Asymmetric amplitude modulations of brain oscillations generate
1236 slow evoked responses. Journal of Neuroscience, 28(31), pp.7781-7787.

1237 Miyakawa, N., Majima, K., Sawahata, H., Kawasaki, K., Matsuo, T., Kotake, N., Suzuki, T., Kamitani, Y.
1238 and Hasegawa, I., 2018. Heterogeneous Redistribution of Facial Subcategory Information Within and
1239 Outside the Face-Selective Domain in Primate Inferior Temporal Cortex. *Cerebral Cortex*, *28*(4), pp.1416-
1240 1431.

1241 Murphy, B., Poesio, M., Bovolo, F., Bruzzone, L., Dalponte, M. and Lakany, H., 2011. EEG decoding of
1242 semantic category reveals distributed representations for single concepts. *Brain and language*, *117*(1),
1243 pp.12-22.

1244 Namazi, H., Ala, T.S. and Bakardjian, H., 2018. Decoding of steady-state visual evoked potentials by
1245 fractal analysis of the electroencephalographic (EEG) signal. Fractals, 26(06), p.1850092.

1246 Norman, K.A., Polyn, S.M., Detre, G.J. and Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern
1247 analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), pp.424-430.

1248 Pincus, S.M. and Huang, W.M., 1992. Approximate entropy: statistical properties and
1249 applications. *Communications in Statistics-Theory and Methods*, *21*(11), pp.3061-3077.

1250 Pouryazdian, S. and Erfanian, A., 2009. Detection of steady-state visual evoked potentials for brain-
1251 computer interfaces using PCA and high-order statistics. In World Congress on Medical Physics and
1252 Biomedical Engineering, September 7-12, 2009, Munich, Germany (pp. 480-483). Springer, Berlin,
1253 Heidelberg.

1254 Preißl, H., Lutzenberger, W., Pulvermüller, F. and Birbaumer, N., 1997. Fractal dimensions of short EEG
1255 time series in humans. Neuroscience letters, 225(2), pp.77-80.

1256 Procaccia, I., 1988. Universal properties of dynamically complex systems: the organization of chaos.
1257 Nature, 333(6174), pp.618-623.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1258 Pulini, A.A., Kerr, W.T., Loo, S.K. and Lenartowicz, A., 2019. Classification accuracy of neuroimaging
1259 biomarkers in attention-deficit/hyperactivity disorder: Effects of sample size and circular analysis.
1260 Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4(2), pp.108-120.

1261 Qin, Y., Zhan, Y., Wang, C., Zhang, J., Yao, L., Guo, X., Wu, X. and Hu, B., 2016. Classifying four-
1262 category visual objects using multiple ERP components in single-trial ERP. *Cognitive*
1263 *neurodynamics*, *10*(4), pp.275-285.

1264 Racine, R., 2011. Estimating the Hurst exponent. *Zurich: Mosaic Group*.

1265 Rasoulzadeh, V.E.S.A.L., Erkus, E.C., Yogurt, T.A., Ulusoy, I. and Zergeroğlu, S.A., 2017. A comparative
1266 stationarity analysis of EEG signals. Annals of Operations Research, 258(1), pp.133-157.

1267 Richman, J.S. and Moorman, J.R., 2000. Physiological time-series analysis using approximate entropy
1268 and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, *278*(6),
1269 pp.H2039-H2049.

1270 Ritchie, J.B., Tovar, D.A. and Carlson, T.A., 2015. Emerging object representations in the visual system
1271 predict reaction times for categorization. *PLoS computational biology*, *11*(6).

1272 Rossion, B., Gauthier, I., Tarr, M.J., Despland, P., Bruyer, R., Linotte, S. and Crommelinck, M., 2000. The
1273 N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects:
1274 an electrophysiological account of face-specific processes in the human brain. Neuroreport, 11(1), pp.69-
1275 72.

1276 Rouder, J.N., Morey, R.D., Speckman, P.L. and Province, J.M., 2012. Default Bayes factors for ANOVA
1277 designs. *Journal of Mathematical Psychology*, *56*(5), pp.356-374.

1278 Rousselet, G.A., Husk, J.S., Bennett, P.J. and Sekuler, A.B., 2007. Single-trial EEG dynamics of object
1279 and face visual processing. Neuroimage, 36(3), pp.843-862.

1280 Rupp, K., Roos, M., Milsap, G., Caceres, C., Ratto, C., Chevillet, M., Crone, N.E. and Wolmetz, M., 2017.
1281 Semantic attributes are encoded in human electrocorticographic signals during visual object
1282 recognition. *NeuroImage*, *148*, pp.318-329.

1283 Sammer, G., 1999. Working memory load and EEG-dynamics as revealed by point correlation dimension
1284 analysis. International journal of psychophysiology, 34(1), pp.89-102.

1285 Seeliger, K., Fritsche, M., Güçlü, U., Schoenmakers, S., Schoffelen, J.M., Bosch, S.E. and Van Gerven,
1286 M.A.J., 2018. Convolutional neural network-based encoding and decoding of visual object recognition in
1287 space and time. *NeuroImage*, *180*, pp.253-266.

1288 Shourie, N., Firoozabadi, M. and Badie, K., 2014. Analysis of EEG signals related to artists and nonartists
1289 during visual perception, mental imagery, and rest using approximate entropy. BioMed research
1290 international, 2014.

1291 Simanova, I., Van Gerven, M., Oostenveld, R. and Hagoort, P., 2010. Identifying object categories from
1292 event-related EEG: toward decoding of conceptual representations. PloS one, 5(12).

1293 Stam, C.J., 2000. Brain dynamics in theta and alpha frequency bands and working memory performance
1294 in humans. Neuroscience letters, 286(2), pp.115-118.

1295 Stam, C.J., 2005. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clinical*
1296 *neurophysiology*, *116*(10), pp.2266-2301.

1297 Stepien, R.A., 2002. Testing for non-linearity in EEG signal of healthy subjects. Acta neurobiologiae
1298 experimentalis, 62(4), pp.277-282.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Stewart, A.X., Nuthmann, A. and Sanguinetti, G., 2014. Single-trial classification of EEG in a visual object task using ICA and machine learning. *Journal of neuroscience methods*, *228*, pp.1-14.

Storey, J.D., 2002. A direct approach to false discovery rates. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64(3), pp.479-498.

Subha, D.P., Joseph, P.K., Acharya, R. and Lim, C.M., 2010. EEG signal analysis: a survey. Journal of medical systems, 34(2), pp.195-212.

Szczepański, J., Amigó, J.M., Wajnryb, E. and Sanchez-Vives, M.V., 2003. Application of Lempel–Ziv complexity to the analysis of neural discharges. Network: Computation in Neural Systems, 14(2), pp.335-350.

Taghizadeh-Sarabi, M., Daliri, M.R. and Niksirat, K.S., 2015. Decoding objects of basic categories from electroencephalographic signals using wavelet transform and support vector machines. *Brain topography*, *28*(1), pp.33-46.

Tong, F. and Pratte, M.S., 2012. Decoding patterns of human brain activity. Annual review of psychology, 63, pp.483-509.

Tononi, G. and Edelman, G.M., 1998. Consciousness and complexity. Science, 282(5395), pp.1846-1851.

Torabi, A., Jahromy, F.Z. and Daliri, M.R., 2017. Semantic category-based classification using nonlinear features and wavelet coefficients of brain signals. *Cognitive Computation*, *9*(5), pp.702-711.

Tzovara, A., Murray, M.M., Plomp, G., Herzog, M.H., Michel, C.M. and De Lucia, M., 2012. Decoding stimulus-related information from single-trial EEG responses based on voltage topographies. *Pattern Recognition*, *45*(6), pp.2109-2122.

Van Gerven, M., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., Haselager, P., Vuurpijl, L., Gielen, S. and Desain, P., 2009. The brain–computer interface cycle. *Journal of neural engineering*, *6*(4), p.041001.

Vaziri-Pashkam, M. and Xu, Y., 2017. Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. Journal of Neuroscience, 37(36), pp.8767-8782.

Vidal, J.R., Ossandón, T., Jerbi, K., Dalal, S.S., Minotti, L., Ryvlin, P., Kahane, P. and Lachaux, J.P., 2010. Category-specific visual responses: an intracranial study comparing gamma, beta, alpha, and ERP response selectivity. Frontiers in human neuroscience, 4, p.195.

Vidaurre, D., Myers, N.E., Stokes, M., Nobre, A.C. and Woolrich, M.W., 2019. Temporally unconstrained decoding reveals consistent but time-varying stages of stimulus processing. Cerebral Cortex, 29(2), pp.863-874.

Voloh, B., Oemisch, M. and Womelsdorf, T., 2020. Phase of firing coding of learning variables across the fronto-striatal network during feature-based learning. Nature communications, 11(1), pp.1-16.

Wairagkar, M., Zoulias, I., Oguntosin, V., Hayashi, Y. and Nasuto, S., 2016, June. Movement intention based Brain Computer Interface for Virtual Reality and Soft Robotics rehabilitation using novel autocorrelation analysis of EEG. In 2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob) (pp. 685-685). IEEE.

Wang, C., Xiong, S., Hu, X., Yao, L. and Zhang, J., 2012. Combining features from ERP components in single-trial EEG for discriminating four-category visual objects. *Journal of neural engineering*, *9*(5), p.056013.

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1340 Wang, Y., Wang, P. and Yu, Y., 2018. Decoding English alphaAlphabet letters using EEG phase
1341 information. *Frontiers in neuroscience*, *12*, p.62.

1342 Watrous, A.J., Deuker, L., Fell, J. and Axmacher, N., 2015. Phase-amplitude coupling supports phase
1343 coding in human ECoG. Elife, 4, p.e07886.

1344 Williams, M.A., Dang, S. and Kanwisher, N.G., 2007. Only some spatial patterns of fMRI response are
1345 read out in task performance. *Nature neuroscience*, *10*(6), pp.685-686.

1346 Wong, K.F.K., Galka, A., Yamashita, O. and Ozaki, T., 2006. Modelling non-stationary variance in EEG
1347 time series by state space GARCH model. Computers in biology and medicine, 36(12), pp.1327-1335.

1348 Woolgar, A., Dermody, N., Afshar, S., Williams, M.A. and Rich, A.N., 2019. Meaningful patterns of
1349 information in the brain revealed through analysis of errors. *bioRxiv*, p.673681.

1350 Zellner, A. and Siow, A., 1980. Posterior odds ratios for selected regression hypotheses. *Trabajos de*
1351 *estadística y de investigación operativa*, *31*(1), pp.585-603.

1352

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figures:

Hamid Karimi-Rouzbahani et al., "Temporal codes provide additional category-related information in object category decoding: a systematic comparison between informative EEG features".

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1

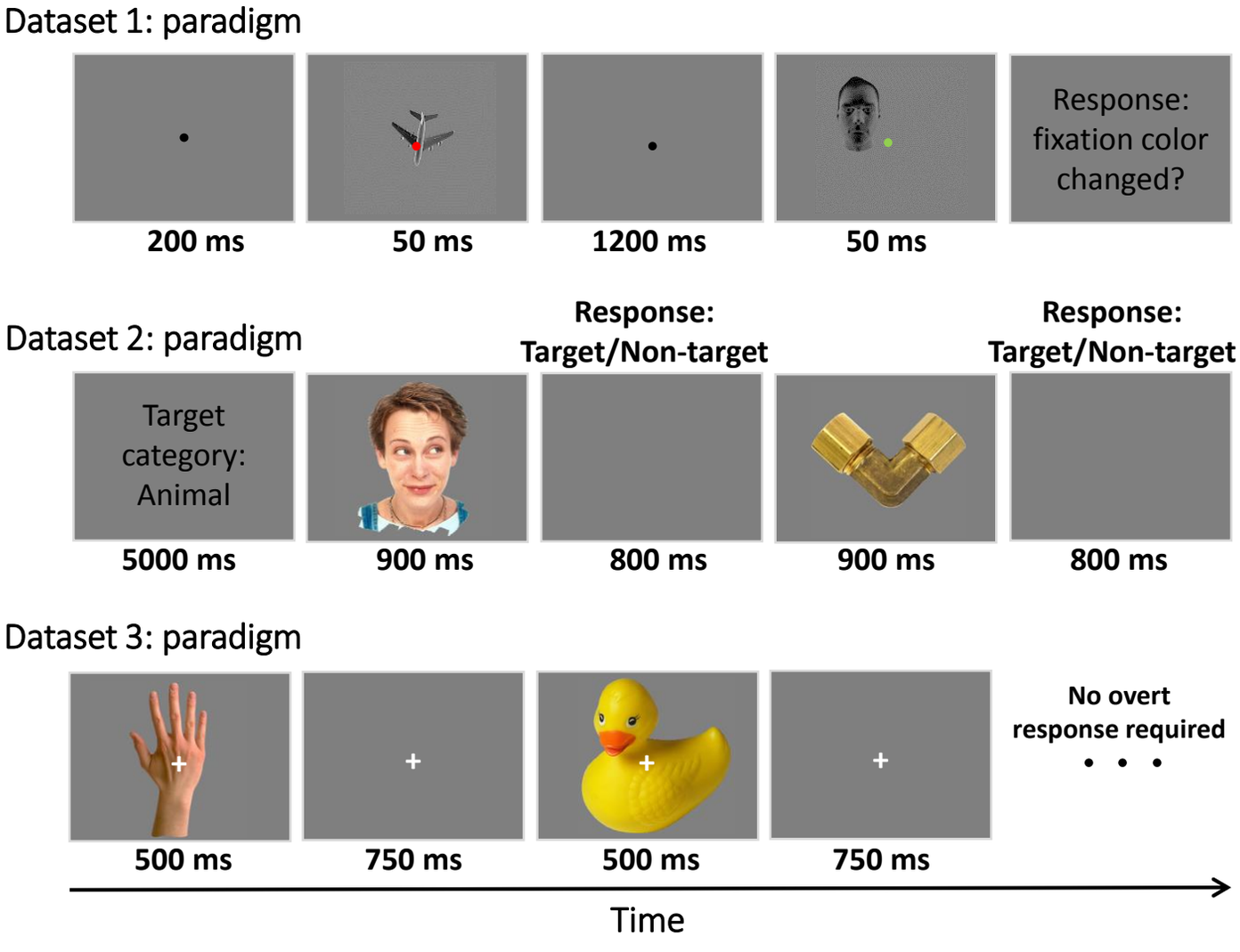


Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

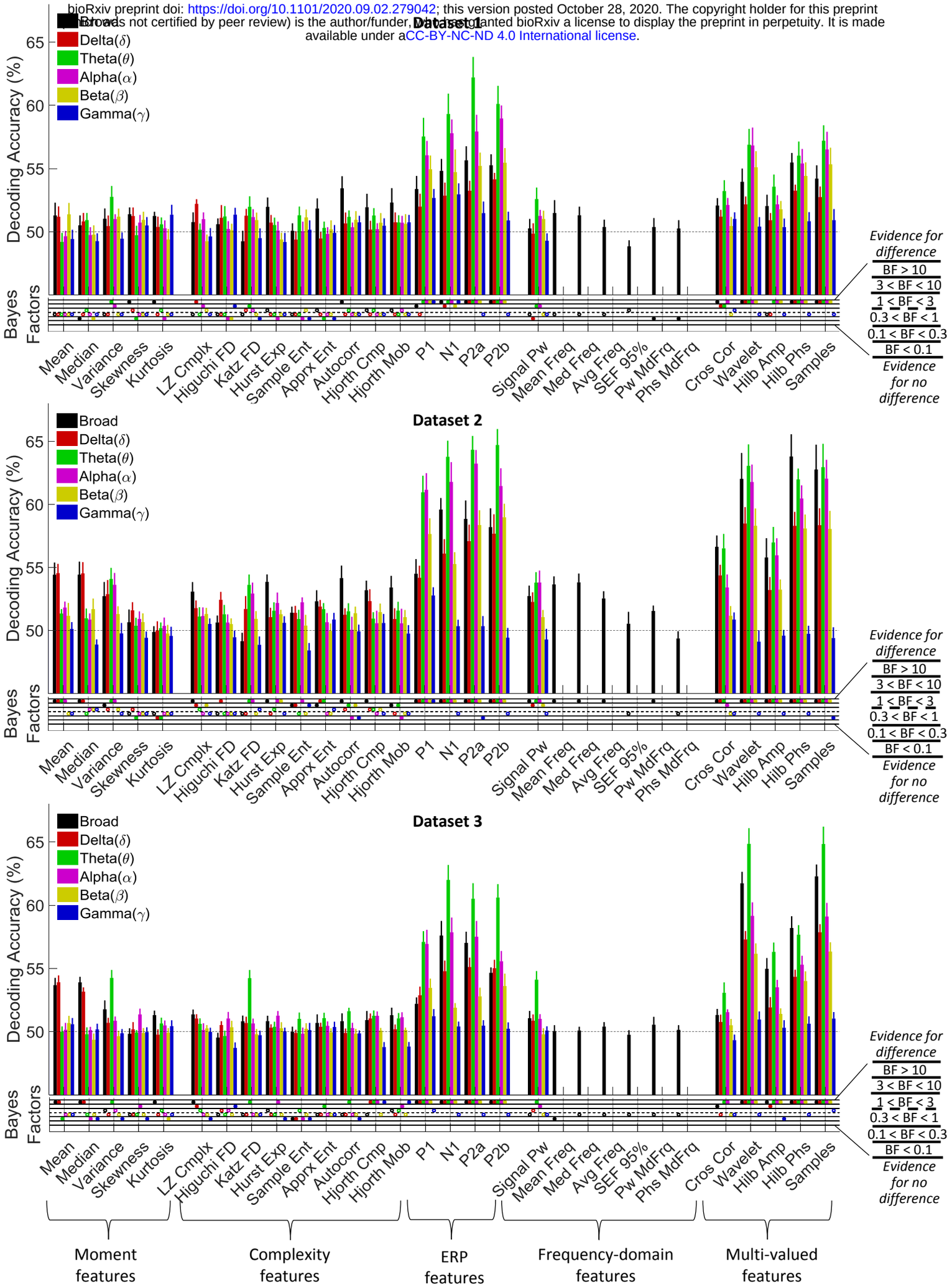


Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

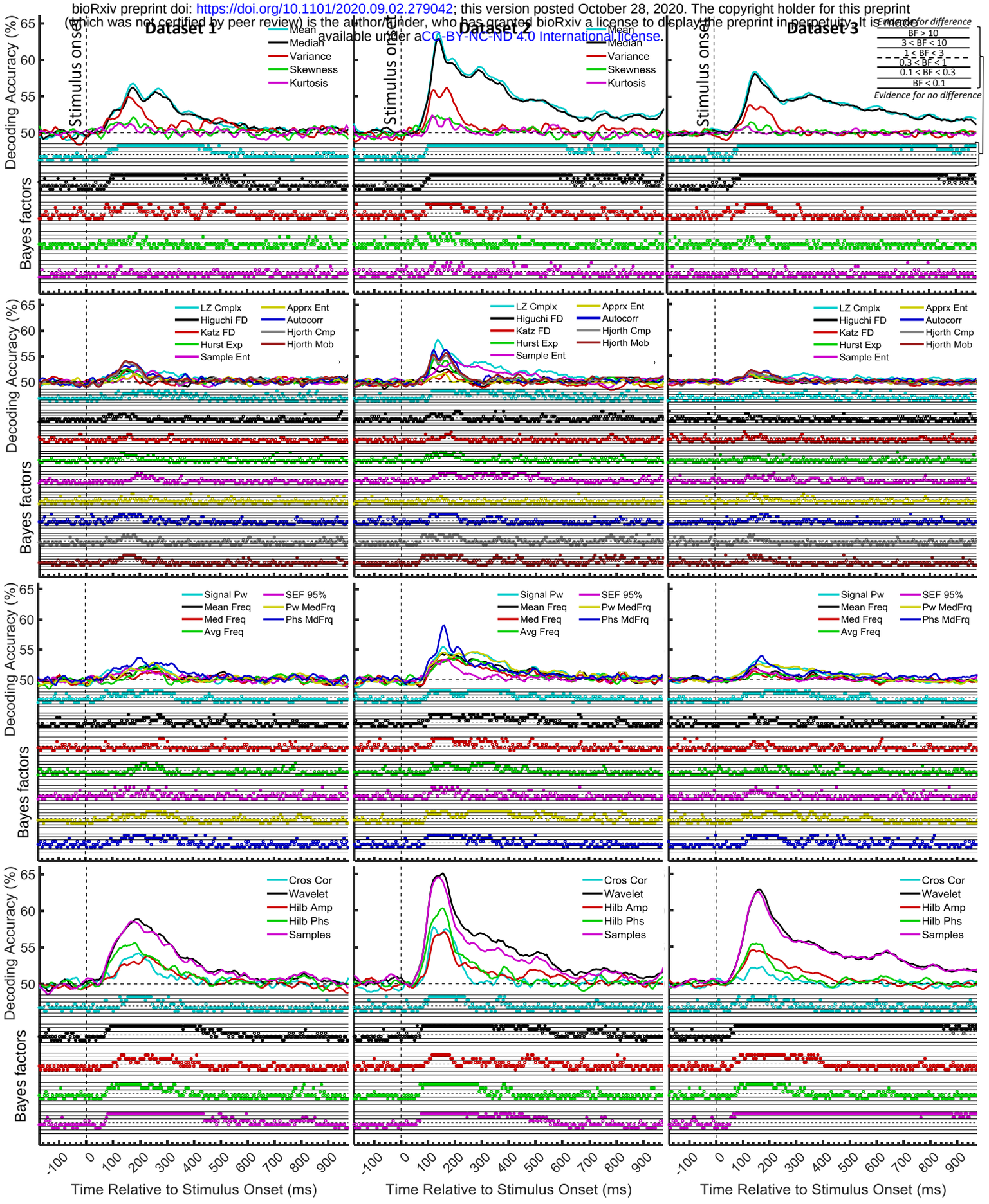


Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

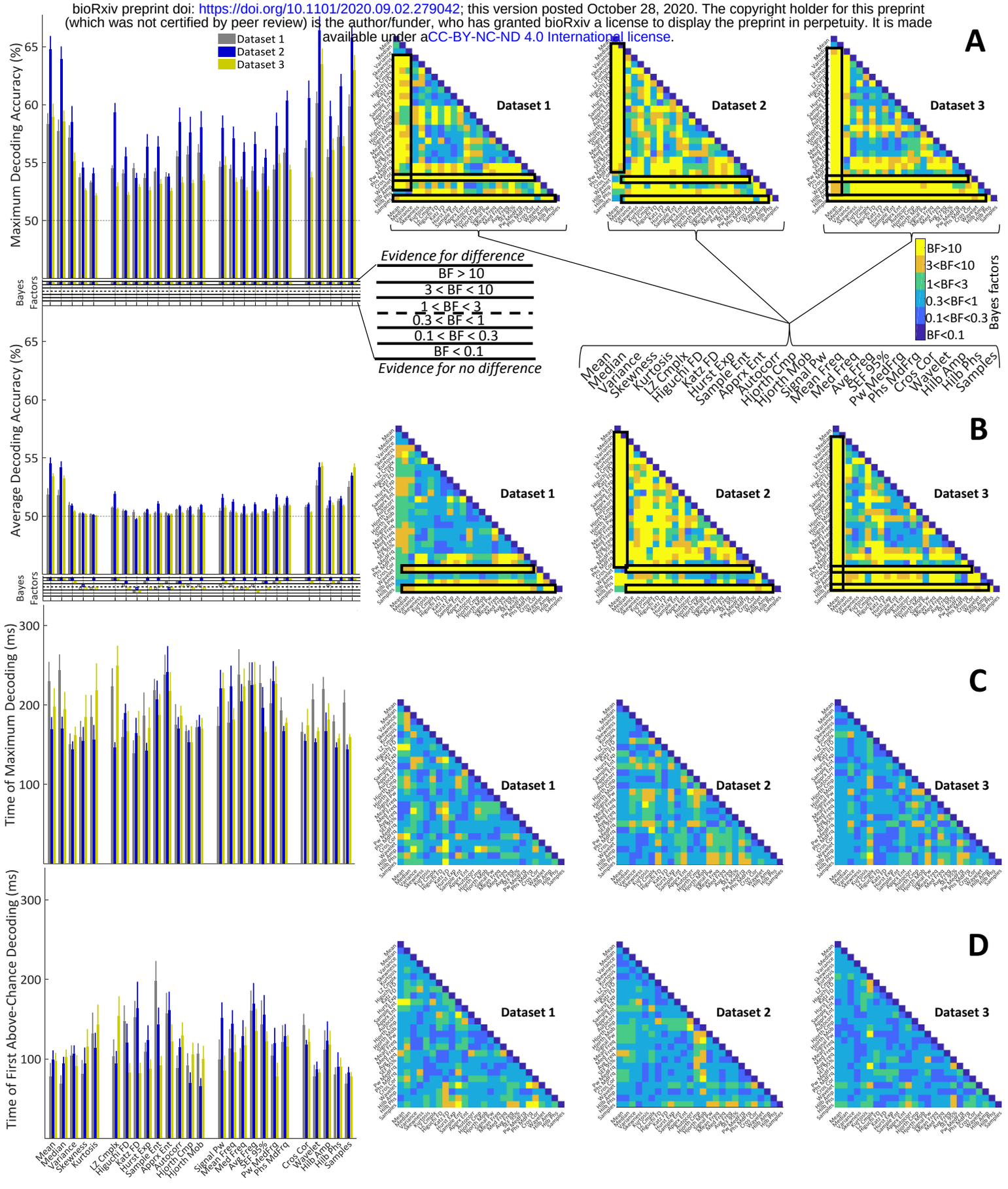
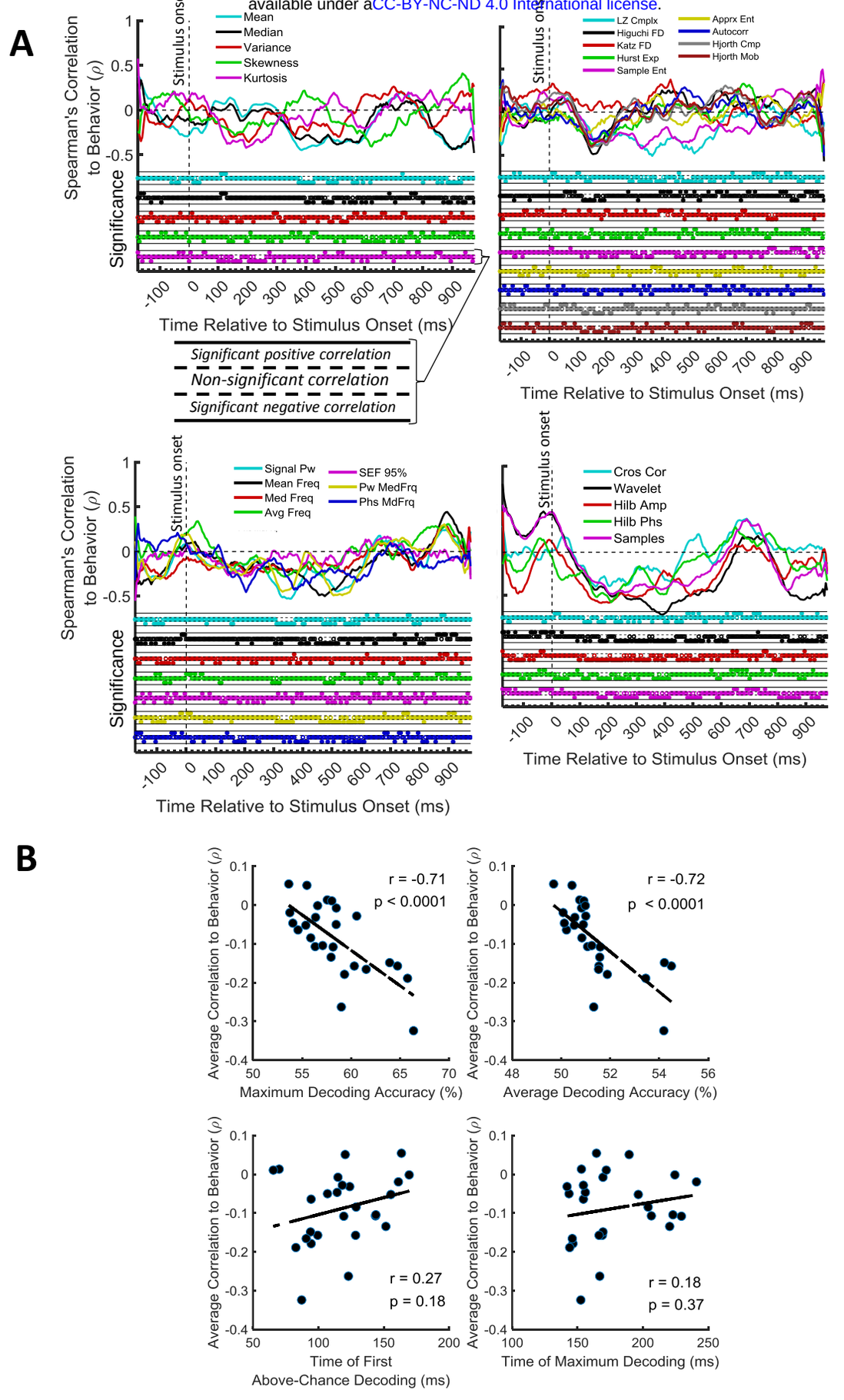


Figure 5

bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

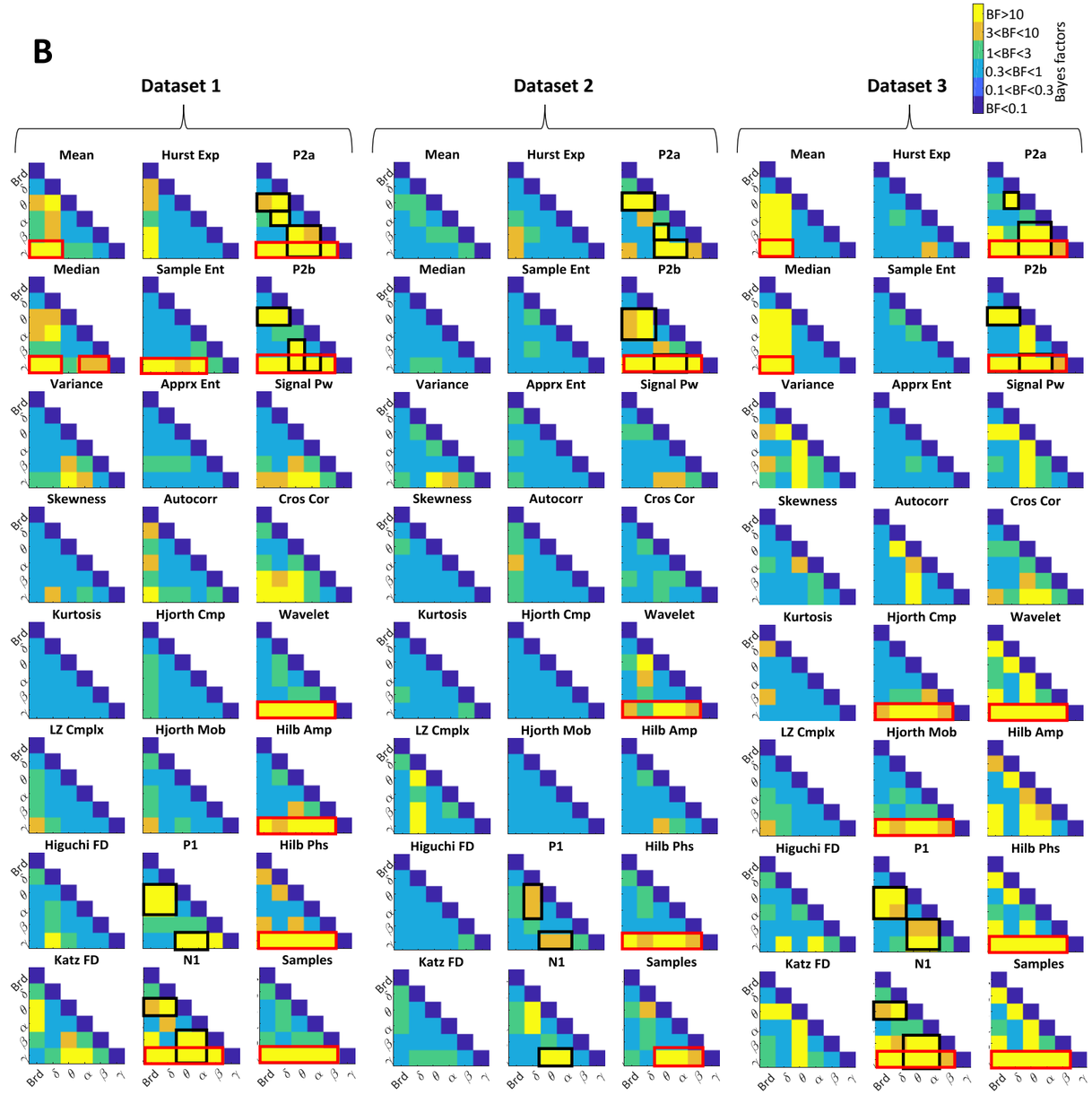
Supplementary Materials:

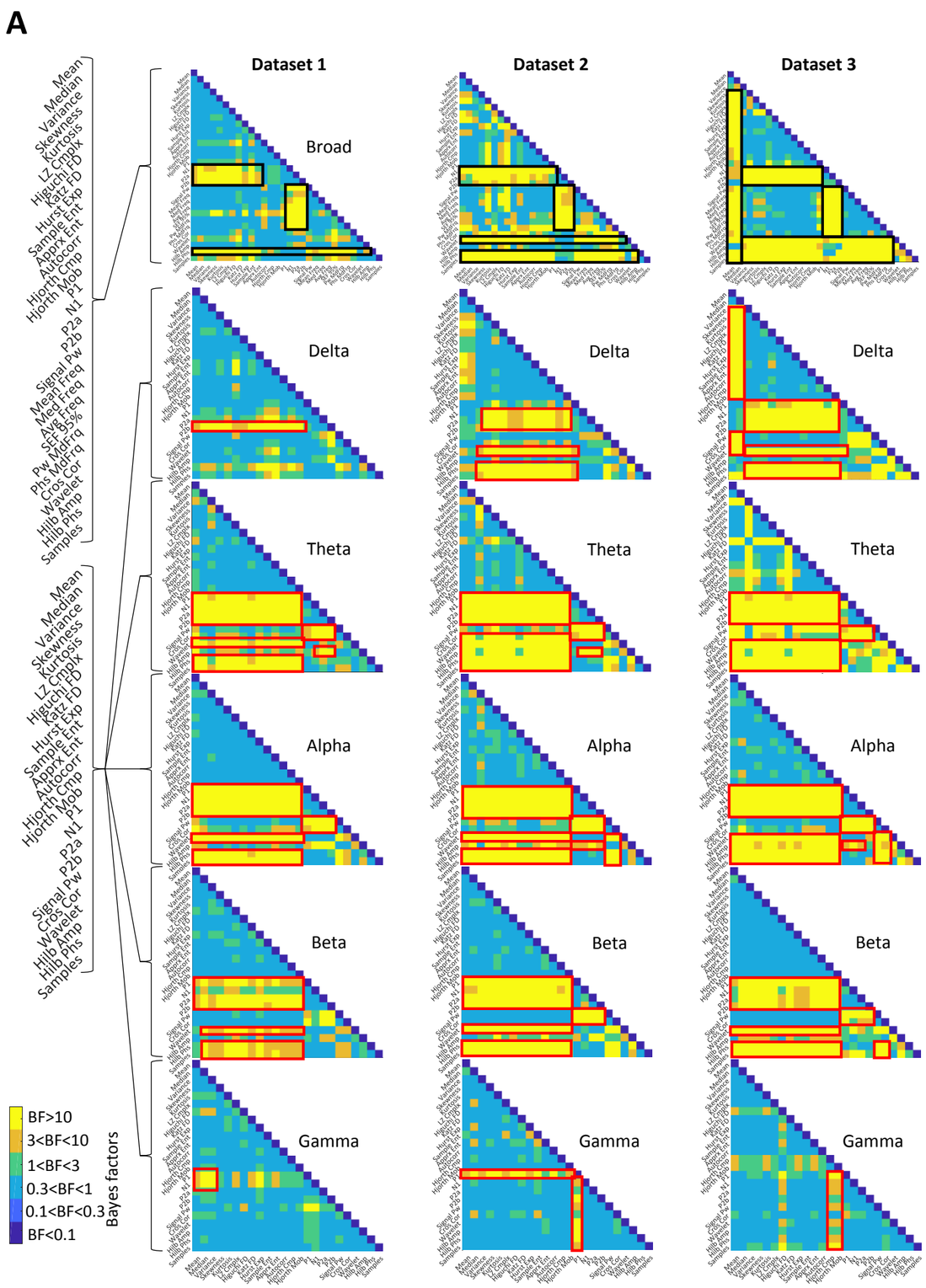
Hamid Karimi-Rouzbahani et al., "Temporal codes provide additional category-related information in object category decoding: a systematic comparison between informative EEG features".

Supplementary Figure 1

(A) Bayes factor matrices comparing the whole-trial decoding results across different features for each frequency band and dataset separately. Matrices show different levels of evidence for existing difference (moderate 3<BF<10 Orange, strong BF>10 Yellow), no difference (moderate 0.1<BF<0.3, light blue; strong BF<0.1, dark blue), or insufficient evidence (1<BF<3 green; 0.3<BF<1 Cyan) for either hypotheses. Black and red boxes show moderate or strong evidence for higher decoding values for specific features compared as explained in the text. For example, for Dataset 1, there is insufficient evidence for difference between decoding values of most features in the Gamma band as indicated by the light blue color in most cells. However, there is moderate to strong evidence that Mean and Median features are different from N1 and P1 as indicated by yellow color and the decoding accuracies in Figure 2. (B) Bayes factor matrices comparing the whole-trial decoding results across different frequency bands for each feature and dataset separately. Black and red boxes show moderate or strong evidence for higher decoding values for specific features compared as explained in the text.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.07.240804; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.





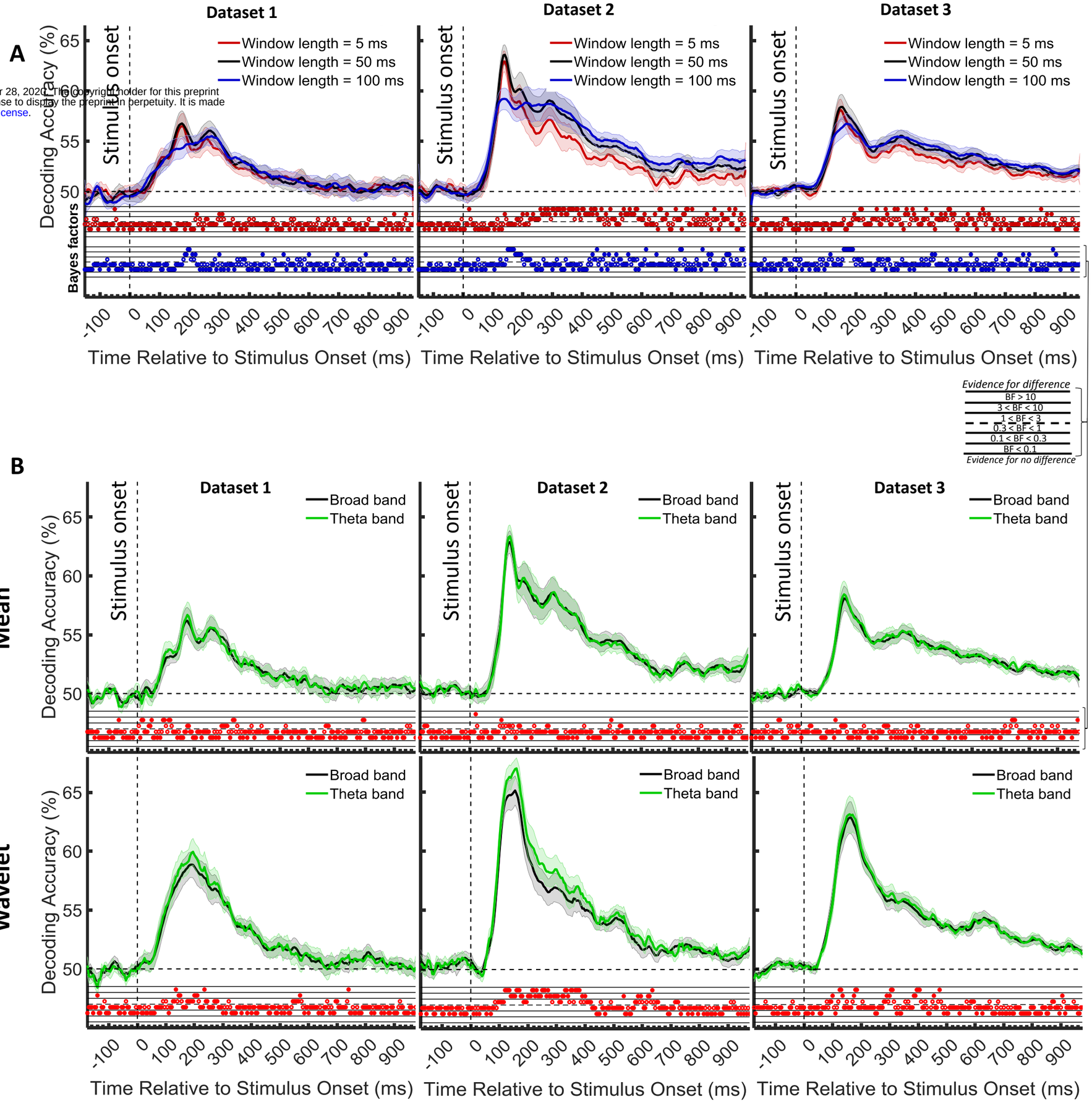
bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint which was not certified by peer review is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Text 2

We selected the window length of 50 ms for our time-resolved analyses because it was neither too long to hide the true temporal dynamics of information processing in the brain, nor too short to avoid the accurate calculation of sample-dependent (e.g. complexity and multi-valued) features. To assure that we did not miss the true obtainable dynamic range (amplitude) of accuracies, we compared category decoding obtained from time windows of 5 (i.e. which was the case in most previous studies all of which relied on signals' mean (Grootswagers et al., 2017; Karimi-Rouzbahani et al., 2017b) and 100 ms with that used here from 50 ms time windows. Consistently across the three datasets, results showed that the highest decoding accuracies were obtained for the 50 ms time windows, both in terms of maximum and average decoding accuracy after the stimulus onset. Interestingly, lengthening the time windows decreased the maximum decoding but increased the decoding accuracies in the later stages of the processing (i.e. from 200 ms onwards; probably after initial hard-wired processing of visual stimuli). This may suggest that later stages of category processing (probably involving feedback/recurrent processing; which are activated by the longer presentation time in datasets 2 and 3), take longer processing times, therefore captured better using longer time windows.

Supplementary Figure 2

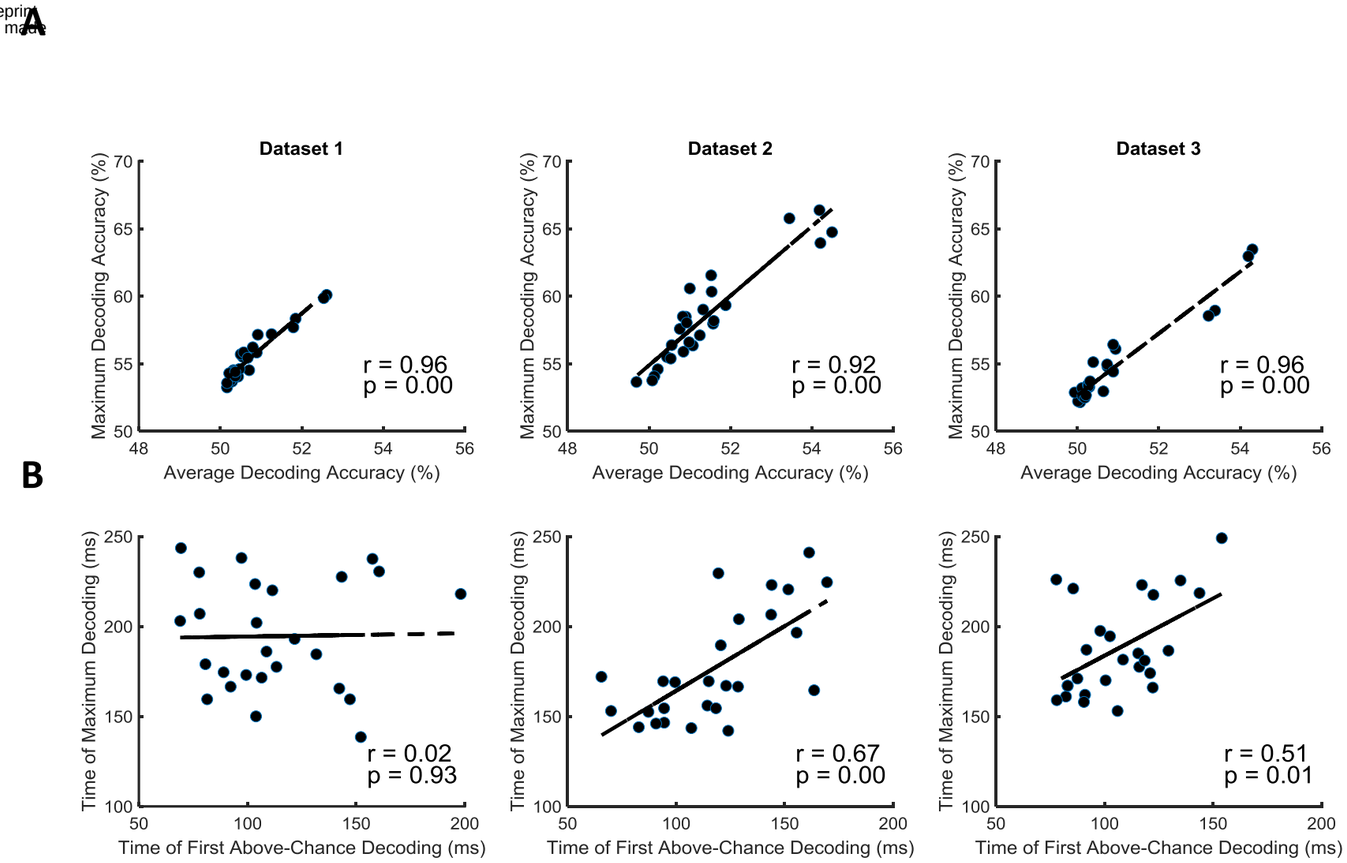
(A) Comparison of decoding accuracies using different length for the sliding time window. The bottom section shows the Bayes factor evidence for the difference between the 50 ms window and the other two window lengths. (B) Comparison of decoding accuracies using different frequency bands for the Mean (top) and Wavelet (bottom) features. Each column shows the results for one dataset. Top section in each panel shows the decoding accuracies across time and the horizontal dashed lines on the top panel refer to chance-level decoding. Filled circles in the Bayes Factors show moderate/strong evidence for either difference or no difference between the decoding curves and empty circles indicate insufficient evidence for either hypotheses. Thick lines show the average decoding accuracy across participants (error bars Standard Error across participants).



bioRxiv preprint doi: https://doi.org/10.1101/2020.09.02.279042; this version posted October 28, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Text 3

The amplitude parameters of maximum and average decoding showed relatively similar patterns across features. This seemed to be the case for the timing parameters of the time of maximum and the time of first above-chance decoding values too. To quantitatively see if there was any correlation within the amplitude and timing parameters, we calculated their correlations across features and for each dataset separately. Results showed significant correlation (Pearson's r>0.9, p<0.01) between maximum and average decoding values across the features and for all three datasets. There was also significant (Pearson's r>0.5, p<0.05) correlation between the time of first above-chance and maximum decoding for datasets 2 and 3, but not for dataset 1 (r=0.02, p=0.93), which might have been because of the lower decoding values in dataset 1 compared to the other datasets making the correlations noisier.



Supplementary Figure 3

Correlation between the pairs of amplitude (A) and timing (B) parameters of the time-resolved decoding (i.e. maximum and average decoding accuracy and time of first and maximum decoding) for the set of N=26 individual features. The slant line shows the best linear fit to the distribution of the correlation data.