Received: 1 September 2019 | Revised: 14 August 2021 | Accepted: 28 August 2021
DOI: 10.1111/ejn.15440

RESEARCH REPORT

EJN European Journal of Neuroscience FENS
WILEY

Task-dependent neural representations of visual object categories

Amirhossein Farzmahdi[1] | **Fatemeh Fallah**[2] | **Reza Rajimehr**[3] | **Reza Ebrahimpour**[2]

[1]School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
[2]Department of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran
[3]MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

Correspondence
Amirhossein Farzmahdi, School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran.
Email: a.farzmahdi@ipm.ir

Reza Ebrahimpour, Department of Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran.
Email: rebrahimpour@sru.ac.ir

Funding information
Institute for Research in Fundamental Sciences; Shahid Rajaee Teacher Training University, Grant/Award Number: 30309

Edited by: Gregor Thut

Abstract

What do we perceive in a glance of an object? If we are questioned about it, will our perception be affected? How does the task demand influence visual processing in the brain and, consequently, our behaviour? To address these questions, we conducted an object categorisation experiment with three tasks, one at the superordinate level ('animate/inanimate') and two at the basic levels ('face/body' and 'animal/human face') along with a passive task in which participants were not required to categorise objects. To control bottom-up information and eliminate the effect of sensory-driven dissimilarity, we used a particular set of animal face images as the identical target stimuli across all tasks. We then investigated the impact of top-down task demands on behaviour and brain representations. Behavioural results demonstrated a superordinate advantage in the reaction time, while the accuracy was similar for all categorisation levels. The event-related potentials (ERPs) for all categorisation levels were highly similar except for about 170 ms and after 300 ms from stimulus onset. In these time windows, the animal/human face categorisation, which required fine-scale discrimination, elicited a differential ERP response. Similarly, decoding analysis over all electrodes showed the highest peak value of task decoding around 170 ms, followed by a few significant timepoints, generally after 300 ms. Moreover, brain responses revealed task-related neural modulation during categorisation tasks compared with the passive task. Overall, these findings demonstrate different task-related effects on the behavioural response and brain representations. The early and late components of neural modulation could be linked to perceptual and top-down processing of object categories, respectively.

KEYWORDS
behavioural responses, event-related potentials (ERPs), levels of object categorisation, N170, task-related information

---

Abbreviations: EEG, electroencephalography; ERP, event-related potential; FDR, false discovery rate; ms, millisecond; RT, reaction time.

1 | INTRODUCTION

In visual object categorisation, meaningful levels of abstraction are constructed to organise visual stimuli. For example, an object can be categorised at the superordinate level (animal), basic level (dog) or subordinate level (hound). Bottom-up (scene-dependent) (Hong et al., 2016; Isik et al., 2013; VanRullen, 2007) and top-down (task-dependent) (Bar et al., 2006; Çukur et al., 2013; Harel et al., 2011; Harel et al., 2014; Vaziri-Pashkam & Xu, 2017) processes control how an object is represented in the brain and perceived by individuals. Thus, our daily life's visual object categorisation levels might depend on the external stimulus and our behavioural goals. Regarding the presence of these two types of processes, our goal was to keep one of these resources identical and alter the other to find out how it affects distinct perceptual and neural dynamics. Using object categorisation tasks provides an intriguing opportunity to explore how the task context can affect object representation processes in the brain and behaviour.

In the late 1970s, Rosch et al. (1976) developed a theory based on a free-naming experiment that the basic level is what subjects would naturally name/label an object. The basic level is defined as the level in which intermediate abstractions, such as human faces, animal faces, cars, human bodies, birds, dogs, tables and fruits, are driven. The subordinate level indicates subdivisions of basic level categories like individual human faces and specific breeds of dogs. The superordinate level is described by the general concepts such as animals. Several empirical studies suggest that the relative timing of visual information processing is linked to specific object categorisation levels. Seminal work by Rosch and others (Mervis & Rosch, 1981; Rosch et al., 1976) demonstrated that humans are faster in object category verification task at the basic level compared with the superordinate and subordinate levels. Further, Morris and Murphy (1990) applied a set of converging operations in the context of event categorisation and found that basic level labels provided faster responses. However, in contrast to the proposed hypothesis, several findings have challenged the basic level advantage by suggesting that the superordinate level information has an advantage compared with the classic basic level entry for object categorisation (Löw et al., 2003; Macé et al., 2009; Martinovic et al., 2008; Praß et al., 2013; Sugase et al., 1999; Wu et al., 2015). For instance, Macé et al. (2009) showed that humans are faster to recognise an object is an animal than it is a dog. The related literature demonstrated a long-lasting controversy over the entry-level of object categorisation.

Several studies attempted to reconcile this contradiction between superordinate and basic level advantage. Mohan and Arun (2012) found that the similarity among different object categories and between category members can account for categorisation speed. A recent study (Mack & Palmeri, 2015) examined exposure duration and category trial context as factors that might explain the difference in categorisation tasks. Besides, Sofer et al. (2015) provided a computational-level explanation for how differences in behavioural responses across object categorisation tasks could originate from natural variations in perceptual processes. Using machine-learning algorithms, they trained classifiers and calculated quantitative measures by comparing distances between individual images and categorisation boundaries. They suggested that the resulting perceptual discriminability measure could account for observed differences in behavioural responses across categorisation tasks. Thus, the stimulus set's heterogeneity, task design and perceptual discriminability may contribute to behavioural differences across categorisation tasks.

Most of these observations considered the speed of access to categorisation levels by using different image sets. However, the variations between image sets might be an alternative explanation for the contradiction of two advantages or differences between the tasks. Previous studies showed examples of stimulus sets that can affect the level of advantage, such as the inbuilt bias toward mammals or more familiarity with household pets (Wu et al., 2015). Another possible factor is the homogeneity of images in one category (for example, motorcycles) compared to other categories (Poncet & Fabre-Thorpe, 2014). These properties, along with other contributing factors like position, scale, in-plane rotation, in-depth rotation, occlusion, lighting, background and, more importantly, different combinations of these variables, could potentially affect task-related brain response and behaviour. For example, increasing the difficulty level in object variations reduces human accuracy and increases reaction time (RT) (Ghodrati et al., 2014). Prior studies used a set of images with a planned categorical structure that could be similarly hierarchically grouped (Carlson et al., 2013; Cichy et al., 2014). The number and type of exemplars varied across categories. These variations in stimuli could result in different behavioural outcomes. Thus, one way to avoid any impact of stimulus variations and biases at distinct perceptual levels is to analyse the same set of images at the superordinate level and basic level tasks.

Despite such strong evidence that confirmed the impact of object categorisation tasks on behavioural responses, it is not fully revealed how task-dependent information influences brain representations. An increasing number of studies argue for contributions of behavioural goals and task-dependent signals in object processing by showing the location or timing of task-related information modulations in the brain (Çukur et al., 2013; Emadi & Esteky, 2014; Erez & Duncan, 2015; Harel et al., 2014; Hebart et al., 2018; Karimi-Rouzbahani et al., 2019; Vaziri-Pashkam & Xu, 2017). For example, Hebart et al. (2018) described the spatiotemporal dynamics of task and object processing by performing four different tasks (two low-level perceptual targets and two high-level conceptual dimensions) on the same stimuli. They found that the task information exhibits a late (after 150 ms) top-down modulation on object representations between low-level perceptual and high-level conceptual tasks in occipitotemporal and parietal cortex. Groen et al. (2016) found that tasks alter late visual information when participants' attention (full versus reduced attention) is manipulated. At later time points, the enhanced attention contributed to the prolonged ERP signatures of scene processing. Overall, these results suggested that task information affects the late component of visual information. Although these studies provided evidence of the task effect on brain representations, the role of task information over different levels of object categorisation (suporordinate and basic levels) and corresponding neural representations remains unclear.

Regarding neural representations of object categories, Kriegeskorte et al. (2008) found the top level of distinction between the neural representation of animate and inanimate objects in the IT cortex. At the next level, faces and bodies form the subclusters within animate objects with dedicated regions/modules within the IT cortex. Subsequently, faces might be separated into different categories, such as human and animal faces. Both human and animal faces are primarily processed within the face regions, and they can be distinguished only based on a fine-scale separation in the pattern of neural responses. There are variations among neural substrates involved in the processing of different object categories. However, from perceptual hierarchical abstractions, the animate/inanimate task takes place at the superordinate level, and both face/body and human/animal face tasks occur at the basic level (Dehaqani et al., 2016). Thus, at the neural dynamics, both animate/inanimate and face/body tasks rely on the information in macroscopically distinct neural structures (Kanwisher & Yovel, 2006; Kanwisher et al., 1997; Downing et al., 2001; Peelen & Downing, 2005; Schwarzlose et al., 2005), while human/animal face task relies on information in specific neural structures for processing faces (Tsao et al., 2003). Regarding this difference between the perceptual levels and neural dynamics of object categorisation, we expected task-related information associated with these levels in different ways.

In this study, the fixed factor was the bottom-up information, while the top-down information was altered across different object categorisation tasks. In these tasks, subjects were instructed to categorise objects into superordinate categories, basic-animate (basic-A) categories or basic-face (basic-F) categories. By controlling all input variations using the same target stimuli within different tasks, we investigated different perceptual levels through 'go/no-go' visual categorisation tasks. In parallel, electroencephalography (EEG) signals were recorded to determine how the task-dependent information influences the human brain response's temporal and spatial dynamics based on neural representations of object categories. Furthermore, we used a passive task to examine neural representations when participants were exposed to the same set of target stimuli without a corresponding behavioural task. Comparing object categorisation tasks is important but not sufficient, since a similar stimulus may be encountered when no relevant behavioural task is available. The differences between passive and three categorisation tasks may help us identify when task-related information affects the temporal dynamics of neural representations. More specifically, at which time-points, we should expect the presence of task-related information.

2 | MATERIALS AND METHODS

2.1 | Participants

Twelve volunteers participated in this study (eight males; 10 right-handed; mean age, 25 years; range, 24–27 years). All participants had normal or corrected to normal vision and provided written informed consent. The Iran University of Medical Sciences approved all experimental protocols.

2.2 | Stimuli

We used coloured object images taken from a set used in previous studies (Kiani et al., 2007; Kriegeskorte et al., 2008) as well as online sources. The stimuli were chosen to be as varied as possible and isolated on a grey background to reduce the scene context influence on object categorisation. Some sample images from the categories are shown in Figure 1a. Images were all 175 × 175 pixels sustaining around $7^\circ \times 7^\circ$ of the visual angle. The selected images consisted of two main categories namely, 'animate' and 'inanimate', each of which contained

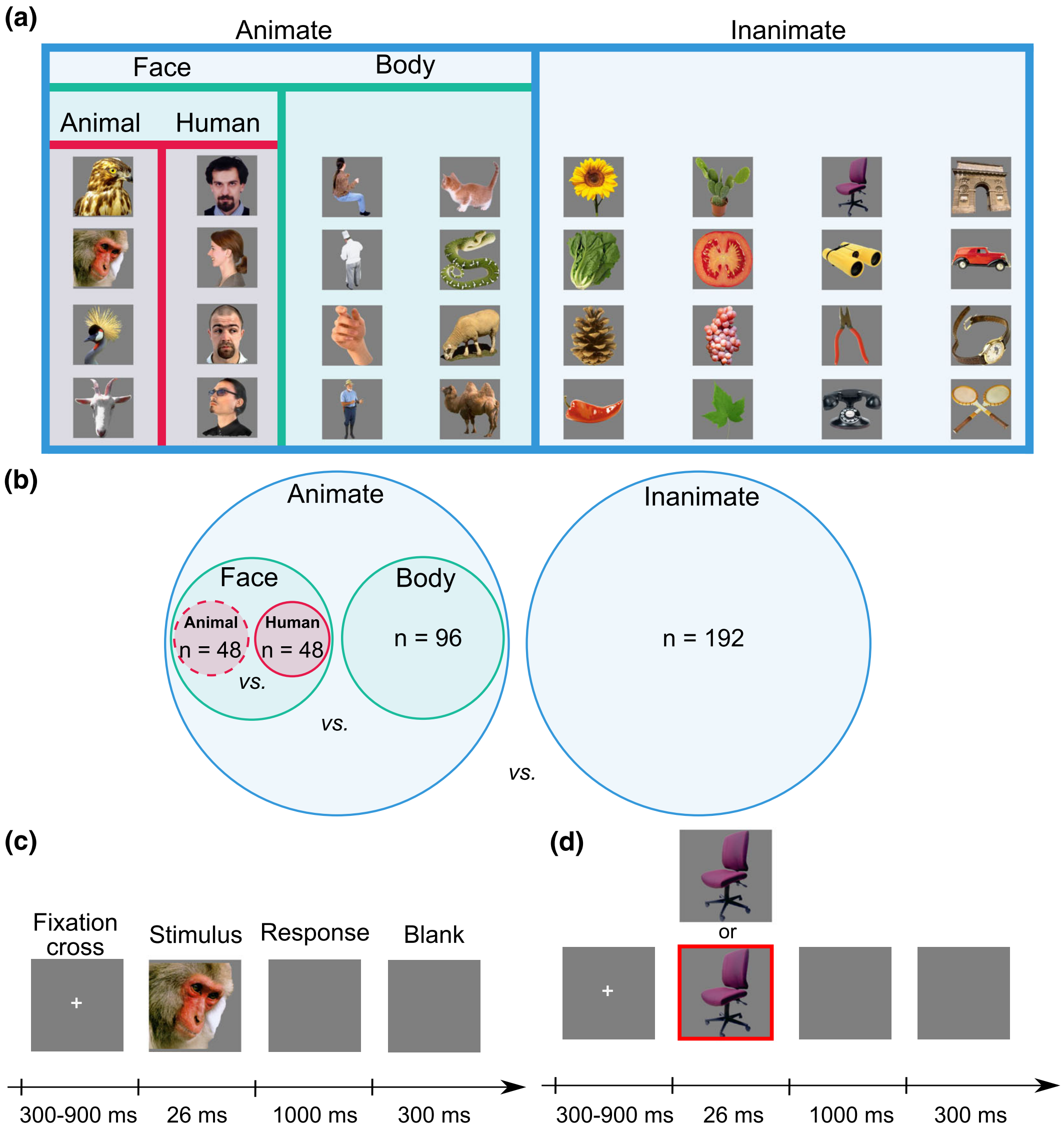

FIGURE 1 Experimental paradigm. (a) Samples of image sets are used in the experiments. Three categorisation tasks are separated with vertical lines according to the stimulus category. The target and non-target categories are illustrated above each set of images. The object categorisation tasks are indicated in colours; blue: superordinate (animate vs. inanimate), green: basic-A (face vs. body), red: basic-F (animal face vs. human face). (b) Each circle represents an object category. The number of images in each category is mentioned at the centre, varied according to the category. The dashed circle represents the target stimuli. (c) This section is an illustration of the main experiment. Each trial started with a fixation cross (with a random variable duration from 300 to 900 ms). Then, a stimulus was presented for a brief time (26 ms), followed by a blank screen delay of 1300 ms in the first 1000 milliseconds of the blank screen, participants had to press or release the space bar to indicate the category of the object according to the tasks. The consecutive trials were separated by a variable interstimulus interval that ranged from 1600 to 2200 ms. (d) This section is an illustration of the passive task. Each trial started with the presentation of a fixation cross (with a random variable duration from 300 to 900 ms). Then, a stimulus was presented for 26 ms (the stimulus enclosed by a red frame randomly in 10% of trials), followed by a blank screen delay of 1300 ms. Participants were required to respond to stimuli held in red frames during the first 1000 ms of the blank screen

192 exemplars (Figure 1b). The animate category included human faces, animal faces, human bodies and animal bodies (48 stimuli in each subset). The inanimate category included 192 images chosen from natural objects such as fruits and vegetables and artificial objects like tools.

2.3 | Procedure

Participants sat approximately 57 cm in front of a monitor (resolution, 800 × 600 pixels; vertical refresh rate, 75 Hz) in a dimly lit room. We used a set of 24 trials for each categorisation level, including 12 target trials and 12 non-target trials, to familiarise participants with the tasks. The images that we used in these training trials were not used in the testing trials. The trained participants were informed about the level of categorisation at the beginning of each block. When starting a trial, participants had to keep pressing the keyboard's spacebar (no-go response). A white fixation cross spanning 0.76° × 0.76° of the visual angle was presented (300–900 ms) at the center of the light grey screen, followed by the stimulus flashed for 26 ms. Participants had to release their fingers from the pressed key as accurately and as quickly as possible when the target image was shown. The duration of pressing the spacebar determined the Go response latency. A maximum of 1000 ms was allowed to respond, and a response after that was considered a no-go response. After the response time window, a blank light grey screen was shown for 300 ms. Thus, the next trial started after an intertrial interval of 1600–2200 ms (Figure 1c).

In the passive experiment, participants were asked to respond to trials in which the images were held in a red frame (Figure 1d). The red frame was superimposed on 10% of images randomly. The passive task block was given before the training phase to prevent participants from having information about the levels of categorisation, whereas they performed the other blocks after training. The order of the blocks was arranged randomly across participants. The participants were not made aware of the target stimuli.

The experiment consisted of seven blocks of 96 trials in which the same number of target and non-target images were randomly distributed. In four blocks, the target was animate (vs. inanimate); in two blocks, the target was face (vs. body); and in one block, the target was the animal face (vs. human face). The passive task, which had no specified categorisation target, was also presented in one block of 96 trials. We analysed only 48 animal face images that were identical in each categorisation level and the passive task. The animal faces as task-related targets could be regarded as animal faces (basic-F level), as faces (basic-A level) or even as animate objects (superordinate level). Each participant completed seven blocks of 96 trials, for a total of 672 trials and 72 training trials.

2.4 | EEG recordings and preprocessing

Brain electrical activity was recorded from 32 electrodes (impedance < 5 kΩ) mounted in an elastic cap (eWave32, produced by ScienceBeam; www.sciencebeam.com) and located at standard positions in accordance with the international 10–20 systems. Signals were acquired with the right mastoid reference and the ground electrode placed along the midline (AFz). The data were digitised at a sampling rate of 1000 Hz. For the analysis, we used a notch filter (50 Hz) to remove power-line noise, as well as a band-pass filter (0.1–100 Hz) to eliminate DC and high-frequency noise (FIR filter with 6 dB roll-off per octave). We split the data into 1100 ms trials, starting 100 ms before the onset of a stimulus and ending 1000 ms after stimulus onset, which was the maximum allowed response time. Since the median response times for object categorisation tasks are less than 600 ms, we consider time points ranging from −100 to 600 ms following stimulus onset. Each trial was baseline-corrected to 100 ms pre-stimulus activity. The common average reference was used to adjust the signal at each electrode by subtracting the average of all electrodes.

Data were processed using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) in Matlab (version 2019b, The Mathworks, Natick, MA). To remove artefacts, trials in which frontal electrodes' voltage changed by more than ±75 μV during the entire period were rejected. Eyeblink artefacts were excluded from the signals using an independent component analysis implemented in EEGLAB (runica algorithm). The artefactual components were selected using the ADJUST plugin (Mognon et al., 2011). On average, 1.33 trials were excluded from the analyses per subject (ranging from one to three trials per subject). All target trials (i.e., animal face), including incorrect trials, were included in subsequent analysis. Given the low number of error (less than 2% of all trials), these would unlikely affect the results. For the visual presentation, the waveforms were low-pass filtered at 30 Hz.

Our method of comparing ERP signals between two conditions was to measure the differences per participant and then average them across all participants.

2.5 | Brain topographic maps

The maps were generated from the average of ERP voltage values across subjects in a condition. The grand average activities for between-electrode areas were calculated by an interpolation method implemented in the EEGLAB. The resulted maps were superimposed on the scalp. Ten time points were selected to cover the entire time course. The time points included early ERP components (such as P100 and N170) and subsequent time points after 170 ms from the stimulus onset with the same time interval (70 ms).

2.6 | Decoding analyses

The multivariate decoding analyses with all 31 electrodes were performed using the Neural Decoding Toolbox

(Meyers, 2013). Decoding across all possible pairs of conditions (superordinate, basic-A, basic-F, and passive) was performed on each subject's data separately. All six possible pairs of conditions were decoded to determine whether at each time point the EEG signals contained information to distinguish between two conditions without making assumptions about the underlying representation. The target stimuli were the same in all conditions, but the task demands were different. Therefore, the decoding accuracy could be affected by task effects. To assess the amount of information in the EEG signals, we trained a linear classifier (Support Vector Machine [SVM]) (Chang & Lin, 2011) to determine the difference between conditions in a time-resolved manner. As input to the classifier, we considered the average in 5-ms non-overlapping time bins for each electrode. The accuracy was assessed by leave-one-out cross-validation, where the classifier was trained on all-but-one exemplar trials per condition (train on 47 × 2, test 2). The entire decoding procedure was repeated 100 times for all pair conditions. Finally, we averaged the decoding accuracy across the cross-validation splits in each of the 100 runs.

2.7 | Statistical testing

To assess the participants' accuracy and RTs, we performed repeated-measures ANOVAs (levels of categorisation) and applied paired *t*-tests as a post hoc analysis to determine the relationship between the levels of object categorisation. All reported *t*-test *p*-values were adjusted using Bonferroni correction (three pairwise comparisons, $p < 0.05/3$). The results are given as median and 25th and 75th percentiles.

At each time point, the difference between event-related potentials (ERPs) of two conditions was statistically evaluated using pairwise *t*-tests and obtaining one *p*-value. Then, to correct the results for multiple comparisons (31 electrodes × 701 time points), false discovery rate (FDR) correction was used (Storey, 2002). A *p*-value of less than 0.05 was considered significant. To estimate differences between the two levels, we measured the 95% confidence intervals at each time point using a non-parametric bootstrap sampling method in the LIMO toolbox (Pernet et al., 2011). More specifically, we created 10,000 bootstrapped samples by sampling the participants' ERPs with a replacement approach (Efron & Tibshirani, 1994). We computed the exact difference for each bootstrap sample as the original data, resulting in bootstrap estimates of difference amplitudes between two levels.

To evaluate the significance of EEG decoding signals (Figures 4c and 5d), at each time point, we performed a Wilcoxon's signed-rank test between decoding accuracy and the permutation-based chance value for all conditions (see Combrisson & Jerbi, 2015). For each time point, we randomly exchanged the class labels of original observations and calculated classification accuracy at each permutation (100 random permutations of labels, while a permutation-based test with a higher number of permutations was computationally not feasible). Next, the accuracy thresholds that correspond to the 99% percentile of the distribution (i.e., $p < 0.01$) were derived, and these new set of thresholds were used in Wilcoxon's signed-rank test. Then, to conduct multiple comparison corrections by controlling the FDR (220 time points for 5-ms non-overlapping time bins), FDR correction was applied. A *p*-value lower than 0.05 was considered statistically significant.

3 | RESULTS

3.1 | Behavioural results

3.1.1 | Effects of task demand on response time and accuracy

We first investigated the effect of task demand on behavioural outcomes as one of the primary objectives of this study. Behavioural data revealed that subjects were highly accurate at all categorisation tasks, and the percentage of incorrect responses was very low (on average, the error rate was 1.6%). The subjects' mean accuracy when target images were considered animate was 99.48 ± 1.29%, if they were seen as an animal face was 98.96 ± 1.66% and when they were perceived as a face was 96.70 ± 3.08% (Figure 2a). Despite the relatively lower accuracy of basic-A (on average about 2.5%) compared with the two other categorisation tasks, this difference was not statistically significant ($F_{(2,11)} = 4.1$, $p > 0.05$). In contrast, a comparison of the median RTs for the three types of task demands revealed considerable differences between them. When target images were categorised as animate, the RTs were significantly faster than basic-A level ($t_{11} = 5.8$, $p = 3.4 \times 10^{-4}$) and basic-F level ($t_{11} = 5.9$, $p = 3.1 \times 10^{-4}$) categorisation (Figure 2b). In addition, the superordinate level versus the basic level advantage was preserved when we investigated human faces as the target images in these two categorisation levels ($t_{11} = 3.7$, $p = 0.009$). Moreover, the median response times of participants at the two basic levels were similar ($t_{11} = 1.3$, $p = 0.6$), as were the distributions of RTs (Figure 2b,c). To determine whether re-exposure to target stimuli would prompt the familiarity effect, we assessed the RT due to the presentation order

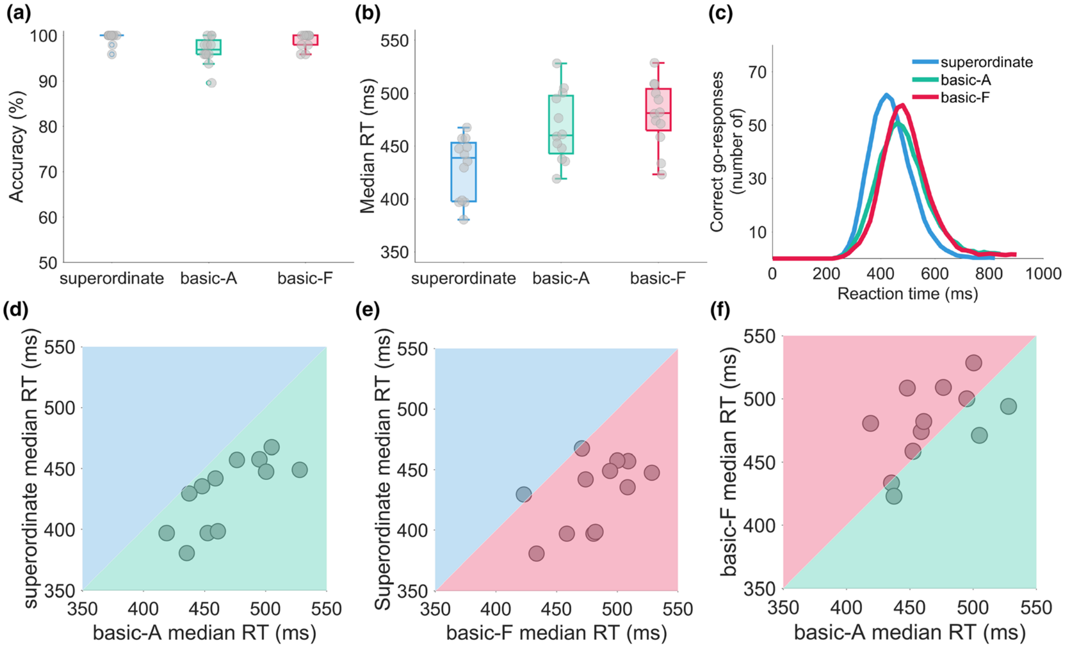

FIGURE 2 Behavioural results. (a) Each boxplot shows the median (filled line), 25th and 75th percentiles (bottom and top of the box) and whiskers (non-outlier extremes) of accuracy across participants for three object categorisation tasks (one at the superordinate level and two at the basic levels). (b) Same format as (a) but for the reaction times (RTs) of participants in these tasks. (c) Represents RT distributions obtained through computing correct go responses within 20-ms time bins across different conditions. (d–f) The median RTs are compared between the object categorisation tasks. Each dot represents one participant. Blue, green and red colours specify superordinate level, basic-A level and basic-F level, respectively

in which the stimuli were shown to subjects. There were no significant differences between RTs in different conditions (all $p > 0.19$).

The differences between categorisation levels were not restricted to median RTs. As illustrated in Figure 2c, the temporal advantage of the superordinate level could be seen from the earliest responses. The animate RT distribution was shifted toward lower latencies (i.e., it is shifted to the left) compared with the face and animal face responses. The stated effects were observed in a large majority of participants, as shown in Figure 2d,e. Based on the median RTs, it can be concluded that participants reacted faster to animate rather than faces (Figure 2d) and animal faces (Figure 2e). However, no advantage was found between the two basic levels due to the scattered median RTs on either side of the borderline in Figure 2f.

3.2 | ERP results

3.2.1 | Comparison of ERP signals across object categorisation tasks

We also recorded ERP signals from subjects performing different categorisation tasks to find out whether the task demands differentially affect the face-specific N170 component (Eimer, 2011). Results suggested that the neural modulation occurred during specific time windows (around 170 and 300 ms) after stimulus onset in different categorisation tasks. The ERP signals of two selected electrodes, with the largest amplitude around 170 ms after stimulus onset, are illustrated in Figure 3a. These two electrodes (i.e., P7 and P8) were situated bilaterally in the occipitotemporal areas. Furthermore, we assessed the differential brain activity between every two tasks by subtracting their ERP signals from each other (Figure 3b–d). Despite the lack of statistical significance in the differences between the superordinate and basic-A levels (Figure 3b) in the selected electrodes, the basic-F level was substantially different from the other two levels (Figure 3c,d).

3.2.2 | Effect of task-related information on neural representations

Since the target images appeared identically across all the visual object categorisation experiments, task-related information seems to be responsible for the observed differences between levels of categorisation. The advantage of using identical images in this study was that there were no differences between the visual information in

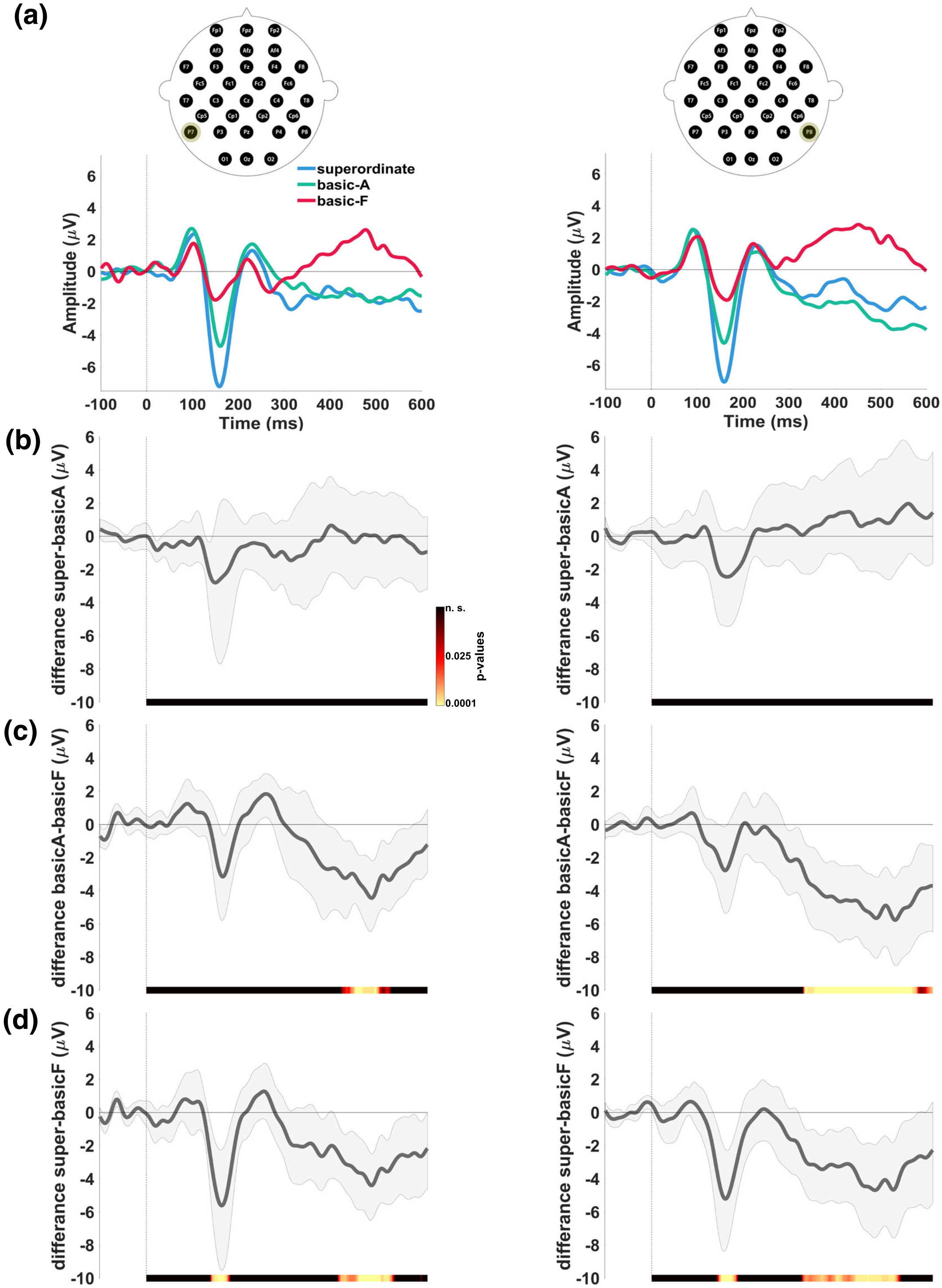

FIGURE 3 Event-related potential signals in object categorisation tasks. Panel (a) illustrates the position of two electrodes, which are selected based on previous studies to report the N170 component (P7 and P8). The analyses of each electrode's signals are shown in the column corresponding to the same electrode. The target stimuli elicit grand-averaged event-related potentials (ERPs) at distinct object categorisation tasks. Blue, green and red colours specify superordinate level, basic-A level and basic-F level, respectively. (b–d) The difference of ERP waveforms between categorisation levels. (b) Superordinate level versus basic-A level. (c) Basic-A level versus basic-F level. (d) Superordinate level versus basic-F level. The shaded areas indicate the 95% confidence interval using the bootstrap method. The colour bar indicates the colour-coding of *p*-values, which are obtained from paired *t*-test (false discovery rate [FDR]-corrected). Responses are aligned with stimulus onset (time 0), which are shown by vertical dash lines

low-level features (e.g., spatial frequency composition, luminance energy, and direction of orientation). Furthermore, this approach eliminated the effect of high-level transformations that altered the objects' appearance (e.g., pose, lighting, and occlusion). Thus, it turns out that the differences between neural representations are associated with task demands. However, other possibilities such as the role of face expectation cannot be completely ruled out due to limitations of the current experimental design. In the discussion, alternative explanations are explored in detail.

The scalp topographic map was also used to further investigate differences in brain activities between levels of categorisation over time. As Figure 4a demonstrated, the maps were similar between the superordinate and basic-A levels (Figure 4a, first and second rows) at different time points. Nevertheless, the maps of basic-F were different from those of the other two levels mainly at the 170 ms and after 310 ms post-stimulus time points. There was a weak negative bilateral occipitotemporal amplitude for the basic-F level compared with the superordinate and basic-A levels, coupled with a frontal positivity. In addition, a pronounced positive activity emerged after the 310-ms time point across centroparietal sites. According to these results, task-related information primarily affected ERP amplitudes.

To assess the neural correlates of object categorisation, we compared the participants' ERP signals for superordinate versus basic-A, basic-A versus basic-F and superordinate versus basic-F for all electrodes at each time point (Figure 4b). The *p*-values were corrected for multiple comparisons (comparisons across all electrodes and 701 time points) with $\alpha = 0.05$. The electrodes were ordered from occipital to frontal and temporal to central. We found no statistical difference between the superordinate and basic-A levels after FDR correction at all electrodes ($p > 0.072$). However, a significant difference was found between the basic-F and the other two levels that could be broadly divided into two time windows—about 170 ms and roughly after 300 ms from the stimulus onset.

We applied a multivariate decoding approach to investigate the multivariate effects of task-related information. The multivariate decoding methods can detect differences in brain activities that are lost during averaging EEG data for univariate analyses (Figure 4c). Decoding analysis was performed using a linear SVM in which the classifier was trained to decode the differences between brain signals in different categorisation tasks when participants were viewing the identical visual stimuli. We then calculated the whole brain average decoding accuracy across subjects at each time point for three paired comparisons (Figure 4c). In agreement with the ERP signals, there was no significant difference in decoding accuracies between the superordinate and basic-A levels. However, the decoding accuracies between the basic-F and two other conditions increased sharply around 170 ms after stimulus onset, followed by a plateau after 300 ms, and a gradual decline toward the end. The significant time points divided into two windows especially when comparing the basic-F with the superordinate, beginning around 140 to 275 ms and several time instances beyond 300 ms.

3.2.3 | Comparison of ERP signals in passive versus object categorisation tasks

This demonstration of task-related neural modulation leads to a question of how the brain representation of visual objects during an irrelevant task differs from when categorising those objects. To address this question, we conducted a passive experiment in which the same set of target stimuli was presented to participants. Their attentiveness during this experiment was preserved by an irrelevant task (see Section 2). The scalp topographic maps during the passive task (Figure 4a, bottom row) were different from those in the categorisation tasks specifically at 170 ms and late time-points (after 310 ms). In line with this finding, differential ERP signals for passive versus three categorisation tasks (Figure 5a,c) revealed different activities mainly emerging at these two specific time windows. Moving from the parietal to the frontal electrodes resulted in a small delay in the start of these significant events (Figure 5b). The significant event in the early time window; however, lasted longer in the frontal areas (Figure 5b). Decoding accuracy between the object categorisation tasks and the passive task (Figure 5d) also yielded peaks at these two specific time windows—around 170 ms after stimulus onset, followed by a gradual increase roughly after 300 ms.

4 | DISCUSSION

A growing set of studies have investigated the temporal dynamics of superordinate, basic and subordinate levels in object categorisation tasks (Bowers And & Jones, 2008; Grill-Spector & Kanwisher, 2005; Johnson & Olshausen, 2005; Macé et al., 2009; Poncet & Fabre-Thorpe, 2014; Wu et al., 2015). However, these studies used a stimulus set which varied in several aspects, including the number of stimuli per level, the target and non-target stimuli and some other uncontrolled parameters such as pose, lighting, rotation, occlusion, familiarity and homogeneity. Variations in stimuli could result in

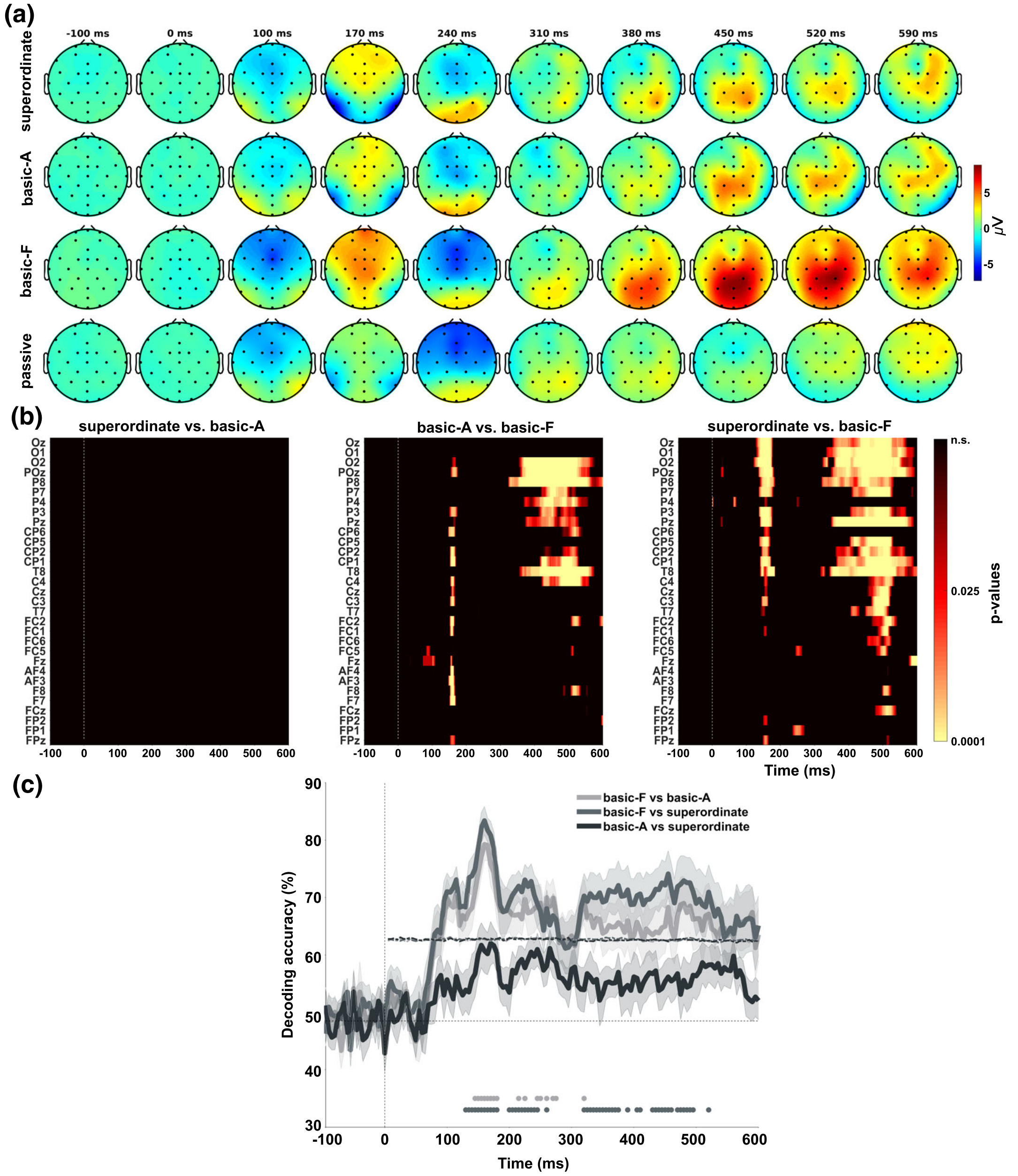

FIGURE 4 Event-related potential and decoding results in object categorisation tasks. (a) Group-level topographical maps for object categorisation and passive tasks over the whole brain. Time series of topographic maps from −100 to 590 ms with respect to the stimulus onset. (b) Paired *t*-test *p*-values for the event-related potential (ERP) differences between object categorisation tasks over all electrodes. *P*-values are corrected for multiple comparisons (comparisons across all electrodes and 701 time points). Electrodes are stacked up across the *y*-axis, and time is shown across the *x*-axis. The electrodes' order is based on their spatial positions from posterior to anterior. From left to right, the comparison is between superordinate versus basic-A, basic-A versus basic-F and superordinate versus basic-F. The colour bar indicates *p*-values. Responses are aligned with stimulus onset (time 0) and indicated by vertical dashed lines. (c) Decoding of differences between object categorisation tasks over the whole brain. The circles indicate the time points at which the decoding accuracies are significantly different from the average permutation-based chance level (grey dot-dashed lines). *P*-values are corrected for multiple comparisons. The shaded areas indicate the standard error across subjects

different behavioural outcomes. For example, Wu et al. (2015) showed a bias in saccading toward an object depending on how it is paired with another object. Participants were better at saccading to dogs (or cats) when paired with birds than saccading to birds when paired with dogs (or cats). This bias could result from an inbuilt bias toward mammals or more familiarity with household pets like dogs and cats. Thus, variations in bottom-up information could potentially affect task-related information.

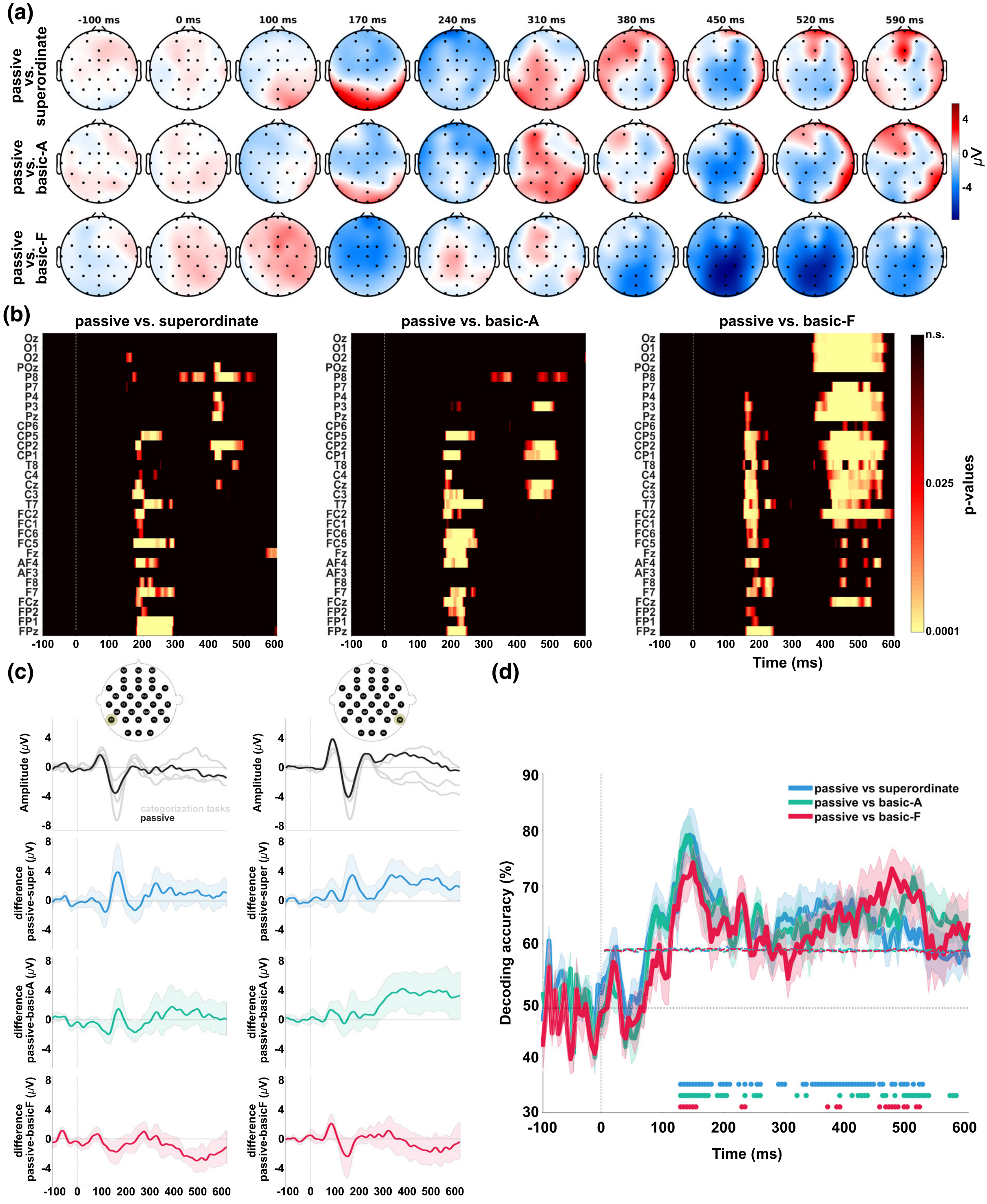

FIGURE 5 Legend on next page.

FIGURE 5 Event-related potential and decoding results of differences between the passive and object categorisation tasks. (a) Group-level topographical maps for differences between the passive and object categorisation tasks over the whole brain. Time series of topographic maps from −100 to 590 ms with respect to the stimulus onset. (b) Paired *t*-test *p*-values for the event-related potential (ERP) differences between object categorisation levels and the passive task over all electrodes. *P*-values are corrected for multiple comparisons (comparisons across all electrodes and 701 time points). Electrodes are stacked up across the *y*-axis, and time is shown across the *x*-axis. The electrodes' order is based on their spatial positions from posterior to anterior. From left to right, the comparison is between passive versus superordinate, passive versus basic-A and passive versus basic-F. The colour bar indicates *p*-values. Responses are aligned with stimulus onset (time 0) and indicated by vertical dashed lines. (c) Grand-averaged ERPs, along with the difference in ERP waveforms between passive and categorisation levels. The inset illustrates two electrodes, which are the same as shown in Figure 3 (i.e., P7 and P8 electrodes). The signals of each electrode are shown in the column corresponding to the same electrode. The shaded areas indicate the 95% confidence interval using the bootstrap method. The corresponding *p*-values are reported in (b). Responses are aligned with stimulus onset (time 0), which are shown by vertical dash lines. (d) Decoding between the passive and object categorisation tasks over the whole brain. The circles indicate the time points at which the decoding accuracies are significantly different from the average permutation-based chance level (coloured dot-dashed lines). *P*-values are corrected for multiple comparisons. The shaded areas indicate the standard error across subjects. Blue, green and red curves specify the decoding of superordinate, basic-A and basic-F versus passive, respectively

In the current study, we aimed to analyse human behaviour coupled with EEG recordings to explore the role of top-down information processing. For this purpose, we made use of the object categorisation experiments at two levels (i.e., superordinate and basic) while the target stimuli were kept the same in these tasks. Since the content and number of stimuli varied across tasks, one could argue that sensory information did not match between them. The argument could stand if we examined the entire stimuli set for each task; however, this study investigated only a subset of stimuli that was shared across all the tasks. Furthermore, to minimise the effects of other stimuli on the target stimuli, we designed the tasks so that brief presentations were followed by long interstimulus intervals. Thus, we can conclude that other stimuli had a negligible impact on our target stimuli.

We also explored the reasons for the differences in brain signals and behavioural responses across different kinds of object categorisation tasks. Our findings suggested that the identical sensory stimuli exhibited different behavioural responses and ERP patterns depending on the task contexts, indicating that object categorisation could happen at distinct perceptual and neural levels. We showed that both face/body and human/animal face tasks, which belong to the broad basic level categorisation tasks, were similar at the perceptual levels (RT profile), although they were different in terms of neural dynamics of object categorisation. A potential explanation for this finding is that each of the face and body categories is processed within distinct regions/modules of the IT cortex. However, both human and animal faces are processed within face regions, and only fine-scale differences in the neural patterns could distinguish between them.

On the other hand, the animate/inanimate task that belongs to the superordinate level categorisation task showed different RT profiles compared with the basic level tasks, although its neural representation (ERP signals) was similar to the face/body task. Both animate/inanimate and face/body tasks relied on macroscopic-level neural structures, while the human/animal face task relied on finer-scale neural structures. Furthermore, we provided evidence that the top-down signal could affect brain responses at particular time windows. The amplitude of object representations varied depending on which tasks were compared. The amplitudes of basic-F ERPs were different from those of the other tasks both at the early (around 170 ms) and late (over 300 ms) phases following stimulus onset. Previous studies demonstrated that the effect of task context on the late components of the neural signal might reflect different neural mechanisms. Hebart et al. (2018) performed different tasks on the same stimuli to show the dynamics of object representation. In addition to the task design differences, a major difference between our study and Hebart et al. (2018) is the definition of the task. The four tasks described by Hebart et al. (2018) target two low-level perceptual dimensions of images (tilt and colour) and two high-level conceptual dimensions (content and size). In addition, no comparison between the conceptual tasks was reported. However, our study has set up three distinct conceptual object categorisation tasks (one superordinate level task and two basic level tasks) and reported the differences between them. Groen et al. (2016) reported the late effect of task on scene processing by manipulating attention. Our results go further by providing new evidence about the early (perceptual) and late (top-down) effects of task demands on neural representations of object categories. Task-specific modulations in both early and late stages of visual processing are probably originated from a top-down signal. This signal may enhance task-relevant features that can lead to better performance. This interpretation is in line with the attentional enhancement of objects' features in the occipitotemporal cortex (Jehee et al., 2011). Furthermore, early components can be considered as a signature of the 'feedforward sweep' of visual information processing (Thorpe et al., 1996). The top-down signal can change brain representations that affect feedforward information processing (Gilbert & Li, 2013). Lack of this signal in the passive task might be a reason for seeing differences between the early components of passive and object categorisation tasks. Overall, our results suggest that task-related information affects both behavioural and neural representations.

4.1 | Superordinate level advantage

Our results revealed the superordinate level advantage in which the median RT was lower compared with the median RT in basic levels of object categorisation. The accuracy was high and very similar across conditions, suggesting that participants' accuracy was not influenced by task-related information. We verified that the superordinate level advantage occurs even if participants encounter exactly the same set of stimuli across different task contexts (Figure 2). The superordinate level advantage has also been reported in previous studies (Fabre-Thorpe, 2011; Macé et al., 2009; Poncet & Fabre-Thorpe, 2014; Wu et al., 2015). The superordinate level advantage was preserved when we analysed human faces as the target images in the superordinate and basic-A levels of categorisation. These results suggest that the superordinate level advantage is robust across different object categories and not the result of heterogeneity within a particular object category. In conclusion, animacy is perceived earlier than the face, regardless of whether it is an animal face or a human face.

Other potential explanations for the difference observed in the median RTs between the superordinate level task and the basic level tasks may be related to task difficulty, visual attention or motor preparation processes. However, the accuracy of subjects in the three tasks was similar, suggesting that the task difficulty and visual attention were balanced across tasks. These behavioural trends indicated that different areas of the brain may be engaged for different categorisation tasks, but it does not mean one is more difficult than another. In addition, an identical 'go/no-go' paradigm was used in the three tasks. Thus, it is unlikely that the difference between median RTs is originated from motor preparation processes.

4.2 | The difference between object categorisation and passive viewing

In the absence of any tasks, there might be a default representation for the processing of object categories. This default processing has been emphasised by studies showing that focused attention is not required to process natural scene categories (Li et al., 2002; Poncet et al., 2012). Thus, we can assume that the target images in our passive experiment are processed through access to the default representation. The ERPs and multivariate decoding analyses revealed clear signs of task-related neural modulation during object categorisation tasks compared with the passive task—though the visual stimuli were identical across all these tasks. These significant effects, from the ERP signal perspective, generally occurred in two time windows. The earliest time window started between 200 and 250 ms after the stimulus onset, though it began as early as 170 ms in some electrodes. The second significant time window started after 300 ms. Visually inspecting the topographical maps of the passive and categorisation tasks (Figure 4a) illustrated differences at the later time-points, especially after 310 ms from stimulus onset. These differences may relate to task-related information in the object categorisation tasks compared with the passive task. In the passive experiment, participants did not have prior knowledge about the presented stimuli, which produced much less brain activity compared with the object categorisation experiments. Task-related information leads to distinct brain signals during categorisation tasks compared with when there is no relevant task. These significant differences can be observed within two main time windows. However, the exact time points depend on both the electrode positioning and the type of task to be compared.

It is important to note that the difference at the early component was more dominant in the frontal electrodes compared with the occipital ones (Figure 5b), which could indicate that the modulatory signals are originated from downstream areas in the frontal lobe. Task-related neural modulation might be a general phenomenon happening across many object categorisation tasks. However, it is also possible that this modulation is weak or undetectable in some tasks, meaning that the default representation would be sufficient for processing of object categories in those tasks.

4.3 | Repetition suppression and perceptual expectation effects

One could argue that the main difference between neural modulations of object categorisation tasks was originated from the repetition of face stimuli in our experiments. When a face is preceded by another face, the absolute ERP's amplitude in particular spatiotemporal windows is reduced (Grill-Spector et al., 2006). However, the repetition suppression effects could not fully explain our results. First, despite the fact that all stimuli in the basic-F condition were from the face category, the target stimuli were selected from animal faces with a diverse range of shape geometries (such as birds, monkeys and sheep), appearances and visual transformations, which were randomly mixed with non-target stimuli. As shown previously, the repetition suppression effects usually happen when identical or highly similar human face stimuli are presented successively (see Schweinberger & Neumann, 2016 for review). Second, several studies have shown that the processing of human faces by humans might be different from the processing of other species' faces on face novelty preference (Pascalis & Bachevalier, 1998) or face categorisation tasks (Dufour et al., 2004). When we look at the faces of other species, we process them at the basic categorical level, and individuation is unlikely relevant for us, whereas the human faces can be processed at the individual level (Tanaka, 2001; Tanaka & Taylor, 1991; Haxby & Gobbini, 2011).

Furthermore, it has been shown that the N170 (or M170) component is strongly modulated by inversion (or repetition) of human faces but not animal faces (Haan et al., 2002; Itier et al., 2011; Schweinberger et al., 2007). Thus, the repetition suppression effects reported for human faces may not be generalised to animal faces. Finally, there were differences in the task design between a typical repetition suppression task and our tasks. In previous studies, the stimulus adaptation was measured over long stimulus durations and short interstimulus intervals (e.g., see Kloth et al., 2010; Kloth & Schweinberger, 2010). In one study, the stimulus adaptation was measured for three successive human face stimuli presented briefly (14 ms) (Amihai et al., 2011). In this study, the interstimulus interval was short (557 ms), and all the stimuli were presented within an oval-shaped aperture. However, in our task, a brief stimulus presentation (26 ms) was followed by a long interstimulus interval (~2 s). This task design in our experiments could reduce the potential effects of repetition suppression.

In a control analysis, we investigated the potential effects of repetition suppression. In the basic-A level task, we divided the animal face trials into two groups based on whether the previous trial was face or body. Then, the ERP signal was computed for these two groups. If our effects were related to the N170 face adaptation, we would expect a higher N170 modulation when a face was presented before an animal face (compared with when a body was presented before an animal face). However, the results indicated no significant difference between these two groups of trials.

Another potential explanation for our findings could be related to the different perceptual expectations between the blocks (i.e., the face expectation at the basic-F level compared with the face/body expectations at the basic-A level and animate/inanimate expectations at the superordinate level) (Esterman & Yantis, 2010). As the expectations' effects have been reported for different object categories (Aranda et al., 2010; Puri & Wojciulik, 2008), it is not unreasonable to consider the category expectation for other objects rather than faces, which might cancel out each other effects. Besides, Esterman and Yantis (2010) provided evidence that category expectation facilitated perceptual discrimination. Participants were faster to categorise faces or houses when the category matched with their expectations. If the face expectation effect could thoroughly explain our findings, we would anticipate a faster RT in the basic-F condition than the other two levels. However, our results demonstrated opposite effects with slower response time for the basic-F level that support the importance of task demands compared with the category expectation effect. Finally, the expectation-based modulations of neural activity generally occur before the stimulus presentation (e.g., in the form of 'baseline shift') (Kastner et al., 1999). Our reported effects are in specific time windows after stimulus onset when the expectation-related signals have been largely decayed.

For faces, there seem to be expectations across basic-A and basic-F categorisation tasks; participants expect face images in basic-F and face/body images in basic-A. Thus, if face expectation plays an essential role, we might see similar patterns in the results. However, the results do not imply similar patterns. For instance, the target images of the basic-A and superordinate conditions have similar ERP signals while their face expectation are different. It is important to note that our results cannot completely rule out the contribution of perceptual expectation due to the possibility of its confounding factor. Our interpretation is that, based on the task design, stimulus set, behavioural results and brain representations, the perceptual expectation and repetition suppression effects might have less essential roles compared with the task-related information.

Our findings showed that brain representations and behavioural responses could vary according to the tasks. Further experiments are required to uncover the mechanisms underlying task-related neural modulations. For example, in a behavioural experiment, subjects can be asked to describe a set of identical stimuli with as much detail and specificity as possible across different task contexts. Such an experiment could reveal which visual object features play an important role according to the task demand. Neural representations in different task conditions can also be investigated using deep neural networks such as Alexnet (Krizhevsky et al., 2012) and ResNet (He et al., 2016). For example, deep networks with various objective functions could be trained on the same input dataset. By visualising the learned features in the networks (see Nguyen et al., 2019 for review), one could investigate which features are important in different tasks.

5 | CONCLUSIONS

In conclusion, our study highlights the critical role of task-related information in object categorisation. The behavioural results showed that the participants' RT was influenced by task demands; a superordinate level advantage was observed in the RT profiles. The ERP results indicated that the neural modulations in the animate/inanimate and face/body conditions were similar to each other but different from the human face/animal face condition. This difference could be related to the nature of neural representations in these conditions. In both animate/inanimate and face/body tasks, each category is processed in a distinct network of cortical areas. However, human/animal faces are primarily processed within face areas, and they can be distinguished only based on a fine-scale separation in the pattern of neural responses. Finally, the differences between passive and three categorisation tasks helped us identify the temporal dynamics of task-related neural modulations. More specifically when we should expect to see task-related information.

ACKNOWLEDGEMENTS

This work was partially supported by the Shahid Rajaee Teacher Training University under contract number 30309 and the Institute for Research in Fundamental Sciences (IPM)-School of Cognitive Sciences (SCS).

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTIONS

A.F., F.F. and R.E. designed the research. F.F. and A.F. collected and analysed the data. A.F., F.F., R.R. and R.E. discussed the results. A. F, R.R. and F. F wrote the manuscript.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/ejn.15440.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding authors, R. E. and A. F., upon request.

ORCID

Reza Ebrahimpour https://orcid.org/0000-0002-7013-8078

REFERENCES

Amihai, I., Deouell, L. Y., & Bentin, S. (2011). Neural adaptation is related to face repetition irrespective of identity: A reappraisal of the N170 effect. *Experimental Brain Research*, *209*, 193–204. https://doi.org/10.1007/s00221-011-2546-x

Aranda, C., Madrid, E., Tudela, P., & Ruz, M. (2010). Category expectations: A differential modulation of the N170 potential for faces and words. *Neuropsychologia*, *48*, 4038–4045. https://doi.org/10.1016/j.neuropsychologia.2010.10.002

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämäläinen, M. S., Marinkovic, K., Schacter, D. L., & Rosen, B. R. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences*, *103*, 449–454.

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*, 1–1. https://doi.org/10.1167/13.10.1

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*, 27.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*, 455–462. https://doi.org/10.1038/nn.3635

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. *Journal of Neuroscience Methods*, *250*, 126–136. https://doi.org/10.1016/j.jneumeth.2015.01.010

Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, *16*, 763–770. https://doi.org/10.1038/nn.3381

Dehaqani, M. R. A., Vahabie, A. H., Kiani, R., Ahmadabadi, M. N., Araabi, B. N., & Esteky, H. (2016). Temporal dynamics of visual category representation in the macaque inferior temporal cortex. *Journal of Neurophysiology*, *116*(2), 587–601. https://doi.org/10.1152/jn.00018.2016

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473. https://doi.org/10.1126/science.1063414

Dufour, V., Coleman, M., Campbell, R., Petit, O., & Pascalis, O. (2004). On the species-specificity of face recognition in human adults. *Cahiers de Psychologie Cognitive-Current Psychology of Cognition*, *22*, 315–333.

Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC press. https://doi.org/10.1201/9780429246593

Eimer, M. (2011). The face-sensitive N170 component of the event-related brain potential. *The Oxford Handbook of Face Perception*, *28*, 329–344.

Emadi, N., & Esteky, H. (2014). Behavioral demand modulates object category representation in the inferior temporal cortex. *Journal of Neurophysiology*, *112*, 2628–2637. https://doi.org/10.1152/jn.00761.2013

Erez, Y., & Duncan, J. (2015). Discrimination of visual categories based on behavioral relevance in widespread regions of frontoparietal cortex. *Journal of Neuroscience*, *35*, 12383–12393. https://doi.org/10.1523/JNEUROSCI.1134-15.2015

Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, *20*, 1245–1253. https://doi.org/10.1093/cercor/bhp188

Fabre-Thorpe, M. (2011). The characteristics and limits of rapid visual categorization. *Frontiers in Psychology*, *2*, 243.

Ghodrati, M., Farzmahdi, A., Rajaei, K., Ebrahimpour, R., & Khaligh-Razavi, S.-M. (2014). Feedforward object-vision models only tolerate small image variations compared to human. *Frontiers in Computational Neuroscience*, *8*, 74.

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Grill-Spector, K., & Kanwisher, N. (2005). Visual Recognition. *Psychological Science*, *16*, 152–160. https://doi.org/10.1111/j.0956-7976.2005.00796.x

Groen, I. I., Ghebreab, S., Lamme, V. A., & Scholte, H. S. (2016). The time course of natural scene perception with reduced attention. *Journal of Neurophysiology*, *115*, 931–946. https://doi.org/10.1152/jn.00896.2015

Haan, M. D., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209. https://doi.org/10.1162/089892902317236849

Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences*, *111*, 962–971.

Harel, A., Ullman, S., Harari, D., & Bentin, S. (2011). Basic-level categorization of intermediate complexity fragments reveals top-down effects of expertise in visual perception. *Journal of Vision*, *11*, 18–18. https://doi.org/10.1167/11.8.18

Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. J. Calder, G. Rhodes, M. H. Johnson & J. V. Haxby (Eds.), *Handbook of Face Perception* (pp. 93–110). Oxford: Oxford University Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition*. arXiv preprint arXiv:1512.03385.

Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. *eLife*, *7*, e32816. https://doi.org/10.7554/eLife.32816

Hong, H., Yamins, D. L., Majaj, N. J., & DiCarlo, J. J. (2016). Explicit information for category-orthogonal object properties increases along the ventral stream. *Nature Neuroscience*, *19*, 613–622. https://doi.org/10.1038/nn.4247

Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2013). The dynamics of invariant object recognition in the human visual system. *Journal of Neurophysiology*, *111*, 91–102.

Itier, R. J., Van Roon, P., & Alain, C. (2011). Species sensitivity of early face and eye processing. *NeuroImage*, *54*, 705–713. https://doi.org/10.1016/j.neuroimage.2010.07.031

Jehee, J. F., Brady, D. K., & Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *Journal of Neuroscience*, *31*, 8210–8219. https://doi.org/10.1523/JNEUROSCI.6153-09.2011

Johnson, J. S., & Olshausen, B. A. (2005). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*, 4.

Bowers And, J. S., & Jones, K. W. (2008). Short article: Detecting objects is easier than categorizing them. *Quarterly Journal of Experimental Psychology*, *61*, 552–557.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *361*, 2109–2128. https://doi.org/10.1098/rstb.2006.1934

Karimi-Rouzbahani, H., Vahab, E., Ebrahimpour, R., & Menhaj, M. B. (2019). Spatiotemporal analysis of category and target-related information processing in the brain during object detection. *Behavioural Brain Research*, *362*, 224–239. https://doi.org/10.1016/j.bbr.2019.01.025

Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*(4), 751–761. https://doi.org/10.1016/S0896-6273(00)80734-5

Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *Journal of Neurophysiology*, *97*, 4296–4309. https://doi.org/10.1152/jn.00024.2007

Kloth, N., & Schweinberger, S. R. (2010). Electrophysiological correlates of eye gaze adaptation. *Journal of Vision*, *10*, 17–17. https://doi.org/10.1167/10.12.17

Kloth, N., Schweinberger, S. R., & Kovács, G. (2010). Neural correlates of generic versus gender-specific face adaptation. *Journal of Cognitive Neuroscience*, *22*, 2345–2356. https://doi.org/10.1162/jocn.2009.21329

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*, 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*, 1097–1105.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*, 9596–9601.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213.

Löw, A., Bentin, S., Rockstroh, B., Silberman, Y., Gomolla, A., Cohen, R., & Elbert, T. (2003). Semantic categorization in the human brain: Spatiotemporal dynamics revealed by magnetoencephalography. *Psychological Science*, *14*, 367–372. https://doi.org/10.1111/1467-9280.24451

Macé, M. J. M., Joubert, O. R., Nespoulous, J. L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: You spot the animal faster than the bird. *PLoS ONE*, *4*, e5927. https://doi.org/10.1371/journal.pone.0005927

Mack, M. L., & Palmeri, T. J. (2015). The dynamics of categorization: Unraveling rapid categorization. *Journal of Experimental Psychology: General*, *144*, 551–569. https://doi.org/10.1037/a0039184

Martinovic, J., Gruber, T., & Müller, M. M. (2008). Coding of visual object features and feature conjunctions in the human brain. *PLoS ONE*, *3*, e3781. https://doi.org/10.1371/journal.pone.0003781

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89–115.

Meyers, E. (2013). The neural decoding toolbox. *Frontiers in Neuroinformatics*, *7*, 8.

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, *48*, 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Mohan, K., & Arun, S. P. (2012). Similarity relations in visual search predict rapid visual categorization. *Journal of Vision*, *12*, 19–19. https://doi.org/10.1167/12.11.19

Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition*, *18*, 407–418. https://doi.org/10.3758/BF03197129

Nguyen, A., Yosinski, J., & Clune, J. (2019). Understanding neural networks via feature visualization: A survey. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (pp. 55–76). Springer. https://doi.org/10.1007/978-3-030-28954-6_4

Pascalis, O., & Bachevalier, J. (1998). Face recognition in primates: A cross-species study. *Behavioural Processes*, *43*, 87–96. https://doi.org/10.1016/S0376-6357(97)00090-9

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*, 603–608. https://doi.org/10.1152/jn.00513.2004

Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A toolbox for hierarchical LInear MOdeling of ElectroEncephaloGraphic data. *Computational Intelligence and Neuroscience*, *2011*, 3.

Poncet, M., & Fabre-Thorpe, M. (2014). Stimulus duration and diversity do not reverse the advantage for superordinate-level representations: The animal is seen before the bird. *European Journal of Neuroscience*, *39*, 1508–1516. https://doi.org/10.1111/ejn.12513

Poncet, M., Reddy, L., & Fabre-Thorpe, M. (2012). A need for more information uptake but not focused attention to access basic-level representations. *Journal of Vision*, *12*, 15–15. https://doi.org/10.1167/12.1.15

Praß, M., Grimsen, C., König, M., & Fahle, M. (2013). Ultra rapid object categorization: Effects of level, animacy and context. *PLoS ONE*, *8*, e68051. https://doi.org/10.1371/journal.pone.0068051

Puri, A. M., & Wojciulik, E. (2008). Expectation both helps and hinders object perception. *Vision Research*, *48*, 589–597. https://doi.org/10.1016/j.visres.2007.11.017

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience*, *25*, 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005

Schweinberger, S. R., Kaufmann, J. M., Moratti, S., Keil, A., & Burton, A. M. (2007). Brain responses to repetitions of human and animal faces, inverted faces, and objects—An MEG study. *Brain Research*, *1184*, 226–233. https://doi.org/10.1016/j.brainres.2007.09.079

Schweinberger, S. R., & Neumann, M. F. (2016). Repetition effects in human ERPs to faces. *Cortex*, *80*, 141–153. https://doi.org/10.1016/j.cortex.2015.11.001

Sofer, I., Crouzet, S. M., & Serre, T. (2015). Explaining the timing of natural scene understanding with a computational model of perceptual categorization. *PLoS Computational Biology*, *11*, e1004456. https://doi.org/10.1371/journal.pcbi.1004456

Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *64*, 479–498. https://doi.org/10.1111/1467-9868.00346

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature*, *400*, 869–873. https://doi.org/10.1038/23703

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General*, *130*, 534–543. https://doi.org/10.1037/0096-3445.130.3.534

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*, 457–482. https://doi.org/10.1016/0010-0285(91)90016-H

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522. https://doi.org/10.1038/381520a0

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. (2003). Faces and objects in macaque cerebral cortex. *Nature Neuroscience*, *6*, 989–995. https://doi.org/10.1038/nn1111

VanRullen, R. (2007). The power of the feed-forward sweep. *Advances in Cognitive Psychology*, *3*, 167–176. https://doi.org/10.2478/v10053-008-0022-3

Vaziri-Pashkam, M., & Xu, Y. (2017). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. *Journal of Neuroscience*, *37*, 8767–8782. https://doi.org/10.1523/JNEUROSCI.3392-16.2017

Wu, C.-T., Crouzet, S. M., Thorpe, S. J., & Fabre-Thorpe, M. (2015). At 120 msec you can spot the animal but you don't yet know it's a dog. *Journal of Cognitive Neuroscience*, *27*, 141–149. https://doi.org/10.1162/jocn_a_00701

How to cite this article: Farzmahdi, A., Fallah, F., Rajimehr, R., & Ebrahimpour, R. (2021). Task-dependent neural representations of visual object categories. *European Journal of Neuroscience*, 1–18. https://doi.org/10.1111/ejn.15440