The Effect of Task on Object Processing revealed by EEG decoding

Hoi Ming Ken Yip[1]
Leo Y. T. Cheung[2]
Yetta Kwailing Wong[2]
Alan C.-N. Wong[1*]

[1]Department of Psychology
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

[2]Department of Educational Psychology, Faculty of Education
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong

Running head: TASK EFFECT ON OBJECT PROCESSING

Keywords: object recognition, attention, top-down effect, task, MVPA

Corresponding author:
Alan C.-N. Wong
334 Sino Building
Department of Psychology
The Chinese University of Hong Kong, Shatin, Hong Kong
Email: alanwong@cuhk.edu.hk
Phone: +852-3943-6505

Abstract

At which phase(s) does task demand affect object processing? Previous studies showed that task demand affects object representations in higher-level visual areas but not so much in earlier areas. There are, however, limitations in those studies concerning the relatively weak manipulation of task due to the use of familiar real-life objects, and/or the low temporal resolution in brain activation measures such as fMRI. In the current study, observers categorized images of artificial objects in one of two orthogonal dimensions, shape and texture. Electroencephalogram (EEG), a technique with higher temporal resolution, and multivariate pattern analysis (MVPA) were employed to reveal object processing across time under different task demands. Results showed that object processing along the task-relevant dimension was enhanced starting from a relatively late time (~230ms after image onset), within the time range of the event-related potential (ERP) components N170 and N250. The findings are consistent with the view that task exerts an effect on object processing at the later phases of processing in the ventral visual pathway.

1. Introduction

Object recognition and categorization depend greatly on the goal of the observer, and yet its underlying mechanism has not been well understood. On the one hand, feedforward theories portrait object processing as a hierarchy in the ventral visual pathway, with representations increasing in receptive field size, complexity, and invariance going from the earlier areas like the primary visual cortex (V1) to later areas in the inferior temporal cortex (IT) (DiCarlo et al., 2012; Grill-Spector et al., 2001; Kriegeskorte et al., 2008). Several computational models of object representation proposed that the processing from V1 to IT is largely feedforward and not influenced by top-down factors such as task (Riesenhuber & Poggio, 2002; Serre et al., 2005; Serre et al., 2007; DiCarlo et al., 2012).

On the other hand, the role of top-down factors in visual object representations is conceivable in more recent perception models with an element of predictive coding (Friston, 2010; Hohwy, 2013; Clark, 2013; Lupyan 2015; Macpherson, 2017; O'Callaghan et al., 2017; Hohwy, 2017; Vance & Stokes, 2017; Ransom et al., 2017). These models share the same basic principle of prediction error minimization: higher-level areas generate predictions and send them to lower-level areas, which compare the predictions with the incoming information and send the prediction errors (difference between the two) back to the higher-level areas for adjusting the predictions. This process is repeated at both high (e.g., whole object representations) and low levels (e.g., edge orientations). Crucially, low-level predictions can

be influenced by top-down factors such that experience, knowledge, and expectations in the form of predictions passed to the lower-level areas (Lupyan, 2015). Accordingly, task can affect object representations even in the earlier phases of visual processing.

Previous studies have examined the effect of task demand on object representations in the visual system. A number of event-related potential (ERP) studies have hinted at the effect of task on face representations (Valdés-Conroy et al., 2014; Itier, & Neath-Tavares, 2017; Goffaux et al., 2003). In these studies, participants were asked to perform different tasks on the same faces, and the difference in the amplitude of ERP component across tasks was examined. For example, Goffaux et al. (2003) found an enhanced amplitude in N170 while participants were categorizing faces in the gender dimension, compared with the familiarity dimension. This effect occurred only in filtered images with low but not high spatial frequency information intact.

In recent years, multivariate pattern analysis (MVPA) has been increasingly used to examine how visual object representations unfold in time (for a detailed introduction of the technique, see Grootswagers et al., 2017; Haynes, 2015; Schwarzkopf & Rees, 2011). A specific type of MVPA is decoding, which involves the use of machine learning algorithms to infer, based on the pattern of brain activations across space (e.g., channels for EEG, voxels for fMRI) and time, which experimental condition is being introduced. Classification accuracy indicates the success of such inference and thus how differently the brain responds

to two conditions. The technique is useful for examining how the brain represents different objects, as object representations can be regarded as the spatio-temporal patterns of activations corresponding to different objects. Hence, the degree to which the activation patterns between objects are distinguishable under different task contexts can reflect the effect of task demand on object representations.

MVPA complements traditional neuroimaging analyses by offering additional information. First, it takes the variability of neural responses across sensors or brain regions into consideration, while traditional, univariate analysis focuses on the response averaged across a smaller, more localized set of channels or voxels. Second, findings of the traditional, activation-based approach are often difficult to interpret. For example, when two conditions differ in the magnitude of an ERP component, it is difficult to interpret this as necessarily stronger activation for one condition vs. the other, given the extraneous factors (e.g., changes in dipole position) that contribute to the direction of amplitude difference. With both fMRI and ERP, interpreting amplitude differences is also tricky, as higher activity in one condition than another could indicate a higher neural engagement or lower processing efficiency for that same condition. In contrast, MVPA is information-based in that it informs us how much information brain activations carry for differentiating between different conditions, and so interpreting its findings is relatively straightforward. It is therefore more informative to use MVPA to examine the effect of task demand on object representations.

1.1 MVPA studies of task effect on object processing

In the past few years, several studies applied MVPA to examine task effect on object processing. Most of these studies used functional magnetic resonance imaging (fMRI) and reported mixed findings. For example, by computing the correlations of different objects in either the same task or different tasks, Harel et al. (2014) observed that object identity information differed in the posterior fusiform gyrus (pFs) but not in the lateral occipital cortex (LO). In Bracci et al. (2017), while analyses with the representational dissimilarity matrices (RDM) suggested a task effect in the ventral but not lateral occipitotemporal cortex, the correlation between the response patterns for objects remained similar regardless of whether the task was the same or different in both regions. Kim & McCarthy (2016) reported a task effect on decoding face versus body images in the occipitotemporal fusiform cortex, but this finding came with a counter-intuitive observation that the face-body decoding was *less* accurate when participants were performing a face-body discrimination task compared with an irrelevant emotion discrimination task. Jackson et al. (2017) reported above-chance decoding in the lateral occipital complex (LOC) only when the task was relevant, although the difference in decoding performance between task-relevant and task-irrelevant conditions was not statistically significant. Two other studies did not observe any significant task effect (Long & Kuhl, 2018; Vaziri-Pashkam & Xu, 2017).

Results remained mixed with techniques that are temporally more sensitive. Using magnetoencephalography (MEG), Hebart et al. (2018) observed a task effect on object decoding performance only after 542ms, suggesting that task affects object representation only at a late time. However, using EEG and applying time-frequency analyses, decoding performance was higher in a task-relevant than a task-irrelevant condition during a much earlier time frame (100-600ms after stimulus onset), but it was the case only for one of the two tasks included in that study (Bocincova & Johnson, 2019).

In sum, the current evidence of task influence on object processing remained suggestive. In particular, the timing of this task effect remains unclear. The majority of previous studies adopted fMRI, with low temporal resolution in the order of seconds. It is thus possible that the task effect found in these studies reflected late feedback of high-level processing after object representations have been formed. While this would still be an evidence for the task effect, the effect may better be regarded as modification of existing object representations rather than influences on the formation of the representations. Previous fMRI studies found that attention and expectation altered responses in primary visual cortex (V1) and claimed that the earliest visual processing was susceptible to top-down modulation (e.g. Kok et al., 2012; Jehee et al., 2011). Accordingly, we should expect a very early effect of attention and expectation on visual processing in an EEG study. Incongruent with this prediction, however, Alilović et al. (2019) found no effect of attention and prediction on the amplitude of C1, the earliest ERP component

that originates from V1 and peaks before 100ms. MVPA failed to show the modulation of attention and prediction until 133ms and 242ms respectively. Whether the effect of task would show a similar time course remains unknown.

The mixed findings in previous studies might also be caused by the use of real-life objects (Kim and McCarthy, 2016; Long & Kuhl, 2018; Harel et al., 2014; Bracci et al., 2017; Vaziri-Pashkam & Xu, 2017; Hebart et al., 2018). In daily life, we often perform multiple tasks on the same object simultaneously (e.g. recognizing the gender and emotion of a face when you want to know if a girl is happy). Participants in those studies may have thus failed to focus on just one dimension as instructed. The resulting weakening of task manipulation might have contributed to the mixed findings in the literature. Consistently, the two studies using artificial stimuli (Jackson et al., 2017; Bocincova & Johnson, 2019) did show some support for the task-dependent object representations.

1.2 The current study

The current study aimed at advancing understanding of the effect of task on object processing in two ways. First, we used EEG in order to more precisely measure when the task effect emerges. Second, we introduced a stronger manipulation of task by using novel, artificial objects that varied in two orthogonal dimensions. This should increase the likelihood that participants attended to only one dimension as instructed.

We applied MVPA to the EEG response of participants performing a verification task

on the objects based on one of two possible dimensions (shape and texture). If there is an effect of task on object representations, then we should expect a higher object decoding accuracy when the decoded dimension (e.g., texture) is task-relevant than when the dimension is task-irrelevant, i.e., the participant's task was to judge an object along the same dimension (e.g., texture) compared with a different dimension (e.g., shape). This would point to a higher quality of information carried in the activation patterns in a relevant task context. In addition, we should also expect that the decoding model trained in one task context would generalize better to data in the same task context than a different one. This would correspond to a difference in the type of information utilized for object processing in different task contexts. The time at which such task-related decoding differences occur would indicate the time-course of task effect.

2. Method

2.1 Participants

Forty college students with normal or corrected-to-normal vision were recruited. Four participants were excluded due to low behavioral accuracy (<90%) or excessive artifacts (>20% of total trials rejected in artifact rejection; see below). Data of the remaining 36 participants (17 males, mean age = 21.83 years, SD = 3.47) were included in all subsequent analyses. All participants gave written consent according to the guidelines of the Survey and Behavioral Research Ethics Committee of the Chinese University of Hong Kong and the

Joint Chinese University of Hong Kong-New Territories East Cluster Clinical Research Ethics Committee, and received monetary compensation for their participation.

2.2 Materials and stimuli

Artificial objects were generated and modified based on the computer-generated artificial objects (‘Greebles’ and ‘Ziggerins’) used in previous studies (Figure 1; Gauthier & Tarr, 1997; Wong et al., 2009). The objects were in an identical brown color and varied in two dimensions: shape and texture. The luminance levels were controlled across the object groups. The objects were sized such that a circle enclosing each object would span a visual angle of 4° X 4°.

A PC with Matlab (Mathworks, Inc.) and Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) was used to present stimuli and record behavioral responses. Stimuli were presented on a 27” LCD monitor with a refresh rate of 60Hz. The background was grey and the viewing distance was 70cm.

2.3 Procedure

A verification task was used in this study (Figure 2). On each trial, a cue (“large”, “small”, “round” or “cubic”) was presented for 500ms, followed by a black fixation cross for 1000ms. Then an object was presented for 1000ms. Participants were asked to press “1” on the keyboard if the object matched with the cue and “2” if it did not. Trials with “large” or “small” cues required judgment about the texture (i.e., a texture task) and trials with “round”

or “cubic” cues required judgment about the shape (i.e., a shape task). Participants were instructed to respond as accurately and as quickly as possible within the 1000ms when the object appeared, and the object remained present after the response. Following the presentation of the object, a white fixation cross was presented for 2000ms. Participants were instructed to blink only during this period.

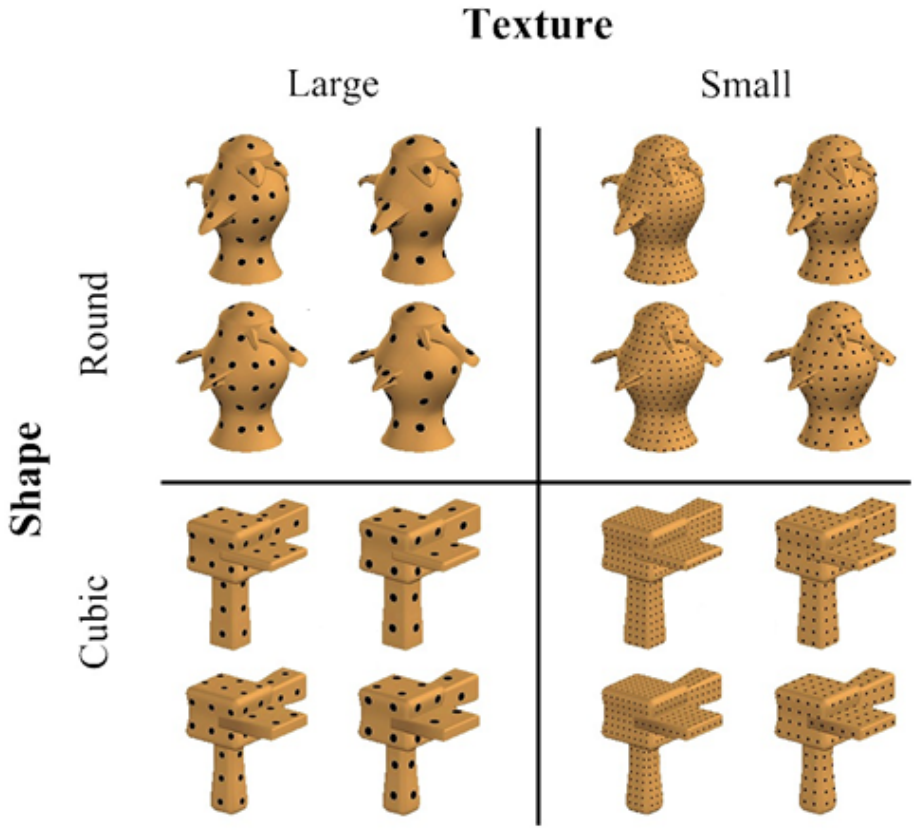


Figure 1. The four groups of objects varying along the shape and texture dimensions. The brightness was higher in this figure than in the actual experiment for better visualization.

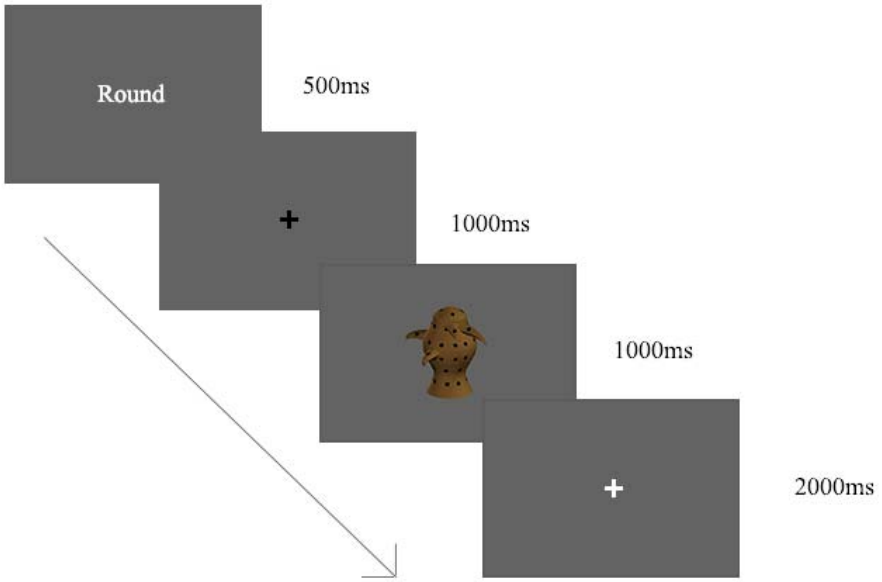


Figure 2. The experimental paradigm. The task was to judge whether the object matched with the cue.

There were 32 blocks and 32 trials in each block, leading to 1024 trials in total. Participants were advised by the experimenter to get at least 30 out of 32 trials correct in each block and received feedback on the number of correct trials after each block. Each mapping of cue and object groups (4 cue X 4 object groups = 16 pairs) was presented twice in a block. The selection of objects within each object group was random. The trial sequence within a block was pseudo-randomized with the following criteria: (i) The same cue did not repeat for more than 3 consecutive trials; (ii) The same object group did not repeat for more than 3 consecutive trials; Sixty-four different trial sequences were generated and 32 of them were randomly selected for each participant. There was a 12-trial short practice before the experiment. Participants were required to obtain 90% accuracy to proceed. Otherwise they

would repeat the practice until they attained the required accuracy.

2.4 EEG recording and Preprocessing

EEG data were recorded on a 128-channel eego system (ANT Neuro, Enschede, The Netherlands) with a 1000-Hz sampling rate. CPz was used as the online reference and the ground electrode was placed close to the left mastoid. Electrode impedance was kept under 20kΩ during the experiment.

Preprocessing was conducted with EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) in MATLAB (Mathworks, Natick, MA). Data were re-referenced off-line to the average reference. Band-pass filtering (0.01–30Hz) was applied to the continuous EEG data, which were then divided into 1000-ms epochs starting at 200ms before and ending at 800ms after the presentation of the object. Since the signals near the end of the stimulus presentation was not of interest, the last 200ms of stimuli presentation (i.e., 800-1000ms) was not used for artifact rejection to avoid rejecting epochs due to artifacts during this time window. Incorrect trials were also excluded from further analyses. Epochs with ocular artifacts were removed by visual inspection and by the moving-window peak-to-peak function in ERPLAB on VEOG, HEOG and the channels selected in the decoding analysis with a threshold of 100 μV, a window size of 200ms and a step size of 50ms. On average, 4.57% and 6.24% of the trials were rejected due to incorrect responses and artifacts respectively.

2.5 Multivariate Pattern Analysis (MVPA)

MVPA was implemented in MATLAB with custom-made codes and LIBSVM (Chang & Lin, 2011) using a linear support vector machine (c = 1). Classification was conducted in each participant separately with a sliding time window approach, using a window size of 50ms and a step size of 10ms. A subset of 67 posterior channels (Figure 3) was selected for the MVPA analysis based on the results of a pilot study, which showed that the general decoding accuracy was better with these channels than with all channels, regardless of conditions. Principle Component Analysis (PCA) was applied to the 50 (time points) X 67 (channels) features in each time window to reduce dimensionality. Principle components (PCs) explaining >99% variance were kept (average number of PCs = 83.25, SD=13.79) and used as input to the classifier. A ten-fold cross validation was then conducted. The trials for classification were randomly assigned to ten bins of roughly equal number of trials and equal proportion of the two classes. Nine bins were pooled to train the classifier and the remaining one bin was used to test its classification accuracy. This step was repeated ten times so that every bin would serve as the test bin once. This cross validation procedure was repeated five times. The resulting accuracy of a classifier in a time window was the average accuracy of these five iterations of ten-fold cross validation.

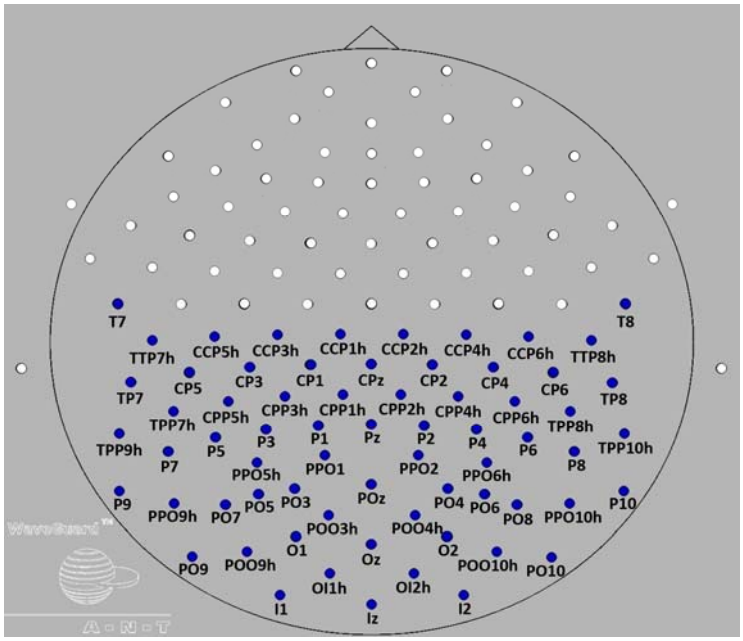


Figure 3. Channels used in MVPA were highlighted in blue.

Two major analyses were conducted to examine the effect of task on object processing. The first analysis involved comparison between the performance of task-relevant and task-irrelevant classifiers (Table 1). Task-relevant classifiers refer to the classifiers trained to discriminate an object dimension (e.g., texture) congruent with the participants' task demand (e.g., as in a texture task trial). In contrast, task-irrelevant classifiers refer to the classifiers trained to discriminate an object dimension (e.g., texture) incongruent with the participants' task demand (e.g., as in a shape task trial). If the task demand could render the object representations more separable along the dimension required by the task (e.g., a larger difference between the representations of Large-Round and Small-Round objects when participants were performing a texture vs. shape task), then the task-relevant classifiers would demonstrate higher average decoding accuracy than the task-irrelevant classifiers.

Table 1. The three types of classifiers involved in the current study. Dimension values relevant to the classification were highlighted in bold.

Types of Classifiers	*Type of trials classifiers trained on*	*Type of trials classifiers tested on*
Task-relevant classifiers		
A) Texture classification		
Large-Round vs. **Small**-Round	Texture task	Texture task
Large-Cubic vs. **Small**-Cubic	Texture task	Texture task
B) Shape classification		
Large-**Round** vs. Large-**Cubic**	Shape task	Shape task
Small-**Round** vs. Small-**Cubic**	Shape task	Shape task
Task-irrelevant classifiers		
A) Texture classification		
Large-Round vs. **Small**-Round	Shape task	Shape task
Large-Cubic vs. **Small**-Cubic	Shape task	Shape task
B) Shape classification		
Large-**Round** vs. Large-**Cubic**	Texture task	Texture task
Small-**Round** vs. Small-**Cubic**	Texture task	Texture task
Across-task classifiers		
A) Texture classification		
Large-Round vs. **Small**-Round	Texture task	Shape task
Large-Cubic vs. **Small**-Cubic	Texture task	Shape task
B) Shape classification		
Large-**Round** vs. Large-**Cubic**	Shape task	Texture task
Small-**Round** vs. Small-**Cubic**	Shape task	Texture task

Another major analysis concerned whether the classifiers trained on one task can be generalized to another task. Performance of task-relevant and across-task classifiers was compared. Both types of classifiers were trained in the object dimension congruent with that of the participants' task demand, however they were tested on data in different tasks. Task-relevant classifiers were tested on data in the same task with which the classifiers were

trained, while across-task classifiers were tested on data in a different task from which the classifiers were trained, e.g., the Large-Round vs. Small-Round classifiers were trained on the texture task trials but tested on the shape task trials. If the features used for distinguishing the same objects differ between tasks, then what the classifiers learnt in one task would not be fully generalized to another task. This would result in a lower average decoding accuracy in the across-task classifiers than the task-relevant classifiers.

Apart from the decoding accuracy, probability estimates generated in the LIBSVM algorithm (Chang & Lin, 2011) were also examined. The algorithm used Platt scaling to estimate the probability that a classifier would favor one class in each trial. The result pattern was similar and was reported in the Appendix.

To perform significance testing on decoding accuracies, we used a non-parametric sign permutation test with a maximum cluster size method to correct for multiple comparisons (Nichols & Holmes, 2002; Hebart et al. 2018). First, we calculated a difference score at each time window for each participant. To test if the decoding accuracies of a classifier were above chance, we subtracted 0.5 from the accuracies. To test if the decoding accuracies of two classifiers were different, we subtracted the accuracies of one from another. Two-tailed tests with participants as the random factor were conducted to test whether the difference scores were significant. A permutation procedure was used in which each participant was randomly assigned a value of 1 or -1 and the difference scores at each time

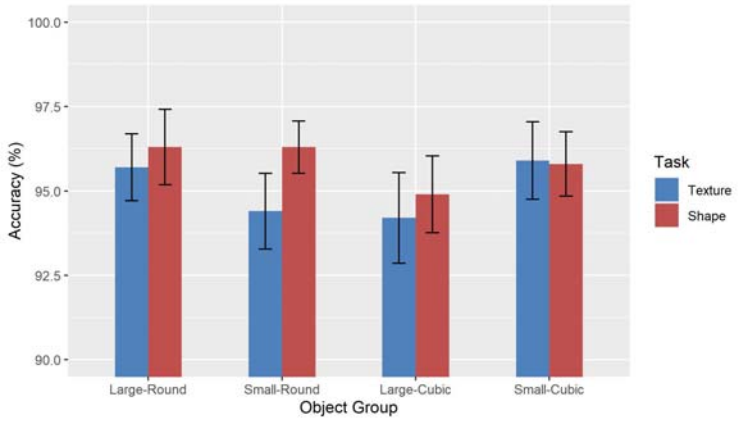
window was multiplied by this value. For each time window, a t-value was computed with the mean and SD of this score across participants. Repeating the procedure for 10,000 times, a distribution of 10,000 t-values was obtained for every time window. The primary threshold was set at the 2.5th and 97.5th percentile of this distribution. To obtain the secondary cluster-size threshold, we searched for clusters of significant time windows. Neighboring time windows would form a cluster only if all their t-values exceeded the primary threshold. The maximum cluster size of each permutation set was recorded, and it formed a distribution of maximum cluster sizes. The 97.5th percentile of this distribution served as a secondary cluster-size threshold. Finally, both thresholds were applied to the difference score of each time window. The difference score in a time window would be regarded as statistically significant if the t-value of the empirical data was higher than that of the primary threshold and if the size of the cluster encompassing this time window was larger than the secondary cluster-size threshold.

3. Results

3.1 Behavioral Results

Overall responses were accurate (>93%) and fast (570ms). Figure 4 shows the accuracy and response time (RT) for each object group in each task. We performed a 2 (task: texture, shape) X 4 (object groups: large & round, small & round, large & cubic, small & cubic) repeated measure analyses of variance (ANOVA) to test if there were significant

differences in the accuracy and response time (RT) between task and object groups. Incorrect trials were excluded from the analysis of RT. Participant responses were less accurate and slower for the texture task than the shape task. This was indicated by a significant main effect of task in accuracy, $F(1,35) = 4.606$, $p = .039$, $\eta_p^2 = .116$, and in RT, $F(1,35) = 49.343$, $p < .001$, $\eta_p^2 = .585$. The main effect of object group was significant in accuracy, $F(3,33) = 4.149$, $p = .013$, $\eta_p^2 = .274$, and in RT, $F(3,33) = 9.973$, $p < .001$, $\eta_p^2 = .476$. Post-hoc pairwise comparisons (Bonferroni corrected) revealed a significantly higher accuracy for Large-Round than Large-Cubic objects, $p = .009$. Pairwise comparisons also revealed faster responses for Large-Round than the other object groups, $p \leq .032$. The interaction of task and object group was significant in accuracy, $F(3,33) = 3.462$, $p = .027$, $\eta_p^2 = .239$, and not in RT, $F(3,33) = 1.352$, $p = .275$, $\eta_p^2 = .109$. Subsequent t-tests comparing accuracies of the same object groups in the two tasks found that the accuracy of Small-Round in the texture task was significantly lower than that in the shape task $p = .001$.



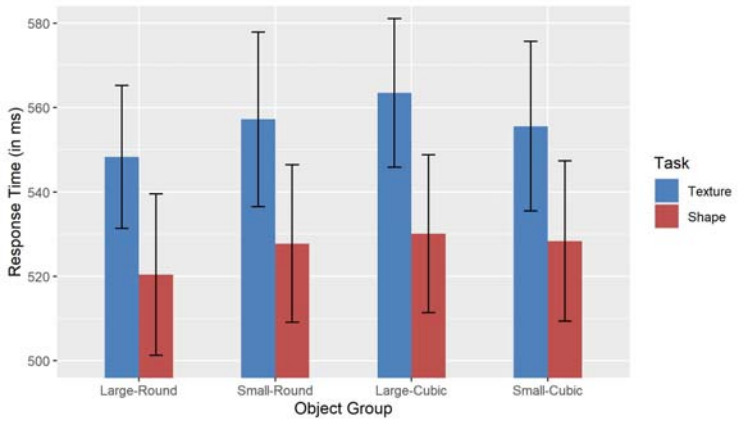


Figure 4. Behavioral accuracy and response time for each object group in each task. Error bars represent 95% confidence interval of the mean estimate.

3.2 EEG Results

Figure 5 shows the grand average ERP collapsed across conditions and participants for different time points and channels. Occipitotemporal channels (e.g., O1, O2, PO7, PO8) showed clear deflections corresponding to P1, N170 and N250 components.

In the first analysis, performance of task-relevant classifiers was compared with that of task-irrelevant classifiers to test if participants' task demand would influence object processing (Figure 6). Performance of both classifiers became significantly above chance very early in time (task-relevant: starting at the 60-110ms time window; task-irrelevant: starting at 50-100ms) and remained significant until the end of the epoch. Performance peaked at 130-180ms in both classifiers. Accuracies of task-relevant classifiers were significantly higher than task-irrelevant classifier at 250ms-340ms (indicated by the green dots). Separate analyses for each decoding (texture or shape decoding) showed that the task effect was significant for texture decoding only (Figure 6, bottom panel). Two clusters were

identified at 320-440ms and 410-580ms.

In the second analysis, performance of task-relevant classifiers was compared with that of across-task classifiers to test how well the trained classifiers generalized to another task (Figure 7). The across-task decoding accuracies were significantly above chance (indicated by red dots) since 50-100ms and lasted until the end of the epoch. Performance peaked at 130-180ms for both classifiers. Accuracies of task-relevant classifiers were significantly higher than that of across-task classifiers in several time clusters (230ms-340ms, 300-410ms, 410-550ms, 550-640ms, indicated by green dots). Separate analyses for each task showed that task effect was significant for both texture (300-410ms, 380-570ms) and shape decoding (200-280ms, 250-330ms) (Figure 7, bottom panel).

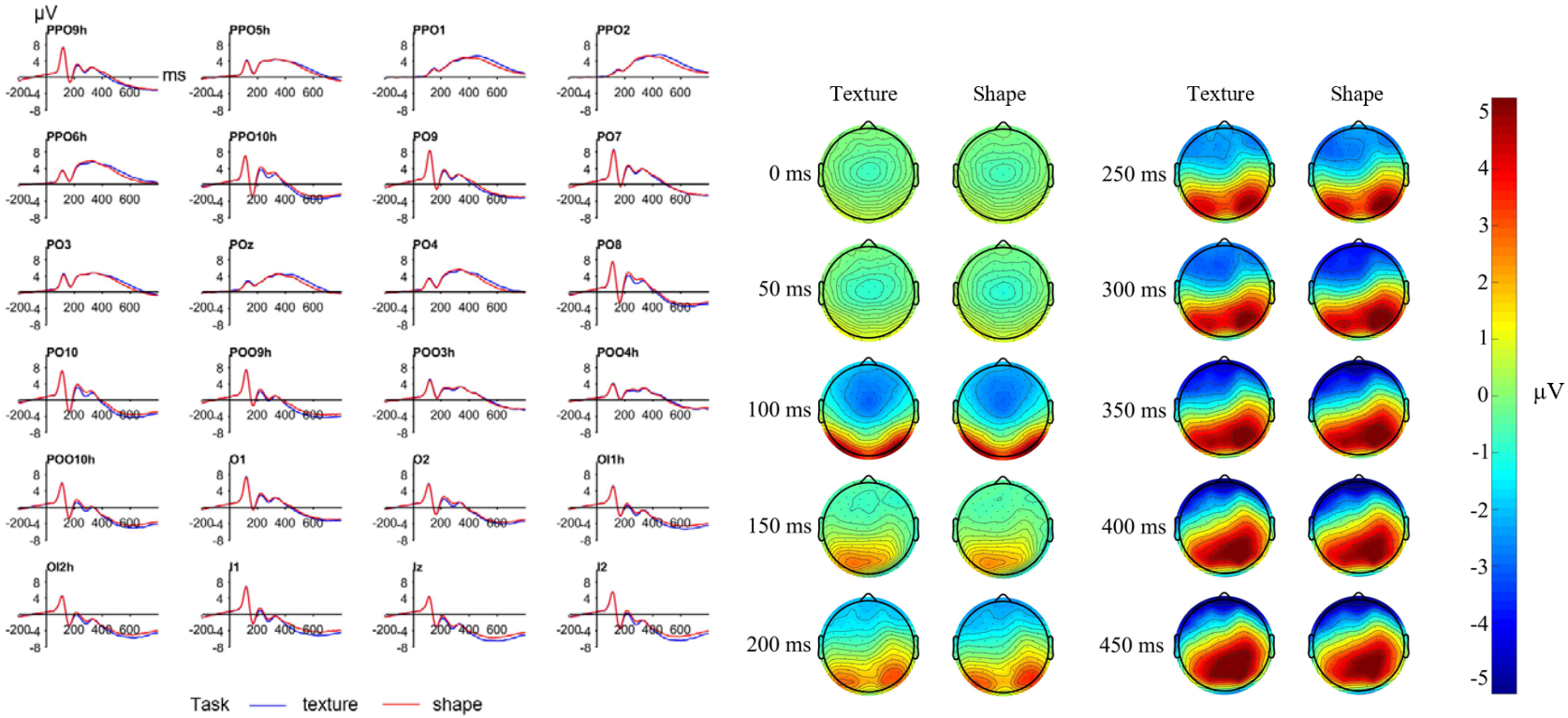


Figure 5. Grand average ERP waveform plots in selected channels (left) and scalp maps (right).

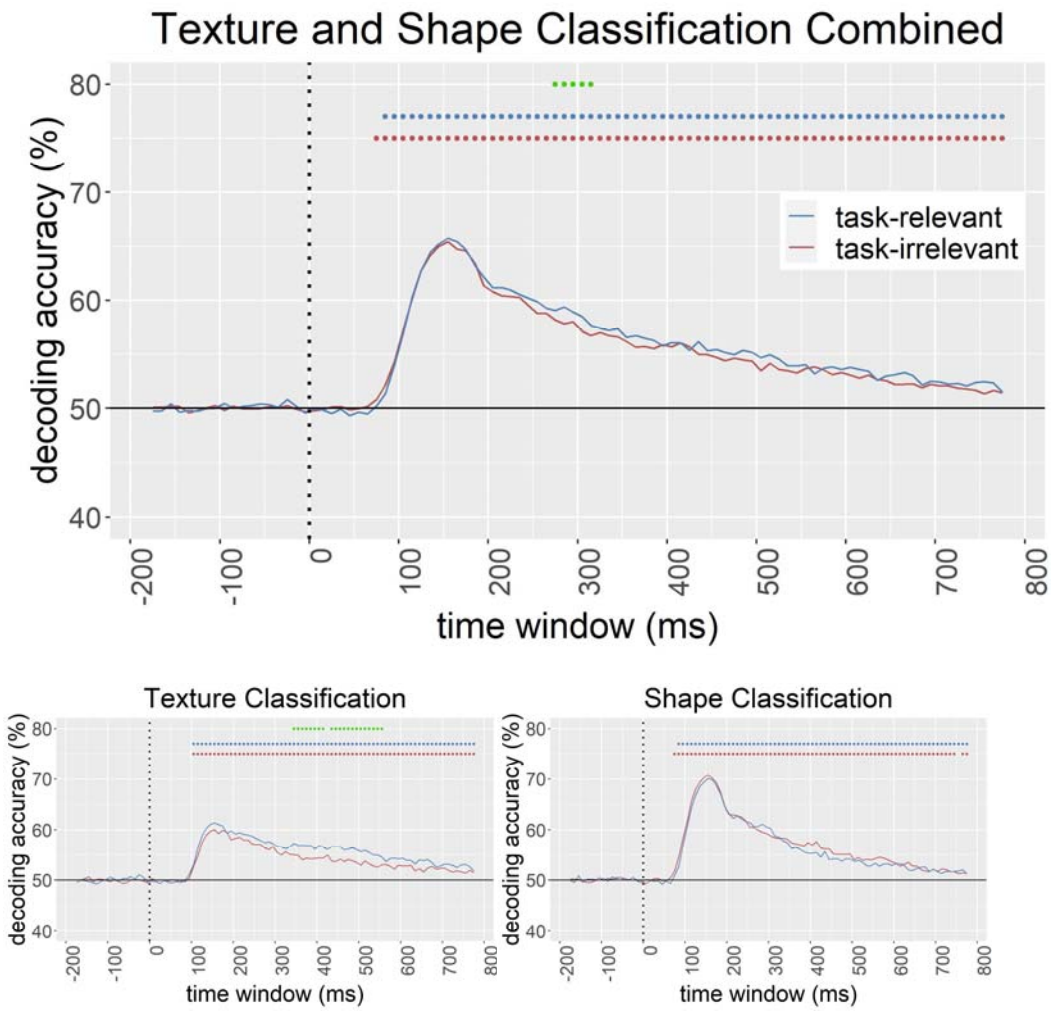


Figure 6. Decoding accuracies of task-relevant and task-irrelevant classifiers for texture and shape classification combined (top), for texture classification only (bottom left), and for shape classification only (bottom right). The *x*-axis depicts the time at which each 50-ms time window used for decoding was centered. The dots at the top indicate time windows with significant differences: blue dots refer to the time windows that were significantly above chance for task-relevant classifiers; red dots refer to the time windows that were significantly above chance for task-irrelevant classifiers; green dots refer to the time windows that were significantly different between task-relevant and task-irrelevant classifiers.

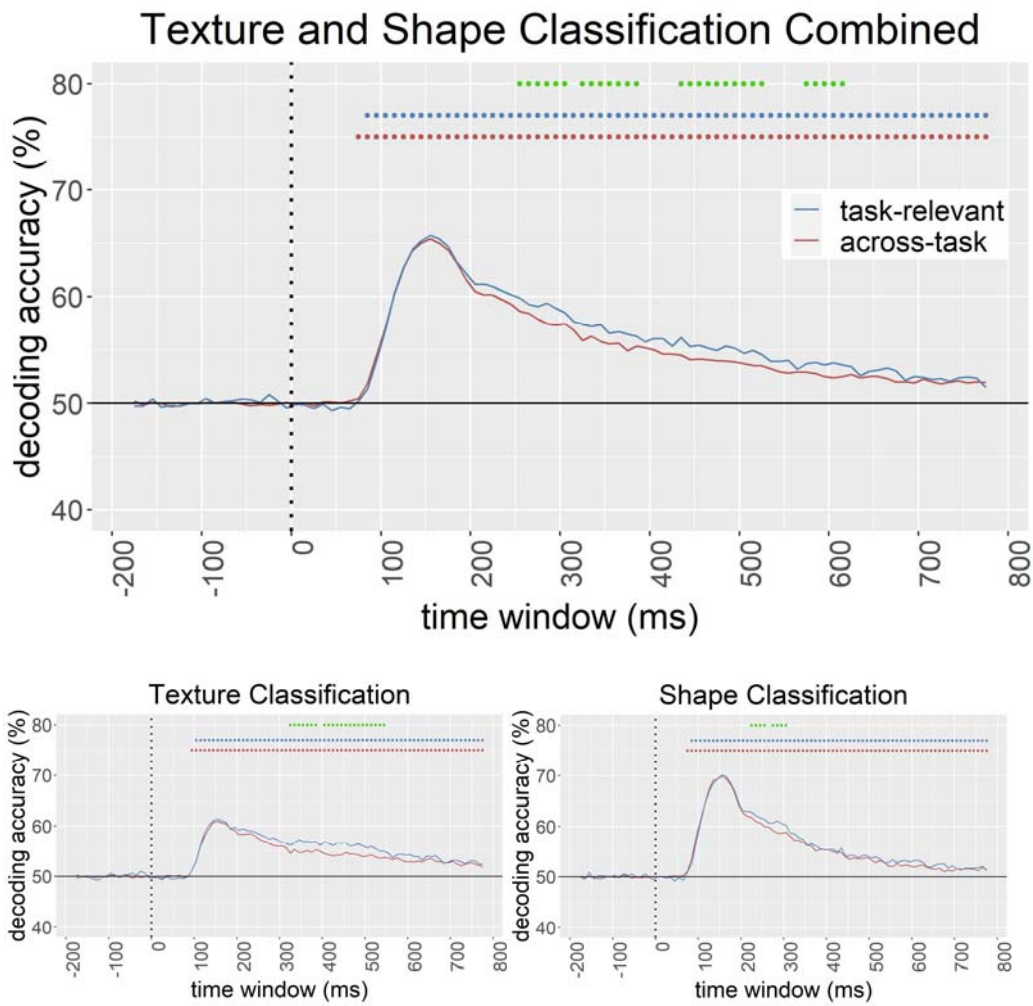


Figure 7. Decoding accuracies of task-relevant and across-task classifiers for texture and shape classification combined (top), for texture classification only (bottom left), and for shape classification only (bottom right). The *x*-axis depicts the time at which each 50-ms time window used for decoding was centered. The dots at the top indicate time windows with significant differences: blue dots refer to the time windows that were significantly above chance for task-relevant classifiers; red dots refer to the time windows that were significantly above chance for across-task classifiers; green dots refer to the time windows that were significantly different between task-relevant and across-task classifiers.

4. Discussion

The current study examined the time course of the effect of task on object processing with EEG. We presented artificial stimuli that could be categorized along two dimensions (texture and shape) and asked participants to perform object judgment based on one dimension during each task. Two findings were obtained from the multivariate analyses of the EEG activation patterns. First, the decoding accuracy of task-relevant classifiers was higher than that of task-irrelevant classifiers at 250-340ms after stimulus onset. This suggests that object representations along task-relevant dimension were more distinguishable than task-irrelevant dimension during this period. In other words, more information was available for separating objects along a particular dimension when it was consistent with the dimension on which the participant was focusing. This task effect, however, was significant only when texture was being decoded and not during shape decoding. A potential reason is that distinguishing the two shapes (round vs. cubic) was easier than distinguishing the two textures (large vs. small dots). Both the behavioral and decoding performance was better for shape than texture discrimination. The shape differences as manifested in this study may be so obvious that even an irrelevant task demand would only cause a limited impairment to decoding performance.

The second finding was that classifiers trained on a certain task performed better when they were tested on the same task than on a different task. And the task effect was found for

both texture decoding and shape decoding. Comparison of the task-relevant classifiers and the across-task classifiers showed a lack of full generalization at 230ms-340ms, 300-410ms, 410-550ms and 550-640ms. Therefore, there existed some qualitative differences in the object representations in the two tasks during this period. In other words, the diagnostic features for an object dimension differed across tasks.

The time-course of the task effect overlapped with the duration of the N170 and N250, two important ERP components in the study of object recognition. The N170 is well-known for its selectivity for face processing (e.g. Bentin et al., 1996) for expert recognition of non-face objects (Tanaka & Curran, 2001; Wong et al., 2005). The N250 is sensitive to the familiarity of individual faces and objects (Schweinberger et al., 2002; Schweinberger et al., 2004; Scott et al., 2006). As a result, the decoding accuracies within the time range of N170 and N250 are likely to reflect mainly object representations in the high-level areas of the ventral visual pathway.

Interestingly, task effect occurred after the peak of decoding accuracy. An intriguing possibility is that the peak decoding accuracy (130-180ms for all analyses) corresponds to relatively task-independent representations in areas such as the lateral occipital cortex (LO), while the subsequent task effect reflects task-dependent representations in the more anterior areas like the fusiform gyrus. Several lines of research suggest that this may be a sensible interpretation. First, EEG source localization studies have identified the fusiform gyrus as a

potentially major source of the N170 (Herrmann et al., 2005; Deffke et al., 2007) and the N250 (Schweinberger et al., 2002; Schweinberger et al., 2004; Scott et al., 2006). Second, Cichy & Pantazis (2017) used fMRI-MEG/EEG fusion technique to study the representational structure of objects in time and space, and found correspondence between representations revealed by EEG decoding at different time points and representations revealed by fMRI at different areas along the ventral visual pathway. Specifically, representations at the primary visual cortex (V1) corresponded the best with MEG/EEG representations at about 100ms, while representations at the inferotemporal cortex (IT) corresponded the best with MEG/EEG representations at close to 300ms. Third, Harel et al. (2014) found that the representations changed across tasks in the posterior fusiform area (pFs) but not LO. Lastly, some studies suggest that decoding of object category could be better in LO than pFs (Julian et al., 2017; Hebart et al., 2018). For example, Julian et al. (2017) found that object category decoding was significant in LO but not in pFs. Results of the fMRI-MEG-model fusion in Hebart et al. (2018) also showed that responses in LO was dominantly explained by the category model at the early time points, while responses in pFs were not that clearly correlated with the category model. Perhaps LO showed the best shape descriptions of objects, and when top-down information was combined with these shape descriptions in the pFs, the representations became less shape-dependent and the decoding accuracy decreased. Taken together, it is reasonable to suspect that the peak of decoding accuracy corresponds to the activities in LO,

which offers best information for the decoding but is task independent, and the time range of N170 and N250 corresponds to the activities in the more anterior sites, which still held the object representations for individuation but could be influenced by task.

There were some time windows (230-280ms, 310-410ms, 410-550ms and 550-640ms) in which decoding accuracy was higher for task-relevant classifiers than across-task classifiers but not task-irrelevant classifiers. What happened in these time windows? The lower accuracy of across-task classifiers than task-relevant classifiers showed that qualitatively different information (i.e., different weighted combinations of data at different channels and timepoints) were used for classification in different tasks. As a result, what the classifiers learned in one task did not generalize as well to another task during testing. The similar decoding performance for task-relevant and task-irrelevant classifiers at these time points indicated that representations were not more distinguishable in task-relevant dimensions. Putting these together, we can conclude that while the object representations changed qualitatively across task, the amount of information that helps discrimination between objects did not change with the task demand in these time windows.

Two previous studies examined the time course of task effect with MEG/EEG. Hebart et al. (2018) found that the decoding accuracies of within-task-type classifiers in perceptual tasks and conceptual tasks differed in 542-833ms. Specifically, object representations were more separable from each other during conceptual tasks than perceptual tasks only after

542ms but not earlier. The authors claimed that it was due to the more in-depth processing in conceptual tasks. Unlike the conceptual vs perceptual comparison in their design, the two tasks in the current study were not designed to differ in terms of the depth of processing. We also asked whether the object representations would be more separable, but our focus was on the difference between the task-relevant and task-irrelevant dimensions. Our analysis showed an earlier effect (250-340ms), reflecting a possible task effect at the high-level visual areas. Hebart et al. (2018) also compared the decoding accuracies of within task type classifiers and across task type classifiers and found no difference between the two. The different findings could be explained by the different stimuli used in the two studies. Hebart et al. (2018) used a set of real-world objects, and due to our rich experience interacting with them for different purposes at the same time, it might be hard to focus only on one required dimension. In the current study, however, artificial stimuli were employed, allowing participants to more easily focus on only one dimension in each task as instructed.

Bocincova & Johnson (2019) found that classifier performance in the decoding of a part's orientation was higher when it was task-relevant than irrelevant during 100-600ms, whereas we only found this effect of task relevance in 250-340ms. Therefore, task apparently had an earlier influence on the formation of representations in their study. One potential reason for the difference is that a blocked design was used in their study and participants were motivated to remember only the task-relevant feature in each block. In the current study,

participants focused on different dimensions in each trial depending on the cue. Another important difference is that Bocincova & Johnson (2019) conducted time-frequency analysis and input the power of frequencies to the classifiers due to their focus on visual working memory. The time-frequency decomposition involved temporal smearing, and thus it is uncertain how much of the effect found starting at 100ms was contributed by activations afterwards.

In conclusion, we found that object representations were more separable along task-relevant than irrelevant dimension, and involved qualitatively different information across different task contexts. Task effects occurred after the peak of decoding accuracy and within the time range of N170 and N250 components, corresponding to object representations at the anterior parts of the ventral visual pathway. Similar techniques can be adopted to investigate how object representations are affected by different cognitive processes such as belief, desire, and concepts (e.g. Pylyshyn, 1999; Raftopoulos, 2014; Firestone & Scholl, 2016; Newen & Vetter, 2017; Teufel & Nanay, 2017).

References

Alilović, J., Timmermans, B., Reteig, L. C., Van Gaal, S., & Slagter, H. A. (2019). No evidence that predictions and attention modulate the first feedforward sweep of cortical information processing. *Cerebral Cortex*, *29*(5), 2261-2278. https://doi.org/10.1093/cercor/bhz038

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, *8*(6), 551-565. https://doi.org/10.1162/jocn.1996.8.6.551

Bocincova, A., & Johnson, J. S. (2019). The time course of encoding and maintenance of task-relevant versus irrelevant object features in working memory. *Cortex*, *111*, 196-209. https://doi.org/10.1016/j.cortex.2018.10.013

Bracci, S., Daniels, N., & Op de Beeck, H. (2017). Task context overrules object-and category-related representational content in the human parietal cortex. *Cerebral Cortex*, *27*(1), 310-321. https://doi.org/10.1093/cercor/bhw419

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436. https://doi.org/10.1163/156856897x00357

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, *2*(3), 27. https://doi.org/10.1145/1961189.1961199

Cichy, R. M., & Pantazis, D. (2017). Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. *NeuroImage*, *158*, 441-454. https://doi.org/10.1016/j.neuroimage.2017.07.023

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, *36*(3), 181-204. https://doi.org/10.1017/s0140525x12000477

Cohen Kadosh, K., Henson, R. N., Cohen Kadosh, R., Johnson, M. H., & Dick, F. (2010). Task-dependent activation of face-sensitive cortex: an fMRI adaptation study. *Journal of Cognitive Neuroscience*, *22*(5), 903-917. https://doi.org/10.1162/jocn.2009.21224

Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *Neuroimage*, *35*(4), 1495-1501. https://doi.org/10.1016/j.neuroimage.2007.01.034

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*(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415-434. https://doi.org/10.1016/j.neuron.2012.01.010

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the

evidence for “top-down” effects. *Behavioral and brain sciences*, *39*. https://doi.org/10.1017/s0140525x15000965

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, *11*(2), 127. https://doi.org/10.1038/nrn2787

Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision research*, *37*(12), 1673-1682. https://doi.org/10.1016/s0042-6989(96)00286-6

Goffaux, V., Jemel, B., Jacques, C., Rossion, B., & Schyns, P. G. (2003). ERP evidence for task modulations on face perceptual processing at different spatial scales. *Cognitive Science*, *27*(2), 313-325. https://doi.org/10.1207/s15516709cog2702_8

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision research*, *41*(10-11), 1409-1422. https://doi.org/10.1016/s0042-6989(01)00073-6

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of cognitive neuroscience*, *29*(4), 677-697. https://doi.org/10.1162/jocn_a_01068

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*, 201312567. https://doi.org/10.1073/pnas.1312567111

Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: principles, pitfalls, and perspectives. *Neuron*, *87*(2), 257-270. https://doi.org/10.1016/j.neuron.2015.05.025

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

Herrmann, M. J., Ehlis, A. C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain topography*, *18*(2), 77-85. https://doi.org/10.1007/s10548-005-0277-7

Hohwy, J. (2013). *The predictive mind*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199682737.001.0001

Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, *47*, 75-85. https://doi.org/10.1016/j.concog.2016.09.004

Itier, R. J., & Neath-Tavares, K. N. (2017). Effects of task demands on the early neural processing of fearful and happy facial expressions. *Brain research*, *1663*, 38-50. https://doi.org/10.1016/j.brainres.2017.03.013

Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Feature-selective attention in frontoparietal cortex: multivoxel codes adjust to prioritize task-relevant

information. *Journal of cognitive neuroscience*, *29*(2), 310-321. https://doi.org/10.1162/jocn_a_01039

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*(22), 8210-8219. https://doi.org/10.1523/jneurosci.6153-09.2011

Julian, J. B., Ryan, J., & Epstein, R. A. (2016). Coding of object size and object category in human visual cortex. *Cerebral Cortex*, *27*(6), 3095-3109. https://doi.org/10.1093/cercor/bhw150

Kim, N. Y., & McCarthy, G. (2016). Task influences pattern discriminability for faces and bodies in ventral occipitotemporal cortex. *Social neuroscience*, *11*(6), 627-636. https://doi.org/10.1080/17470919.2015.1131194

Kok, P., Jehee, J. F., & De Lange, F. P. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265-270. https://doi.org/10.1016/j.neuron.2012.04.034

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., ... & Bandettini, P. A. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126-1141. https://doi.org/10.1016/j.neuron.2008.10.043

Long, N. M., & Kuhl, B. A. (2018). Bottom-up and top-down factors differentially influence stimulus representations across large-scale attentional networks. *Journal of Neuroscience*, 2724-17. https://doi.org/10.1523/jneurosci.2724-17.2018

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. https://doi.org/10.3389/fnhum.2014.00213

Lupyan, G. (2015). Cognitive penetrability of perception in the age of prediction: Predictive systems are penetrable systems. *Review of philosophy and psychology*, *6*(4), 547-569. https://doi.org/10.1007/s13164-015-0253-4

Macpherson, F. (2017). The relationship between cognitive penetration and predictive coding. *Consciousness and cognition*, *47*, 6-16. https://doi.org/10.1016/j.concog.2016.04.001

Newen, A., & Vetter, P. (2017). Why cognitive penetration of our perceptual experience is still the most plausible account. *Consciousness and cognition*, *47*, 26-37. https://doi.org/10.1016/j.concog.2016.09.005

Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping*, *15*(1), 1-25. https://doi.org/10.1002/hbm.1058

O'Callaghan, C., Kveraga, K., Shine, J. M., Adams Jr, R. B., & Bar, M. (2017). Predictions penetrate perception: Converging insights from brain, behaviour and disorder. *Consciousness and cognition*, *47*, 63-74.

https://doi.org/10.1016/j.concog.2016.05.003

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*, 437-442. https://doi.org/10.1163/156856897x00366

Pylyshyn, Z. (1999). Is vision continuous with cognition?: The case for cognitive impenetrability of visual perception. *Behavioral and brain sciences*, *22*(3), 341-365. https://doi.org/10.1017/s0140525x99002022

Raftopoulos, A. (2014). The cognitive impenetrability of the content of early vision is a necessary and sufficient condition for purely nonconceptual content. *Philosophical Psychology*, *27*(5), 601-620. https://doi.org/10.1080/09515089.2012.729486

Ransom, M., Fazelpour, S., & Mole, C. (2017). Attention in the predictive mind. *Consciousness and cognition*, *47*, 99-112. https://doi.org/10.1016/j.concog.2016.06.011

Riesenhuber, M., & Poggio, T. (2002). Neural mechanisms of object recognition. *Current opinion in neurobiology*, *12*(2), 162-168. https://doi.org/10.1016/s0959-4388(02)00304-5

Schwarzkopf, D. S., & Rees, G. (2011). Pattern classification using functional magnetic resonance imaging. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(5), 568-579. https://doi.org/10.1002/wcs.141

Schweinberger, S. R., Huddy, V., & Burton, A. M. (2004). N250r: a face-selective brain response to stimulus repetitions. *Neuroreport*, *15*(9), 1501-1505. https://doi.org/10.1097/01.wnr.0000131675.00319.42

Schweinberger, S. R., Pickering, E. C., Burton, A. M., & Kaufmann, J. M. (2002). Human brain potential correlates of repetition priming in face and name recognition. *Neuropsychologia*, *40*(12), 2057-2073. https://doi.org/10.1016/s0028-3932(02)00050-7

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of cognitive neuroscience*, *18*(9), 1453-1465. https://doi.org/10.1162/jocn.2006.18.9.1453

Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G., & Poggio, T. (2005). *A theory of object recognition: computations and circuits in the feedforward path of the ventral stream in primate visual cortex.* MASSACHUSETTS INST OF TECH CAMBRIDGE MA CENTER FOR BIOLOGICAL AND COMPUTATIONAL LEARNING.

Serre, T., Oliva, A., & Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the national academy of sciences*, *104*(15), 6424-6429. https://doi.org/10.1073/pnas.0700622104

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological science*, *12*(1), 43-47.

https://doi.org/10.1111/1467-9280.00308

Teufel, C., & Nanay, B. (2017). How to (and how not to) think about top-down influences on visual perception. *Consciousness and cognition*, *47*, 17-25. https://doi.org/10.1016/j.concog.2016.05.008

Valdés-Conroy, B., Aguado, L., Fernández-Cahill, M., Romero-Ferreiro, V., & Diéguez-Risco, T. (2014). Following the time course of face gender and expression processing: a task-dependent ERP study. *International Journal of Psychophysiology*, *92*(2), 59-66. https://doi.org/10.1016/j.ijpsycho.2014.02.005

Vance, J., & Stokes, D. (2017). Noise, uncertainty, and interest: Predictive coding and cognitive penetration. *Consciousness and cognition*, *47*, 86-98. https://doi.org/10.1016/j.concog.2016.06.007

Vaziri-Pashkam, M., & Xu, Y. (2017). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. *Journal of Neuroscience*, *37*(36), 8767-8782. https://doi.org/10.1523/jneurosci.3392-16.2017

Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type? *Psychological Science*, *20*(9), 1108-1117. https://doi.org/10.1111/j.1467-9280.2009.02430.x

Wong, A. C., Gauthier, I., Woroch, B., Debuse, C., & Curran, T. (2005). An early electrophysiological response associated with expertise in letter perception. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(3), 306-318. https://doi.org/10.3758/cabn.5.3.306

Appendix

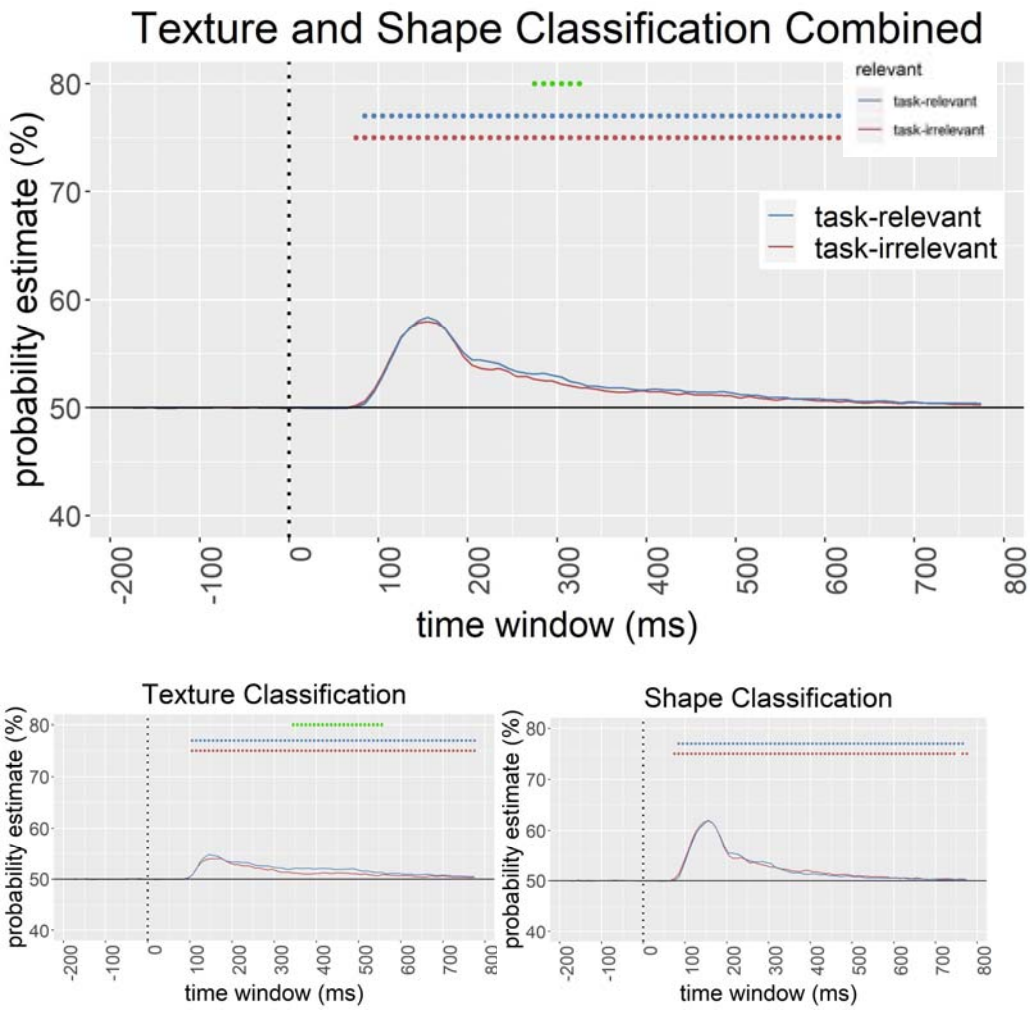


Figure A1. Probability estimates of task-relevant and task-irrelevant classifiers for texture and shape classification combined (top), for texture classification only (bottom left), and for shape classification only (bottom right). The *x*-axis depicts the time at which each 50-ms time window used for decoding was centered. The dots at the top indicate time windows with significant differences: blue dots refer to the time windows that were significantly above chance for task-relevant classifiers; red dots refer to the time windows that were significantly above chance for task-irrelevant classifiers; green dots refer to the time windows that were significantly different between task-relevant and task-irrelevant classifiers.

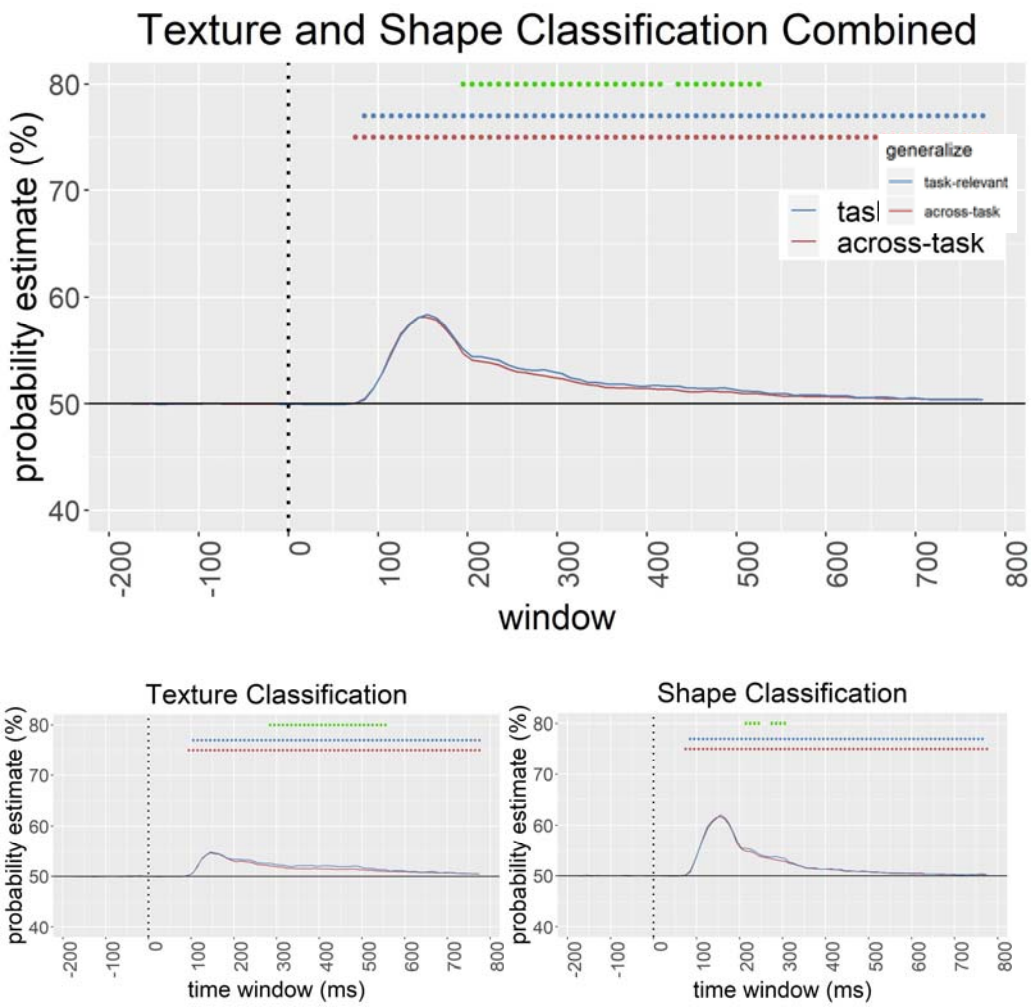


Figure A2. Probability estimates of task-relevant and across-task classifiers for texture and shape classification combined (top), for texture classification only (bottom left), and for shape classification only (bottom right). The *x*-axis depicts the time at which each 50-ms time window used for decoding was centered. The dots at the top indicate time windows with significant differences: blue dots refer to the time windows that were significantly above chance for task-relevant classifiers; red dots refer to the time windows that were significantly above chance for across-task classifiers; green dots refer to the time windows that were significantly different between task-relevant and across-task classifiers.