

Dynamic object-based encoding and automatically prioritized position encoding in visual working memory

Zitong Lu^{1,2,3}, Hui Chen⁴, Yixuan Ku^{1,2,3} *

1. Guangdong Provincial Key Laboratory of Social Cognitive Neuroscience and Mental Health, Department of Psychology, Sun Yat-sen University, Guangzhou, China.

2. Peng Cheng Laboratory, Shenzhen, China.

3. Shanghai Key Laboratory of Brain Functional Genomics, Shanghai Changning-ECNU Mental Health Center, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China.

4. Department of Psychology, Zhejiang University, Hangzhou, China.

* correspondence: Yixuan Ku, kuyixuan@mail.sysu.edu.cn

ResearcherID: D-4063-2018

ORCID: 0000-0003-2804-5123

Running title: Dynamic object-based encoding in VWM

ABSTRACT

Whether visual working memory (VWM) retains information from different features independently or as a unified object representation remains debated. Although behavioral studies have provided evidence that the unattended feature of an object could also be encoded in VWM, there still lacks neural supports for the object-based encoding. Here participants were asked to perform three VWM tasks (orientation-position, position-orientation, or orientation-color). They needed to recall the former feature as a task-relevant feature while the latter was a task-irrelevant feature. Electroencephalogram (EEG) was recorded and decoded to track the representation of the task-relevant and task-irrelevant visual features dynamically. Orientation could be decoded more accurately as a task-relevant feature than a task-irrelevant feature. However, position could be decoded equally well in in task-relevant or task-irrelevant conditions, and the feature decoding accuracy of task-irrelevant position was even higher than the task-relevant orientation. Meanwhile, task-irrelevant color was not prioritized over task-relevant orientation. There existed significant correlation between task-relevant and task-irrelevant features decoding accuracies across three VWM tasks, supporting the object-based encoding hypothesis. The decoding accuracies of task-irrelevant features in all three conditions decreased in the second half of sessions, while the correlation between relevant and irrelevant feature decoding accuracies only decreased in orientation-color condition. Altogether these results suggested dynamic object-based encoding processes in VWM, yet, position, as a prioritized feature, could be encoded automatically.

SIGNIFICANCE STATEMENT

Visual working memory is connecting perception and high-level cognitive functions. Understanding what the fundamental encoding unit is in visual working memory and how visual features and objects are represented in human brain are essential questions in the field. Although behavioral studies suggested that task-irrelevant features were encoded in visual working memory, it was still unclear how they were assembled with task-relevant features. Here, electroencephalogram (EEG) decoding revealed a dynamic object-based encoding mechanism in visual working memory that the task-irrelevant feature was encoded together with task-relevant features at first but dropped later. Meanwhile, position as a prioritized feature was automatically encoded. Using decoding methods is a way to probe dynamic changes of representations of different features in perception and memory.

INTRODUCTION

Visual working memory (VWM) is one of the most critical functions for human cognition (Luck & Vogel, 2013). It is positively correlated with fluid intelligence and determines learning efficiency, mental calculation, reasoning, language learning, and other cognitive abilities. Many studies have revealed cognitive and neural mechanisms underlying VWM (Christophel et al., 2017; Luck & Vogel, 2013). Nevertheless, how memories of visual features are represented dynamically in the brain remains debated. The object-based coding hypothesis that the object is the primary storage unit for maintaining VWM (Gao et al., 2011; Hyun et al., 2009; Luck & Vogel, 1997; Luria & Vogel, 2011; Quinlan & Cohen, 2011; Shen et al., 2013; Vogel et al., 2001; Yin et al., 2012). The brain can encode both the task-relevant and task-irrelevant features of an object. In contrast, the feature-based coding hypothesis contends that all features are independently stored in the number-limited sub-system (Maunsell & Treue, 2006; Olson & Jiang, 2002; Wheeler & Treisman, 2002). Although previous studies have found several behavioral and neural indicators to support one of two viewpoints, it is challenging to track dynamic memory representations over time.

The traditional studies in neural correlates of VWM are limited to univariate analysis based on single-voxels (Baddeley, 2003; Nee et al., 2013), which is less sensitive to evaluate hidden neural patterns (Norman et al., 2006). Multivariate pattern analysis (MVPA) has been applied to neural data to unveil better the hidden visual representation in our brains (Norman et al., 2006). Interestingly, not only objects and faces from different categories but also features such as orientation, position, color, and emotion can be decoded from functional magnetic resonance imaging (fMRI) (Albers et al., 2013; Ban et al., 2012; Bannert & Bartels, 2013; Johnson & Johnson, 2014; Koch et al., 2020; Koenig-Robert & Pearson, 2019; Lescroart & Gallant, 2019; Liu et al., 2019; Reddy et al., 2010; Schlegel et al., 2013; Stokes et al., 2009; Vetter et al., 2014), electroencephalograph (EEG) (Bae & Luck, 2019a; Fahrenfort et al., 2017; Grootswagers et al., 2019; Hogendoorn & Burkitt, 2018; Hong et al., 2020; Long & Kuhl, 2019; Mares et al., 2020; Noah et al., 2020; Robinson et al., 2019; Shatek et al., 2019; Smith & Smith, 2019; Xie et al., 2020), and magnetoencephalography (MEG) (Carlson et al., 2011; Cichy et al., 2016; Dijkstra et al., 2018; Isik et al., 2014; King et al., 2014; Pantazis et al., 2018; Teichmann et al., 2020) during various cognitive processes, especially VWM (Bae & Luck, 2018, 2019b; Bocincova & Johnson, 2019; Cai et al., 2019; Christophe et al., 2012; Ester et al., 2009, 2013, 2015; Gosseries et al., 2018; Harrison & Tong, 2009; Rose et al., 2016; Serences et al., 2009; Sprague et al., 2016; Wolff et al., 2015, 2017; Xing et al., 2013).

Using the EEG decoding method, we can effectively explore the dynamic representation of different features in the VWM task and further explore how

memories are stored. We propose the following hypothesis: If both the relevant and irrelevant features were successfully decoded, the storage unit of memory was an object. Our brains not only encoded the relevant feature but also automatically encoded the irrelevant feature. If only the relevant feature were successfully decoded, the storage unit of memory was the independent feature. Moreover, if VWM is object-based encoding, it remains unknown whether the coding patterns of both the task-relevant and task-irrelevant features are consistent. We can track the dynamic encoding of memory using a very long experiment in which subjects consistently focus on only one feature of an object in the same task. As the task goes on, if both the relevant and irrelevant features' intensities remain unchanged, there is a strong binding between features to constitute the entire memory object. If only the task-irrelevant feature's coding becomes weak, there are likely to be two different coding patterns of the relevant and irrelevant feature for the memory object in the brain.

Our current study conducted a more than ten-hour experiment with three recall VWM tasks to reveal the memory storage mechanism by EEG decoding. Firstly, successfully decoding both the task-relevant and task-irrelevant features in all three tasks and significant correlations between task-relevant and task-irrelevant features decoding accuracies supported the object-based coding hypothesis. Secondly, our results with high accuracies of position throughout confirmed that position was a powerful feature rather than orientation and color, which was consistent with previous research ([Chen & Wyble, 2015](#); [Pertsov & Husain, 2014](#); [A. M. Treisman & Gelade, 1980](#); [A. Treisman & Zhang, 2006](#); [Wheeler & Treisman, 2002](#)). Task-irrelevant color and orientation were not prioritized over task-relevant orientation, but position could be decoded equally well in both task-relevant and task-irrelevant conditions. Thirdly, benefited by the massive number of trials in our experiments, we divided all trials into two phases (early and late) and found that the decoding accuracies of all task-irrelevant features decreased. There might be two encoding patterns of the relevant feature and the irrelevant feature by phased decoding. Based on these findings, we proposed a dynamic object-based theory that our brains synchronously encoded both the irrelevant and irrelevant features of an object at the beginning. However, the irrelevant feature's coding strength instead of the relevant feature would decrease as the task went on. Besides, the position, as a prioritized feature, has a consistently strong representation regardless of as a task-relevant feature or a task-irrelevant feature, which might play an essential role in this dynamic object-based process.

MATERIALS AND METHODS

Participants

Six students from East China Normal University between the ages of 21 and 24 (3 male and 3 female, age=22.67 \pm 1.49) with normal or corrected-normal visual acuity participated in all the experiments for monetary

compensation. All six subjects had no history of mental illness and were paid after the experiment. Written informed consent was provided by all participants before the experiment. They performed the tasks for over ten hours in total within three sessions (three days of visits).

Stimuli and apparatus

Participants were asked to conduct three kinds of tasks (**Figure 1**). All stimuli were generated by Psychtoolbox for MATLAB. Participants were seated on a chair 63cm far from the screen with their heads fixed using a chin rest.

Task 1 is an Orientation-Position memory task. Participants were asked to memorize the orientation of the grating, ignoring the position information. During the experiment, each trial begun with a grating (visual angle 5°) displayed randomly on the left or right side (50% probability each) of the screen for 0.5s. The grating orientation was set to 23° , 158° , or a random angle different from the two (the ratio of three conditions is 1:1:1). Following a delay period of 2s, a probed grating with a random orientation would be presented at the same position as the grating in-memory display. Participants were requested to recall the memorized grating orientation as precisely as possible by using the mouse to rotate the grating.

Task 2 is a Position-Orientation memory task. Participants were asked to memorize the position of the grating, ignoring the orientation information. During the experiment, each trial began with a grating (visual angle 5°) with the orientation of 23° or 158° randomly (50% probability each) displayed on a fixed position on the left side or a fixed position on the right side or a random position of the screen for 0.5s. The ratio of the three position conditions is 1:1:3. Following a delay period of 2s, a probed grating with the same orientation as the grating in-memory display would be presented at a random position. Participants were requested to recall the memorized grating position as precisely as possible by using the mouse to move the grating.

Task 3 is an Orientation-Color memory task. Participants were asked to memorize the grating orientation like in task 1, ignoring the color information. During the experiment, each trial began with a grating (visual angle 5°) with the color of red or green randomly (50% probability each) displayed on the screen center for 0.5s. The grating orientation was set to 38° , 143° , or a random angle different from the two (the ratio of three conditions is 1:1:1). Following a delay period of 2s, a probed grating with a random orientation would be presented at the center of the same color as the grating in-memory display. Participants were requested to recall the memorized grating orientation as precisely as possible by using the mouse to rotate the grating.

In task 1, the position was undoubtedly a task-irrelevant feature. In task 2, the orientation was a task-irrelevant feature. Moreover, in task 3, the color was a task-irrelevant feature. The inter-trial interval of three tasks was 1 to 1.3s

randomly. Each participant completed at least 15 blocks (15 to 18 for each subject variably) of task 1 and 25 blocks of task 2 and 15 blocks of task 3 (at least 3300 trials at total, 60 trials per block, approximately 450s in each block). Six participants finished at least 20000 trials. Each participant received at least 10 practice trials for each task before the task. Taking in the rest time in the experiment, it will take more than 10 hours to complete. The block order of three tasks is random.

EEG recording and preprocessing

The continuous EEG was recorded using a Brain Products recording system, with 64 channels embedded in a cap according to the international 10/20 system. The impedance on all channels was kept below 10 k Ω during recording. EEG signals were sampled with a frequency of 250 Hz. EEGLAB toolbox ([Delorme & Makeig, 2004](#)) was used to preprocess the data.

All channels were re-referenced offline to the averaged mastoids through the signal recorded from channels TP9 and TP10. EEG data were filtered by 0.1 Hz high-pass and 40 Hz low-pass and then divided into segments ranging from 500ms before to 3000ms after the presentation of the memory display. Independent component analysis (ICA) was then performed on the scalp EEG for each subject to identify and remove components that were associated with blinks ([Jung et al., 2000](#)) and eye movements ([Drisdelle et al., 2017](#)). Select the trials of two fixed orientations in task 1 and task 2 and the trials of two fixed positions in task 3 for subsequent analysis.

Decoding Analysis

The implementation of the decoding part (including classification training and test) was based on Python. We attempted to decode relevant features and irrelevant features in our three tasks, including orientation (as the relevant feature in task 1 and task 2, as the irrelevant feature in task 3), position (as the relevant feature in task 1, as the irrelevant feature in task 3) and color (as the irrelevant feature in task 3).

The ERP decoding procedure was limited to frequencies between 2 and 6 Hz. The EEG data was bandpass filtered at 2-6 Hz by EEGLAB. In all decoding cases, the data were resampled by a 20ms time-window, which contained 5 time-samples. Each participant got six 4-dimensional data matrices for six visual features in three tasks, with dimensions of time (175 time-windows), channels (64 channels), trials (at least 600 trials), conditions (2 conditions, e.g., 2 fixed orientations in task 1).

The linear support vector machine (Linear-SVM) was selected to conduct a time-by-time classification of the basic visual feature based on the signal's spatial distribution over 64 channels for each participant. We binarized the two kinds of labels (for conditions of the relevant feature and conditions of the irrelevant feature) for each task. Consequently, we conducted 6 times (2 kinds

of feature \times 3 tasks) coding process for each participant.

The decoding for each time-point used a five-fold cross-validation procedure in which the data from 80% of the trials (selected at random) were used to train the Linear-SVM. Then the remaining trials were put into the after-training classifier to get the classification performance. This classification step was repeated 100 times for each time-point with a different seed for dividing the trials to get a more reliable result. The final classification accuracy is the average of these ten times for each time-point. Every feature by each kind of EEG signals, each subject, each time-point was decoded independently. The classification part was implemented based on *svm* and *model_selection* modules in scikit-learn (Pedregosa et al., 2011) by Python.

Across-subjects decoding

Unlike the above decoding process for each subject, six subjects' individual differences were not distinguished for cross-subjects decoding. Their data were mixed with classifying the features by time points. In this case, the Linear-SVM classifier for each feature in each task was trained and tested in the same way with a five-fold cross-validation procedure and 100 iterations. The result difference caused by individual memory strategy could be excluded by across-subjects decoding. In the across-subject decoding part, the accuracies of 100 iterations were used for later statistical analysis.

Cross-temporal decoding

The purpose of cross-temporal decoding, an extension of single-time decoding, was to construct the time generalization matrix of different features in different tasks. We input the sample data of a certain time point to the classifier for training and then use this trained classifier to test the data of any other time. We could use this method to analyze all combinations of neural representations between a training time and a test time. Here, the trained classifier at each time point is used to test all the time points from the stimulus presentation to the next 3000 milliseconds. Therefore, we obtained the corresponding time generalization matrix for each computing condition by cross-temporal decoding.

Statistical analysis

Suppose the neural representational pattern of a certain time contains encoding a specific visual feature in the brain. In that case, the decoding accuracy should be greater than chance, which was 50%, because we have two categories in our decoding processes. To compare decoding accuracy for a feature in a task to chance at each decoding time point, we use a permutation test to assess statistical significance. This means that we randomly shuffled decoding accuracies and chance accuracies. Then, we calculated the results for 5000 iterations to draw a distribution. Real accuracies exceeding 95% of the distribution are regarded as significant. *P*-values were calculated from this

permutation distribution.

For the cluster-based permutation test for cross-temporal decoding results, we first extracted each significant cluster. We calculated the clustering statistic as the sum of t -values in each cluster. Then we conducted 5000 permutations to calculate the maximum cluster statistic for each iteration and obtained a distribution of maximum permutation cluster statistic. Finally, we assigned p -values to each cluster of the actual decoding accuracies by comparing its cluster statistic with the permutation distribution.

RESULTS

Orientation and position decoding in task 1 and task 2

Firstly, we decoded the orientation and position in both orientation-position task and position-orientation task time-by-time. In the former, the orientation was the task-relevant feature, and the position was the task-irrelevant feature. In the latter, they were converse. The ERP time-by-time classification performances of both orientation and position in these two tasks are shown in **Figure 2a and 2b**. These decoding results showed that both orientation and position could be successfully decoded in the stimulus period, delay period, and response period, whether as the task-relevant feature or the task-irrelevant feature. However, the position was easier to decode in both tasks than the classification accuracy of orientation and had longer significant periods than orientation.

Although we could hardly detect whether the subject had memorized the irrelevant feature through the behavioral indicator, we successfully used decoder to decode the irrelevant feature. It proved a mechanism of storing object information, including orientation and position in our brains in VWM.

We then calculated the difference in decoding accuracies between the same feature being relevant or irrelevant (**Figure 2c**). The results showed that the decoding performance of the task-relevant orientation was significantly higher than the task-irrelevant orientation during 300-400ms, 900-1100ms, and 2200-2600ms. On the contrary, the decoding performances of the relevant position and the irrelevant position were hardly different. However, the feature decoding accuracy of task-irrelevant position was even higher than the task-relevant orientation (**Figure 3**).

Moreover, we conducted a series of cross-temporal decoding processes on orientation and position in task 1 and task 2. We input ERP data at different time-points to train the classifier and used this trained-classifier at a certain time-point to test the data at all time-points on the time series. If the classifier trained on certain time-point's data could be a good predictor of brain activity at another time, this meant the memory representations at these two time-points were highly consistent. The cross-temporal decoding results not only reflected

the time-by-time results but also reflected the continuity and consistency of feature coding in a task.

Figure 4a shows the temporal generalization decoding results by cross-temporal decoding in task 1 and task 2. Whether as the relevant feature in task 2 or as the irrelevant feature in task 1, the position had strong coding persistence during the delay period. Orientation had coding consistency in the early stage of delay while as the relevant feature in task 1. When orientation was the irrelevant feature in task 2, the decoding results were only significantly higher than the random level during the stimulus period and early delay period. The poor continuity of their coding patterns meant that only the predictor trained and test on very close time had a significantly higher decoding accuracy.

These decoding results demonstrated that there were different coding formats between different irrelevant features in VWM. Position might be a stronger feature than orientation in VWM, which can be automatically prioritized encoded. Whether the position is regarded as a relevant or irrelevant feature, the brain had strong coding for spatial information. In contrast, orientation might be a weaker feature. When orientation is the relevant feature in the task, the brain had more robust coding for orientation information. However, when orientation was the irrelevant feature, its coding became weak.

Orientation and color decoding in task 3

As the task-irrelevant feature in VWM, we could decode position better than orientation. Color is also one of the basic features of vision. Here we would like to test how color is represented in our brain as an irrelevant feature. We hypothesized that color and orientation were features of the same kind and position were a unique feature, so the decoding results of color as the irrelevant feature would be similar to task-irrelevant orientation.

First, we conducted the time-by-time decoding process on orientation and color in task 3. The time-by-time classification performances show that both the task-relevant feature (orientation) and the task-irrelevant feature (color) were coded in this task. Both decoding accuracies were significantly higher than chance during the stimulus period, early delay period, and the response period (**Figure 5**). Then, we conducted the cross-temporal decoding for orientation and color in task 3. **Figure 4b** shows that the cross-temporal decoding result of orientation is similar to orientation as the relevant feature in task 1. The cross-temporal decoding result of color is similar to the result of orientation as the irrelevant feature in task 2.

We found that the brain represented irrelevant features of three different visual features in all three tasks using ERP decoding. Our dynamic decoding results powerfully demonstrate the object-based encoding mechanism in VWM. Although the decoding results of orientation and color in this task were similar to the previous time-by-time results, it could be seen from the cross-temporal

analysis that color, as an irrelevant feature, had very weak coding persistence and consistency. Combining orientation decoding results in task 2 and color decoding results in task 3, we found that both orientation and color are weak features. When these weak features were task-irrelevant in VWM, their coding was not so continuous that they could only be trained and successfully predicted by neural activities at closed time points. However, the position was a more powerful feature. Whether position was the relevant or irrelevant feature, the decoding results were similar with strong continuity and consistency.

Phased decoding in three tasks

To explore whether the coding pattern of visual features, both task-relevant and task-irrelevant features, would change as the experiment repeatedly went on, we conducted the phased decoding for visual features (both relevant and irrelevant features) in three tasks. The trials in each task were divided equally into two parts: early phase and late phase according to the time sequence, and the features were decoded, respectively. The decoding process was the same as above.

We calculated the differences between the decoding results of the two phases (**Figure 6**). It could be seen that the decoding performances of the relevant features in all three tasks were no significant difference between the two phases. However, the decoding performances of the irrelevant features in all three tasks were higher in the early phase than in the late phase in some periods (position in task 1: 600-400ms, 2300-2500ms; orientation in task 2: 300-500ms, 650-700ms, 2600-2700ms; color in task 3: 0-150ms, 350-550ms, 2500-2600ms).

These phase decoding results demonstrated that when the subject memorized one single feature (the relevant feature) in a VWM task, the irrelevant feature would also be encoded from the start. Furthermore, as the experiment continues, we did not see a decrease in synchronization of both the relevant and irrelevant features or a rise of the relevant feature and fall of the irrelevant feature. It is only the coding of the irrelevant feature instead of the relevant feature in the brain that would gradually weaken. These results suggested that the task-relevant feature and the task-irrelevant feature had two independent different coding patterns in our brains. Besides, this reduction of the task-irrelevant feature was more substantial for encoding position information than other features, such as orientation or color, possibly because position was easier to decode.

Correlations between the relevant and irrelevant features

Since we had decoded various features successfully, we further confirmed the object-based hypothesis by correlation analysis. Here, we calculated the correlations (Spearman correlation coefficients) of decoding accuracies between the task-relevant and task-irrelevant features in three tasks. Giving the

decoding accuracies from the sample period to the response period (0-3000ms), we found that the decoding accuracies of the task-relevant and task-irrelevant features were significantly correlated in all three tasks (**Figure 7a**). Then, we computed the correlations between two kinds of accuracies for the two phases (early and late) (**Figure 7b**). The early phase's correlations of all three tasks were significant. The correlations of orientation-position task and position-orientation task in the late phase were also significant. However, it was not significant for the late phase's correlation of orientation-color task. Also, there was no significant difference between the two phases.

On the one hand, the results suggested a synchronous integration mechanism to bind different features into an object in VWM. On the other hand, although there was no significant difference between the two phases, the mean values of the later phase's correlations were lower than the early phase. There might be a weakening of the binding effect, and the consistency of the irrelevant feature to the relevant feature was destroyed in the orientation-color task. Besides, although the task-irrelevant feature's coding strength decreased, the consistency between the relevant and irrelevant features was still high. The sustaining consistency in orientation-position task and position-orientation task just went to show that position played an essential role in maintaining this bound mechanism that the irrelevant feature could be automatically encoded.

Across-subjects decoding

Finally, to test whether different individual memory strategies influenced this object-based encoding phenomenon, we then decoded two corresponding visual features of each task across six subjects. We conducted a similar classification process by inputting all subjects' ERP data ten times. The final accuracy was the average of ten iterations. **Figure 8** shows the results, which are similar to that of single-subject decoding. Either the task-relevant or task-irrelevant feature could be successfully decoded.

It demonstrated that the decoding results of each visual feature were not affected by the differences in individual memory strategies. This mechanism of object-based representation in VWM was common and consistent in the human brain.

DISCUSSION

EEG decoding can be an efficient method to explore how information dynamically represents in perception ([Grootswagers et al., 2019](#); [Hogendoorn & Burkitt, 2018](#); [Mares et al., 2020](#); [Noah et al., 2020](#); [Robinson et al., 2019](#); [Smith & Smith, 2019](#)), mental imagery ([Shatek et al., 2019](#); [Xie et al., 2020](#)), and memory ([Bae & Luck, 2018, 2019b](#); [Bocincova & Johnson, 2019](#); [LaRocque et al., 2013](#); [Wolff et al., 2015](#)). Especially in the VWM task, it is difficult to effectively monitor how the brain encodes different low dimensional features at different times. EEG decoding can be an effective method to track

the representations of object features. In our study, we attempted to decode the visual features time-by-time and compared the difference between classification accuracies and the random level at different time-points to evaluate the encoding of features in the memory process.

In our present study, we successfully decoded the task-relevant features of three different tasks. To our delight, we also successfully decoded all the task-irrelevant features. In some behavioral studies, participants had consistent performances between remembering single features and multi-feature objects in change detection tasks (Delvenne & Bruyer, 2004; Luck & Vogel, 1997; Olson & Jiang, 2002; Riggs et al., 2011). Some research also utilized the surprising test by a single surprise trial (Chen & Wyble, 2016; Shin & Ma, 2016; Swan et al., 2016), finding that participants could memorize the task-irrelevant feature. Our decoding results provided strong neural evidence that the task-irrelevant feature could be automatically encoded in our brains. Thus, the format of memory representations in our brains is object-based.

Some previous studies demonstrated that position as the task-irrelevant feature was automatically encoded into VWM (Jiang et al., 2000; A. Treisman & Zhang, 2006), while other nonspatial features as the task-irrelevant feature had no representation (Serences et al., 2009; Woodman & Vogel, 2008). However, our decoding results with large trials showed that even orientation and color were automatically encoded. In the meantime, the successful decoding of visual features occurred in the sample stage, the delay stage, and the response stage, which indicated that these object-based representations happened from the perception period to the maintenance period to the recall period.

Importantly, comparing the decoding results for different features, we concluded that our brains represented different features differently. The decoding performances were lower for orientation and color than position. For the three different irrelevant features, position was significantly more powerful than orientation and color in terms of time-by-time decoding accuracies and cross-temporal decoding results. Also, coding patterns of orientation and color information hardly had time persistence, while position had strong coding consistency in the whole task. These proved that position was a very powerful feature (Chen & Wyble, 2015; Pertzov & Husain, 2014; A. M. Treisman & Gelade, 1980; A. Treisman & Zhang, 2006; Wheeler & Treisman, 2002), while other visual features such as orientation and color were weak features. This might be the reason why we could hardly find evidence for the representation of task-irrelevant orientation or color by the behavioral performance and the ERP component such as contralateral delay activity (CDA). This super feature had similar strong coding under the two conditions by comparing the results of position as a relevant feature and an irrelevant feature. This indicated that the representation of position might not be affected by the task, and position might be the basis for integrating features into the object in VWM. There is an

automatically prioritized position encoding mechanism in our brains.

Though the phased decoding benefited by the vast number of trials in our experiments, we compared representational differences between the early and late phases. A previous behavioral study ([Swan et al., 2016](#)) found that memory had different variable precisions for the task-relevant and the task-irrelevant features. Different precisions corresponding to different features were influenced by various degrees of attention. If, as it said, as the task progressed, the attention became more focused on the relevant feature, the encoding of the relevant feature in the late phase would be stronger, and the encoding of the irrelevant feature would be weaker. However, our phased decoding results suggested that the irrelevant feature's coding strength would decrease as the task went on. There was no difference in the relevant feature's coding strength between the early and late. This process might not be because participants paid more attention to the task-relevant feature. On the contrary, there might be two independent representational patterns for the relevant and irrelevant features. When the participant only memorized one feature continuously in the same task, the relevant feature's encoding remained unchanged, but the irrelevant feature's encoding gradually weakened.

By calculating the correlation between decoding accuracies of the relevant and irrelevant features, we contended again that VWM was object-based encoding that the relevant and irrelevant features had consistent representations. Although the phased correlation results indicated no significant differences between the two phases, we could see a little tendency for the correlations between the two features to become weak, especially in the orientation-color task. Lacking position changes in the task and position as a unique feature strongly supporting feature integration into object could explain why the correlation between orientation and color was not significant in the late phase.

Profoundly exploring the coding mechanism through a very long experiment and EEG decoding method, we highlighted the dynamic object-based encoding and the automatically prioritized position encoding in VWM. The foundational encoding unit of memory was the object. Thus, the object's task-irrelevant feature was automatically represented and bounded to the task-relevant feature in VWM. Furthermore, the irrelevant feature's coding strength but not the relevant feature would decrease if people continuously focused on one feature of the object in the task. In this dynamic object-based encoding process, the irrelevant feature binding the relevant gradually weakened with the relevant feature's coding staying the same. And position, as a prioritized feature, could be encoded automatically. This helps us to understand further the processing and storage of memory in our brains.

However, there is still a lot more to explore about the binding of object features in VWM. First, in our current research, what we focused on was the

integration of different kinds of features. Some previous studies found that objects containing two features from the same type had difficulty in remembering them in the form of the object (Delvenne & Bruyer, 2004; Olson & Jiang, 2002; Parra et al., 2011; Wheeler & Treisman, 2002; Xu, 2002). Because in this case, the participant would simultaneously view two features from the same type, it would be more difficult to decode than one feature corresponding to one feature type. Hence, we need to design more complex experiments to answer it from the perspective of neural decoding. Second, we should increase the memory load to evaluate the object-based mechanism under more complicated conditions. Thirdly, the storage mechanism of different features of real-world objects might be different with simple features (e.g., orientation, position and color) in memory (Utochkin & Brady, 2019). We need to apply the more complex and meaningful stimuli to explore the binding of object features in VWM. Finally, we also hope to use fMRI with a higher spatial resolution to see which regions dynamic processing the memory in our brains and further explain the dynamic object-based encoding process in the future.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the National Social Science Foundation of China (17ZDA323), the Shanghai Committee of Science and Technology (19ZR1416700) and the Hundred Top Talents Program from Sun Yat-sen University to YK.

REFERENCE

- Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & De Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, 23(15), 1427–1431. <https://doi.org/10.1016/j.cub.2013.05.065>
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4, 829–839. <https://doi.org/10.1038/nrn1201>
- Bae, G. Y., & Luck, S. J. (2018). Dissociable decoding of spatial attention and working memory from EEG oscillations and sustained potentials. *Journal of Neuroscience*, 38(2), 409–422. <https://doi.org/10.1523/JNEUROSCI.2860-17.2017>
- Bae, G. Y., & Luck, S. J. (2019a). Decoding motion direction using the topography of sustained ERPs and alpha oscillations. *NeuroImage*, 184, 242–255. <https://doi.org/10.1016/j.neuroimage.2018.09.029>
- Bae, G. Y., & Luck, S. J. (2019b). Reactivation of Previous Experiences in a Working Memory Task. *Psychological Science*, 30(4), 587–595. <https://doi.org/10.1177/0956797619830398>
- Ban, H., Preston, T. J., Meeson, A., & Welchman, A. E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience*, 15, 636–643. <https://doi.org/10.1038/nn.3046>

- Bannert, M. M., & Bartels, A. (2013). Decoding the yellow of a gray banana. *Current Biology*, 23(22), 2268–2272. <https://doi.org/10.1016/j.cub.2013.09.016>
- 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>
- Cai, Y., Sheldon, A. D., Yu, Q., & Postle, B. R. (2019). Overlapping and distinct contributions of stimulus location and of spatial context to nonspatial visual short-term memory. *Journal of Neurophysiology*, 121(4), 1222–1231. <https://doi.org/10.1152/jn.00062.2019>
- Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, 11(10), 9–9. <https://doi.org/10.1167/11.10.9>
- Chen, H., & Wyble, B. (2015). The location but not the attributes of visual cues are automatically encoded into working memory. *Vision Research*, 107, 76–85. <https://doi.org/10.1016/j.visres.2014.11.010>
- Chen, H., & Wyble, B. (2016). Attribute amnesia reflects a lack of memory consolidation for attended information. *Journal of Experimental Psychology: Human Perception and Performance*, 42(2), 225–234. <https://doi.org/10.1037/xhp0000133>
- Christophe, T. B., Hebart, M. N., & Haynes, J. D. (2012). Decoding the contents of visual short-term memory from human visual and parietal cortex. *Journal of Neuroscience*, 32(38), 12983–12989. <https://doi.org/10.1523/JNEUROSCI.0184-12.2012>
- Christophel, T. B., Klink, P. C., Spitzer, B., Roelfsema, P. R., & Haynes, J. D. (2017). The Distributed Nature of Working Memory. In *Trends in Cognitive Sciences* (Vol. 21, Issue 2, pp. 111–124). Elsevier Ltd. <https://doi.org/10.1016/j.tics.2016.12.007>
- Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-Based Fusion of MEG and fMRI Reveals Spatio-Temporal Dynamics in Human Cortex During Visual Object Recognition. *Cerebral Cortex*, 26(8), 3563–3579. <https://doi.org/10.1093/cercor/bhw135>
- 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>
- Delvenne, J. F., & Bruyer, R. (2004). Does visual short-term memory store bound features? *Visual Cognition*, 11(1), 1–27. <https://doi.org/10.1080/13506280344000167>
- Dijkstra, N., Mostert, P., de Lange, F. P., Bosch, S., & van Gerven, M. A. J. (2018). Differential temporal dynamics during visual imagery and perception. *eLife*, 7. <https://doi.org/10.7554/eLife.33904>
- Drisdelle, B. L., Aubin, S., & Jolicoeur, P. (2017). Dealing with ocular artifacts

- on lateralized ERPs in studies of visual-spatial attention and memory: ICA correction versus epoch rejection. *Psychophysiology*, 54(1), 83–99. <https://doi.org/10.1111/psyp.12675>
- Ester, E. F., Anderson, D. E., Serences, J. T., & Awh, E. (2013). A neural measure of precision in visual working memory. *Journal of Cognitive Neuroscience*, 25(5), 754–761. https://doi.org/10.1162/jocn_a_00357
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience*, 29(48), 15258–15265. <https://doi.org/10.1523/JNEUROSCI.4388-09.2009>
- Ester, E. F., Sprague, T. C., & Serences, J. T. (2015). Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron*, 87(4), 893–905. <https://doi.org/10.1016/j.neuron.2015.07.013>
- Fahrenfort, J. J., Grubert, A., Olivers, C. N. L., & Eimer, M. (2017). Multivariate EEG analyses support high-resolution tracking of feature-based attentional selection. *Scientific Reports*, 7, 1886. <https://doi.org/10.1038/s41598-017-01911-0>
- Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The Perceptual Root of Object-Based Storage: An Interactive Model of Perception and Visual Working Memory. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1803–1823. <https://doi.org/10.1037/a0025637>
- Gosseries, O., Yu, Q., Larocque, J. J., Starrett, M. J., Rose, N. S., Cowan, N., & Postle, B. R. (2018). Parietal-occipital interactions underlying control- and representation-related processes in working memory for nonspatial visual features. *Journal of Neuroscience*, 38(18), 4357–4366. <https://doi.org/10.1523/JNEUROSCI.2747-17.2018>
- Grootswagers, T., Robinson, A. K., Shatek, S. M., & Carlson, T. A. (2019). Untangling featural and conceptual object representations. *NeuroImage*, 202, 116083. <https://doi.org/10.1016/j.neuroimage.2019.116083>
- Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, 458, 632–635. <https://doi.org/10.1038/nature07832>
- Hogendoorn, H., & Burkitt, A. N. (2018). Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. *NeuroImage*, 171, 55–61. <https://doi.org/10.1016/j.neuroimage.2017.12.063>
- Hong, X., Bo, K., Meyyappan, S., Tong, S., & Ding, M. (2020). Decoding attention control and selection in visual spatial attention. *Human Brain Mapping*, 41(14), 3900–3921. <https://doi.org/10.1002/hbm.25094>
- Hyun, J. seok, Woodman, G. F., Vogel, E. K., Hollingworth, A., & Luck, S. J. (2009). The Comparison of Visual Working Memory Representations With Perceptual Inputs. *Journal of Experimental Psychology: Human*

- Perception and Performance*, 35(4), 1140–1160.
<https://doi.org/10.1037/a0015019>
- Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2014). The dynamics of invariant object recognition in the human visual system. *Journal of Neurophysiology*, 111(1), 91–102. <https://doi.org/10.1152/jn.00394.2013>
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of Visual Short-Term Memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26(3), 683–702. <https://doi.org/10.1037/0278-7393.26.3.683>
- Johnson, M. R., & Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Frontiers in Human Neuroscience*, 8(1 FEB), 59. <https://doi.org/10.3389/fnhum.2014.00059>
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology*, 111(10), 1745–1758. [https://doi.org/10.1016/S1388-2457\(00\)00386-2](https://doi.org/10.1016/S1388-2457(00)00386-2)
- King, J. R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two distinct dynamic modes subtend the detection of unexpected sounds. *PLoS ONE*, 9(1), e85791.
<https://doi.org/10.1371/journal.pone.0085791>
- Koch, G. E., Paulus, J. P., & Coutanche, M. N. (2020). Neural Patterns are More Similar across Individuals during Successful Memory Encoding than during Failed Memory Encoding. *Cerebral Cortex (New York, N.Y. : 1991)*, 30(7), 3872–3883. <https://doi.org/10.1093/cercor/bhaa003>
- Koenig-Robert, R., & Pearson, J. (2019). Decoding the contents and strength of imagery before volitional engagement. *Scientific Reports*, 9, 3504.
<https://doi.org/10.1038/s41598-019-39813-y>
- LaRocque, J. J., Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2013). Decoding attended information in short-term memory: An EEG study. *Journal of Cognitive Neuroscience*, 25(1), 127–142. https://doi.org/10.1162/jocn_a_00305
- Lescroart, M. D., & Gallant, J. L. (2019). Human Scene-Selective Areas Represent 3D Configurations of Surfaces. *Neuron*, 101(1), 178-192.e7.
<https://doi.org/10.1016/j.neuron.2018.11.004>
- Liu, S., Yu, Q., Tse, P. U., & Cavanagh, P. (2019). Neural Correlates of the Conscious Perception of Visual Location Lie Outside Visual Cortex. *Current Biology*, 29(23), 4036-4044.e4.
<https://doi.org/10.1016/j.cub.2019.10.033>
- Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. *NeuroImage*, 201, 116001. <https://doi.org/10.1016/j.neuroimage.2019.07.014>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281.

- <https://doi.org/10.1038/36846>
- Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: From psychophysics and neurobiology to individual differences. In *Trends in Cognitive Sciences* (Vol. 17, Issue 8, pp. 391–400). Elsevier Current Trends. <https://doi.org/10.1016/j.tics.2013.06.006>
- Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, 49(6), 1632–1639. <https://doi.org/10.1016/j.neuropsychologia.2010.11.031>
- Mares, I., Ewing, L., Farran, E. K., Smith, F. W., & Smith, M. L. (2020). Developmental changes in the processing of faces as revealed by EEG decoding. *NeuroImage*, 211, 116660. <https://doi.org/10.1016/j.neuroimage.2020.116660>
- Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. In *Trends in Neurosciences* (Vol. 29, Issue 6, pp. 317–322). Elsevier Current Trends. <https://doi.org/10.1016/j.tins.2006.04.001>
- Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2013). A meta-Analysis of executive components of working memory. *Cerebral Cortex*, 23(2), 264–282. <https://doi.org/10.1093/cercor/bhs007>
- Noah, S., Powell, T., Khodayari, N., Olivan, D., Ding, M., & Mangun, G. R. (2020). Neural Mechanisms of Attentional Control for Objects: Decoding EEG Alpha When Anticipating Faces, Scenes, and Tools. *Journal of Neuroscience*, 40(25), 4913–4924. <https://doi.org/10.1523/JNEUROSCI.2685-19.2020>
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. In *Trends in Cognitive Sciences* (Vol. 10, Issue 9, pp. 424–430). Elsevier Current Trends. <https://doi.org/10.1016/j.tics.2006.07.005>
- Olson, I. R., & Jiang, Y. (2002). Is visual short-term memory object based? Rejection of the “strong-object” hypothesis. *Perception and Psychophysics*, 64(7), 1055–1067. <https://doi.org/10.3758/BF03194756>
- Pantazis, D., Fang, M., Qin, S., Mohsenzadeh, Y., Li, Q., & Cichy, R. M. (2018). Decoding the orientation of contrast edges from MEG evoked and induced responses. *NeuroImage*, 180, 267–279. <https://doi.org/10.1016/j.neuroimage.2017.07.022>
- Parra, M. A., Cubelli, R., & della Sala, S. (2011). Lack of color integration in visual short-term memory binding. *Memory and Cognition*, 39(7), 1187–1197. <https://doi.org/10.3758/s13421-011-0107-y>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Cournapeau, D., Passos, A., Brucher, M., Perrot, M., & Duchesnay, M. (2011). Scikit-learn: Machine Learning in Python. In *Machine Learning in Python*.

- Journal of Machine Learning Research* (Vol. 12). Microtome Publishing.
<https://hal.inria.fr/hal-00650905v2>
- Pertsov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Attention, Perception, and Psychophysics*, 76(7), 1914–1924. <https://doi.org/10.3758/s13414-013-0541-y>
- Quinlan, P. T., & Cohen, D. J. (2011). Object-based representations govern both the storage of information in visual short-term memory and the retrieval of information from it. *Psychonomic Bulletin and Review*, 18(2), 316–323. <https://doi.org/10.3758/s13423-011-0064-2>
- Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information during mental imagery. *NeuroImage*, 50(2), 818–825. <https://doi.org/10.1016/j.neuroimage.2009.11.084>
- Riggs, K. J., Simpson, A., & Potts, T. (2011). The development of visual short-term memory for multifeature items during middle childhood. *Journal of Experimental Child Psychology*, 108(4), 802–809. <https://doi.org/10.1016/j.jecp.2010.11.006>
- Robinson, A. K., Grootswagers, T., & Carlson, T. A. (2019). The influence of image masking on object representations during rapid serial visual presentation. *NeuroImage*, 197, 224–231. <https://doi.org/10.1016/j.neuroimage.2019.04.050>
- Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316), 1136–1139. <https://doi.org/10.1126/science.aah7011>
- Schlegel, A., Kohler, P. J., Fogelson, S. V., Alexander, P., Konuthula, D., & Tse, P. U. (2013). Network structure and dynamics of the mental workspace. *Proceedings of the National Academy of Sciences of the United States of America*, 110(40), 16277–16282. <https://doi.org/10.1073/pnas.1311149110>
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207–214. <https://doi.org/10.1111/j.1467-9280.2009.02276.x>
- Shatek, S. M., Grootswagers, T., Robinson, A. K., & Carlson, T. A. (2019). Decoding images in the mind's eye: The temporal dynamics of visual imagery. *Vision (Switzerland)*, 3(4), 53. <https://doi.org/10.3390/vision3040053>
- Shen, M., Tang, N., Wu, F., Shui, R., & Gao, Z. (2013). Robust object-based encoding in visual working memory. *Journal of Vision*, 13(2), 1–1. <https://doi.org/10.1167/13.2.1>
- Shin, H., & Ma, W. J. (2016). Crowdsourced single-trial probes of visual working memory for irrelevant features. *Journal of Vision*, 16(5), 10–10. <https://doi.org/10.1167/16.5.10>
- Smith, F. W., & Smith, M. L. (2019). Decoding the dynamic representation of facial expressions of emotion in explicit and incidental tasks.

- NeuroImage*, 195, 261–271.
<https://doi.org/10.1016/j.neuroimage.2019.03.065>
- Sprague, T. C., Ester, E. F., & Serences, J. T. (2016). Restoring Latent Visual Working Memory Representations in Human Cortex. *Neuron*, 91(3), 694–707. <https://doi.org/10.1016/j.neuron.2016.07.006>
- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, 29(5), 1565–1572.
<https://doi.org/10.1523/JNEUROSCI.4657-08.2009>
- Swan, G., Collins, J., & Wyble, B. (2016). Memory for a single object has differently variable precisions for relevant and irrelevant features. *Journal of Vision*, 16(3), 32–32. <https://doi.org/10.1167/16.3.32>
- Teichmann, L., Quek, G. L., Robinson, A. K., Grootswagers, T., Carlson, T. A., & Rich, A. N. (2020). The Influence of Object-Color Knowledge on Emerging Object Representations in the Brain. *Journal of Neuroscience*, 40(35), 6779–6789. <https://doi.org/10.1523/JNEUROSCI.0158-20.2020>
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
[https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory and Cognition*, 34(8), 1704–1719.
<https://doi.org/10.3758/BF03195932>
- Utochkin, I. S., & Brady, T. F. (2019). Independent Storage of Different Features of Real-World Objects in Long-Term Memory. *Journal of Experimental Psychology: General*, 149(3), 530–549.
<https://doi.org/10.1037/xge0000664>
- Vetter, P., Smith, F. W., & Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. *Current Biology*, 24(11), 1256–1262.
<https://doi.org/10.1016/j.cub.2014.04.020>
- Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 27(1), 92–114. <https://doi.org/10.1037/0096-1523.27.1.92>
- Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, 131(1), 48–64.
<https://doi.org/10.1037/0096-3445.131.1.48>
- Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in Systems Neuroscience*, 9(september), 123.
<https://doi.org/10.3389/fnsys.2015.00123>
- Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, 20, 864–871. <https://doi.org/10.1038/nn.4546>
- Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of

- an object's features in visual working memory. *Psychonomic Bulletin and Review*, 15(1), 223–229. <https://doi.org/10.3758/PBR.15.1.223>
- Xie, S., Kaiser, D., & Cichy, R. M. (2020). Visual Imagery and Perception Share Neural Representations in the Alpha Frequency Band. *Current Biology*, 30(13), 2621–2627.e5. <https://doi.org/10.1016/j.cub.2020.04.074>
- Xing, Y., Ledgeway, T., McGraw, P. V., & Schluppeck, D. (2013). Decoding working memory of stimulus contrast in early visual cortex. *Journal of Neuroscience*, 33(25), 10301–10311. <https://doi.org/10.1523/JNEUROSCI.3754-12.2013>
- Xu, Y. (2002). Limitations of object-based feature encoding in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 458–468. <https://doi.org/10.1037/0096-1523.28.2.458>
- Yin, J., Zhou, J., Xu, H., Liang, J., Gao, Z., & Shen, M. (2012). Does high memory load kick task-irrelevant information out of visual working memory? *Psychonomic Bulletin and Review*, 19(2), 218–224. <https://doi.org/10.3758/s13423-011-0201-y>

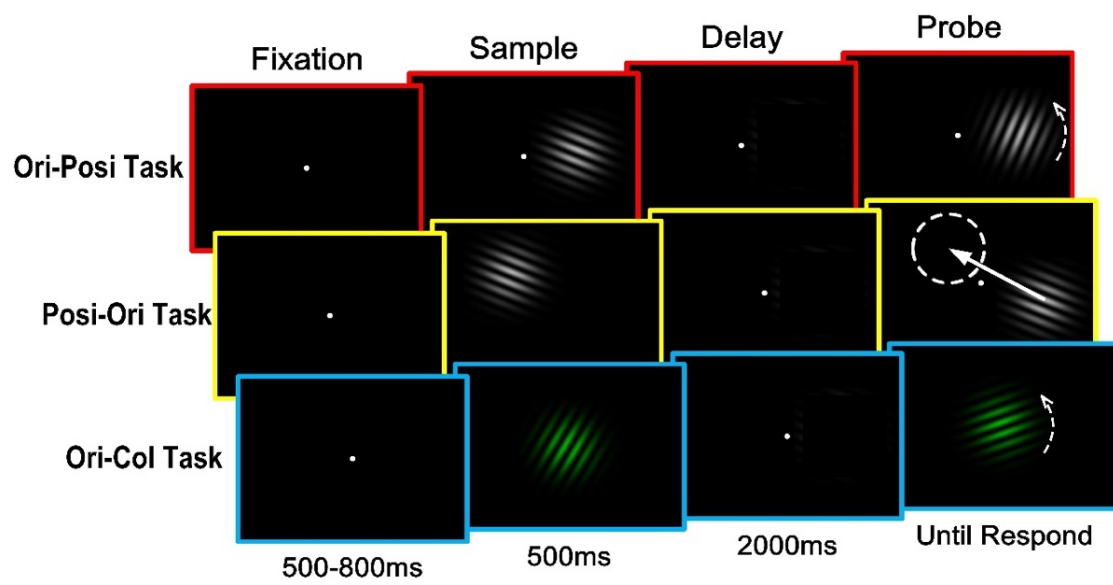


Figure 1. Trial sequence in three tasks.

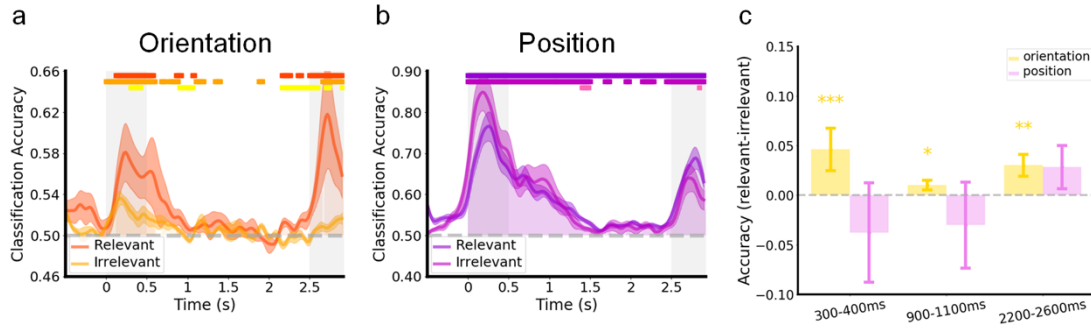


Figure 2. Time-by-time decoding accuracies in task 1 and task 2. a) Decoding results for orientation as the relevant feature in task 1 and as the irrelevant feature in task 2. b) Decoding results for position as the relevant feature in task 2 and irrelevant feature in task 1. The baseline of classification accuracy is 50%. Gray shadow represents the stimulus period and response period. The curve shows the average accuracies and the light shadow of the curve indicates \pm SEM. The square points in the top two lines and light color areas indicate the time point significantly higher than the random level (permutation test, $p < 0.05$). The yellow square points (a) and pink square points (b) indicate the time point that decoding accuracies between relevant condition and irrelevant condition are significantly different (permutation test, $p < 0.05$). c) The differences in decoding accuracies of orientation and position between task-relevant and task-irrelevant during three time periods (300-400ms, 900-1100ms, and 2200-2600ms). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Irrelevant-Position - Relevant-Orientation

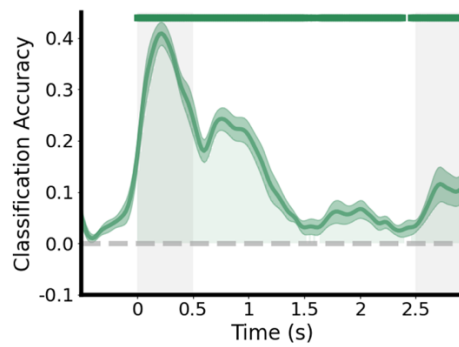


Figure 3. Differences of decoding accuracies between the task-irrelevant position and the task-relevant orientation. The baseline of classification accuracy is 50%. Gray shadow represents the stimulus period and response period. The curve shows the average differences and the light shadow of the curve indicates \pm SEM. The square points indicate the time point that decoding accuracies between the task-irrelevant position and task-relevant orientation are significantly different (permutation test, $p < 0.05$).

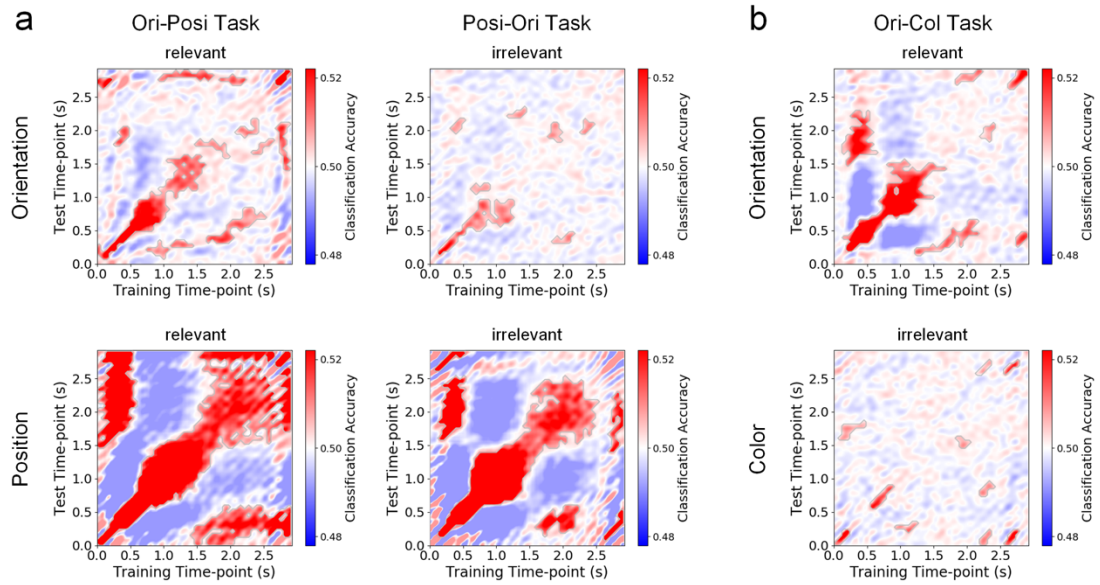


Figure 4. Temporal generalization decoding results. a) Cross-temporal decoding results for orientation and position in task 1 and task 2. b) Cross-temporal decoding results for orientation and color in task 3. These matrices show the average classification accuracy corresponding to a training time-point and a test time-point across all pairwise comparisons. The light grey outline indicates that the average accuracy is significantly higher than chance (cluster-based permutation test, $p < 0.01$).

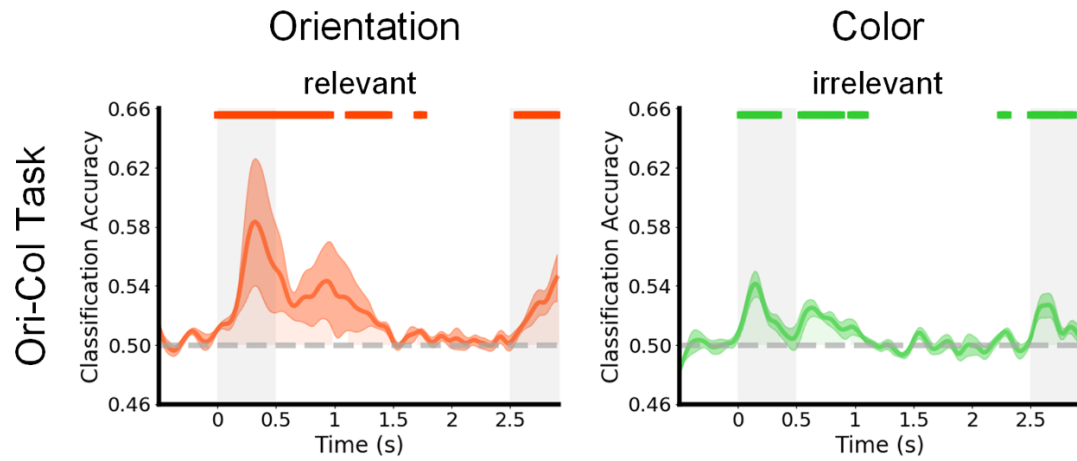


Figure 5. Time-by-time decoding accuracies in task 3. The baseline of classification accuracy is 50%. Gray shadow represents the stimulus period and response period. The curve shows the average accuracies and the light shadow of the curve indicates \pm SEM. The upper square points and light color area indicate the time point significantly higher than the random level (permutation test, $p < 0.05$).

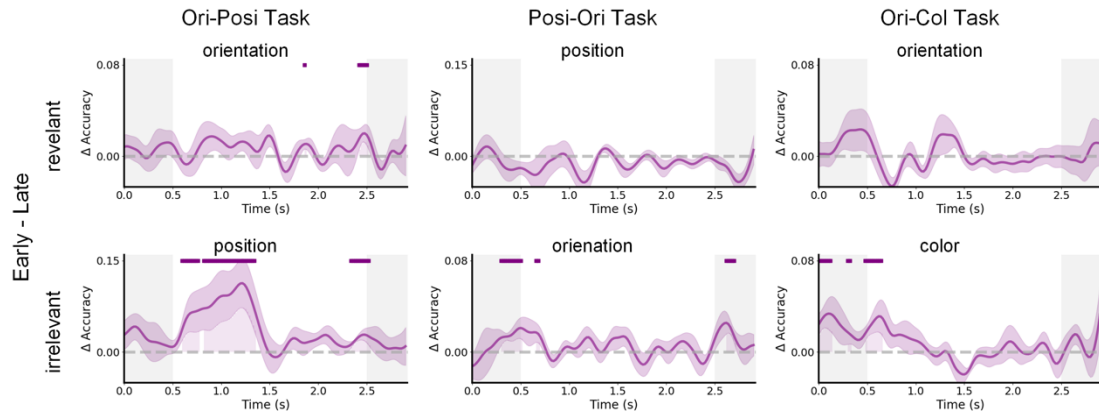


Figure 6. Differences between the classification performances of the two phases. The top shows the result for the task-relevant feature, and the bottom shows the results for the task-irrelevant feature—three tasks from left to right. The curves below show the differences time-by-time. The upper square points and light color area in curves indicate the time point significantly higher than the random level (permutation test, $p < 0.05$).

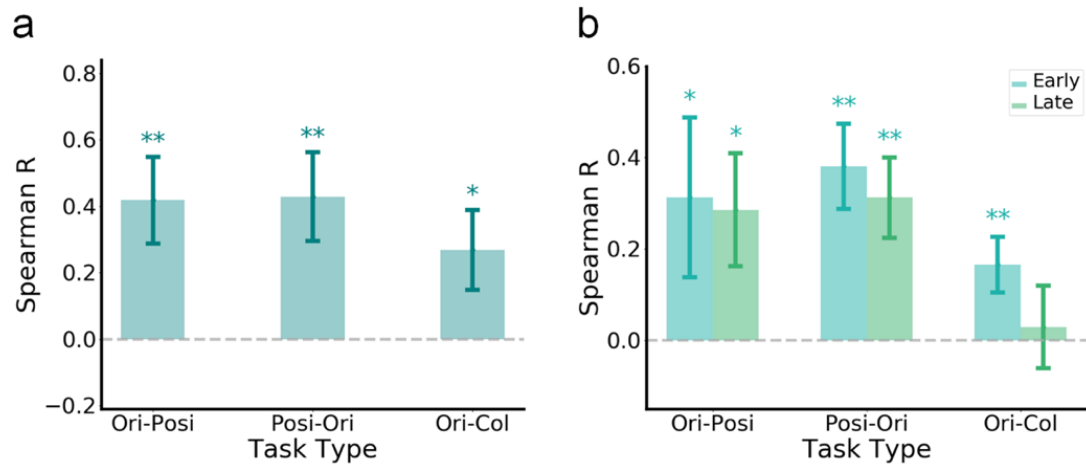


Figure 7. Correlations of decoding accuracies between the task-relevant and task-irrelevant features. a) Correlations of decoding accuracies computed on the whole trials. b) Correlations of decoding accuracies computed on trials of the two phases. The correlation coefficients were calculated based on Spearman Correlation. * $p < 0.05$, ** $p < 0.01$.

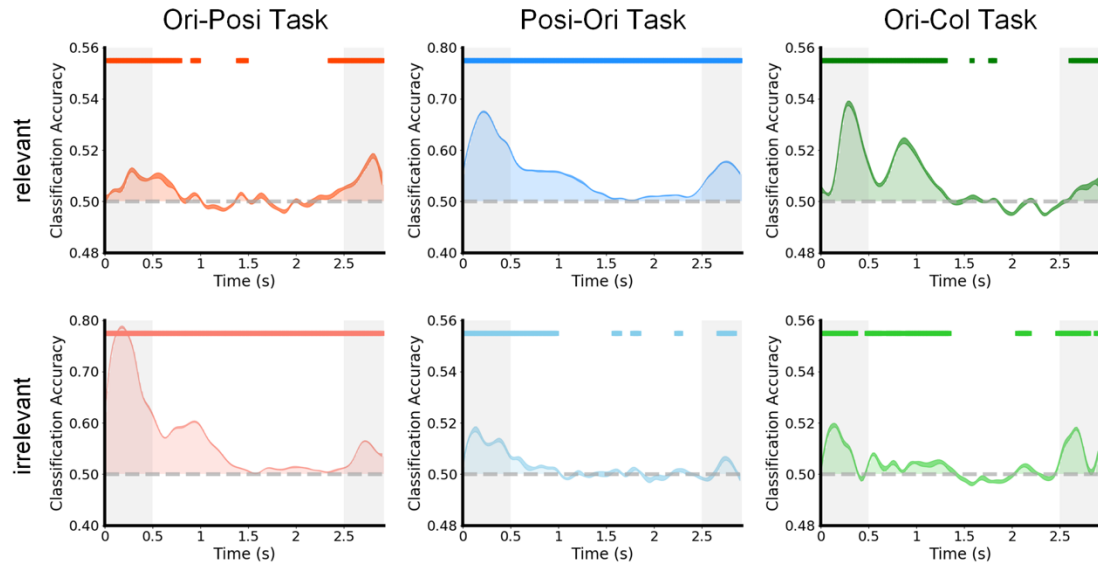


Figure 8. Across-subjects decoding accuracy. The baseline of classification accuracy is 50%. Gray shadow represents the stimulus period and response period. The curve shows the average accuracy and the light shadow of the curve indicates \pm SEM. The upper square points and light color area indicate the time point significantly higher than the random level ($p < 0.001$).