

Decoding Current and Previous Orientation from Small Targets in the Periphery Using Magnetoencephalography

Henry Allen

An Undergraduate Honors Thesis
Submitted to the Cognitive Science Program
University of California, Berkeley

Spring 2021

Advisors: Yuki Murai, PhD and David Whitney, PhD
Readers: David Whitney, PhD and Kevin Weiner, PhD

Abstract

Orientation is one of the most basic visual features encoded in the brain, and many neuroimaging studies have used decoding techniques to reveal how orientation information is represented in the brain (Haynes and Rees, 2005; Kamitani and Tong, 2005; Garcia et al., 2013; Cichy et al., 2015). However, few studies have investigated how orientation decoding is affected by stimulus history, despite research that shows orientation perception is biased towards recent orientation stimuli (Fischer and Whitney, 2014). This bias towards previous stimuli - serial dependence - facilitates perceptual stability, counteracting noisy visual inputs resulting from head movements and lighting changes. In this study, we tested if current and previous orientation information can be decoded from magnetoencephalography (MEG) data. In the experiment, we showed participants small, randomly-oriented Gabor patches in the periphery, with MEG signals recorded concurrently. We decoded current and previous Gabor patch orientations at various timings relative to the stimulus onset. Our model achieved significant mean decoding performance for the current stimulus orientation, and revealed significant decoding structure from 225 to 350 ms after stimulus onset. This structure mirrors the temporal structure of event-related fields (ERFs) in the contralateral visual cortex after stimulus onset, which peaked around 150 to 225ms, on average. We also revealed a significant decoding accuracy for previous orientation at 175ms after stimulus onset. Our results suggest that both current and previous orientation information are encoded in current stimulus ERFs.

Introduction

Detecting the orientation of contrast edges is a fundamental function in the visual cortex (Hubel and Wiesel, 1959, 1968). In the past twenty years, research has shown that the orientation of a stimulus can be decoded from the human brain using MEG and fMRI (Haynes and Rees, 2005; Kamitani and Tong, 2005; Mannion et al., 2009; Cichy et al., 2015). However, recent experiments in visual serial dependence have shown that the perception of stimulus orientation is serially dependent: the perception of the current stimulus is biased towards the orientation of the previous stimulus (Fischer and Whitney, 2014). Visual serial dependence suggests that the orientation perception of the current stimulus is biased towards the orientation of the previous stimulus. To what extent are orientation decoding models affected by stimulus history? Here, we investigated the impact of previous stimuli on orientation decoding with MEG.

Orientation

Orientation is a well understood mechanism in the visual cortex, thanks largely to the works of Hubel and Wiesel. Hubel and Wiesel first discovered cells in the visual cortex that were sensitive to orientation in 1958, when measuring the excitation of neurons in the cat primary visual cortex (Hubel and Wiesel, 1959). Hubel and Wiesel (1962) then revealed the existence of simple and complex receptive fields in the cat visual cortex, both maximally sensitive to a particular orientation. Simple cells have very narrow excitatory receptive fields in which they respond to a single orientation at a single spatial location. These excitatory regions are surrounded by inhibitory regions, which generate inhibitory responses when aligned with the correct orientation. Complex cells have much wider excitatory receptive fields, as they take in inputs from multiple simple cells tuned to a certain orientation in a spatial region. Hubel and Wiesel (1968) further revealed the existence of slabs in the primary visual cortex of a macaque and spider monkey that were made up of these simple and complex cells, each tuned to a small range of orientations. Later work revealed that simple cells were not bar detectors, as Hubel and Weisel supposed, but spatial frequency detectors (De Valois et al., 1982; De Valois et al., 1978).

Orientation columns are organized around loci called pinwheels, and orientation selectivity continuously changes clockwise or counterclockwise around these pinwheels. Yacoub et al. (2008) visualized the pinwheel structure in the human visual cortex, using high-field fMRI. The primary visual cortex is generally retinotopic (Engel et al., 1997), i.e. the orientation columns that are spatially close to each other in the primary visual cortex represent orientations that arrive spatially close to each other

on the retina. This uniform, retinotopic cortical organization lends itself well to decoding analysis because there is a spatial correspondence between orientation and cortical activations that can be exploited by decoding models. Further, because simple cells are receptive to spatial frequencies, we can design stimuli, such as Gabor patches or gratings, that maximally activate simple and complex cells.

Temporal Dependencies in Visual Perception

There is a notable difference between our perception of the world and the signal that arrives at the photoreceptors and early visual areas. Real world visual input typically arrives in a discontinuous and noisy stream. The visual system has to piece together a continuous view of the world from this noisy input stream, so it uses certain bias mechanisms to smooth out our perception of the world. It is important to be able to both understand how perceptual mechanisms affect decoding algorithms and how we can model these perceptual mechanisms in decoding algorithms, as this will allow us to reconstruct a much better representation of a continuous and coherent perception of the world, rather than just reconstructing noisy inputs arriving from the LGN or primary visual cortex.

There are many perceptual mechanisms that the mind uses to create a more continuous and organized perception of the world. Adaptation is among the most studied of these mechanisms, and affects the perception of everything ranging from color (Webster and Mollon, 1997) to faces (Webster and Macleod, 2011). Adaptation suggests that neurons become attuned to regularities in signals, saving energy by firing less when receiving regular signals. This adaptation mechanism relies on an understanding that the world generally remains constant, and it is more efficient for neurons to adapt to the world. However, over-adaptation can have a negative impact on visual perception, as a repulsion after-effect can result from long exposure to a single stimulus. In orientation perception, this manifests as the perceptual repulsion of a new orientation away from the adapted orientation (He and MacLeod, 2001).

Serial Dependence

Another representative temporal dependency in visual perception is serial dependence, a perceptual mechanism that proposes that the present visual perception is systematically biased towards inputs from the recent past (Fischer and Whitney, 2014; Cicchini et al., 2014). This is a reasonable mechanism, as the world generally remains constant from moment to moment, but the visual signal arriving at the retina is highly susceptible to changes in lighting, small head and eye movements, and noise. Therefore, serial dependence systemically biases new, noisy inputs towards a previous, consistent reconstruction of the world. Fischer and Whitney (2014) first showed that a serial dependence bias existed in the perception of orientation, finding that the perceived orientation of the current stimulus was biased up to 10 degrees toward the previous stimulus. This bias peaked when the previous stimulus was 20-30 degrees away from the current stimulus in either direction, decaying as the orientation difference increased. This bias function is systemic and significant, showing that there is a scaled correction in orientation perception towards previously seen orientations. This shows that there is a serial dependence bias at one of the basic levels of visual perception, potentially affecting primary visual cortex neurons.

Serial dependence also occurs in larger ensembles of oriented Gabor patches (Manassi et al., 2017). Manassi et al. (2017) found that the perceived average orientation of a group of oriented Gabor patches was susceptible to the same serial dependence bias function found with single oriented gratings. This ensemble serial dependence suggests that serial dependence is also associated with scene processing, i.e. calculating average orientation. Calculating average orientation, or other statistical methods performed by the visual system, are important in determining the gist of a scene; for serial dependence to appear in these integrative "gist" calculations suggests that there is a systematic bias in scene perception. This would support the theory that serial dependence is used to smooth noisy visual input streams, as the smaller steps between scene gists would result in a more continuous perception. Further, the results of Manassi et al. (2017) confirm results from Fischer and Whitney (2014) that

serial dependence effects are relatively long-lasting, up to 5-10 seconds.

Decoding and Encoding

Decoding and encoding methods are essential to our understanding of the computational, or algorithmic, aspects of mind. Decoding methods attempt to find $\text{Pr}(\text{mental representation} \mid \text{brain activity})$, i.e. finding the likelihood that a certain mental representation was present during some recorded brain activity, while encoding models find $\text{Pr}(\text{brain activity} \mid \text{mental representation})$ (King et al., 2018). Here, we were interested in decoding orientation from magnetoencephalography (MEG) signals, i.e. predicting what stimulus was shown to a subject given some MEG brain activity. This type of analysis is important because it can lead to sensitive and specific neural responses corresponding to a certain mental representation (King et al., 2018). The neural mechanisms behind orientation in the primary visual cortex (Hubel and Wiesel, 1959) are well understood, but decoding analysis of orientation can still lead to a more specific understanding of the perceptual mechanisms, temporal dynamics, and frequency responses underlying orientation perception. In this research, we tested the limits of decoding with small, peripheral, randomly oriented gratings, and used decoding to generate insight into the temporal dynamics of current and previous orientation decoding.

Imaging Methods for Decoding

Imaging techniques are central to decoding analysis; imaging measures the effects of processes in the brain, e.g. the blood-oxygen levels or synchronized electric field responses of neurons, which are then used to make informed guesses about those processes. However, all imaging techniques suffer some weakness, whether it be invasiveness, poor spatial locality, or poor temporal locality. For example, electroencephalography (EEG) is non-invasive and has excellent temporal locality, but has poor spatial locality; functional magnetic resonance imaging (fMRI) is similarly non-invasive, but it has excellent spatial locality and poor temporal locality; invasive methods like single or multi-unit recording cells or ECoG have great locality, but require surgery or animal testing to implement.

Initial successful attempts at decoding orientation were achieved with fMRI (Haynes and Rees, 2005; Kamitani and Tong, 2005). However, the resolution of fMRI (3mm) is lower than the sub-millimeter size of orientation columns in the visual cortex, calling into question whether or not fMRI is decoding orientation, or just discriminating between larger-scale activations related to the stimuli (Cichy et al., 2015). Further, fMRI measures BOLD response, i.e. an indirect measurement of brain activity based on the amount of oxygen delivered to neurons. This introduces both temporal latency and complex convolutions of hemodynamic responses into the decoding, making it harder to interpret decoding results (Cichy et al., 2015). Recent studies have also used MEG and EEG for decoding orientation, which could be a more direct measure of neuronal activity (Cichy et al., 2015; Pantazis et al., 2018; Garcia et al., 2013), and achieved success at discriminating between small sets of large, oriented stimuli.

In this study, we primarily used MEG. MEG measures the weak magnetic fields that are generated by the electrical activity of neuron populations (Senior et al., 2006). Any electrical current will generate a perpendicular magnetic field, and neurons are no exception. The magnetic fields associated with neurons are particularly weak, as field strength falls off quadratically with distance, and the fields must penetrate through the skull. Thus, MEG primarily measures cortical activity, rather than deep brain activity, as magnetic fields associated with deep signals fall off too quickly to be measured. MEG uses incredibly powerful superconductors to pick on these magnetic fields, making the technology incredibly expensive, and very susceptible to noise. It is also clear that MEG does not react to single neuron action potentials, but instead reacts to synchronized firings of large populations of neurons. Thus, it is likely that MEG does not pick up the magnetic field associated with action potentials, which are relatively asynchronous, but the slower post-synaptic potentials associated with aligned populations of dendrites. From these aligned dendritic populations, MEG picks up three types of neuromagnetic fields: an anterior-posterior field, a left-right field, and a vertical field, pointing in or out from the skull.

MEG has similar properties to EEG; it is non-invasive, has good temporal locality, and poor spatial locality. MEG and EEG both suffer from the inverse problem; MEG and EEG electrodes are only able to pick up on electrical activity on the two-dimensional surface area of the scalp, and it is currently impossible to localize the origin of electromagnetic activity that could occur anywhere in the brain, especially after it passes through layers of bone and tissue. Because the spatial resolution of MEG is so poor, it seems unlikely that MEG decoding can directly measure activity at the resolution of cortical orientation columns. Nevertheless, Cichy et al. (2015) and Pantazis et al. (2018) suggest that MEG is sufficient for decoding orientation at these orientation columns. Further, to combat the poor spatial locality of MEG, structural MRI data can be used to provide information about how electrical signals would pass through the skull, allowing the signals to be localized with far greater accuracy. This process is known as source localization. We investigated the performance of source localization augmented decoding in comparison to MEG-only decoding.

MEG is able to measure the direct electromagnetic effects of orientation stimuli with great temporal locality, rather than the secondary BOLD effect found in fMRI. This temporal locality allows us to analyze the initial response to the orientation decoding as it happens, and investigate the MEG response as orientation information moves to higher processing levels. This is particularly useful for something like serial dependence analysis, where perceptual changes may occur at the millisecond level.

Decoding Models

There is a wealth of statistical models that can be used to decode orientation. Orientation decoding models have achieved success using vanilla statistical models, particularly with support vector machines (SVM) in fMRI (Mannion et al., 2009) and MEG studies (Cichy et al., 2015; Pantazis et al., 2018). These decoding models are well suited to experiments with few stimulus orientation possibilities. Mannion et al. (2009) used SVM to discriminate between two highly differentiable spiral orientation figures, oriented at +45 degrees and -45 degrees. Similarly, Cichy et al. (2015) used only two gratings at +45 and -45 degrees. Pantazis et al. (2018) used 6 different stimuli, but only decoded the stimuli in pairs, effectively giving two stimuli to discriminate at a time. SVMs and other traditional decoding algorithms like logistic regression provide a great baseline to compare performance accuracy to, and generate proven results.

Decoding orientation is particularly tricky because of the circular nature of orientation. With orientation, a Gabor patch oriented at 3 degrees clockwise would look very similar to a Gabor patch oriented at 179 degrees clockwise, for example. This presents a challenge for the loss functions that are typically used in regression or classification problems, which would assign a high loss value to a prediction of 179 degrees when the actual orientation was 3 degrees, despite the angular similarity of the two orientations. This wouldn't be a problem for decoding the number of objects in a scene, for example, as predicting 179 objects in a scene with 3 objects in it should generate a high loss value. This problem also exists in decoding color, where a researcher might try to decode colors picked from a color wheel. The most natural solution to orientation decoding with a circular targets would be a circular regression, in which orientation outputs are modeled as the sine and cosine components of the orientation angle. However, this model can be very difficult to learn, especially for data with high noise and low expected decoding accuracy, like MEG data. Thus, a classification approach with continuous orientations binned into discrete classes can be a good substitute for circular regression, given small enough bins and enough data for each bin.

To solve this problem of circular decoding with a color stimulus, Brouwer and Heeger (2009) applied an Inverted Encoding Model (IEM) to their fMRI data, which first transforms the stimulus orientations into a cosine basis set. The model then performs an encoding step, to predict the BOLD response for the stimulus orientations. A following decoding step generates activations corresponding to each potential orientation that can be decoded. The IEM has also been successfully used in EEG studies to decode orientation (Garcia et al., 2013; Sprague and Serences, 2013; Sprague et al., 2015). Because the IEM gives channel responses as an output rather than orientation predictions, any bias in channel response according to previous stimuli will be apparent. This may be useful for finding

a serial dependence effect on decoding, as channel responses of orientations for the current stimulus may be biased by the relative position of the previous orientation.

Methods

Experimental Design

Participants

A total of 21 participants were recruited in this experiment. 21 participants took part in the MEG experiment, and the same 21 participants took part in the source localization experiment. All participants had normal or corrected-to-normal vision. The experiments were carried out in accordance with the recommendations of the institutional review boards, with written informed consent obtained from all the participants. Three participants were excluded based on poor performance in the behavioral task.

MEG Measurement

In a magnetically shielded room, MEG data was measured using a 360-channel whole head MEG system (Neuromag 360, Elekta), which had 204 planar gradiometers, 102 magnetometers, and 54 axial gradiometers. Magnetic signals were recorded at 1,000 hz. Planar gradiometers (204 channels) were used for the analysis, as they have relatively high signal-to-noise ratios. Planar gradiometers, located in pairs at 102 positions, measure x and y gradients.

Stimulus and Design

In our experiment, participants were shown a Gabor patch at 7 degrees eccentricity in the right visual field at a random orientation between 0 and 180 degrees for 500ms. This Gabor patch was then replaced by a random Gaussian noise mask for 1000ms, and then was followed by an inter-stimulus interval (ISI) of 250ms, in which nothing was shown. In seventy percent of trials, participants were then shown a response bar, still at 7 degrees eccentricity in the right visual field, and were asked to match the orientation of the response bar to the orientation of the stimulus Gabor patch. In the other thirty percent of trials, there was no response bar task, and participants were asked to just observe the peripheral Gabor patch target. This procedure is outlined in Figure 1.

MEG Analysis

Preprocessing

MEG data was preprocessed in Python with the MNE-Python package (Gramfort et al., 2013). Analysis was performed on 204 planar gradiometers (the 102 magnetometers were not used for analysis). Trials were epoched from the onset of a Gabor patch stimulus to 375 milliseconds after the stimulus onset, with 16 time points for each epoch, separated 25 ms apart. An MEG epoch is represented as 204 x 16 data matrix, where each of the 204 rows represents the MEG data for a single gradiometer for a particular behavioral trial, and each column represents the MEG data for all gradiometers at a 25ms time step after stimulus onset. Epochs were rejected for a subject if maximum peaks for any gradiometer exceeded a threshold of $4000 \times 10^{-13} \frac{T}{m}$. MEG data was band-pass filtered from 2-40hz, removing any potential low-frequency artifacts and unused high-frequency data. Independent component analysis (ICA) was conducted for each subject to remove artifacts caused by ECG and EOG artifacts. ICA performs source separation for statistically independent components of signals. The top eighty independent components were visually inspected for each subject, and independent components resembling ECG and EOG artifacts were excluded from the filtered MEG dataset. An example of these components is seen in figure 2. Here, we see that the left component represents artifacts resulting from eye blinks (EOG), while the right component represents ECG artifacts. Six hundred to eight hundred epochs were analyzed for each subject.

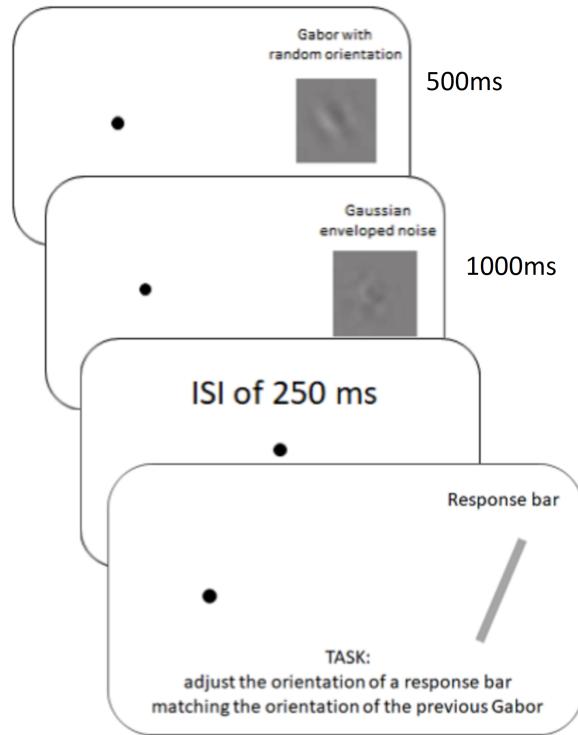


Figure 1: Stimulus time-course. Participants were asked to fixate on the dot, and stimuli were shown at seven degrees eccentricity in the right visual field. In seventy percent of trials, participants were asked to orient a response bar to match the stimulus Gabor patch target.

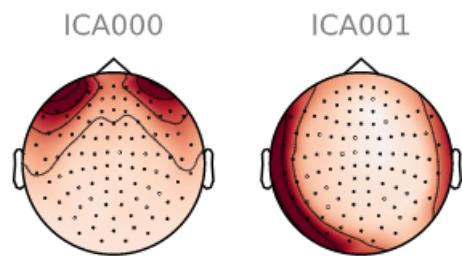


Figure 2: Example ICA Components to be excluded. Left: EOG artifact, Right: ECG artifact

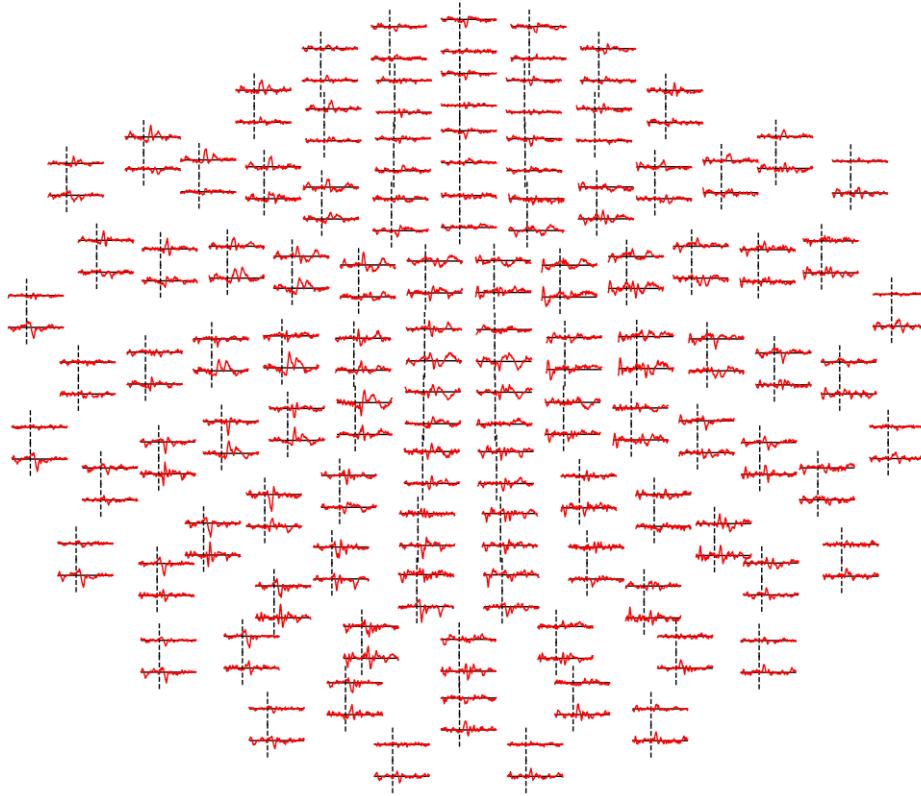


Figure 3: Example of a representative ERF for a subject

ERF Analysis

The averaged evoked responses for each subject were measured to ensure that visual responses appeared in the occipital area as expected. Time-locked epoch responses were averaged over all trial blocks for each subject at each electrode and visually analyzed for aberrations. In a typical subject, shown in Figure 3, evoked responses were observed in left occipital and temporal electrodes around 0.2 seconds after the stimulus was presented. Because the stimulus was only shown in the right visual field, there was limited activation in the right hemisphere.

Source Localization Analysis

In source localization analysis, we estimated an inverse solution, i.e. the sources of underlying neural activations underlying the MEG sensor readings (Gramfort et al., 2013). The goal was to convert from sensor space, or the readings at MEG electrodes, to source space, or the estimated activations on the cortical surface. To generate a source localization estimate of our MEG data, we followed the basic process outlined by MNE-Python (Gramfort et al., 2013). This first required pre-processing the MRI data for each subject. We first computed the cortical surface reconstruction from each subject's anatomical MR image using FreeSurfer's recon-all function (Dale et al., 1999). We then used the skull stripping approach (Segonne, 2004) implemented in FreeSurfer as mri watershed. Next, we performed co-registration with our MRI and MEG data, lining up the MEG sensors with the reconstructed MR image according to fiducial markers on the scalp. This aligns the coordinate systems of the MEG space and MRI space. Finally, we performed the MRI watershed procedure again in the new coordinate system.

We next computed a solution to the forward problem for each subject. The forward problem involves computing the external magnetic field at various sensor locations on the scalp resulting from primary currents (Mosher et al., 1999). MEG electrodes read both primary and secondary currents; primary currents represent microscopic cellular currents that are associated with cognitive processes, while secondary currents are associated with macroscopic electric fields that we are less interested in

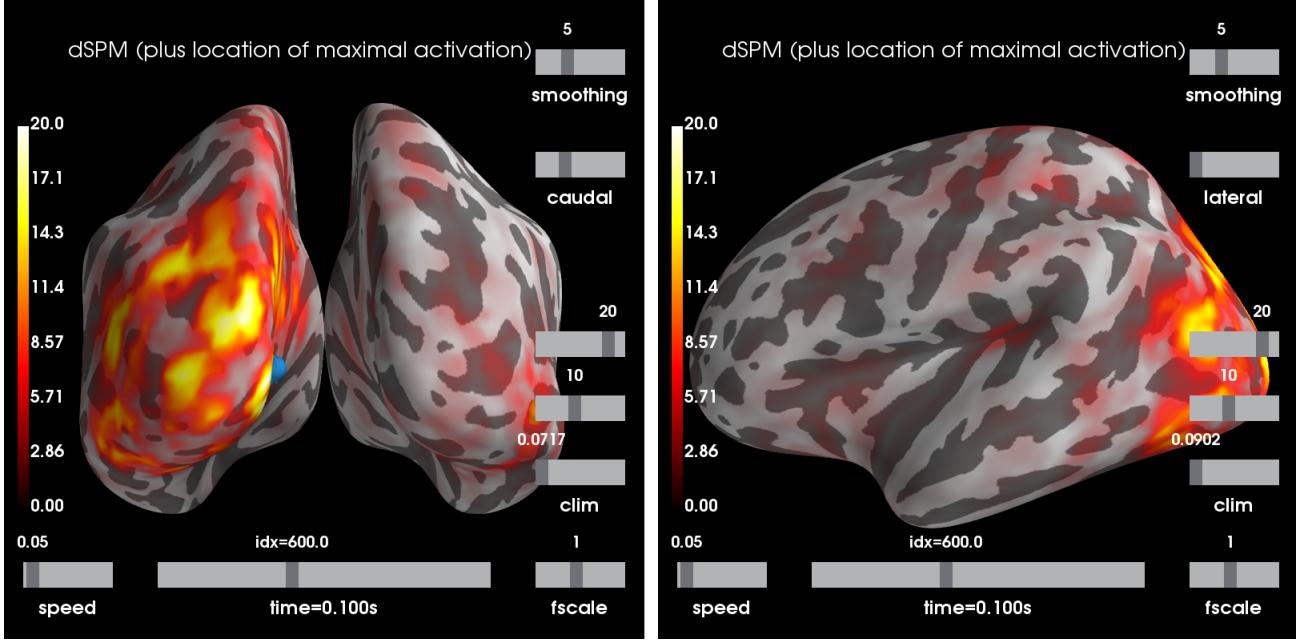


Figure 4: Average source estimate for subject one subject at 100 ms after stimulus onset. Activations originate in the occipital lobe.

(Mosher et al., 1999). To solve the forward problem, we first computed the source space locations for each subject, i.e. the locations of elementary dipole sources that we computed the forward operator at (Gramfort et al., 2013). Each dipole is analogous to a sensor location located on the cortical surface. To compute these locations, we inflated each cortical surface into a sphere, and overlaid an octahedron onto it. We then subdivided this octahedron 6 times (oct6), giving us 4098 dipole locations for each hemisphere (Gramfort et al., 2013). We then calculated the Boundary Element Method (BEM) solution with MNE-python using the linear collocation algorithm (Mosher et al., 1999; Hamalainen and Sarvas, 1989) for each subject. Using the BEM solution, source space, and the evoked response for a subject, we computed the forward solution as implemented in MNE-python.

Once the forward solution was calculated, we were able to estimate the inverse solution with MNE (Gramfort et al., 2013). We first used a weighted minimum norm estimate (Fuchs et al., 1999; Lin et al., 2006) with a loose orientation approach (Lin et al., 2005) to estimate source current densities. This approach used our forward solution and the covariance matrix from our epoched data. Source estimates were then computed from the source current densities using dynamic statistical parameter mapping (dSPM) (Dale et al., 2000). These source estimates contained source amplitudes epoched at stimulus onset for each of the 4098 sources in each hemisphere at 16 timesteps, and could then be incorporated into a decoding pipeline as numpy arrays. An example source estimate can be seen in Figure 4. To validate the results of our source localization estimates of the inverse solution, we plotted the average source space response after stimulus onset for each subject, and compared each source space response to the corresponding sensor space response. We checked that source space responses occurred in the occipital lobe, and ensured that source space peaks occurred at the same time as sensor space peaks.

Decoding Stimulus Orientation from MEG Data

Data Processing

Analysis was performed in Python using decoding libraries from MNE-Python (Gramfort et al., 2013) and the scikit-learn library (Pedregosa et al., 2011). In the experimental design, Gabor patches were shown at random orientations from 0-180 degrees. For decoding analysis, these Gabor patches were binned into n different classes for decoding model classification. Bins were used because it was

difficult to perform a circular regression of orientation angles, especially with relatively few data points. Bin sizes were chosen as a balance between minimizing the angular width of each bin (modeling the original 180 possible Gabor orientations more closely) and maximizing the number of features per bin (giving more trials for each class to improve model performance). Each bin was normalized to have the same number of features by filtering out data points for bins that had more data points than the smallest bin. Decoding models were tested with 4, 8, and 9 bins, and final analysis was performed with 9 bins. Individual models were trained for each subject before calculating group averages to account for individual differences in MEG response.

Gabor patches were binned in bow-tie ranges. These bow-tie ranges put orientations at 0 degrees and 179 degrees in the same orientation bin, as these orientations are nearly identical. For 4 orientation classes, for example, class 0 corresponds to Gabor patches oriented from (0-22.5) degrees and (157.5-180) degrees, class 1 from (22.5-67.5) degrees, class 2 from (67.5-112.5) degrees, class 3 from (112.5-157.5) degrees. If a Gabor patch were to have an orientation of 125 degrees, it would be binned into class 3, and the label of "3" would be used as the classification label in machine learning decoding models.

For each subject, we originally had 600 or 800 trial epochs. We excluded some trials for each subject to ensure that each orientation bin had the same number of trials, as uneven bin distributions could bias some estimators. Decoding model accuracy was calculated with k-fold cross-validation, where the accuracy metric was the number of correct labels predicted by the model out of the total number of labels predicted. Model accuracy was then compared with the results of a permutation test, in which we shuffled around the training and test labels of the data set to create a null distribution with which to test our null hypothesis. Models were run using both senior space (MEG electrodes) and source space (source estimates) data. There were many more MRI vertices than MEG electrodes, so the weight matrices for the source localization models were much larger than the electrode models. Thus, hyperparameters were tuned differently for the MEG electrode models and source localization models.

The final data had the shape $\mathbf{X} = (n \text{ epochs} \times m \text{ sensors/sources} \times t \text{ time steps})$ and $\mathbf{y} = (n \text{ epochs})$, $1 \leq y_i \leq 9$, $y_i \in N$, where \mathbf{X} represents the trial epochs for each subject, and \mathbf{y} represents corresponding orientation bins for each trial.

Permutation Test

A permutation test was used to determine the statistical significance of the various decoding experiments. Here, features will refer to the $n \times m \times t$ MEG data matrix with n epochs, m electrodes, and t time steps, and the ω labels refer to the classification bins of the stimulus Gabor orientations, where label i is the label corresponding to the feature data for epoch i . In this test, each model was trained 10 times with randomly selected training and test trials to determine the performance of the experimental model. The same analysis was conducted another 100 times with shuffled labels, i.e. the permutation test. In each of these 100 iterations, the labels were shuffled, separately from the corresponding feature vector. This label shuffling removes any structured relationship between features and labels. Thus, these permuted trials will give us a null distribution for our experiment, where the null hypothesis suggests that there is no structured relationship between the MEG features and the labels. To calculate a p-value, the performance of each of the $n = 100$ permutation trials is compared against the $m = 10$ experimental trials, as follows.

$$\rho = \frac{1}{m} \sum_{i=1}^m \frac{1}{n} \sum_{j=1}^n f(exp_i, perm_j)$$

$$f(exp_i, perm_j) = 1 \text{ if } perm_j \geq exp_i, 0 \text{ otherwise}$$

exp_i = performance of experimental trial i

$perm_j$ = performance of permutation trial j

This can be thought of as the proportion of permutation trials that outperform a particular experimental trial, averaged over all of the experimental trials. A low p-value suggests that the experimental

trials are on the tail end of the null distribution, suggesting that there may be a structured relationship between the features and labels of the dataset, disputing the null hypothesis.

Sliding Logistic Regression Model

The first machine learning model used was a sliding logistic regression model, which calculated a separate accuracy for each time step of our MEG input signal. For 16 time steps, we calculated 16 different logistic regression weights, and updated these weights using a gradient descent procedure while training. The model was regularized with an elastic net regularization term, which acts as a combination of L1 and L2 regression. This regularization term helped enforce sparsity within our model and prevented weights that were too large, combining the traditional benefits of L1 and L2 regularization. The model also used a "select K best" routine that selected the K best features from the model, based on their contributions to the accuracy. We then tuned the hyperparameters for these additions to our model, namely the size of K , the ratio between L1 and L2 loss for elasticnet, and the regularization parameter C . The hyperparameters were tuned to get the best cross-validation accuracy.

Here, logistic regression models $\text{Pr}(\text{orientation of stimulus } (y_i) = \omega | X_i)$. We used multinomial logistic regression for multi-class prediction with logistic regression. The logistic regression model was used for its simplicity and interpretability. The model was relatively easy to implement, especially with the machine learning packages that are already built into scikit-learn (Pedregosa et al., 2011) and mne-python (Gramfort et al., 2013). The model had only one layer, making it relatively fast to run the model and tune the model parameters. This also meant that we could interpret the weights easily, as there was a one-to-one correspondence between a particular weight and feature from an MEG electrode or MRI vertex. This told us that if a particular weight had a high value, the corresponding electrode/vertex was weighted highly in the model.

Sliding Support Vector Machine

We then implemented a similar sliding SVM model. We again used the scikit-learn (Pedregosa et al., 2011) implementation of the SVM. Our model used relatively high regularization, i.e. a small C parameter in the SVM model, as we found the SVM prone to overfitting. We again used the "select K best" routine to select the K best features in our model. Model weights were updated using a gradient descent procedure. Model accuracy was generated with 5-fold cross-validation.

Inverted Encoding Model

The final decoding model was the IEM (Brouwer and Heeger, 2009, 2011; Garcia et al., 2013; Sprague and Serences, 2013; Sprague et al., 2015). The IEM calculates the predicted responses of each classification channel based on the MEG data, rather than giving a class prediction. This model is well-suited for orientation decoding because it uses a cosine basis set to transform the input data, modeling the inputs better than a linear scale from 0 to 179. In the cosine basis set (shown in Figure 5a), a perfect channel response centered at 160 degrees would propagate equally to the right and left and wrap around the unit circle, having a similar response at 140 degrees and 180 degrees, and 120 degrees and 20 degrees. This can be seen in Figure 5b.

To compute the predicted channel response for our MEG data, we first constructed a hypothetical channel response C , a matrix of number of epochs \times number of orientation channels. This was computed by matrix multiplying a stimulus mask matrix by the cosine basis set. The size of the stimulus mask matrix was number of trials \times 180, where row i had value 1 at the degree of trial i 's stimulus orientation, and 0 elsewhere. In the problem formulation of the IEM, there was the MEG input data, B_{train} with shape (number of epochs \times number of MEG electrodes), which was related to the channel response set, C_{train} , by the following formula:

$$B_{\text{train}} = WC_{\text{train}}$$

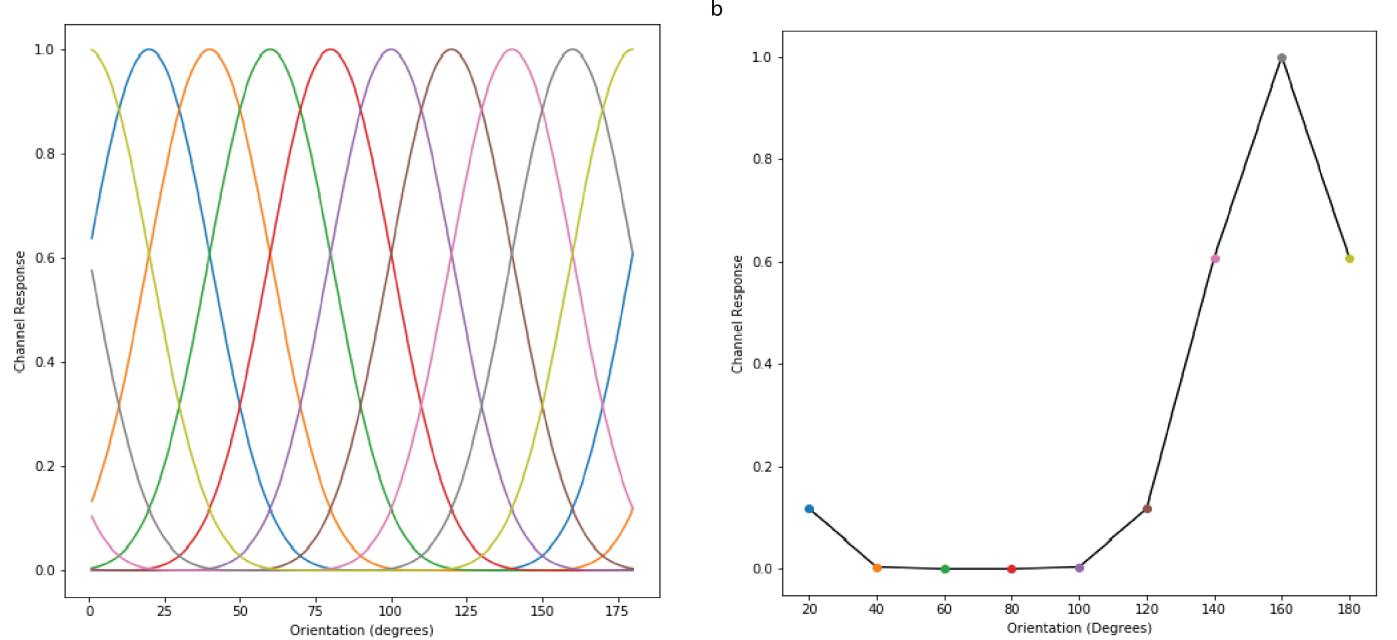


Figure 5: a) An example cosine basis set for 9 orientation classes. Each color represents the basis function for a single orientation channel, where the cosine function peaks at that orientation channel. b) The ideal channel response for a stimulus oriented at 160 degrees.

We first computed $\hat{W} = B_{\text{train}} C_{\text{train}}^T (C_{\text{train}} C_{\text{train}}^T)^{-1}$, the linear least squares estimate for W . We then solved for the predicted channel responses C_{test} of our test set, related to the test MEG data B_{test} by the formula $B_{\text{test}} = WC_{\text{test}}$, which was estimated by least squares as

$$C_{\text{test}} = (\hat{W}^T \hat{W})^{-1} \hat{W}^T B_{\text{test}}$$

We then performed k-fold cross-validation to generate accuracy over all the trials, and performed this analysis at each time step. The channel response is informative on its own, but we can also generate a classification for a trial by taking the argmax of the channel response (channel corresponding to maximum channel response) for a particular trial.

Previous Orientation Decoding

After decoding the orientation of stimulus t from MEG, we attempted to decode the orientation of the $(t - 1)$ th stimulus from the MEG response data corresponding to the t th trial. If the previous stimulus orientation biases the current stimulus orientation perception, then it's possible that a decoding model could reverse engineer this bias and decode the orientation of the previous stimulus. This analysis was computed with the same decoding models as the current stimulus decoding, but we shifted the data labels backward by one time step in relation to the data, such that the data at time t was used to predict the label at time $t - 1$. We performed the same permutation test as in previous decoding models to test the significance of this previous trial decoding.

Serial Dependence Analysis

Behavioral Analysis

Before performing decoding analysis for temporal bias (how the previous stimulus affected the decoding of the current stimulus), we performed the serial dependence analysis outlined in Fischer and Whitney (2014) to determine if any serial dependence effect existed in our data. For each subject, we calculated the error for a trial (reported orientation - presented orientation), where positive values indicate errors in the clockwise direction. For each trial, we calculated the difference between the

previous and current trial for each trial, where a positive value signifies that the previous orientation was more clockwise than the current orientation. We plotted the error on the y-axis and the previous-current trial difference on the x-axis. We then fit the error plot with a derivative of Gaussian (DoG) curve, allowing us to measure the amplitude of the serial dependence effect. This curve is modeled by the function, $y = (abc)xe^{-(bx)^2}$, where x is the relative orientation of the previous trial, a is the amplitude of the DoG curve, b is the DoG curve width, and $c = \frac{\sqrt{2}}{e^{-0.5}}$.

Error bars were generated by bootstrapping the DoG curve fit 5000 times, sampling with replacement. To generate P values for this DoG curve, a permutation test was again used, generating 100,000 DoG curves with shuffled data labels (relative previous orientation) for each iteration. The amplitudes of the permuted DoG fits were compared against the measured amplitude of the participant DoG fits, and the P value was taken as the proportion of permuted amplitudes that were greater than or equal to the measured subject amplitude. This analysis was performed using the lmfit library for python (Newville et al., 2014).

Channel Response Bias Analysis

We next investigated how the relative previous stimulus orientation affected the decoding of the current stimulus. We again calculated the relative previous stimulus orientation (previous orientation - current orientation) for each trial, where positive values indicate that the previous orientation was more clockwise than the current one. We then ran the same k-fold cross-validation decoding process on our MEG data with the IEM. We generated the channel responses for each orientation for each trial. We could then bin these channel responses and class probabilities by relative previous orientation to reveal any relationship between model predictions and relative previous orientation. Each bin was normalized to have the same number of entries. We then centered each channel response such that the actual label of the data lined up with the same label for all trials. For example, if we aligned all responses to channel 5, and trial i had a label of channel 3, we rolled the channel response for i two places to the right.

We first performed a pixel-wise analysis of channel response bias. We subtracted the mean channel response from all bins, giving us the deviation of each bin from the mean. If there was any impact of serial dependence on decoding, we would expect to see channel response biased away from the mean in certain bins. To test how significantly each bin deviates from the mean, we performed another permutation test, this time shuffling the bin that each channel response falls into. We then calculated a P value for each channel response orientation in each relative previous orientation bin. This P value was calculated as the proportion of permutations that had greater deviation from the mean than the measured channel response for said bin.

We then fit a Gaussian function to each relative previous orientation bin at each time step, as the pixel-wise analysis was not necessarily suited for a more subtle bias structure. We performed a permutation test by shuffling the relative previous orientation bin for each subject. We fit a Gaussian to each shuffled relative previous orientation bin. To generate P values for this analysis, we compared the absolute values of the decoded Gaussian means and the permutation Gaussian means.

Results

Evoked Responses

MEG ERFs

We first analyzed ERFs time-locked to the onset of the Gabor stimulus. The stimuli were displayed in the right visual field, and we found that the ERFs were strongest in the left posterior electrodes, contralateral to the stimulus presentation (Figure 6). Evoked responses peaked from 150-200 ms in these posterior electrodes. Although ERFs recorded at an electrode do not necessarily originate from cortical activity directly beneath that electrode, the high activity in the left posterior areas and the lack of activity in the frontal areas and right hemisphere suggest that our MEG signal originated from the left visual area.

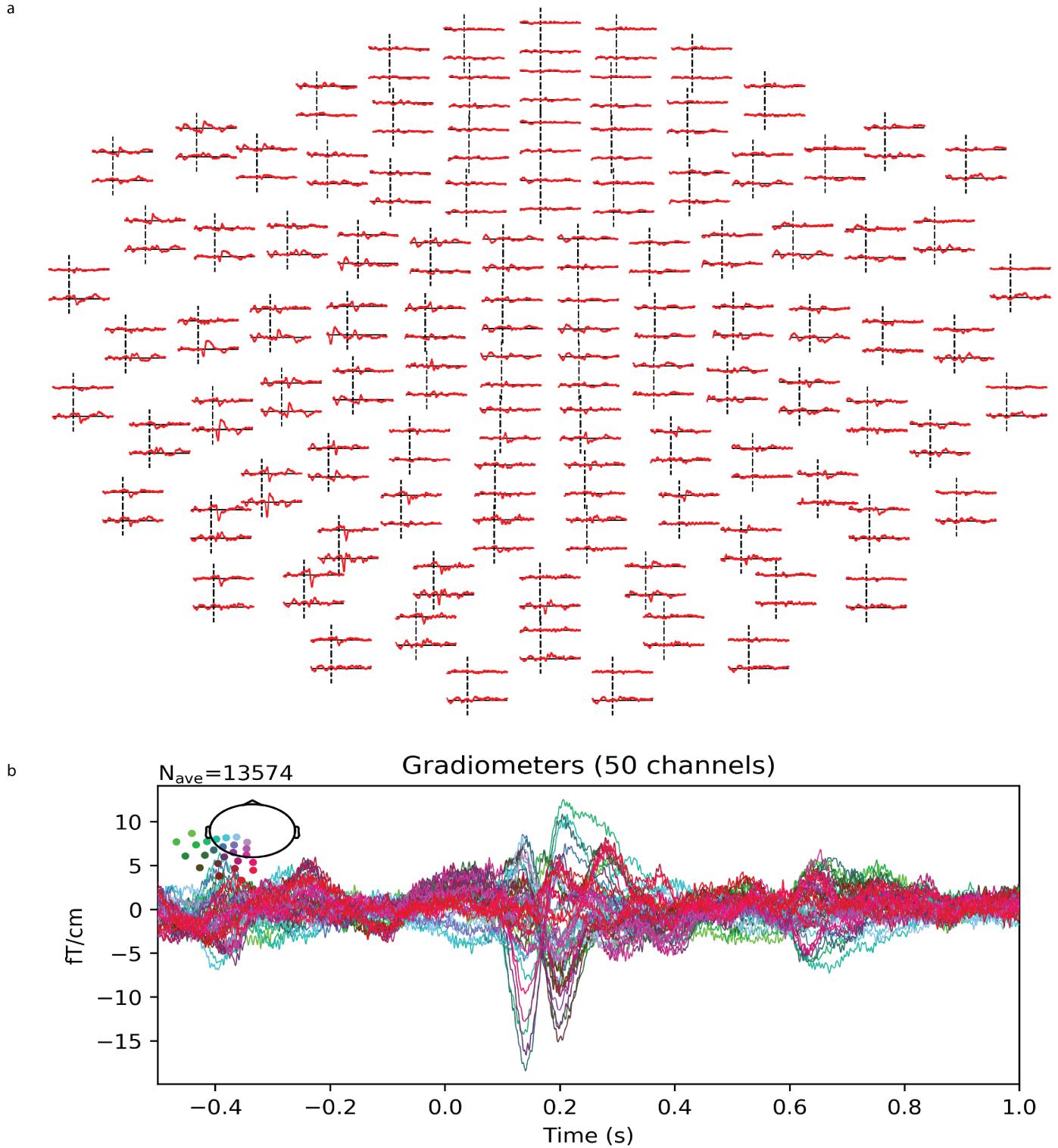


Figure 6: a) Stimulus-onset locked evoked response fields (ERFs) for all electrodes averaged across all subjects. ERFs are strongest in the left posterior electrodes, and ERF peaks begin at 150ms. b) Average ERFs for left occipital and temporal electrodes, from -0.5 to +1.5 seconds.

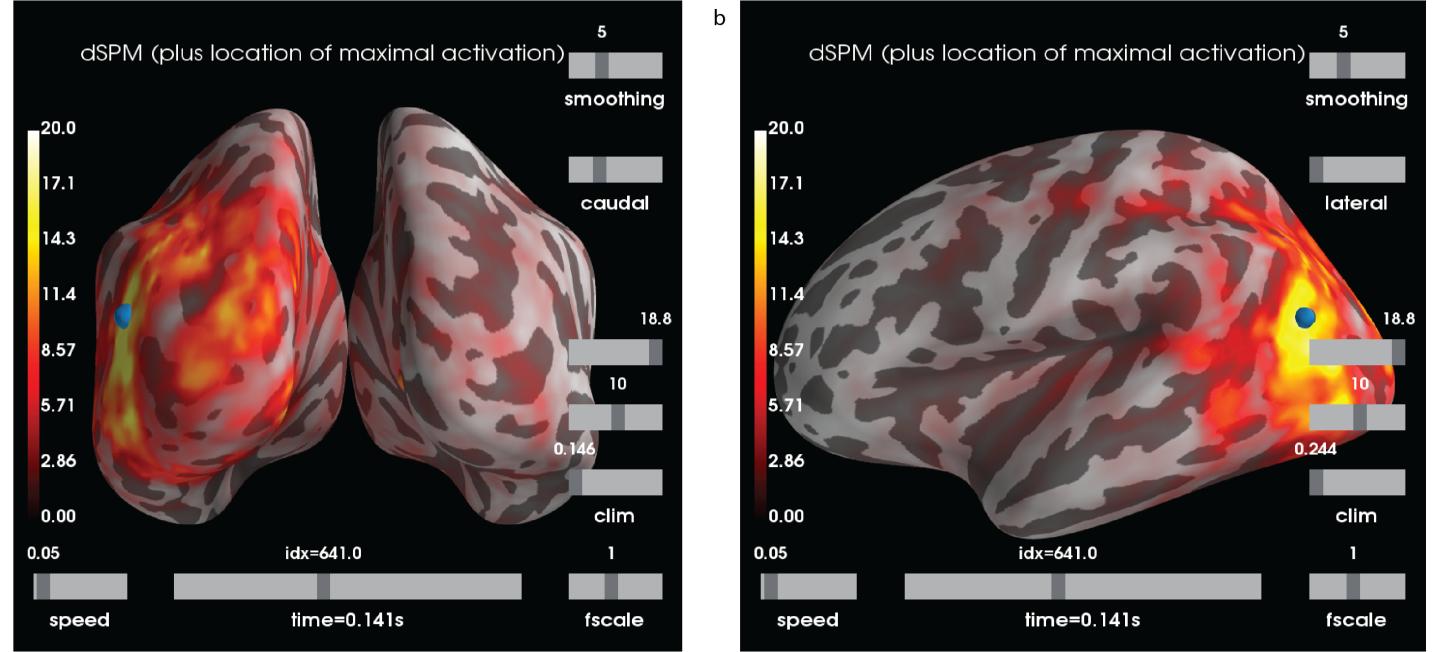


Figure 7: Peak evoked response for stimulus-locked source estimate data. Activity occurs mainly in the left posterior cortical areas, around the visual cortex. This maximum activity peak occurs at 150ms after stimulus onset.

Source Estimates

After processing the MEG epochs, we performed a source localization analysis for each subject. Figure 7 shows the averaged source estimate (morphed to the default Freesurfer "fsaverage" subject) for all subjects at the max evoked response time, around 150 ms. Similar to the ERFs, source space activations are localized to left and posterior cortical areas, particularly the visual cortex.

Current Orientation Decoding

Logistic Regression

Our first decoding analysis was performed with multi-class logistic regression. Here, we binned Gabor stimuli into 9 groups (i.e. 9 ranges of orientations). We then performed a 5-fold cross-validation on each subject to compute accuracies over 16 time steps from 0 to 375ms milliseconds. We also computed the mean accuracy over all 16 time steps. Each cross-validation was performed 10 times with shuffled data to generate decoding results. One hundred permutation tests were run with permuted data labels to generate permutation results. Figure 8 shows the logistic regression accuracy computed at each time step. Only one time step had accuracy significantly above the permutation accuracy, at 175 ms after stimulus onset. From 225 to 325 ms, we observed a peaking structure where accuracy increased over each step, but none of these time steps were individually significant. The mean accuracy over all time steps was significant, but with very low accuracy (Figure 8. Variance was very low, as mean accuracies were averaged across all 18 subjects before permutation test comparison. These results suggest that the logistic regression was able to decode orientation, but no significant decoding structure was revealed between time steps.

We also performed logistic regression decoding with our source localized data. We did not find any significant time steps, nor did we find the mean accuracy to be significant (Figure 9). This might suggest that the source estimation process introduced noise to the decoding process that reduced decoding accuracy.

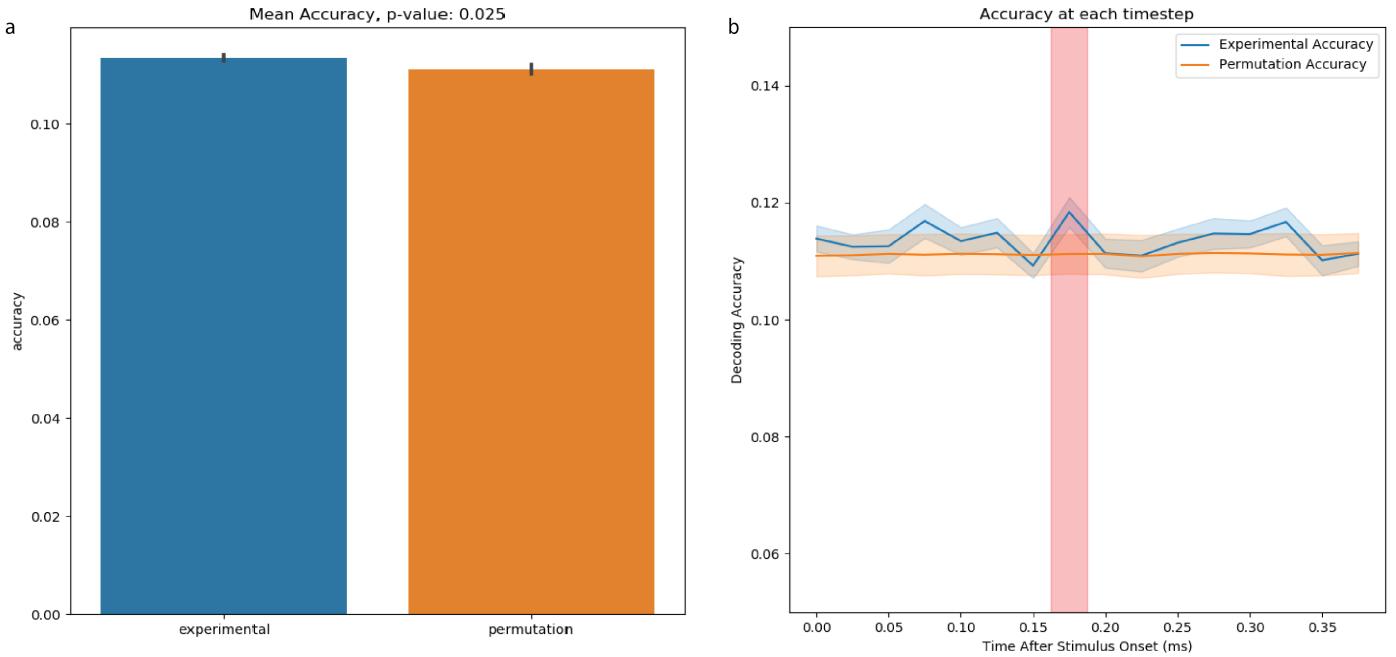


Figure 8: a) 5-fold cross-validation accuracy computed at 16 time steps, with standard deviation bars. Accuracies are averaged across 18 subjects. b) Mean accuracy over all time steps. The experimental model achieves modest accuracy gains over the mean permutation accuracy. The red highlighted section at 175ms denotes a time step with accuracy significantly above the permutation test decoding accuracy. Shaded regions indicate standard deviation of accuracy.

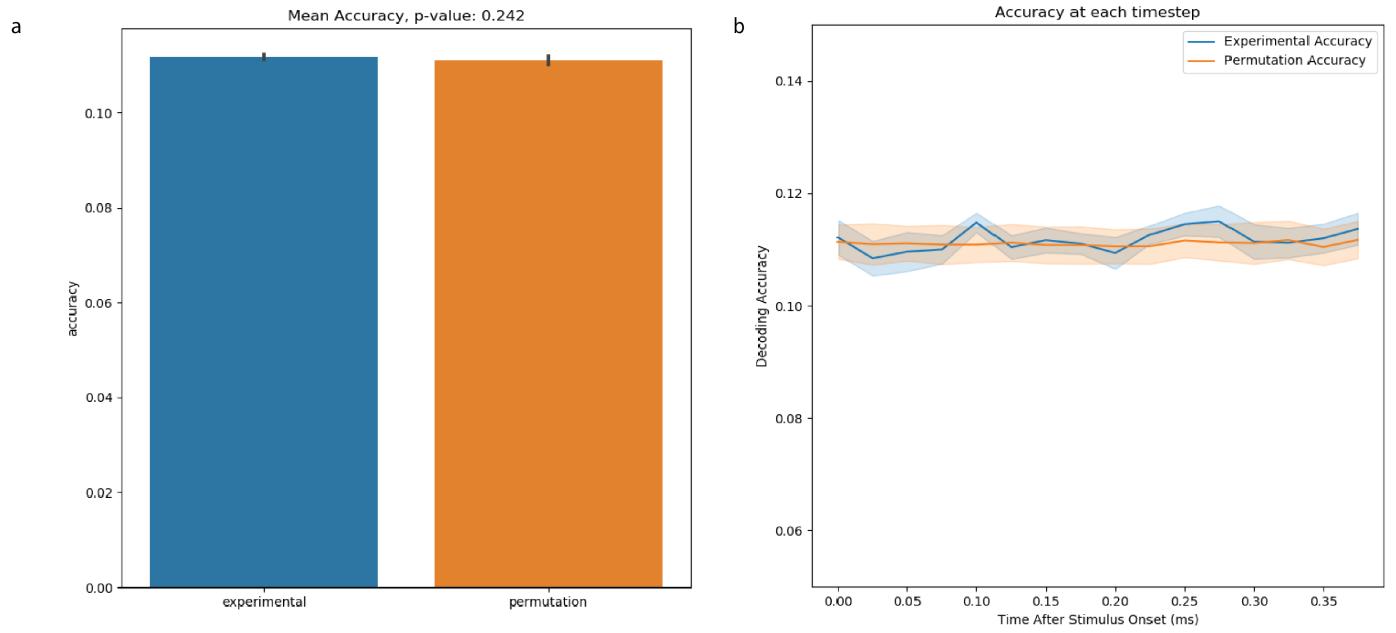


Figure 9: a) Mean accuracy across 16 time steps for 18 subjects using the logistic regression model with source space inputs, compared to permutation test accuracy at 16 time steps. b) Mean accuracy over all time steps for the logistic regression model with source space inputs, compared to a permutation test accuracy. No time steps were significant.

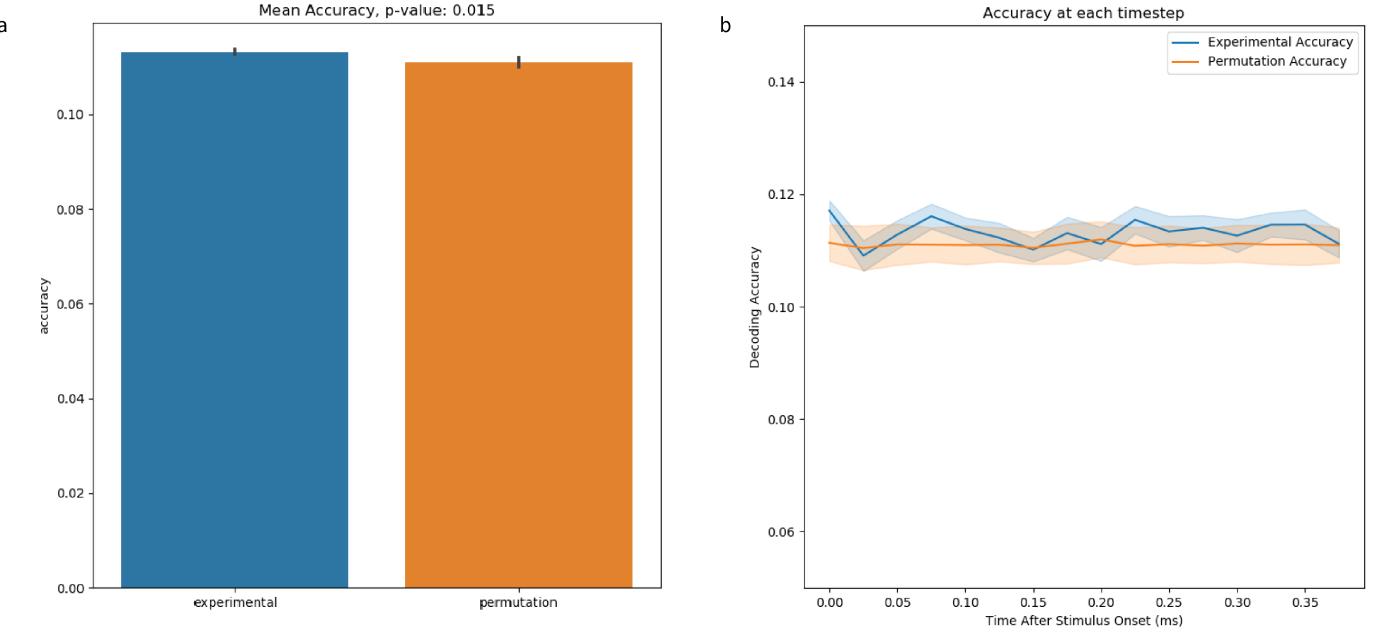


Figure 10: a) Mean accuracy over 16 points for a support vector machine compared to a permutation test. Black bars indicate decoding accuracy standard deviation. b) Accuracy at 16 time steps for a support vector machine. Shaded regions around the line plots indicate standard deviation. No significant accuracies are achieved at any time step.

Support Vector Machine

Similar results were achieved using the same decoding process with an SVM. SVMs are the most commonly used model in multi-variate pattern analysis, or decoding, but we didn't observe any performance increases over the logistic regression model here. In Figure 10, there were no individual time points with significant decoding accuracy. Further, while the decoding accuracy was generally above the permutation accuracy at each time step, it also remained within the standard deviation of the permutation time steps. We observed an upward trend in decoding accuracy beginning around 225 ms, similar to the trend we observed in the logistic regression model, but it was again not significant, based on decoding accuracy. Similar to the logistic regression, there was a small, but significant, increase in decoding accuracy in the mean accuracy ($p < 0.05$, permutation test; Figure 10). The variance for mean accuracy was again very low, as the mean accuracy was averaged across 18 subjects before the permutation test.

Inverted Encoding Model

We finally looked at the inverted encoding model (IEM) for decoding current stimulus orientation, and achieved much better results in comparison to the more traditional models. Rather than generating a prediction, the IEM predicts a channel response for each of the orientation channels. This channel response is a prediction of how much a particular orientation is associated with a set of MEG data. In Figure 11, we first looked at the mean channel response of our experimental IEM in comparison to a permutation test. The mean channel response, centered at 90 degrees for visualization, had a significant bell-curve structure centered at 90 degrees, while the permutation channel response was flat. This indicates that the IEM successfully predicted the highest channel response at the displayed orientation, and that the channel response decreased as a function of distance from the stimulus orientation. This might indicate that similar orientations have a similar encoding. These results are in line with findings from other IEM decoding papers, particularly Garcia et al. (2013). However, we reproduced these results with a wider variety of possible orientations and smaller, peripheral Gabor targets. When we looked at the channel response at each time step, we found that the

channel response was mostly flat until 175 ms, at which point it increased until it peaked at 225ms, and then decreased again until 375ms. This indicates that channel responses increased after evoked responses, which we found to peak from 150-175ms in the left posterior electrodes.

We then examined the decoding accuracy achieved by the IEM, using the orientation corresponding to the maximum channel response as the orientation prediction for an epoch. We first ran this decoding at each time step, as seen in Figure 11. Here, we achieved significant decoding accuracy at 3 different time steps, and the decoding accuracy was generally higher than decoding accuracy for the previous models. The accuracy started to increase much earlier, at 125ms, and a sustained peak was observed from 225-350ms. We next computed the accuracy achieved by the mean channel response. Note that unlike previous decoding models, where we computed a mean accuracy by averaging the accuracies achieved over multiple time steps, here we computed the mean channel response and then computed the accuracy of the mean channel response as the percentage of trials where the orientation for the maximum channel response matched the orientation of the stimulus. In Figure 11, the mean channel response achieved an accuracy significantly above the permutation test. This indicates that the IEM is incredibly successful at decoding orientation from MEG, and decoding accuracy increases following evoked responses in MEG, revealing a significant decoding structure as a function of time.

Previous Orientation Decoding

We then examined if the models from the decoding analysis had any power predicting the previous stimulus orientation from the MEG response to the current stimulus. We first investigated this previous stimulus decoding with the SVM and logistic regression, but did not find any significant results in mean accuracy or time step accuracy. This does not indicate that residual effects from the previous stimulus were not there, just that these models were unable to detect these effects. The SVM and logistic regression model already had slim accuracy improvements over a chance accuracy, so this result is expected.

However, we found very interesting results with the IEM model (Figure 12). Though there was no significance found in the mean channel response accuracy, there was significant decoding accuracy at 175ms ($p < 0.05$), and increased channel responses from 150-200ms. We also observed a bell curve structure similar to the one found in the current stimulus decoding, but it had a lower amplitude. This indicates there was some significant residual effect left by the previous stimulus, though no significant bias structure was revealed through this analysis. Further, the peak decoding accuracy in previous stimulus decoding occurred just before the time points when current stimulus decoding accuracy became significant. This could indicate that there was a bias toward the prior stimulus shortly after the current stimulus was processed, but this bias disappeared as the current stimulus became encoded in perception.

Serial Dependence Analysis

Behavioral Analysis

Before examining serial dependence effects on decoding accuracy, we checked for a serial dependence effect in the experimental task. In Figure 13, we observed a significant serial dependence effect, with response orientations peaking at 2.036 degrees bias towards the previous stimulus orientation ($P < 0.001$). This effect size was largest when the 1-back stimulus was 20 degrees away from the current stimulus. This effect size is consistent with previously reported serial dependence in orientation perception (Fritsche et al., 2017; Manassi et al., 2017). These results indicate that subjects' perception of the current stimulus was pulled towards their perception of the previous target.

Channel Response Temporal Bias Analysis

We finally analyzed how serial dependence might directly affect IEM decoding performance by comparing channel responses to different relative previous orientations (previous orientation - current orientation). Here, we binned channel responses according to relative previous orientation with 15

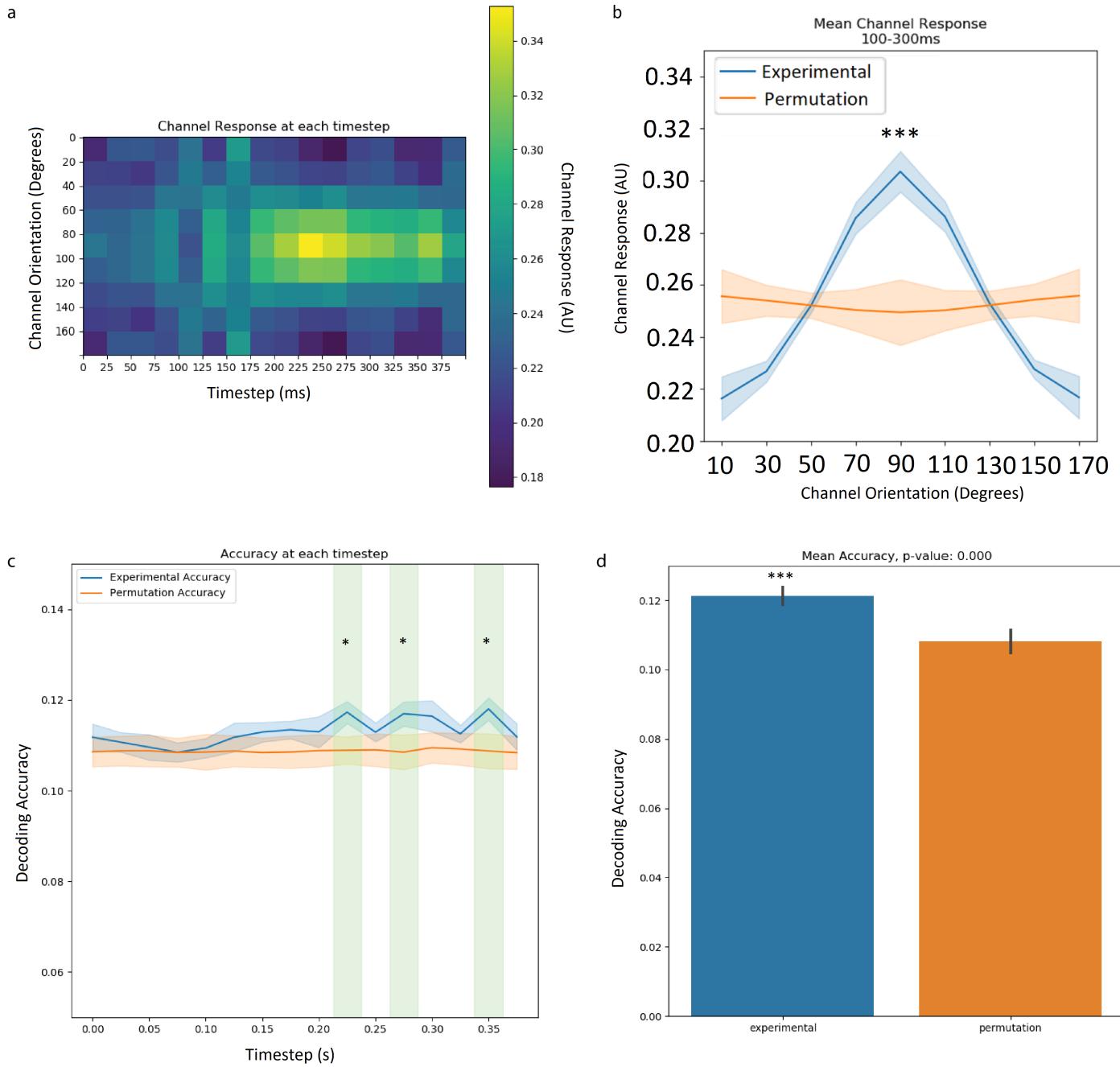


Figure 11: a) IEM channel response at 16 time steps from 0-400ms. Higher channel response values are in yellow, while lower channel response values are in blue. Channel responses begin to increase at 175ms, forming the bell curve structure found in the mean channel response. Channel responses peak at 225ms, and slowly decrease until 375ms. b) Mean channel response for the IEM compared to a permutation test. Here, there is a significant structure in the actual channel responses in comparison to the permutation test. All channels were centered at 80 degrees for visualization. c) IEM decoding accuracy at 16 time steps. Time steps shaded green represent significant decoding time steps ($p < 0.05$). Experimental accuracy increases beginning at 100ms after stimulus onset, and peaks around 225-350ms. d) IEM accuracy for the mean channel response. Accuracy is significantly above the permutation accuracy, with $p < 0.005$.

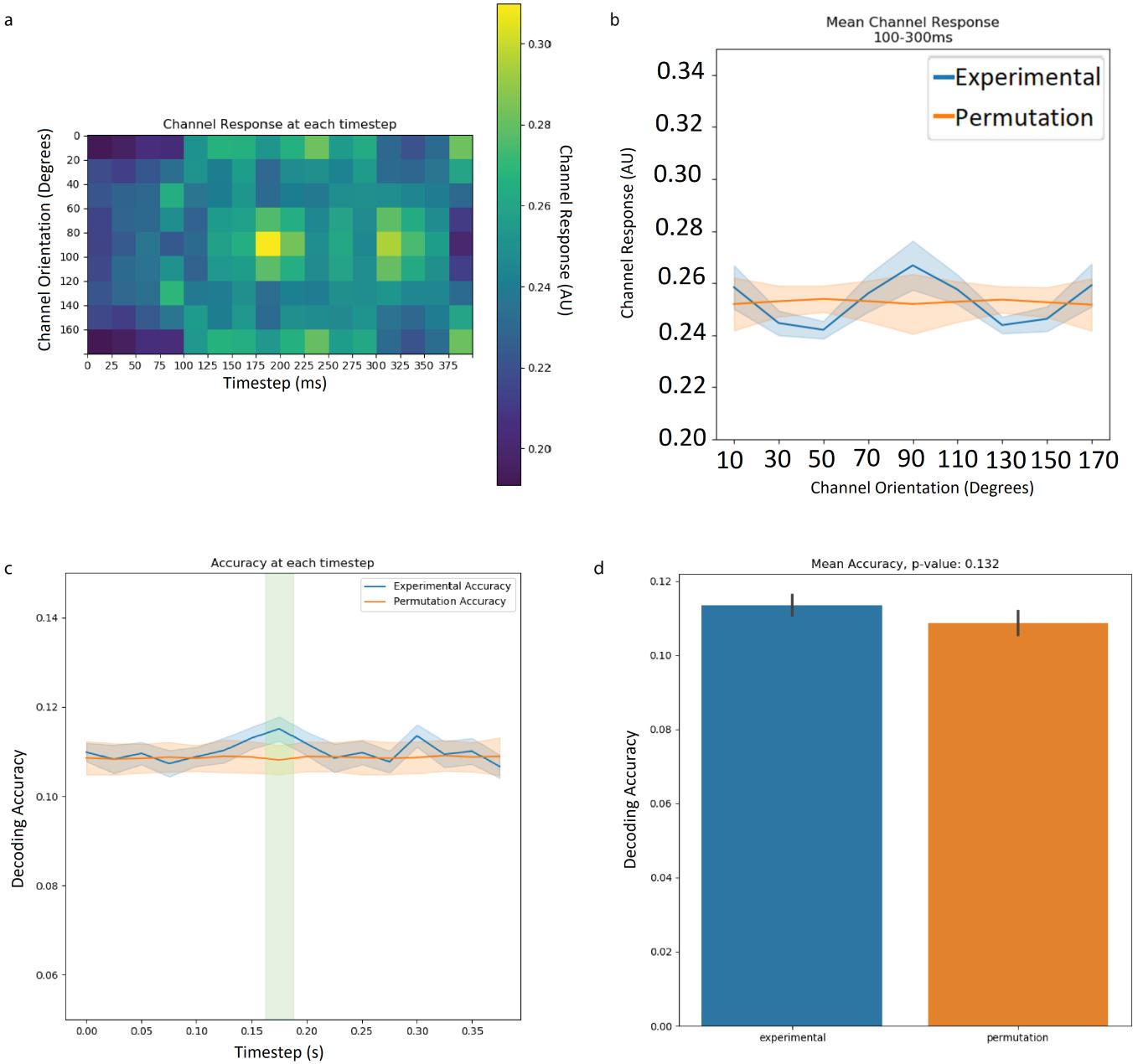


Figure 12: a) IEM channel response for previous stimulus orientation. Higher channel responses are in yellow, lower are in blue. Channel responses peak at 175ms. Peak channel responses are lower in previous decoding (0.3 vs. 0.35). b) IEM mean channel response for previous stimulus decoding compared to a permutation test. c) IEM previous decoding accuracy over 16 timesteps, averaged across 18 subjects. The accuracy at 175 ms is significant with $p < 0.05$, (marked in green). d) IEM previous trial stimulus decoding accuracy for the mean channel response, averaged across 18 subjects. Decoding accuracy is not significant in comparison to a permutation test.

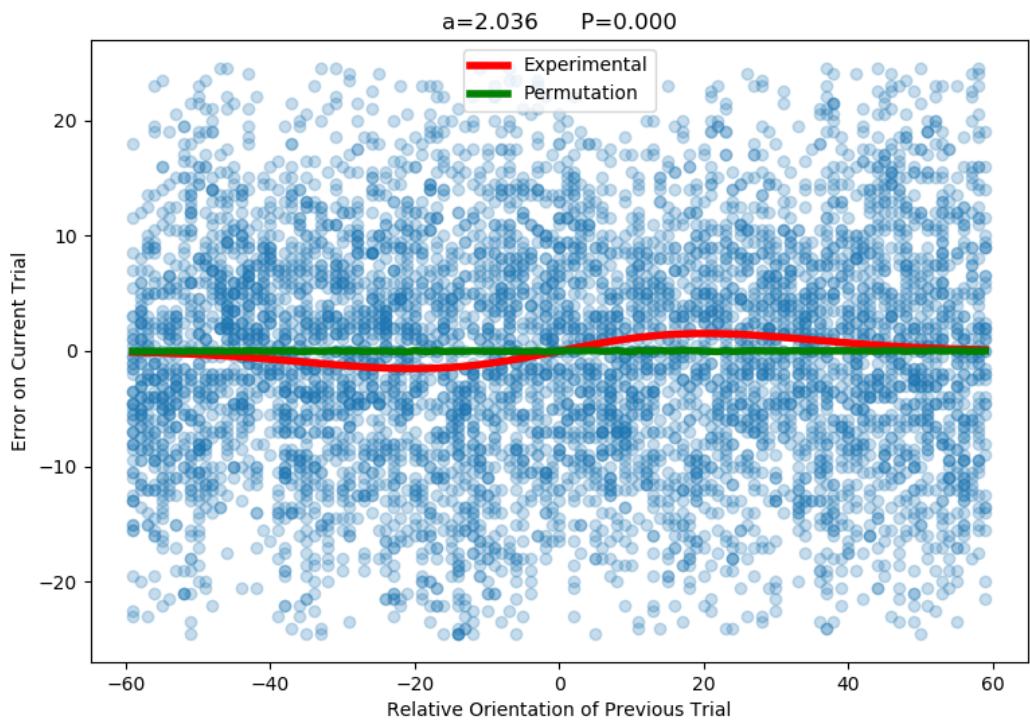


Figure 13: A comparison of relative previous orientation (previous orientation - current orientation) and response error on current trial (responded orientation - actual orientation). Blue dots represent the individual subject trial data points. The red curve indicates a derivative of Gaussian function fit to the data points. The derivative of Gaussian had a half-amplitude of 2.036 degrees, achieving a p-value < 0.001 according to a permutation test. The green curve shows the permutation fit of the derivative of Gaussian curve, which is flat. Trials with error > 25 degrees were considered outliers and were excluded from testing.

bins, centered at 0 degrees. We calculated the mean channel response for each bin, and compared each bin to randomly shuffled permutations of the bins. We repeated this for each of the 16 time steps, but did not find any bins with significant or systemic channel response bias towards or away from the previous orientation. For example, in Figure 14, the binned channel response at 175ms, where we had the highest previous decoding accuracy, had no significant channel responses. We were particularly interested in channel responses from -20 to +20 degrees relative previous orientation, where serial dependence effects would be strongest. There was actually channel response bias away from previous orientations in this range, but it was not significant according to this test. This lack of effect might suggest that bins were too large to gain meaningful information from, or that the orientation bins we used for decoding were too far apart to generate a meaningful serial dependence effect in decoding, considering a small serial dependence effect size. Further, this might suggest that we did not have enough trials for this particular analysis, as we split one tuning function into 15 bins.

We then performed this analysis by comparing the mean of a Gaussian fit to channel responses at each bin to a permutation test. There was no significance at any time step, with P values around 0.4-0.6 for all relative previous orientation bins. At 175ms after stimulus onset, where we observed significant previous orientation decoding, there was no significance (Figure 15).

Discussion

The orientation of small, peripheral Gabor patch targets can be decoded from MEG

Previous orientation decoding studies (Haynes and Rees, 2005; Kamitani and Tong, 2005; Garcia et al., 2013; Cichy et al., 2015; Pantazis et al., 2018) have used very large gratings extending from fovea to periphery, and have shown that models can discriminate orientation with high accuracy from EEG, MEG, and fMRI. We focused on using smaller Gabor targets at seven degrees eccentricity in the right visual field. This should be a harder decoding task, as there is less visual cortex area dedicated to the periphery. This harder task pushes our understanding of orientation decoding. Our orientation decoding was further complicated by the continuous range of orientations we used. Many previous studies discriminated between only two orientations, usually -45 degrees and +45 degrees (Haynes and Rees, 2005; Cichy et al., 2015; Garcia et al., 2013), though Kamitani and Tong (2005) experimented with decoding eight different orientations. Pantazis et al. (2018) used a larger stimulus set, with 6 different orientations, each 30 degrees apart. However, Pantazis et al.'s study only performed pairwise decoding, discriminating between two out of the six stimuli at a time.

The present study aimed to decode current and previous stimulus orientation using small, peripheral, and continuous orientations, which are more realistic and potentially more psychophysically useful. Similar to previous serial dependence studies, we used 180 different orientations, ranging evenly from 0 to 179 degrees. To solve the continuous orientation problem, we binned these orientations into nine different orientation classes for use in our multi-class decoding models (Haynes and Rees, 2005; Brouwer and Heeger, 2009, 2011; Garcia et al., 2013; Cichy et al., 2015). This decoding method of binned orientations with a variety of models was successful. Our finding reinforces previous MEG research (Cichy et al., 2015; Pantazis et al., 2018) suggesting that orientation information can be decoded by MEG.

Further, we reveal that smaller, peripheral Gabor patches can be decoded. Real world stimuli are much more complex than the large, oriented gratings used in previous MEG decoding studies. The smaller Gabor patch stimuli that we used are closer to modeling real-world stimuli and push the limits of MEG decoding. Future MEG studies may wish to examine the temporal dynamics of decoding natural images, for example. If natural images are represented by compositions of spatial frequency patches of varying size and location throughout the visual cortex, as suggested by Olshausen and Field (1997), then natural image decoding could perhaps be performed by correlating MEG responses to different compositions of orientation detectors, ranging in size, position, and spatial frequency. Our research expands the breadth of these oriented spatial frequency targets that we are able to decode, making a task like natural image decoding with MEG more feasible in the future.

Nevertheless, accuracy is low - though consistent with decoding literature - and we had to bin

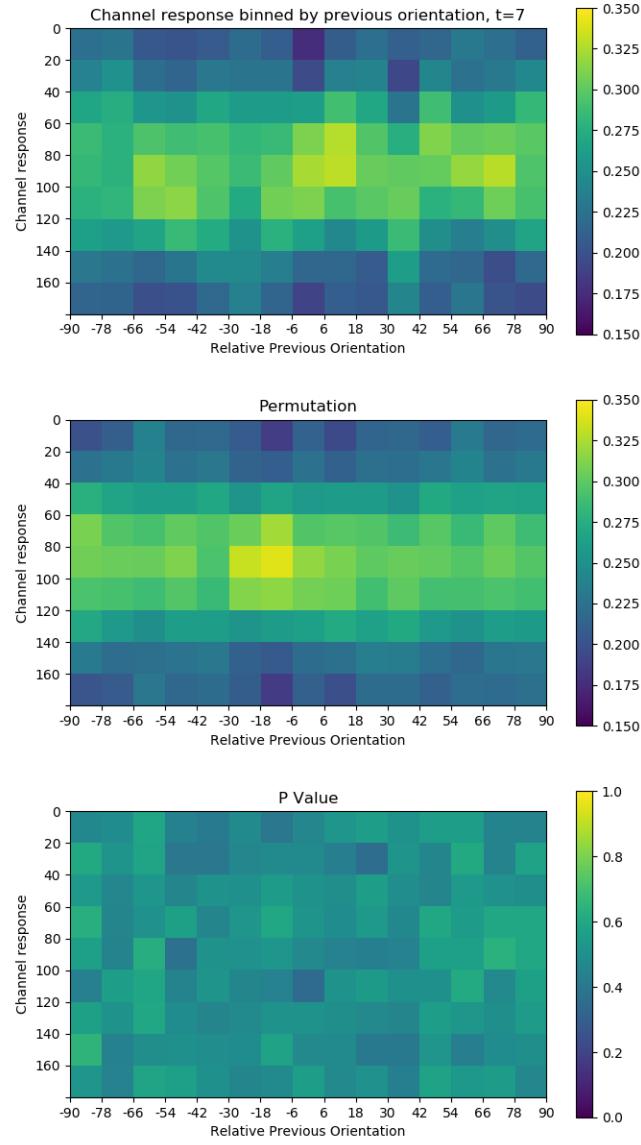


Figure 14: A comparison of IEM channel responses binned by relative previous orientation at 175ms. The top panel shows the binned channel responses. Blocks that are more yellow represent higher channel responses, while blue blocks have lower channel response. The middle panel shows a permutation test of the binned channels, in which bins were assigned according to shuffled relative previous orientation. The bottom panel shows the p values for each channel response block at each relative previous orientation. A low p value at a specific orientation for a bin would indicate that the block is significantly different than the average channel response for that orientation. Here, most p values hover around 0.4 to 0.6.

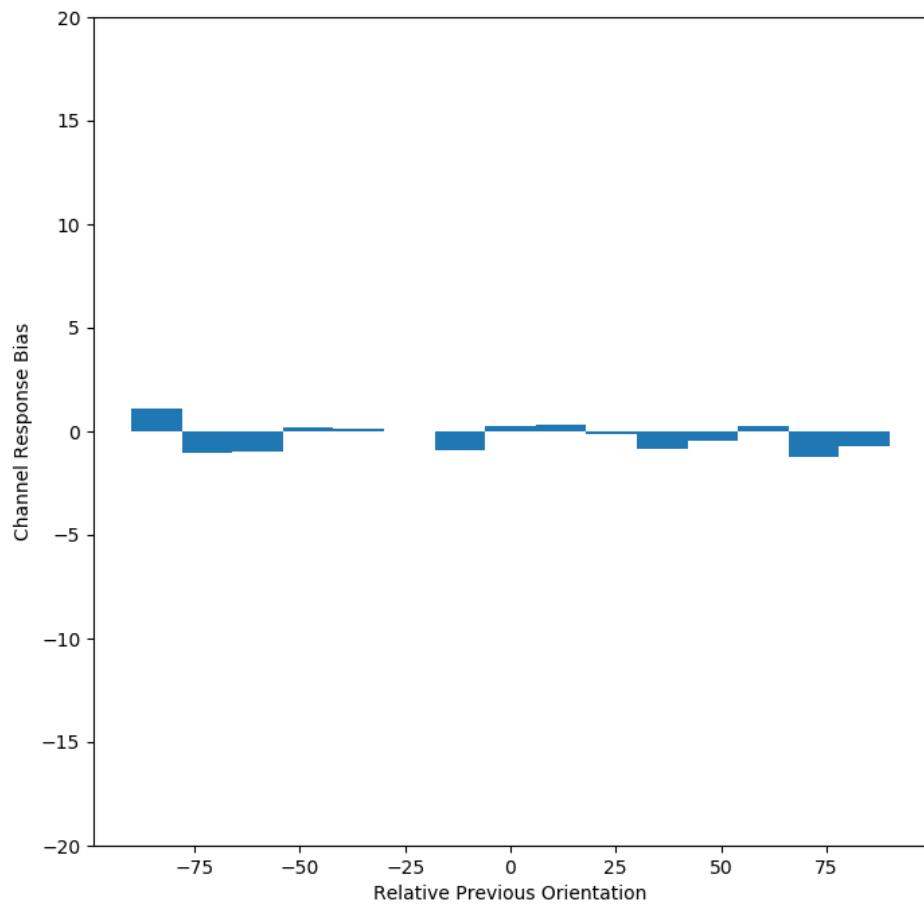


Figure 15: Mean of Gaussian functions best-fit to channel responses binned by relative previous orientation at 175ms. No bin has significant deviation from 0 degrees mean in comparison to a permutation test.

orientations to achieve significant decoding results. In order to improve accuracy and specificity with small, continuous, peripheral orientations, future studies would likely benefit from a larger training set. We had 600-800 trials for each subject, which has been shown to be sufficient for binary discrimination tasks. However, it is possible that decoding with small, continuous targets and MEG data requires more data in order to prevent overfitting in more complex models and remove variability.

Source localization decoding was not significant

We were unable to decode stimulus orientation from source-localized activities using the logistic regression model. This could have resulted from poor model hyperparameter tuning, as the hyperparameter tuning process was much more laborious due to the much longer training times. We may have had success using source localization with the IEM, but we decided to focus on sensor space decoding after poor initial results with logistic regression source space decoding. It's also possible that the source localization model introduced additional noise to the decoding, rather than reducing noise as intended. The main step that could have introduced noise was the co-registration process, which was where we aligned the MRI skull structure with the MEG electrodes. This involved some subjective visual estimation of how things should line up. We did verify that evoked source estimation potentials originated in the visual area, but this verification would have been biased by how we assumed the source localization results should look.

In this research, we mostly focused on the temporal results of the decoding, rather than the spatial results. With this in mind, it was not necessary to use source localization to augment our data set, as the sensor space data is sufficient for temporal investigations of decoding. The source localization process ended up using lots of computational resources and researcher time, especially for the initial cortical segmentation steps and the final model training steps. However, if future studies were able to reduce the bias introduced in the co-registration step, source localization could be incredibly useful in studies seeking to investigate decoding over time and space.

Traditional decoding models have low performance in binned multi-class orientation decoding

Many orientation decoding studies (Kamitani and Tong, 2005; Haynes and Rees, 2005; Cichy et al., 2015) have used machine learning models such as SVMs, linear discriminant analysis (LDA), and logistic regression for classification. With the exception of Kamitani and Tong (2005), which used LDA for multi-class decoding, many early decoding studies focused on binary classification. We tested the SVM and logistic regression models in our multi-class decoding problem, and found that mean decoding results were significant in both cases, but with low accuracy. Although low accuracy is typical in MEG decoding studies, these results were not informative beyond confirming that decoding the current orientation was feasible. We chose MEG for its great temporal resolution, but we did not find any significant structure over time in either the SVM or logistic regression decoding. We found one time point with significance for logistic regression decoding, but this does not indicate any significant structure over time. We found no time points of significance for the SVM. Further, we did not find any significance when decoding previous orientation from current MEG responses with either model.

This indicates that SVM and logistic regression are not as well suited for multi-class orientation decoding as they are for two-class orientation decoding. In our use case, where we had continuous orientations that were binned to a set of fixed orientations, orientations on the margins of these bins could have introduced additional noise into our model, resulting in poorer decoding performance. Additionally, the low accuracy of our SVM and logistic regression models could be attributed to the circular nature of the orientation targets. For example, a Gabor patch oriented at 180 degrees is more similar to a Gabor patch oriented at 20 degrees than it is to a Gabor patch at 140 degrees. This is hard for an unmodified logistic regression model to encode in a loss function. Prior circular decoding research has had more success with LDA (Kamitani and Tong, 2005) and the IEM (Brouwer and Heeger, 2009, 2011; Garcia et al., 2013; Sprague and Serences, 2013; Sprague et al., 2015). LDA

could give useful predictions about how close to the margins a certain orientation might be, and future studies may attempt to use LDA for orientation decoding. However, we focused on using the IEM.

IEM channel response has a significant bell-curve structure

We primarily approached decoding with the IEM (Brouwer and Heeger, 2009, 2011; Sprague and Serences, 2013; Garcia et al., 2013; Sprague et al., 2015), which models orientation channels with a set of cosine basis functions. We first investigated the mean channel response, and found that mean channel response had a significant bell-shaped structure, and that decoding with the mean channel response was significant. This bell-curve structure is similar to the orientation tuning functions of V1 neurons in orientation columns (De Valois et al., 1978), suggesting that our IEM has tuning similar to V1 neurons. This models orientation decoding much more naturally than logistic regression or SVMs. This improved upon the logistic regression and SVM mean accuracy, and showed that channel response drops off as a function of distance from stimulus orientation.

IEM channel response has significant structure over time

While the traditional decoding models failed to show significant structure over time, we had much more success showing a temporal structure with the IEM. The IEM results showed three significant decoding time steps from 225ms to 350ms. We had also found that the ERPs peak from 150 to 200ms, suggesting that MEG decoding accuracy increases after the ERF peaks. Further, we found a bell curve structure over time in channel response, beginning at 175ms and continuing until 350ms, lining up with our decoding results. We notice that channel response has a continuous, peaking structure here, with channel responses increasing from 175ms to a peak at 225ms, then decreasing continuously to around 325ms, before a small peak at 350ms. These results are mostly consistent with other MEG decoding papers (Cichy et al., 2015; Pantazis et al., 2018), which found decoding accuracy to peak from around 150-300ms.

However, we were unable to find any significant decoding information in the first 150ms after stimulus onset, while a previous MEG study found visual information encoded as soon as 50ms after stimulus onset (Cichy et al., 2015). Our model likely suffers from more noise than the Cichy et al.'s model, as we used smaller targets with more orientations. The peak decoding accuracy time points are generally consistent across models, but the easier discrimination task in these previous studies may have been able to significantly decode orientation sooner. These studies also revealed a decoding structure over a longer period of time, up to 1000ms after stimulus onset. Our study only decoded up to 375ms after stimulus onset. We may have been able to reveal more significant information by fitting a Gaussian function to decoded channel responses and estimating the stimulus orientation as the peak orientation of the Gaussian function, as this would have played to the strengths of the IEM, rather than the all-or-none decoding accuracy scheme. Further, we may have been able to use the model outlined in van Bergen et al. (2015) to better model the noise of our MEG data, and revealed a more significant structure at more time points. Here, we were limited by the computational resources that this analysis would have taken.

Future analysis could also investigate the spatial locality of this decoding structure. In our source localized data, we found that MEG waves began moving from the primary visual cortex towards the anterior areas around 150ms. A more detailed look at decoding with source localized data might be able to correlate this movement of activities with IEM weights to give us a detailed look at decoding in both space and time.

Previous orientation can be decoded by the IEM

Previous research has shown that traces of previous experiences or stimuli are found in fMRI frontal eye fields (Papadimitriou et al., 2016) and in EEG (Bae and Luck, 2019). Recent research also showed that it is possible to decode previous stimulus orientation from current stimulus-related fMRI from visual areas (Sheehan and Serences, 2021). Our results showed that we could significantly decode

previous stimulus orientation in MEG at one out of 16 time steps. Interestingly, our results showed that this significant decoding occurred at 175 ms, during or after ERF peaks, but before peaks in current stimulus decoding. Further, we show that previous orientation decoding seems to increase in the time steps before 175ms, and then decrease back to chance decoding by 225ms. This hints at a significant structure over time in previous orientation decoding. We may reveal more significant structure if we were to fit Gaussian functions to our channel responses and compare Gaussian amplitude and mean to the permutation responses. Further, higher frequency decoding analysis with more time steps could reveal a more robust structure over time.

Because we did not find any significant decoding in the 200ms before current stimulus onset, or in the 150 ms following current stimulus onset, this result suggests that information about the previous stimulus is stored in current stimulus ERFs. However, this information could be related to adaptation or serial dependence, or could be an afterimage of the previous stimulus that does not affect perception. Given that previous orientation decoding peaks just before current orientation decoding, it is possible that this previous orientation information is used to bias current orientation perception towards previous orientations, as serial dependence would suggest. We could also be decoding negative aftereffects of the previous stimulus, which might push current channel response decoding away from previous orientations.

The success of previous orientation decoding also gives us some potential insight into improving decoding performance. Previous orientation has a significant effect on current MEG responses; it may be possible to use knowledge of previous stimuli to improve current decoding performance. This could manifest in the use of a recurrent neural network architecture, for example.

Subjects showed a serial dependence effect in behavioral responses

Our behavioral task asked participants to orient a response bar to match the orientation of a Gabor patch they had been shown in the right visual field. We found that participant responses were significantly and systemically biased towards the orientation of the previous stimulus displayed. This bias had a maximum magnitude of about 2 degrees at 20 degrees relative previous orientation (previous - current orientation). These results are consistent with past literature in serial dependence (Fischer and Whitney, 2014; Cicchini et al., 2014; Liberman et al., 2014; Manassi et al., 2017; Kiyonaga et al., 2017), and suggest that the participants' current perception of the world is biased towards their previous perception of the world.

Serial dependence has been shown to affect multiple components of visual perception. Corbett et al. (2011) and Cicchini et al. (2014) showed that this serial dependence effect persists to numerosity, where the perception of the number of objects in a scene is dependent on the number of objects in the previous scene displayed. The same serial dependence effects are also noticeable in face perception (Liberman et al., 2014), suggesting that higher-level features are affected by serial dependence bias. Kiyonaga et al. (2017) also noted the parallels between working memory serial dependence and perceptual serial dependence. Future studies may investigate the potential effects of serial dependence on orientation decoding in greater detail. If serial dependence has an effect on orientation decoding, these effects may apply to not just orientation decoding, but face, numerosity, or color decoding. Further, the exact neural mechanisms behind serial dependence are unknown. Future orientation decoding studies with MEG could reveal the temporal dynamics of serial dependence, and future studies with fMRI could help reveal the spatial locality of serial dependence.

Results show no significant relationship between previous orientation and current channel response

A recent fMRI study by Sheehan and Serences (2021) has found a repulsion effect in channel response compared to relative previous orientation, rather than the attraction effect that would be consistent with serial dependence. We performed multiple analyses investigating any relationship between current channel response and relative previous orientation, but were ultimately unable to find a significant relationship. We binned current channel responses by relative previous orientation,

but found that our results were not significant by a pixel-wise permutation test of channel response. We further investigated results by fitting a Gaussian to channel responses at each time point and comparing the mean to a permutation distribution, but again found no significant results. Although we found a significant serial dependence effect in the behavioral results, decoding results would suggest that channel response is not biased by relative previous orientation. Nevertheless, there is still much investigation to be done. Further investigation could look at time points after 375ms, when a serial dependence response may appear. Further, source localization analysis may be used to attempt decoding by specific brain regions, similar to fMRI studies. Decoding of primary visual cortex may show less of a temporal channel response bias than higher visual cortex areas, for example.

A circular regression may be more suitable for this analysis. An approach similar to Sheehan and Serences (2021) with MEG may also reveal significant structure with higher temporal resolution in our decoded orientations compared to relative previous orientation. This analysis uses van Bergen et al.'s modification of the IEM, and performs a circular correlation of the model bias, which we did not perform. Future studies could also investigate how the individual differences in serial dependence effect size relate to performance in previous orientation decoding and orientation bias in current orientation decoding. In our data, we found that some subjects had large serial dependence effects, up to ten degrees bias, while others showed no effect or even a mild repulsion. An individual differences study could establish a much stronger connection between decoding and serial dependence than we were able to make.

Limitations of the IEM

Although we found statistically significant results for our current and previous orientation decoding with the IEM, we have to be careful of our interpretation of the results. Gardner and Liu (2019) shows that linear transformations of the IEM basis function results in linear transformations of the decoded channel responses, essentially reconstructing some arbitrary response. This would suggest that our channel responses have a bell-curve structure because the inputs were transformed to have a bell-curve structure, and not necessarily because of the inherent shape of the channel responses. Certainly, the bell-curve transformation is a much better estimation of how similar orientations are encoded in the brain than some arbitrary linear transformation. Further, the structure of our channel responses is still significant in comparison to a permutation test, suggesting that our decoding accuracy interpretation is plausible. However, we have to be cautious when performing analysis that specifically relies on the shape of our IEM channel responses, such as the investigation of channel response bias binned by relative previous orientation.

van Bergen et al. (2015) proposed a modified version of the IEM, which uses the initial transformation and weights estimation. However, it also models individual voxel (the algorithm was designed for fMRI) and total model noise. It then estimates log likelihoods for each output channel using maximum likelihood estimation. This removes the potential bias introduced by the typical IEM. Future investigations of channel response bias due to perceptual mechanisms like serial dependence would likely benefit from the use of this model, rather than the traditional IEM. If we had modeled the noise from our MEG sensors, we may have been able to find bias from the previous orientation in the current channel response.

Conclusion

We found that the IEM allowed us to significantly decode the orientation of small, continuous targets in the periphery. Our results suggest that current stimulus orientation decoding peaks around 225ms, after ERF peaks in occipital and temporal electrodes. These results were consistent with previous research that used large, fixated, fixed orientation targets. We were further able to decode the orientation of the previous stimulus target from current stimulus related MEG responses. We found that this decoding accuracy peaked just before current stimulus decoding. Finally, although we found a serial dependence bias effect in our behavioral data, we were unable to find any significant bias in current orientation decoding based on the relative orientation of the previous stimulus. Future

studies may expand upon the methods we used here to uncover any relationship between previous orientation and current orientation decoding.

References

- Bae, G. and Luck, S. (2019). Reactivation of previous experiences in a working memory task. *Psychological Science*, 30(4):587–595. Funding Information: We thank Aaron Simmons for assistance with data collection. This research was made possible by Grant No. R01MH076226 from the National Institute of Mental Health.
- Brouwer, G. J. and Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *Journal of Neuroscience*, 29(44):13992–14003.
- Brouwer, G. J. and Heeger, D. J. (2011). Cross-orientation suppression in human visual cortex. *Journal of Neurophysiology*, 106(5):2108–2119. PMID: 21775720.
- Cicchini, G. M., Anobile, G., and Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences*, 111(21):7867–7872.
- Cichy, R. M., Ramirez, F. M., and Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *NeuroImage*, 121:193–204.
- Corbett, J. E., Fischer, J., and Whitney, D. (2011). Facilitating stable representations: Serial dependence in vision. *PLOS ONE*, 6(1):1–7.
- Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., and Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fmri and meg for high-resolution imaging of cortical activity. *Neuron*, 26(1):55–67.
- De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1978). Cortical cells: Bar and edge detectors, or spatial frequency filters? In Cool, S. J. and Smith, E. L., editors, *Frontiers in Visual Science*, pages 544–556, Berlin, Heidelberg. Springer Berlin Heidelberg.
- De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):545–559.
- Engel, S. A., Glover, G. H., and Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, 7(2):181–192.
- Fischer, J. and Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, 17(5):738—743.
- Fritzsche, M., Mostert, P., and de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, 27(4):590–595.
- Fuchs, M., Wagner, M., Köhler, T., and Wischmann, H.-A. (1999). Linear and nonlinear current density reconstructions. *Journal of Clinical Neurophysiology*, 16(3):267–295.
- Garcia, J., Srinivasan, R., and Serences, J. (2013). Near-real-time feature-selective modulations in human cortex. *Current Biology*, 23(6):515–522.
- Gardner, J. L. and Liu, T. (2019). Inverted encoding models reconstruct an arbitrary model response, not the stimulus. *eNeuro*, 6(2).

- Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., and Hämäläinen, M. (2013). Meg and eeg data analysis with mne-python. *Frontiers in Neuroscience*, 7:267.
- Hamalainen, M. S. and Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions on Biomedical Engineering*, 36(2):165–171.
- Haynes, J.-D. and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience*, 8(5):686–691.
- He, S. and MacLeod, D. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature*, 411:473–6.
- Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591.
- Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *The Journal of Physiology*, 160(1):106–154.
- Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1):215–243.
- Kamitani, Y. and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5):679–685.
- King, J.-R., Gwilliams, L., Holdgraf, C., Sassenhagen, J., Barachant, A., Engemann, D., Larson, E., and Gramfort, A. (2018). Encoding and decoding neuronal dynamics: Methodological framework to uncover the algorithms of cognition. working paper or preprint.
- Kiyonaga, A., Scimeca, J. M., Bliss, D. P., and Whitney, D. (2017). Serial dependence across perception, attention, and memory. *Trends in Cognitive Sciences*, 21(7):493–497.
- Liberman, A., Fischer, J., and Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology*, 24(21):2569–2574.
- Lin, F.-H., Belliveau, J. W., Dale, A. M., and Hämäläinen, M. S. (2005). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, 27(1):1–13.
- Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., and Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in meg source localization by depth-weighted minimum-norm estimates. *NeuroImage*, 31(1):160–171.
- Manassi, M., Liberman, A., Chaney, W., and Whitney, D. (2017). The perceived stability of scenes: Serial dependence in ensemble representations. *Scientific Reports*, 7:1971.
- Mannion, D., McDonald, J., and Clifford, C. (2009). Discrimination of the local orientation structure of spiral glass patterns early in human visual cortex. *NeuroImage*, 46(2):511–515.
- Mosher, J. C., Leahy, R. M., and Lewis, P. S. (1999). Eeg and meg: forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, 46(3):245–259.
- Newville, M., Stensitzki, T., Allen, D. B., and Ingargiola, A. (2014). LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python.
- Olshausen, B. A. and Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research*, 37(23):3311–3325.

- Pantazis, D., Fang, M., Qin, S., Mohsenzadeh, Y., Li, Q., and Cichy, R. M. (2018). Decoding the orientation of contrast edges from meg evoked and induced responses. *NeuroImage*, 180:267–279.
- Papadimitriou, C., White, Robert L., I., and Snyder, L. H. (2016). Ghosts in the Machine II: Neural Correlates of Memory Interference from the Previous Trial. *Cerebral Cortex*, 27(4):2513–2527.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- Segonne, F. (2004). A hybrid approach to the skull stripping problem in mri. *NeuroImage*.
- Senior, C., Russell, T., and Gazzaniga, M. S. (2006). *Methods in Mind*. Mit Press.
- Sheehan, T. C. and Serences, J. T. (2021). Sensory readout accounts for adaptation. *bioRxiv*.
- Sprague, T. C., Saproo, S., and Serences, J. T. (2015). Visual attention mitigates information loss in small- and large-scale neural codes. *Trends in Cognitive Sciences*, 19(4):215–226.
- Sprague, T. C. and Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, 16(12):1879–1887.
- van Bergen, R., Ma, W., Pratte, M., and Jehee, J. (2015). Sensory uncertainty decoded from visual cortex predicts behavior. *Nature neuroscience*, 18.
- Webster, M. A. and Macleod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1571):1702–1725.
- Webster, M. A. and Mollon, J. (1997). Adaptation and the color statistics of natural images. *Vision Research*, 37(23):3283–3298.
- Yacoub, E., Harel, N., and Uğurbil, K. (2008). High-field fmri unveils orientation columns in humans. *Proceedings of the National Academy of Sciences*, 105(30):10607–10612.