Temporal uncertainty affects the visual processing of predicted stimulus properties

Sanjeev Nara [1] Mikel Lizarazu [1] Craig G Richter [1] Diana C Dima [3] Mathieu Bourguignon [1,2] Nicola Molinaro [1]

1. *Basque Center for Cognition, Brain and Language , San Sebastian, Spain*
2. *Laboratoire de Cartographie fonctionelle du Cerveau, UNI – ULB Neuroscience Institute, Université libre de Bruxelles (ULB), Brussels, Belgium*
3. *John Hopkins University, Baltimore , USA*

KEYWORDS: predictive processing, temporal predictability, visual perception, magnetoencephalography, time-resolved decoding

Corresponding author:
Sanjeev Nara
Basque Center on Cognition, Brain and Language
Mikeletegi Pasealekua, 69
20009 Donostia, Gipuzkoa, Spain
email: s.nara@bcbl.eu
phone: +34 943309300

Abstract

Predictive processing has been proposed as a fundamental cognitive mechanism to account for how the brain interacts with the external environment via its sensory modalities. The brain processes external information about the content (i.e. "*what*") and timing (i.e., "*when*") of environmental stimuli to update an internal generative model of the world around it. However, the interaction between "*what*" and "*when*" has received very little attention when focusing on vision. In this magnetoencephalography (MEG) study we investigate how processing of feature specific information (i.e. "what") is affected by temporal predictability (i.e. "when"). In line with previous findings, we observed a suppression of evoked neural responses in the visual cortex for predictable stimuli. Interestingly, we observed that temporal uncertainty enhances this expectation suppression effect. This suggests that in temporally uncertain scenarios the neurocognitive system relies more on internal representations and invests less resources integrating bottom-up information. Indeed, temporal decoding analysis indicated that visual features are encoded for a shorter time period by the neural system when temporal uncertainty is higher. This supports the fact that visual information is maintained active for less time for a stimulus whose time onset is unpredictable compared to when it is predictable. These findings highlight the higher reliance of the visual system on the internal expectations when the temporal dynamics of the external environment are less predictable.

Introduction

Our interaction with the external environment is largely shaped by our internal expectations (Clark, 2013; Mechelli et al., 2004; Mumford, 1992). Predictive processing theories propose that an internal model of the surrounding environment is used to generate inferences about the external causes of the environmental energy impacting our senses (Spratling, 2017). Supporting this view, predictable visual stimuli trigger reduced amplitude of visual responses and are processed with shorter latencies in the visual cortex when compared to unpredictable stimuli (Hogendoorn and Burkitt, 2018). It is worth to underscore that the perception is spatiotemporal in nature. In primary sensory cortices, sensations are decomposed into their fundamental features (such as edges, orientation, colour, shape), as shown by studies that decoded such representational properties from neural signals (Carlson et al., 2019; Pantazis et al., 2018). These features are the building blocks that determine the perceptual content of the stimulation, i.e. *what* it is. Importantly, in real life situations environmental stimuli are temporally dynamic. Suppose you see a car coming toward you on the road. Even after determining all the *what* information features available (e.g., it is a pink Cadillac), it is critical to estimate *when* you and the car will intersect to avoid a collision. While certain stimuli are temporally regular (and hence predictable), temporal uncertainty in the environment is very high.

Studies on visual perception have paid relatively low attention to temporal predictability (*when*) compared to the study of the stimulus content predictability (*what*) (Demarchi et al., 2019; Kok et al., 2017) .The role of timing for predictive

processing in vision has been mainly studied for the perception of objects in motion. These studies showed that the visual system requires a certain amount of time to process incoming sensory information (Blom et al., 2020; Maunsell and Gibson, 1992). When considering the perception of moving objects, this amount of time could cause potential problems for the neurocognitive system as it rapidly needs to extrapolate the object trajectory to then rapidly plan action. Predictive processing models propose a compensatory mechanism to solve this problem, based on the generation of predictions about incoming stimuli before perceptual processing. Research on the perception of moving objects thus underscores that *what* and *when* stimulus properties are strongly interwoven and shape human perception. Here we aim at evaluating how predictive processing of the "*what* properties" is affected by manipulation of the "*when* properties". Specifically, we test if temporal uncertainty enhances predictive processing of stimulus content. Indeed, predictive processing is central to support perception of uncertain stimuli.

In the present MEG study, participants were presented with a series of consecutive Gabor patches (referred henceforth to as entrainers) that were introduced prior to a target Gabor for which participants performed a spatial frequency judgment task. We here focused on visual responses to the entrainers whose sequence of orientations were either predictable or not, and whose time onset could be either fixed or jittered with respect to the other entrainers. Firstly, we expect the neural responses to the entrainers to gradually decrease in amplitude along the sequence if the orientation is predictable, an effect defined as "expectation suppression" (Grill-

Spector et al., 2006). Secondly, if temporal uncertainty is handled by the cognitive system by enhancing reliance on internal expectations, expectation suppression should be higher for onset-jittered compared to onset-stable entrainers. Thirdly, we expect target orientation to be increasingly decodable (with multivariate pattern analysis, MVPA) (Pantazis et al., 2018; King et al., 2016; Cichy et al., 2014) along the entrainers' sequence for orientation predictable conditions. Finally, higher reliance on internal expectations should trigger less detailed stimulus-related predictions and, consequently, less efficient decoding of predicted target orientation for time-jittered (compared to time-stable) entrainers.

Methods

Participants

From the initial set of twenty participants, we considered in this experiment sixteen participants (7 females; age range: 19–31; M = 24.8; SD = 3.6). Two participants were excluded from the study as they did not complete the whole experiment and two others were excluded from the study due excessive motion artifacts in the data. The ethical committee and the scientific committee of the Basque Center on Cognition, Brain and Language (BCBL) approved the experiment (following the principles of the Declaration of Helsinki). Participants gave written informed consent and were financially compensated. The participants were recruited from the BCBL Participa website (https://www.bcbl.eu/participa/). Participants were free from any neurological or psychological disorders, and had normal or corrected to normal vision.

Experimental procedure

A series of Gabor patches were presented with variable properties of orientation and spatial frequency – measured in cycles per degree of visual angle (CPD). Stimuli were back-projected on a screen placed 60 cm from the participants' nasion. The Gabors were presented in the centre of the screen with a gray background. Each trial began with a fixation cross (black color) followed by four sequential Gabor patches (entrainers) presented for 200 ms. After a longer interstimulus interval, a fifth Gabor

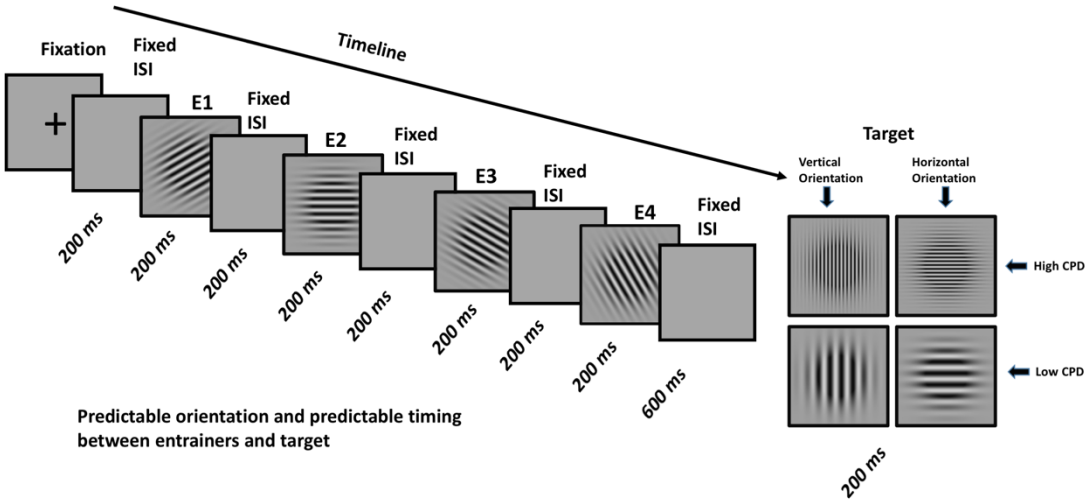
(target) was presented for 200 ms and participants were required to respond if the target was of a higher or lower spatial frequency than the entrainers.

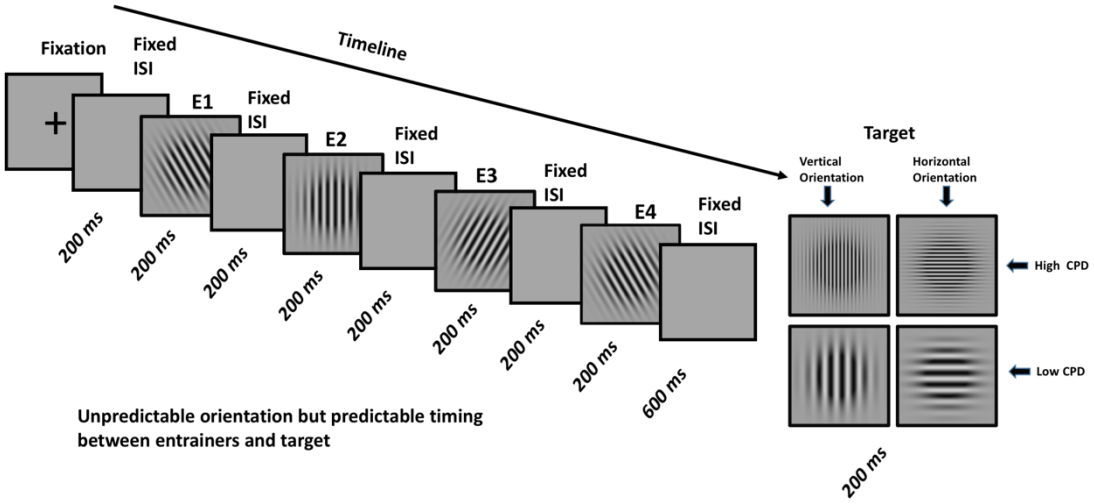
The entrainers had an intermediate spatial frequency (40 CPD), while the target could have either a higher (60 CPD) or a lower (20 CPD) spatial frequency with respect to the entrainers. Four properties of these sequences were experimentally manipulated (Figure 1): a) the orientation of the target could be either horizontal or vertical; b) the target spatial frequency could be higher or lower compared to the spatial frequency of the entrainers; c) the orientation of the target could be either predictable based on the previous entrainers (i.e., scaled in fixed angle steps of 15 or 30 degree) or unpredictable; d) the timing between the entrainers and the target could be predictable (i.e., fixed interstimulus interval of 200 ms for first three entrainers and 600 ms between entrainer 4 and target) or unpredictable (varying interstimulus intervals ranging between 70-330 ms for the first three entrainers and 450-770 ms between entrainer 4 and the target).

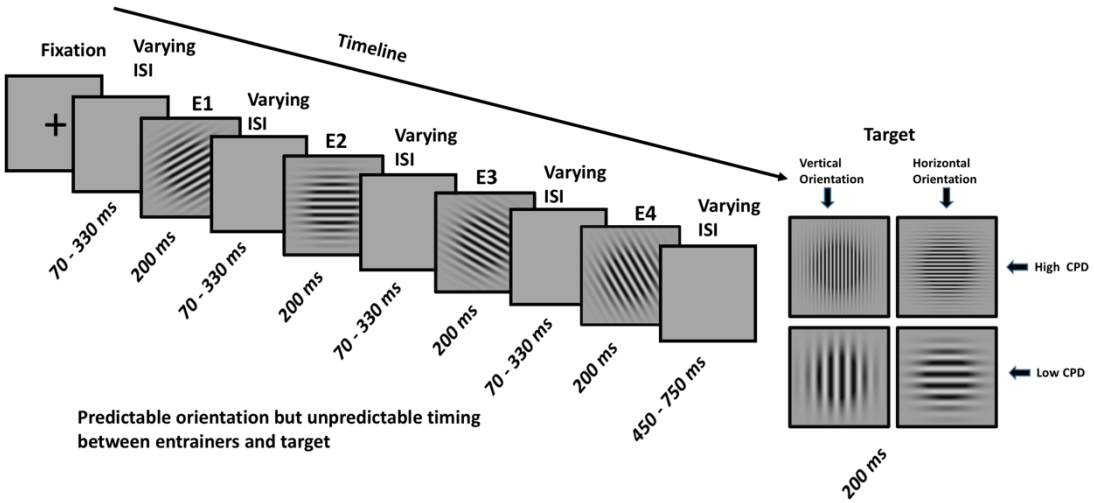
The Participants' task was to determine whether the target had a higher or a lower spatial frequency with respect to the entrainers. Participants responded by pressing a button with their left or right hand, where the hand pairing with the response was counterbalanced across participants.

Depending on the timing and the orientation of the entrainers and the target, trials were divided into four conditions (Figure1): *what* + *when*, *when*, *what* and *random*. In the *what* + *when* condition, the entrainers had a predictable timing and the orientation of the target Gabor was also predictable based on the previous entrainers.

In the *when* condition the timing between the entrainers and target was predictable but the target orientation was unpredictable. In the *what* condition the target orientation was predictable but the timing between the entrainers and target was unpredictable. In the *random* condition both the target orientation and the timing between the entrainers and target were unpredictable. A total of 160 trials were presented for each condition (80 horizontal targets and 80 vertical targets & randomly assigned to 80 high and 80 low spatial frequency targets), leading to a total of 640 trials per participant. 80 localizer trials for horizontal and vertical targets were also acquired. For these stimuli the participants were asked to simply fixate at the centre of the screen. The participants were given a short optional break (the participants press the button to continue) every 12 trials, and a mandatory long time break every 60 trials (the MEG operator presses a button from the operating console).







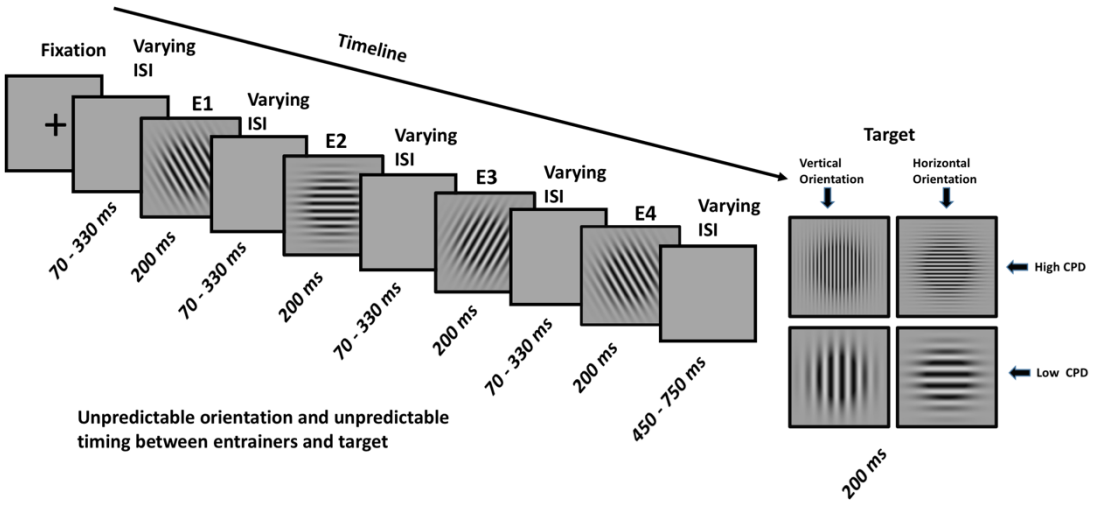


Figure 1: Stimuli design. A) Predictable orientation and predictable timing (*what* + *when* information). B) Unpredictable orientation but predictable timing (*when* information). C) Predictable orientation but unpredictable timing (*what* information). D) Unpredictable orientation and unpredictable timing (*random*).

**** Abbreviations: E1 – Entrainer 1, E2 – Entrainer 2, E3 – Entrainer 3, E4 – Entrainer 4, ISI – Inter Stimulus Interval**

Data acquisition and preprocessing

MEG data were acquired in a magnetically shielded room using the whole-scalp MEG system (Elekta-Neuromag, Helsinki, Finland) installed at the BCBL (http://www.bcbl.eu/bcbl-facilitiesresources/meg/). The system is equipped with 102 sensor triplets (each comprising a magnetometer and two orthogonal planar gradiometers) uniformly distributed around the head of the participant. Head position inside the helmet was continuously monitored using four Head Position Indicator (HPI) coils. The location of each coil relative to the anatomical fiducials

(nasion, left and right preauricular points) was defined with a 3D digitizer (Fastrak Polhemus, Colchester, VA, USA). This procedure is critical for head movement compensation during the data recording session. Digitalization of the fiducials plus ~300 additional points evenly distributed over the scalp of the participant were used during subsequent data analysis to spatially align the MEG sensor coordinates with T1 magnetic resonance brain images acquired on a 3T MRI scanner (Siemens Medical System, Erlangen, Germany). MEG recordings were acquired continuously with a bandpass filter of 0.01-330 Hz and a sampling rate of 1 kHz. Eye movements were monitored with two pairs of electrodes in a bipolar montage placed on the external chanti of each eye (horizontal electrooculogram (EOG)) and above and below right eye (vertical EOG). Similarly, electrocardiogram (ECG) was recorded using two electrodes, placed on the right side of the participant's abdomen and below the left clavicle.

Continuous MEG data were pre-processed off-line using the temporal Signal-Space-Separation (tSSS) method (Taulu & Simola, 2006) which suppresses external electromagnetic interference. MEG data were also corrected for head movements, and bad channel time courses were reconstructed in the framework of tSSS. Subsequent analyses were performed using Matlab R2014b (Mathworks, Natick, MA, USA).

Behavioural data

The accuracy of participant responses were calculated by comparing the response from each participant (higher / lower spatial frequency) and the true spatial frequency of the presented Gabor. The response time (RT) was calculated for all the four conditions (*what + when, when, what and random*). Trials with response times longer than 1500 ms were considered to be outliers and were removed from the analysis. The mean RT and standard deviation was computed for each experimental condition.

Sensor level ERFs

MEG trials were corrected for jump and muscle artifacts with an standard automated scripts published on the fieldtrip software website . Heartbeat and EOG artifacts were identified using Independent Component Analysis (ICA) and linearly subtracted from the MEG recordings. The ICA decomposition (30 components extracted per participant) was performed using the FastICA algorithm implemented in the Fieldtrip toolbox (Oostenveld et al., 2011). ICA components having maximum correlation with the EOG and ECG were automatically removed using a coherence threshold of 40 % with the MEG signals . On average two components were removed per participant. The artifact-free data were bandpass filtered between 0.5 and 45 Hz. Trials were extracted time-locked to each of the entrainers (entrainer 1, 2, 3 and 4) and the target, from xx to xx ms. The trial segments were grouped together for each entrainer and target, and then averaged to compute the ERFs. For each planar

gradiometer pair, ERFs were quantified at every time point as the Euclidean norm of the two gradiometer signals. Baseline correction was also applied to the evoked data based on the 400 ms of data prior to the onset of the fixation cross presented at the beginning of each trial.

Sensor-level statistics first relied on an ANOVA employed to explore the influence of our experimental factors on the ERFs. To this aim we extracted ERF amplitudes in a set of five occipital sensors in a time window classically associated with the initial visual evoked response (85-135 ms post-stimulus). A three-way repeated measures ANOVA was computed in JASP (Team, 2020) with these amplitude values as dependent variables and the following factors: entrainer (four levels, corresponding to each entrainer), *what* (two levels: predictable/unpredictable orientation of the target), and *when* (two levels: predictable/unpredictable timing of the entrainers and target). Significant main effects and interactions were further investigated with pairwise t-tests.

To evaluate a possible differential distribution of the experimental effects across sensors that could not be accounted for by the ANOVA, pairwise comparisons between conditions were performed using a cluster-based permutation test (Maris & Oostenveld, 2007). Four different comparisons were carried out. In the first comparison, we contrasted the ERFs for the *when* and the *what* + *when* conditions. This comparison evaluated the effect of orientation predictability when the timing of the entrainers and the target were predictable. In the second comparison, we compared the ERFs for the *random* and the *what* conditions. This comparison

evaluated the effect of orientation predictability with temporal uncertainty between entrainers and the target. These two comparisons mainly focused on the orientation predictability, i.e., on the *what* manipulation and highlight the main effect of expectation suppression. We then compared the ERFs for the *what* + *when* and the *what* conditions. Here we directly contrasted the two orientation predictable conditions to evaluate the effect of temporal predictability on stimulus predictability. The last comparison contrasted the ERFs of the *when* and the *random* conditions. This comparison was performed to analyse the effect of temporal predictability in absence of orientation predictability. In all cases, differences between conditions were analysed using cluster-based permutation testing (Maris & Oostenveld, 2007). A randomization distribution of cluster statistics was constructed for each subject over time and sensors and used to evaluate whether conditions statistically differ over participants. In particular, t-values were computed for each sensor (combined gradiometers) and time point during the 0-270 ms time window, and were clustered based on t-values that exceeded a t-value corresponding to the 99.5th percentile of Students t-distribution, i.e. a two-tailed t-test at an alpha of 0.01, and were both spatially and temporally adjacent. Cluster members were required to have at least two neighboring channels that also exceeded the threshold to be considered a cluster. The sum of the t-statistics in a sensor cluster was then used as the cluster-level statistic, which was then tested by permuting the condition labels 1000 times.

Source level ERF

MEG-MRI co-registration was performed using MRIlab (Elekta Neuromag Oy, version 1.7.25). Individual T1-weighted MRI images were segmented into scalp, skull, and brain components using the segmentation algorithms implemented in Freesurfer (Martinos Center of Biomedical Imaging, MQ) (Dale et al., 1999). The source space was defined as a regular 3D grid with a 5 mm resolution and the lead fields were computed using a single-sphere model for 3 orthogonal source orientations. The lead fields at each grid point was reduced to its two first principal components. Whole brain source activity was estimated using a linearly constrained minimum variance (LCMV) beamformer approach (Veen et al., 1997). Both planar gradiometers and magnetometers were used for inverse modelling. The covariance matrix used to derive LCMV beamformer weights was estimated from the pre- and post-stimulus data in the pre-stimulus (from 400 ms prior to fixation cross onset to 400 ms after the presentation of the target).

For each condition and entrainer, the LCMV beamformer was applied to the evoked data (with baseline correction) in the time period 85–125 ms post-stimulus and in the pre-stimulus interval. This post-stimulus interval was chosen because it is the time period containing the ERF peak amplitude across participants at the sensor level. Brain maps containing source activity were transformed from the individual MRIs to the standard Montreal Neurological Institute (MNI). For that, we applied a non-linear transformation using the spatial-normalization algorithm implemented in

Statistical Parametric Mapping (SPM8) (Friston et al., 1994). Transformed maps were further averaged across participants. Freesurfer's *tksurfer* tool was used to visualize the brain maps in MNI space. For each condition and entrainer, we obtained the source value and the MNI coordinates of local maxima (sets of contiguous voxels displaying higher source activation than all other neighboring voxels (Bourguignon et al., 2018).

Source activity was compared between conditions (e.g., *when* vs. *what* + *when, random* vs. *what, what vs. what* + *when* and *when vs. random*) by extracting a peak value within a 5mm sphere around the local maxima in the source space. This was performed for every participant and the group comparison were made using t-tests.

MVPA

Time-resolved within-subjects multivariate pattern analysis was performed to decode the features (i.e., the orientation and spatial frequency) of all the Gabors (i.e., E1, E2, E3, E4 and T) from the MEG data. This within-subject classification has an advantage over other methods: the classification algorithm may leverage individual subject specific characteristics in neural patterns since the classifiers do not need to generalize across different subjects. For E1, E2 and E3, data were segmented from 50 ms prior to 250 ms after the onset of the entrainers. The time interval between E4 and the target was longer than the time interval between the rest of the entrainers. For this reason, for E4, the data was segmented from 50 ms prior to 600 ms after the onset of the entrainer. For the target, the data was segmented from -400 ms to 550

ms. The data were classified separately for both orientation and spatial frequency of the Gabor using a linear support vector machine (SVM) classifier with L2 regularization and a box constraint of 1. The classifiers were implemented in Matlab using the LibLinear package (Fan et al., 2008) and the Statistics and Machine Learning Toolbox (Mathworks, Inc.). We performed a binary classification of the orientation of each Gabor depending on the orientation and the spatial frequency of the subsequent target. In other words, the class labels (i.e., horizontal vs. vertical, higher vs. lower spatial frequency) were derived from the target orientation: if the target orientation was horizontal then all the preceding Gabor orientations in the corresponding condition were labelled as horizontal and vice-versa. The same for higher vs. lower spatial frequency of the target. The data were down sampled to 200 Hz prior to the classification. Pseudo-trials were generated to improve the SNR by averaging trials over bins of 10, without overlap (Dima and Singh, 2018) . This pseudo-trial generation was repeated 100 times based on random ordering of the data to generate trials with a higher signal to noise ratio. The data were then randomly partitioned using 5-fold cross-validation. The classifier was trained on 4 folds and tested on the remaining fold, with this process repeated until each fold had been left out of the training once. The procedure of generating pseudo-trials, dividing the data into 5 folds, and training and testing classifiers at every time point was repeated 25 times; classification accuracies were then averaged over all these instances to yield more stable estimates. To improve classification, we also performed multivariate noise normalization (Guggenmos et al., 2018). The time-resolved error covariance

between sensors was calculated based on the covariance matrix of the training set and used to normalize both the training and test sets in order to down-weight MEG channels with higher noise levels. Cluster corrected sign permutation tests (one-tailed) (Dima et al., 2018) were applied to the accuracy values obtained from the classifier with cluster-defining threshold $p < 0.05$, corrected significance level i.e., cluster-alpha $p < 0.01$.

Results

Behavioural results

Table 1: Accuracy and reaction time (RTs) of all the four conditions.

	what + when	*when*	*what*	*random*
Accuracy (Mean ± SD %)	96.56 ± 3.49	96.00 ± 3.80	95.89 ± 4.71	96.56 ± 3.33
Reaction Time (Mean ± SD ms)	727.5 ± 204	741.4 ± 204	737.4 ± 203	752 ± 207 ms

Table 1 presents the accuracy and reaction time (RTs) of all the four conditions.We found no significant differences in behavioural accuracy. For the RTs, the *what + when* condition showed faster responses while *random* condition responses were numerically slower. The data from table 1 indicate that temporal predictability elicits faster responses (*what + when* > *what | when* > *random*). We fit a Linear mixed model (*lmer*: R function) considering participants and observations as *random* effects, with the fixed effects of *what* (orientation: predictable or not), *when* (timing:

predictable or not) and their interaction. We observed an effect of *when* (t = -2.794, $p < 0.05$).

Sensor-level MEG results

We first analysed the amplitude of the initial visual evoked response for 5 occipital sensors in an approach considering our whole experimental design. In this ANOVA (Table 1) we observed a significant main effect of entrainer ($p < 0.001$) on the peak amplitudes of the ERFs. Interestingly, this factor interacted with the factor *what* ($p < 0.001$). We also observed a main effect of *what*($p < 0.001$). Importantly, the interaction between the three factors, i.e., entrainer, *what* and *when* was significant ($p < 0.01$).

Table 2: Repeated measure ANOVA considering the factors Entrainer (four levels, one for each entrainer), *what* (two levels: orientation: predictable or not) and *when* (two levels: timing: predictable or not).

	Sum of Squares	df	Mean Square	F	p
Entrainer	2.607e -22	3	8.691e -23	42.299	< .001
what	3.430e -23	1	3.430e -23	65.203	< .001
when	2.969e -24	1	2.969e -24	2.376	0.144
Entrainer ∗ *what*	2.572e -23	3	8.572e -24	18.503	< .001
Entrainer ∗ *when*	8.851e -25	3	2.950e -25	0.860	0.469
what ∗ *when*	3.032e -25	1	3.032e -25	0.833	0.376
Entrainer ∗ *what* ∗ *when*	2.277e -24	3	7.589e -25	5.073	0.004

Note. Type III Sum of Squares

Figure 2 shows the sensor-level ERFs time-locked to the onset of each entrainer (E1, E2, E3 and E4) and target (T) for the *when* and the *what* + *when* conditions. The

amplitude of the ERFs was significantly higher ($p < 0.01$, cluster based permutation test) for the *when* compared to the *what + when* condition for E2, E3 and E4, but not for E1 and the target. The amplitude enhancement for the *when* compared to the *what + when* condition emerged in the 95–105 ms, 96–110 ms and 97–121 ms time interval for the E2, E3 and E4 respectively. These clusters were located in occipital sensors for all the entrainers.

Figure 3 shows the sensor-level ERFs for the *random* versus the *what* conditions. The amplitude of the ERFs was significantly higher ($p < 0.01$) for the *random* compared to the *what* condition for E2, E3 and E4, but not for E1 or the target. The amplitude enhancement for the *random* compared to the *what* condition emerged within the 95–119 ms, 94–123 ms and 96–127 ms time interval for the E2, E3 and E4 respectively. These clusters were also clearly distinguishable in occipital sensors for all the entrainers.

Since both the comparisons are significant from entrainer 2 onward, we compared the two orientation predictable conditions with (*what + when*) and without (*what*) temporal predictability. Figure 4 shows that the initial early evoked activity (0–75 ms) at E1 is similar for both conditions. As we move across entrainers such early differences increase and reach statistical significance ($p <0.001$) but these effects vanish at the target. It thus seems that differential pre-stimulus activity distinguishes the two orientation predictable conditions depending on temporal predictability. Here, it is worth noting that the baseline time period for all the entrainers (E1 – E4) and T was the same (400 ms before the cross fixation at the beginning of the trial).

Since our focus in this analysis is on the early evoked response to the visual stimulus that showed robust expectation suppression effects, the time window for statistical comparison was selected from 75 to 135 ms, corresponding to the initial ERF peak reflecting early visual processing. The amplitude enhancement for the *what* compared to the *what + when* condition was identified in a cluster spanning the 106–124 ms time interval at E4. This cluster was located in occipital sensors. Figure 5 shows the sensor level comparison of the ERFs for the *when* and the *random* conditions.

A difference in the initial activity time-locked to the Gabor patch is evident at E2, E3 and E4 within a time range of 0–75 ms (Figures 4 and 5). This difference is not evident at E1 and T. It is worth noting that all the entrainers (E1, E2, E3 & E4) and the T were baseline corrected using the same data extracted before the fixation cross at the beginning of the trial. It is worth noticing that this early effect might arise because of the temporal predictability affecting ongoing brain activity before the initial visual response to each Gabor. However, to address the questions asked in this study, we focussed on the effect of temporal predictability on only the initial evoked response. Thus, we predefined our time window for statistical comparison in the range of 75–135 ms, and we do not report statistical comparisons outside of this time window. Of notice, the 0–75 ms effect was not different between orientation predictable (Figure 4) and orientation unpredictable (Figure 5) conditions.

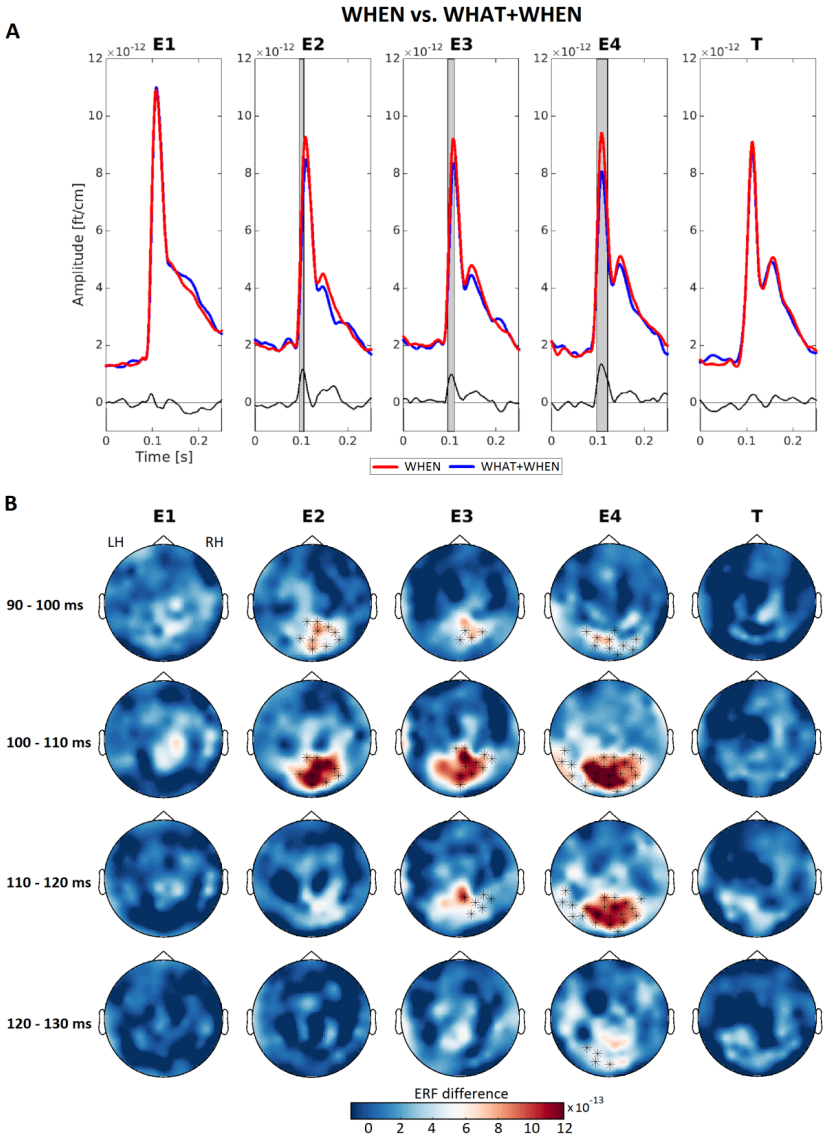


Figure 2: Sensor level ERFs for the *when* and the *what* + *when* conditions. A) For each condition (red, *when*; blue, *what* + *when*) and stimuli (Entrainer 1 (E1), E2, E3, E4 and Target (T)), we show the average of the event related fields (ERFs) in representative channels located above occipital regions (MEG02042/3, MEG2032/3, MEG2342/3, MEG2122/3 and MEG1922/3). Also shown are the ERF difference between the *when* and the *what* + *when* (black line). Grey boxes represent time points where the amplitude of the ERFs was higher ($p < 0.01$, cluster-based permutation test) for the *when* compared to the *what* + *when* condition. B) Sensor maps of the ERF difference between the *when* and the *what* + *when* conditions in temporal windows ([90 – 100], [100 – 110], [110 – 120] and [120 – 130] ms) around the amplitude peak value. Sensors showing significant differences ($p < 0.01$, cluster-based permutation test) are highlighted.

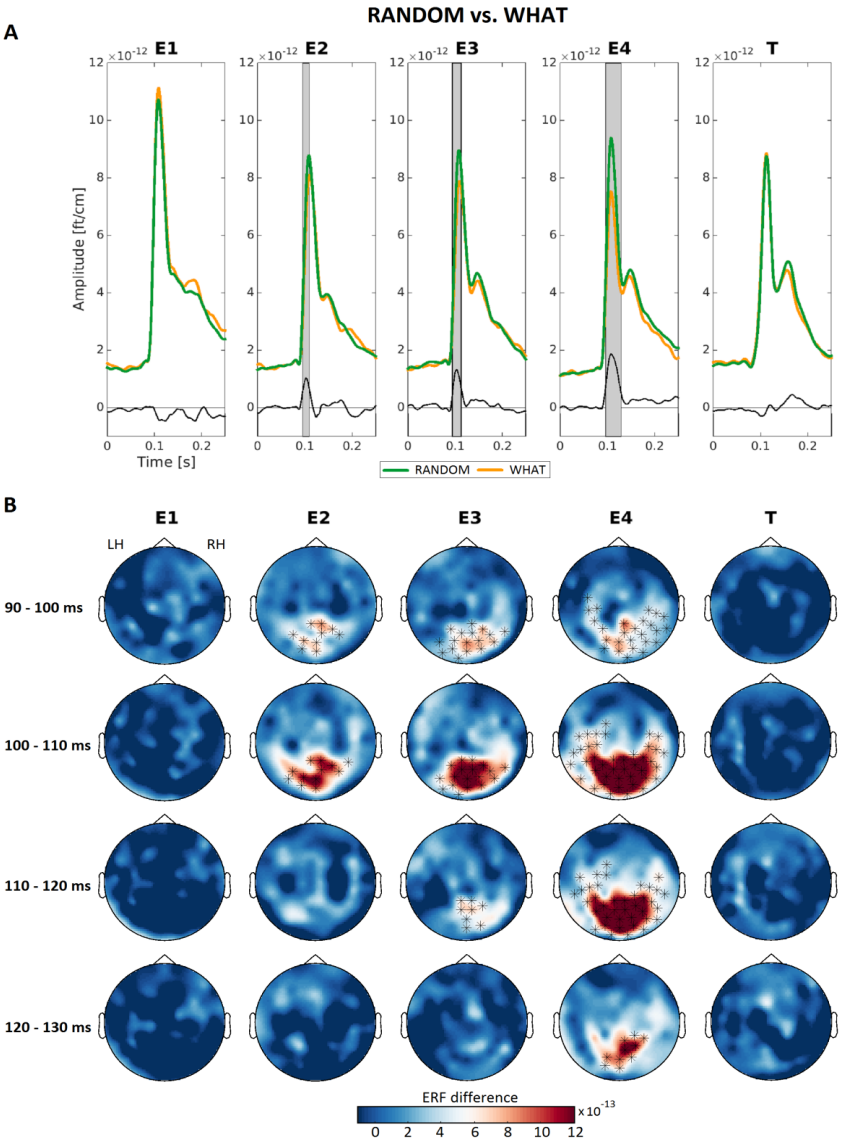


Figure 3: Sensor level ERFs for the *what* and the *random* conditions. For each condition (orange, *what*; green, *random*) and stimuli (E1), E2, E3, E4 and Target (T)), we showed the average of the Event Related Fields (ERFs) Grey boxes represent time points where the amplitude of the ERFs was higher ($p < 0.01$, cluster-based permutation test) for the *random* compared to the *what* condition. B) Sensor maps of the ERF difference between the *when* compared to the *what* + *when* condition in temporal windows ([90 – 100], [100 – 110], [110 – 120] and [120 – 130] ms) around the amplitude peak value. Sensors showing significant differences ($p < 0.01$, cluster-based permutation test) are highlighted.

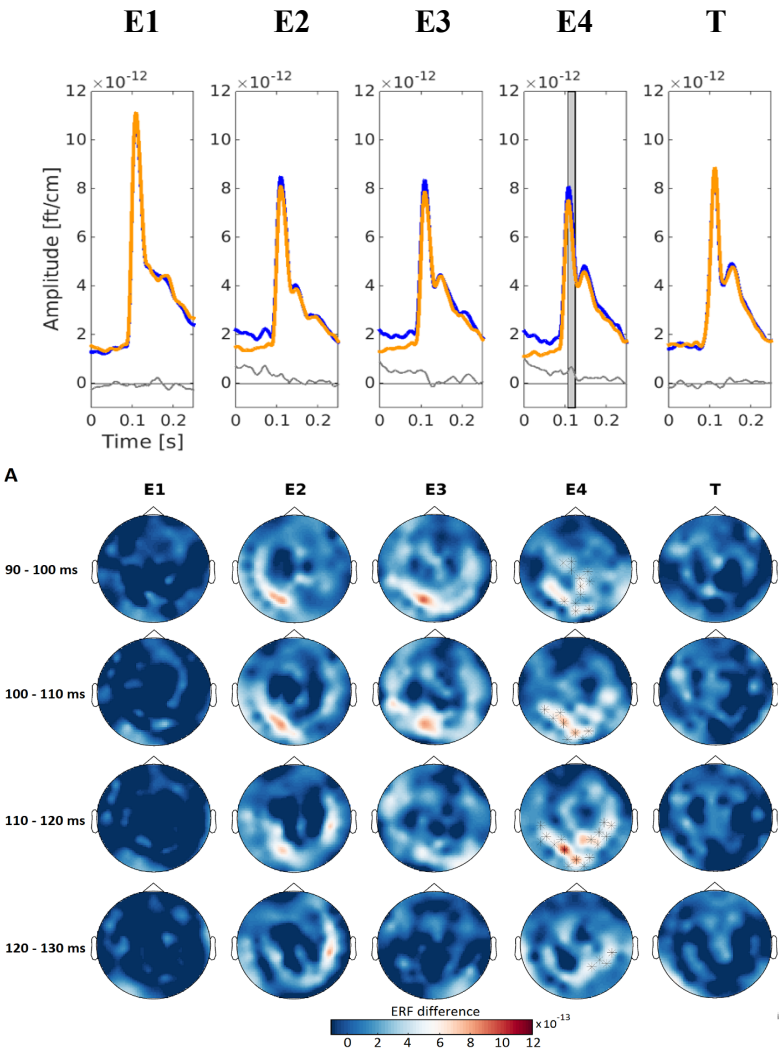


Figure 4: Sens(level ERFs for the *what* + *when* and the *what* conditions. A) For each condition (orange, *what*; blue, *what* + *when*) and stimuli (Entrainer 1 (E1), E2, E3, E4 and Target (T)). Grey boxes represent time points where the amplitude of the ERFs was higher ($p < 0.05$, cluster-based permutation test) for the *what* + *when* compared to the *what* condition. B) Sensor maps of the ERF difference between the *what* + *when* compared to the *what* condition in temporal windows ([90 – 100], [100 – 110], [110 – 120] and [120 – 130] ms) around the amplitude peak value. Sensors showing significant differences ($p < 0.01$, cluster-based permutation test) are highlighted.

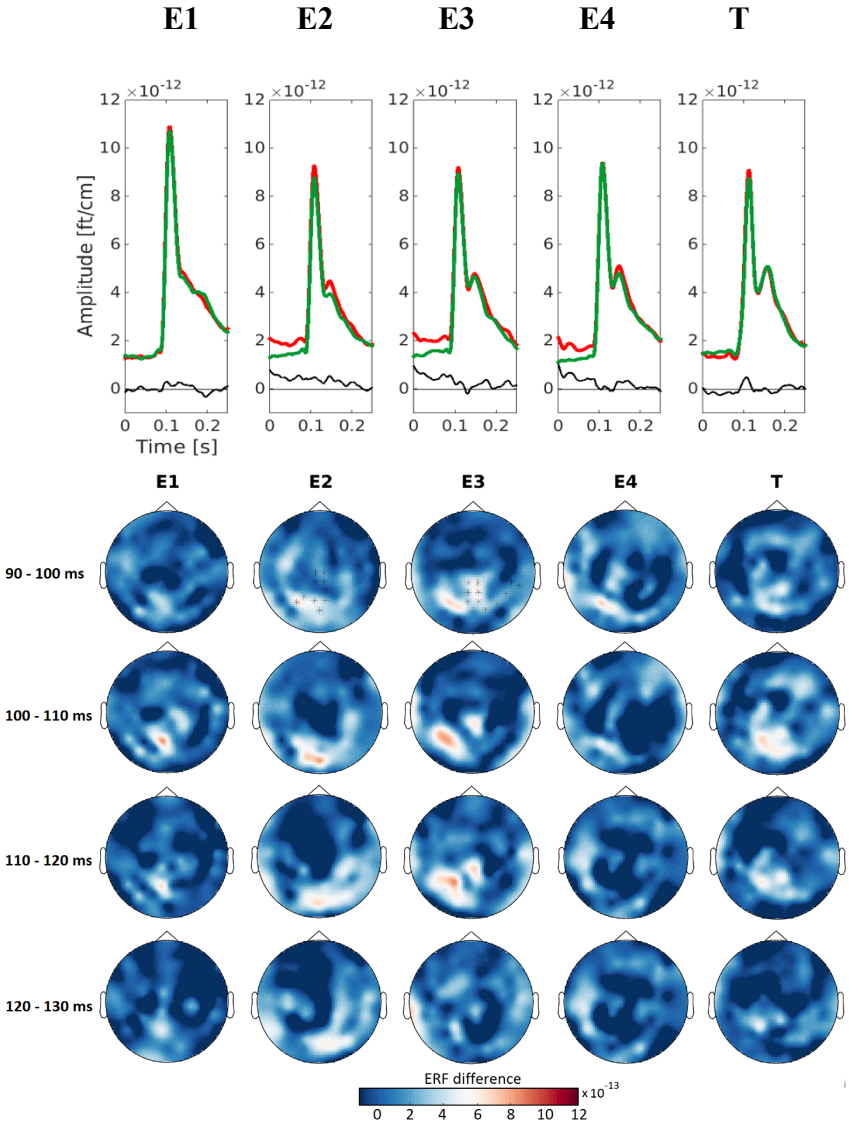


Figure 5: Sensor level ERFs for the *when* and the *random* conditions. A) For each condition (red, *when*; green, *random*) and stimuli (E1), E2, E3, E4 and Target (T).
B) Sensor maps of the ERF difference between the *when* compared to the *random* condition in temporal windows ([90 – 100], [100 – 110], [110 – 120] and [120 – 130] ms) around the amplitude peak value. Sensors showing significant differences ($p < 0.01$, cluster-based permutation test) are highlighted.

Source level ERFs

We next identified the brain regions underlying the significant differences between conditions at the sensor level. Source activity was estimated around the peak amplitude of the ERFs that we observed at the sensor level ERFs, i.e., in the 85–125 ms interval. Whole-brain maps of source activity were created for each condition (*what* + *when*, *when*, *what* and *random*) and Entrainer (E1, E2, E3, E4) (Figure 6). Strongest source activity was observed in bilateral occipital regions for all conditions and entrainers compared to baseline at the group level. The first local maxima emerged in visual association areas (Brodmann Area 18: BA 18) of the left occipital cortex in all conditions and entrainers. The mean value of all the local maxima was [–3 –76 –2]. All the individual local maxima are reported in supplementary table 1.

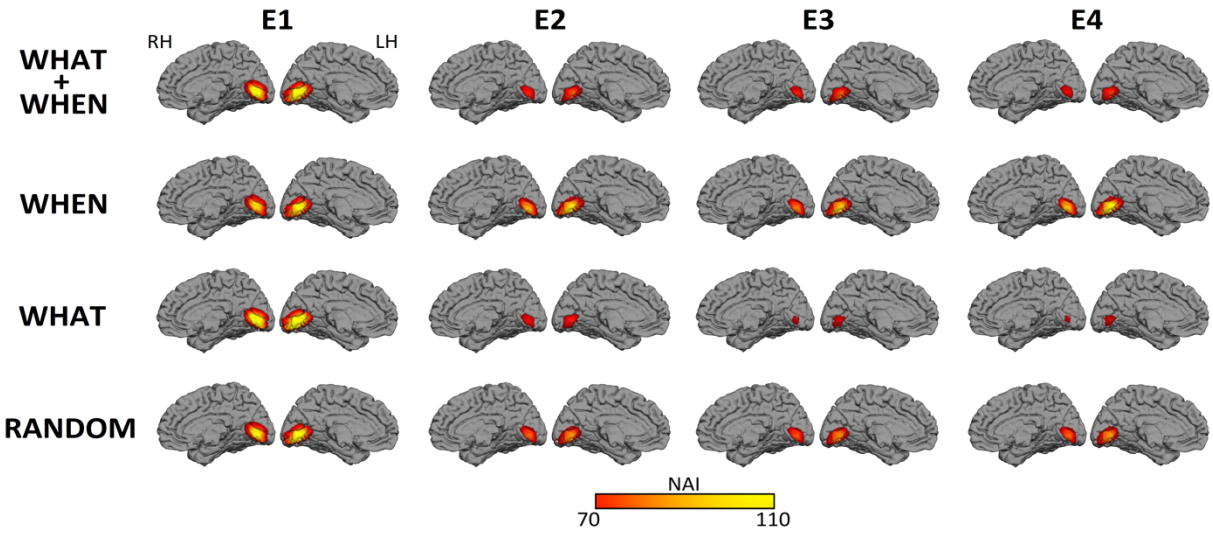


Figure 6: Source level ERFs for all conditions. Brain maps representing source activity values for each condition (*what* + *when*, *when*, *what* and *random*) and Entrainer (E1, E2, E3 and E4). We included a view of the medial surface and the occipital lobe of the left (LH) and the right (RH) hemisphere.

Figure 7 shows source amplitude around the mean location of the participants' peak values ([–3 –76 –2] mm)for each condition and entrainer. Source amplitude was significantly higher for the *when* compared to the *what + when* condition in E3 and E4 (t=2.45, t=4.16 respectively $p < 0.05$), andr for the *random* compared to the *what* condition in E3 and E4 (t=2.64, t=5.20 respectively, $p < 0.05$). Crucially, these values were higher for the *what + when* compared to the *what* conditions at E3 and E4 (t=2.30, t=2.38 respectively, $p < 0.05$), and for the *when* compared to the *random* condition in E2 and E3 (t=2.65, t=2.55 respectively, $p < 0.05$). Overall, the present results confirm the effects observed at the sensor-level, providing a candidate location for the generation of the expectation suppression effects reported.

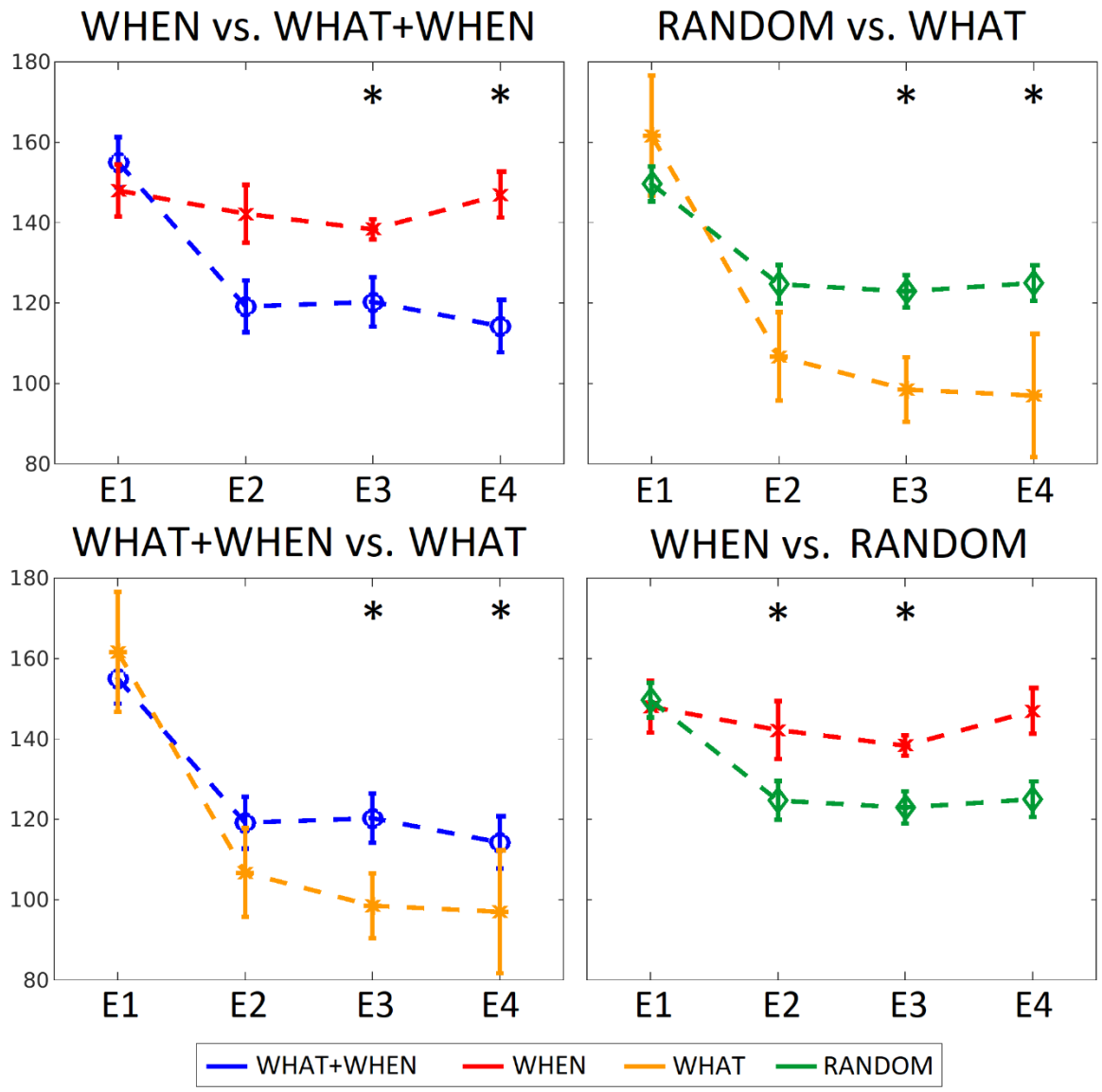


Figure 7: Local maxima activity values for all the conditions. Mean and standard values across participants of the source activation at the first local maxima for each condition (*when*, *what* + *when*, *random* and *what*) and Entrainer (E1, E3, E3 and E4). The asterisks indicate the comparisons that were statistically significant.

MVPA results

MVPA analyses show that only the conditions having predictable orientation (*what* + *when* and *what*) revealed above-chance and statistically significant decoding accuracy values. The decoding accuracy of horizontal versus vertical targets in paired conditions (i.e., *what* + *when* vs. *when* and *what* vs. *random*) is available in

the Supplementary materials (supplementary Figure 1 and 2 respectively). Figure 8 shows the horizontal versus vertical decoding accuracy of the predictable orientation with (*what* + *when*) and without (*what*) temporal predictability.

At E1, we found significantly above chance (50%) decoding starting at 135 ms post stimulus onset and peaking with 56.82 % accuracy at 140 ms (in a significant cluster spanning the 135–150 ms interval) in the *what* + *when* condition. In the *what* condition, decoding accuracy was significantly above chance from 140 ms, peaking with 56.88 % accuracy at 165 ms (significant cluster: 140–165 ms). These effects are short-lived in time and slightly above chance. For this reason we do not consider this a robust finding. It is worth noticing that in conditions lacking the predictable orientation (i.e., *when* and *random*) there were no significant clusters at E1, likely because these conditions are starting from random orientation angles of the Gabor patch.

At E2, we observed significantly above chance decoding clusters in the *what* + *when* condition. The cluster emerged from –20 ms to –10 ms before the presentation of the second entrainer. This cluster was very small in size and peaking with 57.41% accuracy at –15 ms. This effect probably corresponds to a false positive.

At E3, the first cluster in the *what* + *when* condition started at 100 ms, and peaked with 63.07 % accuracy at 120 ms (significant cluster: 100–125 ms). In the *what* condition, decoding accuracy was significant from 125 ms, peaking at 60.20% at 135 ms (significant cluster: 125–160 ms). Accuracy of these effects is higher and could reflect increasing expectations concerning the target orientation.

At E4, the condition *what* + *when* showed one cluster starting at 95 ms, peaking with 70.46 % accuracy at 145 ms (significant cluster 95–215 ms). In the *what* condition, results showed three subsequent significant clusters, the first of them starting at 100 ms peaking with 68.32 % accuracy at 140 ms (significant cluster 100–170 ms). These clusters are robust in both accuracy and duration reflecting the expectation for the orientation of the target gabor.

At the target, the *what* + *when* condition showed the earliest cluster in the time range 95–450 ms peaking with 78.95% accuracy at 105 ms. For the *what* condition, three significant clusters emerged, the first of them in the time range 95–275 ms peaking with 75.29% accuracy at 105 ms.

The MVPA results also highlighted that the spatial frequency was decoded at a chance level or all the four conditions i.e., *what* + *when, when, what* and *random* at the four entrainers (E1, E2, E3 and E4). At the target, results showed that there is a significant cluster ($p < 0.001$) in time range 80 to 550 ms peaking with 97.73 % accuracy at 150 ms for the *what* + *when* condition. For the *what* condition there is a significant ($p < 0.001$) cluster in time range 85 to 550ms peaking with 97.19 % accuracy at 150 ms.

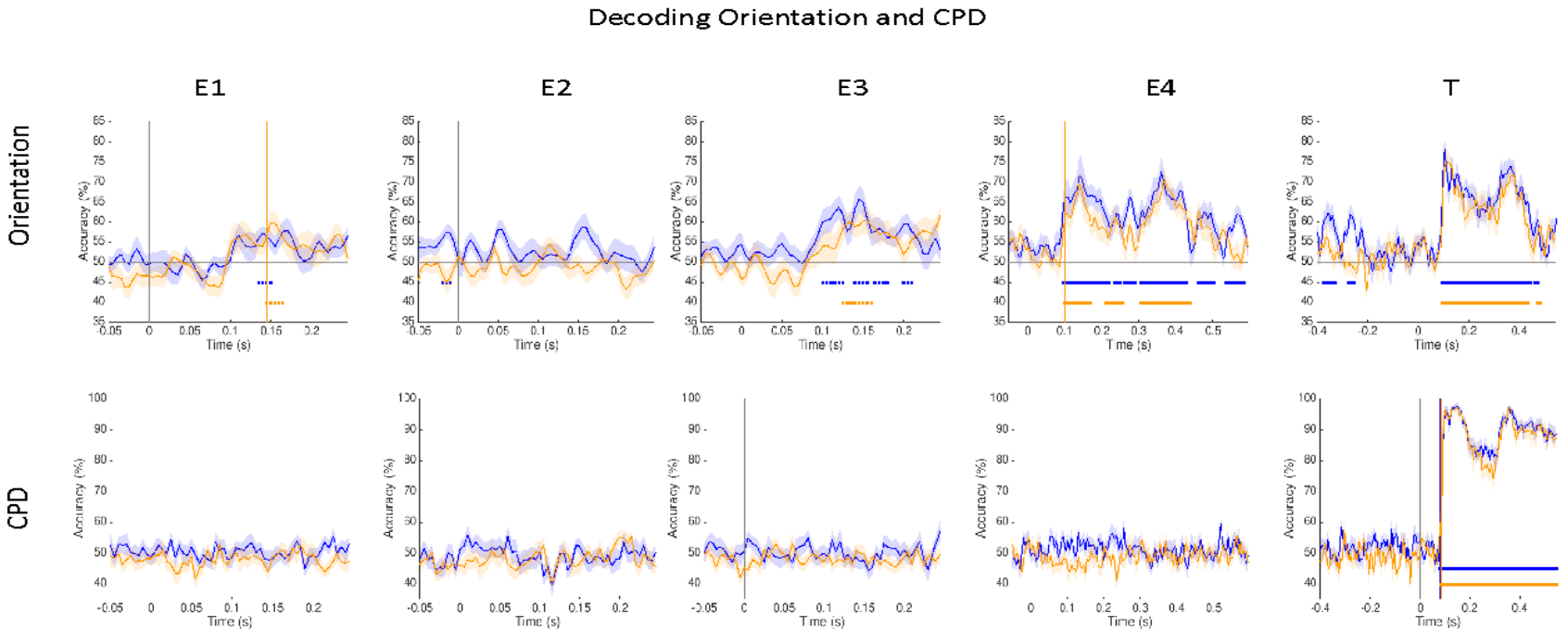


Figure 8: Time-resolved decoding accuracy for the *what* + *when* condition (blue line) and *what* condition (orange line) time-locked to Entrainer 1 (E1), E2, E3, E4 and Target (T). The coloured dots under the curves indicate statistical significance of decoding accuracy.

Discussion

In the present study we elicited robust expectation suppression effects for the processing of visual Gabors. Across a series of four entrainers we observed increasingly large suppression of the visual evoked responses when Gabor orientation was predictable, which was accompanied by increasingly large decodability of the target Gabor orientation. Importantly, these effects were modulated by temporal predictability: expectation suppression of the evoked responses was larger when timing of the entrainers was random and decodability of the visual response was less sustained in time for random timings. These findings indicate that the neurocognitive system invested less resources in visual analysis in temporally uncertain scenarios, possibly due to higher reliance on internal predictions.

Expectation suppression effects

The goal of our study was to establish how temporal uncertainty modulates predictions in the visual domain. Evidence for predictive processing mainly comes from three basic phenomena (Walsh et al., 2020): omission effects (neural response to the omission of an expected stimulus), repetition suppression (neural response to the same repeated stimuli), and expectation suppression (reduced neural response to a stimulus due to expectations generated by a prior cue). Expectation suppression has been studied - in both the visual and auditory domains - using experimental paradigms which generate expectations through predictive cues, predictable

stimulus sequences, and paired associations. The predictability of a stimulus has been usually generated by varying the probability of encountering a matching/mismatching stimulus based on a previous cue (Auksztulewicz et al., 2018; Utzerath et al., 2017). These kinds of manipulations are used in most of the experimental designs which aim to study the prediction errors elicited by a mismatching stimulus based on a previous cue. In our experimental design, we did not induce contextual expectations based on probability manipulations since our four experimental conditions had equal probability. Importantly, we did not employ "mismatching stimuli" in order to maximize the efficiency of the predictive mechanisms (in other words, enhancing precision weight).

Experiments focusing on expectation suppression trigger stimulus expectations using various methods: exposure, multi-day training, manipulating baseline probabilities, active or passive viewing, engaging or excluding attention (Walsh et al., 2020). Some of these studies inform participants about their contextual manipulations. In contrast, in our experiment we made sure participants were not aware of the experimental manipulations present in the study. Given the debated interaction between attention and predictive processing (Kok et al., 2012) we developed an experimental design which aimed to balance across conditions strategic effects on the processing of the Gabor orientation. While the orientation manipulation was noticeable, it is important to underscore that our participants did not report having noted the temporal jitter of the visual stimuli in the temporally unpredictable conditions.

There are several studies in which reduced neural responses for predictable stimuli have been found during passive viewing (Alink et al., 2010), as well as when stimuli are fully task irrelevant (Den Ouden et al., 2009), supporting the idea that expectation suppression may not vary based on task engagement. In contrast to this notion, however, other authors found no effect of expectation suppression on sensory activity when stimuli were unattended (Larsson and Smith, 2012) suggesting that contextually predictable stimuli may not result in the suppression of early visual neural responses (John-Saaltink et al., 2015). In the present experimental design, we observed expectation suppression effects when the orientation of the entrainers was predictable. This was the strongest main effect that we observed that also interacted with the position of entrainers along the sequence, increasing in size. We interpret this effect as increasing expectation developing in the visual circuitry for the orientation of the target Gabor: the stronger the expectation for the target orientation the larger the suppression of the visual response. This effect was significant at both E3 and E4 (Figure 7: top left and top right) and was present in both comparisons (with and without temporal predictability).

This peak effect possibly originated in visual area 2 (V2) which showed reduced activity for predictable stimuli compared to unpredictable stimuli. The source location of the present effect could reflect some sort of top-down activity generated in an extrastriate region projecting to primary visual cortex (V1). This possibility however should be further validated (possibly employing direct brain recordings in non-human primates) with additional connectivity analyses to investigate the

bidirectional interaction between V1 and V2 and determine if the flow of information in the top-down direction is enhanced for the content predictable conditions. The small spatial separation between V1 and V2 makes this comparison infeasible based on MEG recordings.

It is worth noting that this incremental effect was not mirrored in the behavioural responses, that probably reflect later decision processes. In addition, expectation suppression effects evident during the entrainer sequence vanish at the presentation of the target Gabor (where participants had to perform the task). At the target it is possible that resources invested for processing the task-relevant spatial frequency difference of the target Gabor from the entrainers interacted with the on-going neural expectation effects observable at the entrainers, thus washing out the stimulus predictability effects. It would be interesting in future studies to evaluate the processing of the different features of the target gabor (i.e., orientation and spatial frequency) by using a delayed cueing task in which participants are randomly cued after target presentation as to which taks they must perform (orientation or spatial frequency discrimination). Another possibility would be to avoid the use of any task, i.e. passive viewing, to determine if the expectation suppression effect is always preserved under such conditions.

The interaction between stimulus features and temporal predictability

In previous studies, similar experimental manipulations have been designed for testing the interaction between content and temporal predictability in the auditory

domain (Auksztulewicz and Friston, 2016; Todorovic and de Lange, 2012). In these experiments, expectation suppression effects have been observed at later latencies (100-200 ms). In the visual domain we observed a repetition suppression effect in an earlier time interval, i.e., at 85–125 ms. Given the different nature of the perceptual modality and due to the multiple differences between the experimental paradigms involved it is hard to directly compare these different results. More importantly, the analysis of the amplitude of this initial evoked response showed that there was a significant interaction between the *what* and *when* dimensions of visual stimuli. This interaction was mainly driven by the higher neural response to temporally predictable stimuli. Hence, tt seems that the perceptual system generates a greater response to an incoming stimulus whose onset is predictable compared to a stimulus that is not. It could be argued that the visual system is not "capable of preparing" for a temporally unpredictable stimulus. In fact, the effects observed in the initial evoked responses are preceded by a large difference between temporally predictable and unpredictable conditions in the evoked activity (Figures 4 and 5). This difference may be due that brain is differently preparing for the incoming stimuli depending on their temporal predictability.

However, it is worth noticing that at around 100 ms, when the initial visual peak is evident, there is a difference for the temporal contrast of the two content predictable conditions (*what* + *when* vs. *what*, Figure 4, sensor-level results) and not for the unpredictable ones (*when* vs. *random*, Figure 5, sensor-level results). If this effect would have been driven only by the inability of the visual system to prepare for a

visual stimulus (independently from its visual properties), one would expect a similarly larger visual response for both cases. One potential explanation for the larger suppression of the early evoked response in the temporally unpredictable condition (*what* condition) is that the visual system relies more on the internal prediction and less on the external evidence, when temporal uncertainty is higher (and if predictions have been developed). This evidence would support theoretical claims suggesting that predictive mechanisms are central for reducing the uncertainty of the external environment (Clark, 2013).

Stimulus specific neural activity

Since evoked responses did not speak for stimulus specificity in the present study, we used time-resolved multivariate pattern analysis to measure stimulus information that could be decoded from the neural activity. Our results show that the decoding of the Gabor orientation is increasing in magnitude (peak decoding accuracy) across entrainers when the target orientation is predictable. This indicates that stimulus predictability is a crucial factor to enhance the decodability of orientation during the presentation of the entrainers. In fact, when the stimulus is not predictable, decodability is at chance level.

Small but significant time differences emerged based on temporal predictability, with the effect at E4 showing an effect more sustained in time for temporally predictable (*what* + *when*) compared to the temporal unpredictable (*what*) conditions. It is possible that such difference could be ascribed to prolonged

processing by the visual system that invests more resources to analyse the stimulus features when the onset of the visual stimulus is predictable. The most robust effect observed for temporally predictable stimuli indicates that the visual system processes visual information of the orientation of the expected Gabor for a longer time period compared to when it is less temporally predictable. In temporally uncertain conditions the orientation information of the Gabor may be maintained active for less time. This difference mirrors the evoked effects, where we observed higher visual response for the temporally predictable, compared to the temporally unpredictable conditions. Decoding results thus reinforce our hypothesis that stimulus-specific neural processes are recruited more for the processing of the expected/temporally predictable visual stimuli, compared to expected/temporally uncertain visual stimuli.

A side note concerns the decoding of the spatial frequency. This feature was constant across entrainers and conditions thus leading to chance level decoding for all four entrainers. At the target, however, spatial frequency showed very high decoding (around 97–98 %), even higher than orientation decoding. Spatial frequency effects were also evident earlier and lasted longer than the orientation effects. Since our task focused on the difference in spatial frequency between the entrainers and target, the neural system likely maintains active for a longer interval the spatial frequency information of the target compared to the orientation of the target thus obscuring or interfering with the on-going expectation suppression effects due to Gabor orientation.

Conclusions

In the present study we investigated the effect of temporal predictability on visual predictive processing. Our results show that temporal predictability modulates processing of expected visual features. Lack of temporal predictability suppresses visual evoked responses at a large extent compared to temporally predictable visual stimuli, perhaps due to the fact that the brain increases reliance on internal prediction models.

Acknowledgements: This research is supported by the Basque Government through the BERC 2018-2021 program and by the Spanish State Research Agency through BCBL the Severo Ochoa excellence accreditation SEV-2015-0490 and through the project BES-2016-077560 funded by the Spanish Ministry of Economy and Competitiveness (MINECO). S N does acknowledge the support from EMBO short term fellowship. NM has been supported by the Spanish Ministry of Science, Innovation and University (grants PSI2015-65694-P, RTI2018-096311-B-I00), the Agencia Estatal de Investigación (AEI), the Fondo Europeo de Desarrollo Regional (FEDER) and by the Basque government (grant PI_2016_1_0014). The authors acknowledge Asier Zarraza, Ning Mei, David Soto and Lucia Amoruso for their valuable comments on the initial draft of this publication. SN thanks the BCBL lab research staff for their valuable support.

Declaration: The **author**(s) **declare**(s) that there is **no conflict of interest** regarding the publication of this article.

References :

Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., Muckli, L., 2010. Stimulus predictability reduces responses in primary visual cortex. J. Neurosci. 30, 2960–2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Auksztulewicz, R., Friston, K., 2016. Repetition suppression and its contextual determinants in predictive coding. Cortex 80, 125–140. https://doi.org/10.1016/j.cortex.2015.11.024

Auksztulewicz, R., Schwiedrzik, C.M., Thesen, T., Doyle, W., Devinsky, O., Nobre, A.C., Schroeder, C.E., Friston, K.J., Melloni, L., 2018. Not all predictions are equal: "what" and "when" predictions modulate activity in auditory cortex through different mechanisms. J. Neurosci. 38, 8680–8693. https://doi.org/10.1523/JNEUROSCI.0369-18.2018

Blom, T., Feuerriegel, D., Johnson, P., Bode, S., Hogendoorn, H., 2020. Predictions drive neural representations of visual events ahead of incoming sensory information. https://doi.org/10.1073/pnas.1917777117

Bourguignon, M., Molinaro, N., Wens, V., 2018. Contrasting functional imaging parametric maps: The mislocation problem and alternative solutions. Neuroimage 169, 200–211. https://doi.org/10.1016/j.neuroimage.2017.12.033

Carlson, T.A., Grootswagers, T., Robinson, A.K., 2019. An introduction to time-resolved decoding analysis for M/EEG.

Cichy, R.M., Pantazis, D., Oliva, A., 2014. Resolving human object recognition in space and time. Nat. Neurosci. 17, 455–462. https://doi.org/10.1038/nn.3635

Clark, A., 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204. https://doi.org/10.1017/S0140525X12000477

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical Surface-Based Analysis. Neuroimage 9, 179–194. https://doi.org/10.1006/nimg.1998.0395

Demarchi, G., Sanchez, G., Weisz, N., 2019. Automatic and feature-specific prediction-related neural activity in the human auditory system. Nat. Commun. 10, 1–11. https://doi.org/10.1038/s41467-019-11440-1

Den Ouden, H.E.M., Friston, K.J., Daw, N.D., McIntosh, A.R., Stephan, K.E., 2009. A dual role for prediction error in associative learning. Cereb. Cortex 19, 1175–1185. https://doi.org/10.1093/cercor/bhn161

Dima, D.C., Perry, G., Messaritaki, E., Zhang, J., Singh, K.D., 2018. Spatiotemporal dynamics in human visual cortex rapidly encode the emotional content of faces. Hum. Brain Mapp. 39, 3993–4006. https://doi.org/10.1002/hbm.24226

Dima, D.C., Singh, K.D., 2018. Dynamic representations of faces in the human ventral visual stream link visual features to behaviour.

Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J., 2008. LIBLINEAR: A library for large linear classification. J. Mach. Learn. Res. 9, 1871–1874. https://doi.org/10.1145/1390681.1442794

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J. -P, Frith, C.D., Frackowiak, R.S.J., 1994. Statistical parametric maps in functional imaging: A general linear approach. Hum. Brain Mapp. 2, 189–210. https://doi.org/10.1002/hbm.460020402

Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: Neural models of stimulus-specific effects. Trends Cogn. Sci. 10, 14–23. https://doi.org/10.1016/j.tics.2005.11.006

Guggenmos, M., Sterzer, P., Cichy, R.M., 2018. Multivariate pattern analysis for MEG: A comparison of dissimilarity measures. Neuroimage 173, 434–447. https://doi.org/10.1016/j.neuroimage.2018.02.044

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

John-Saaltink, E.S., Utzerath, C., Kok, P., Lau, H.C., De Lange, F.P., 2015. Expectation suppression in early visual cortex depends on task set. PLoS One 10, 1–14. https://doi.org/10.1371/journal.pone.0131172

King, J.R., Pescetelli, N., Dehaene, S., 2016. Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. Neuron 92, 1122–1134. https://doi.org/10.1016/j.neuron.2016.10.051

Kok, P., Mostert, P., De Lange, F.P., 2017. Prior expectations induce prestimulus sensory templates. Proc. Natl. Acad. Sci. U. S. A. 114, 10473–10478. https://doi.org/10.1073/pnas.1705652114

Kok, P., Rahnev, D., Jehee, J.F.M., Lau, H.C., De Lange, F.P., 2012. Attention reverses the effect of prediction in silencing sensory signals. Cereb. Cortex 22, 2197–2206. https://doi.org/10.1093/cercor/bhr310

Larsson, J., Smith, A.T., 2012. FMRI repetition suppression: Neuronal adaptation or stimulus expectation? Cereb. Cortex 22, 567–576. https://doi.org/10.1093/cercor/bhr119

Maunsell, J.H.R., Gibson, J.R., 1992. Visual response latencies in striate cortex of the macaque monkey. J. Neurophysiol. 68, 1332–1344. https://doi.org/10.1152/jn.1992.68.4.1332

Mechelli, A., Price, C.J., Friston, K.J., Ishai, A., 2004. Where bottom-up meets top-down: Neuronal interactions during perception and imagery. Cereb. Cortex 14, 1256–1265. https://doi.org/10.1093/cercor/bhh087

Mumford, D., 1992. On the computational architecture of the neocortex - II The role of cortico-cortical loops. Biol. Cybern. 66, 241–251. https://doi.org/10.1007/BF00198477

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011. https://doi.org/10.1155/2011/156869

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

Spratling, M.W., 2017. A review of predictive coding algorithms. Brain Cogn. 112, 92–97. https://doi.org/10.1016/j.bandc.2015.11.003

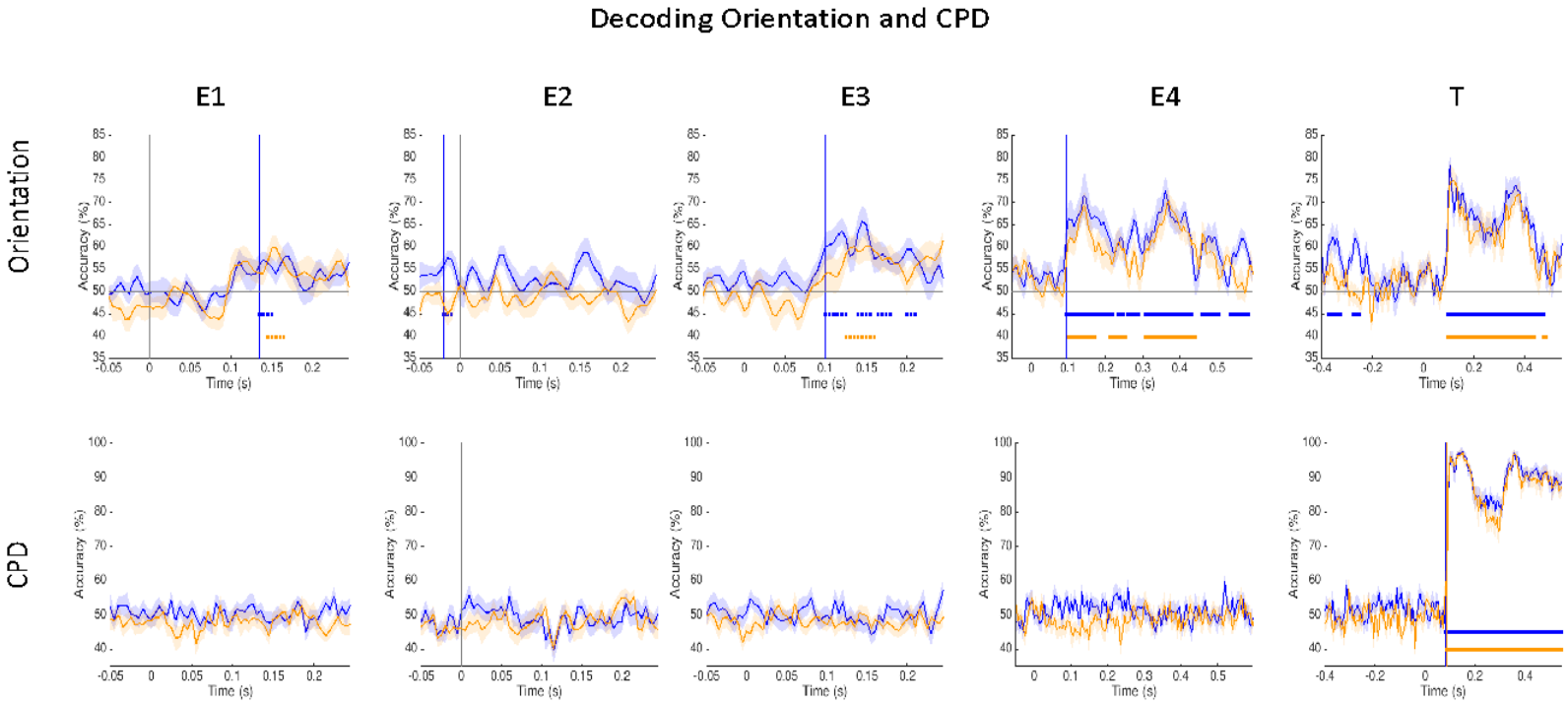
Todorovic, A., de Lange, F.P., 2012. Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. J. Neurosci. 32, 13389–13395. https://doi.org/10.1523/JNEUROSCI.2227-12.2012

Utzerath, C., St John-Saaltink, E., Buitelaar, J., De Lange, F.P., 2017. Repetition suppression to objects is modulated by stimulus-specific expectations. Sci. Rep. 7, 1–8. https://doi.org/10.1038/s41598-017-09374-z

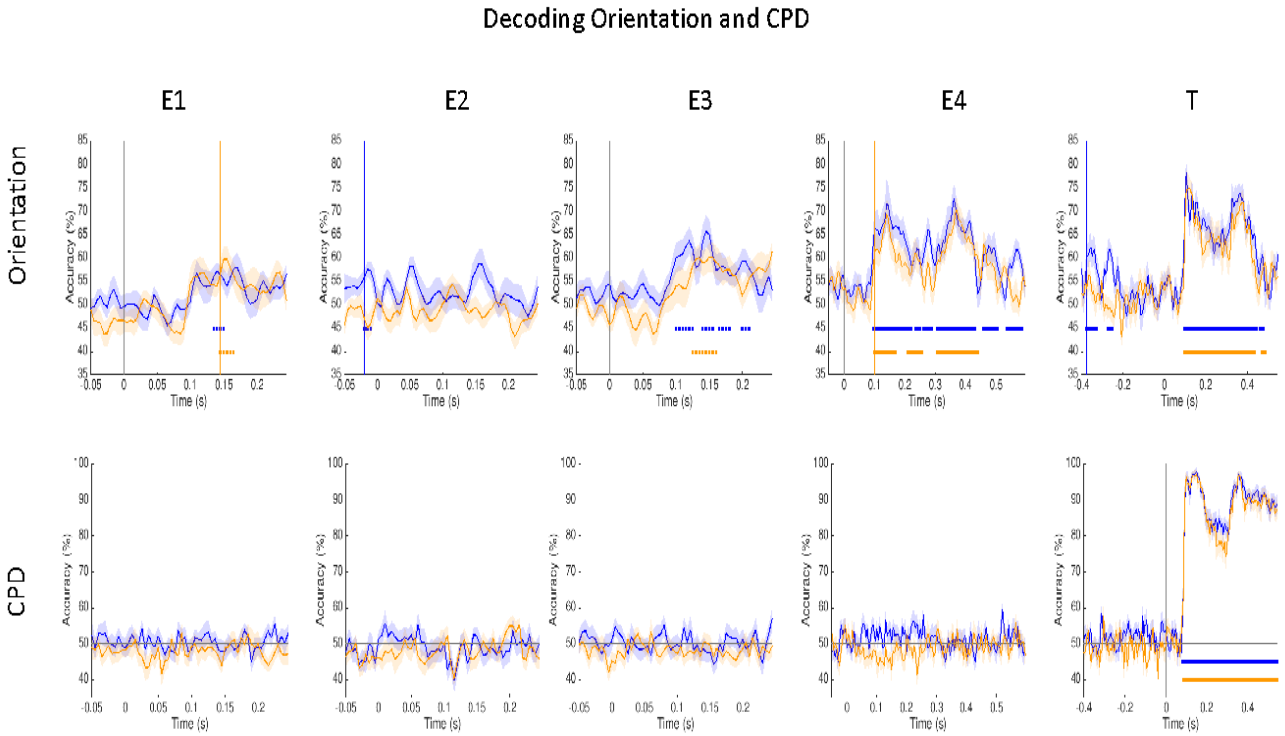
Veen, B.D. Van, Drongelen, W. Van, Yuchtman, M., Suzuki, A., 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. IEEE Transactions on. Biomed. Eng. (NY). 44, 867–880.

Walsh, K.S., Mcgovern, D.P., Clark, A., Connell, R.G.O., 2020. Evaluating the neurophysiological evidence for predictive processing as a model of perception 1–27. https://doi.org/10.1111/nyas.14321

Supplementary figures:



Sup. Figure 1: Time resolved decoding of *what* + *when* and *what* trials timelocked to Entrainer 1(E1), Entrainer 2 (E2), Entrainer 3 (E3), Entrainer 4 (E4) and Target (T). The coloured dots represents the statistical significance of accuracy.



Sup. Figure 2 : Time resolved decoding of *what* only and *random* trials time-locked to Entrainer 1(E1), Entrainer 2 (E2), Entrainer 3 (E3), Entrainer 4 (E4) and Target (T). The coloured dots represents the statistical significance of accu

Sup Table 1: Location of the local maxima expressed in MNI coordinates (x, y, z mm) for each condition and Entrainer.

	E1	**E2**	**E3**	**E4**
what* + *when	–2 –76 –1	–3 –74 –3	–4 –75 –5	–4 –75 –1
when	–3 –76 1	–3 –76 –5	–5 –77 –4	–4 –76 –1
what	–3 –77 0	–2 –76 –3	–3 –76 –3	–3 –75 –1
random	–3 –77 –1	–3 –77 –4	–4 –77 –3	–4 –77 –2