Running head: Repetition suppression and prediction error

Prediction Error and Repetition Suppression Have Distinct Effects on Neural Representations of Visual Information

Matthew F. Tang[1,2], Cooper A. Smout[1,2], Ehsan Arabzadeh[2,3] & Jason B. Mattingley[1,2,4]

1. Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia
2. Australian Research Council Centre of Excellence for Integrative Brain Function, Victoria, Australia
3. Eccles Institute of Neuroscience, John Curtin School of Medical Research, The Australian National University, Canberra, ACT, Australia
4. School of Psychology, The University of Queensland, St Lucia, QLD, Australia

Corresponding author

Matthew F. Tang

Queensland Brain Institute

The University of Queensland

St Lucia, QLD, Australia

Email: m.tang1@uq.edu.au

Abstract

Predictive coding theories argue recent experience establishes expectations in the brain that generate *prediction errors* when violated. Prediction errors provide a possible explanation for *repetition suppression*, in which evoked neural activity is attenuated across repeated presentations of the same stimulus. According to the predictive coding account, repetition suppression arises because repeated stimuli are expected, non-repeated stimuli are unexpected and thus elicit larger neural responses. Here we employed electroencephalography in humans to test the predictive coding account of repetition suppression by presenting sequences of gratings with orientations that were expected to repeat or change in separate blocks. We applied multivariate forward modelling to determine how orientation selectivity was affected by repetition and prediction. Unexpected stimuli were associated with significantly enhanced orientation selectivity, whereas there was no such influence on selectivity for repeated stimuli. Our results suggest that repetition suppression and expectation have separable effects on neural representations of visual feature information.

Introduction

At any moment in time, the brain receives more sensory information than can be responded to, creating the need for selection and efficient processing of incoming signals. One mechanism by which the brain reduces its information processing load is to encode successive presentations of the same stimulus in a more efficient form, a process known as neural adaptation (Fairhall, Lewen, Bialek, & de Ruyter van Steveninck, 2001; Kvale & Schreiner, 2004; Smirnakis, Berry, Warland, Bialek, & Meister, 1997). Such adaptation has been observed across different sensory modalities and species, and has been suggested as a potential mechanism for enhancing the coding efficiency of individual neurons and neuronal populations (Adibi, McDonald, Clifford, & Arabzadeh, 2013; Benucci, Saleem, & Carandini, 2013; Maravall, Petersen, Fairhall, Arabzadeh, & Diamond, 2007). A particular form of neuronal adaptation, known as *repetition suppression*, is characterised by attenuated neural responses to repeated presentations of the same stimulus (Diederen, Spencer, Vestergaard, Fletcher, & Schultz, 2016; Gross, Schiller, Wells, & Gerstein, 1967; Keller et al., 2017; Movshon & Lennie, 1979; Rasmussen, Schwartz, & Chase, 2017). Here we asked whether predictive coding theory, which assumes that sensory processing is influenced by prior exposure, can account for changes in neural representations observed with repetition suppression.

The phenomenon of repetition suppression has been widely exploited to investigate neural representations of sensory information. Repeated exposures allow for more efficient representation of subsequent stimuli, as manifested in improved behavioural performance despite a significant reduction in neural activity (Henson &

Rugg, 2003; Schacter & Buckner, 1998). Repetition suppression paradigms have been used extensively in human neuroimaging because they are commonly considered to be analogous to the single-cell adaptation effects observed in animal models (see Barron, Garvert, & Behrens, 2016 for review). The exact relationship between the effects seen in human neuroimaging studies and animal neurophysiology has, however, yet to be fully established.

The view that repetition suppression observed in human neuroimaging studies reflects neuronal adaptation has recently been challenged by hierarchical predictive coding theories (Auksztulewicz & Friston, 2016; Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008). These theories argue that the brain interprets incoming sensory events based on what would be expected from the recent history of exposure to such stimuli (Friston, 2005; Rao & Ballard, 1999). Predictions are generated within each cortical area, and are bi-directionally propagated from higher to lower areas, including to primary sensory regions, allowing for more efficient representation of expected stimuli. When there is a precise expectation, incoming information can be efficiently represented by recruiting a small pool of relevant neurons (Friston, 2005). When there is a mismatch between an expectation and the stimulus presented, i.e., when there is a *prediction error*, the stimulus is less efficiently represented and thus elicits a larger neural response.

Some of the most compelling evidence for predictive coding comes from human neuroimaging experiments in which the presentation of an unexpected stimulus generates a larger response than the presentation of an expected stimulus. In studies employing electroencephalography (EEG) and magnetoencephalography

(MEG), this effect is known as the *mismatch negativity* (Garrido, Kilner, Stephan, & Friston, 2009; Näätänen, Paavilainen, Rinne, & Alho, 2007; Wacongne et al., 2011)

To test the hypothesis that prediction error can account for repetition suppression effects, Summerfield and colleagues (2008) introduced an experimental paradigm in which the identity of a face stimulus was either repeated in 80% of trials (making the repetition *expected*) or was changed in 80% of trials (making the repetition *unexpected*). There was a greater attenuation of the BOLD response in the fusiform face area when a face repetition was expected, relative to when it was unexpected, suggesting that repetition suppression is reduced by unexpected stimuli. This attenuation of repetition suppression by failures of expectation has also been replicated using fMRI (Larsson & Smith, 2012) and M/EEG, using high-level stimuli such as faces (Summerfield, Wyart, Mareike Johnen, & de Gardelle, 2011), and simple stimuli such as tones (Todorovic & de Lange, 2012; Todorovic, van Ede, Maris, & de Lange, 2011)

A potential reconciliation of the relationship between expectation and repetition suppression comes from work showing that while expectations decrease the overall amount of neural activity, they can also yield sharper representations of sensory stimuli (Kok, Jehee, & de Lange, 2012). This work goes beyond conventional neuroimaging approaches, which typically only measure overall levels of neural activity (Buckner et al., 1998; Kourtzi & Kanwisher, 2001; Tootell, Reppas, Dale, & Look, 1995). Such amplitude changes could in principle be produced by one or more different types of change in the underlying neural representations. For instance, both sharpening, where the response to only unpredicted features is

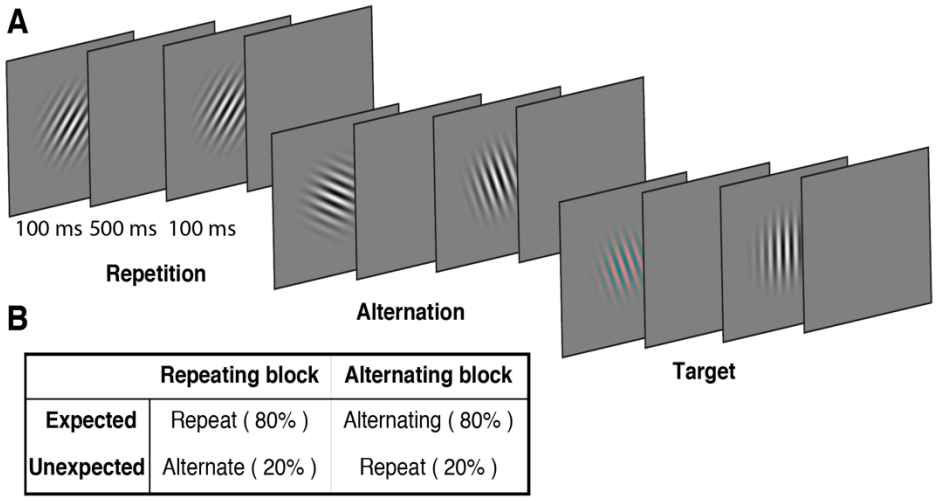
suppressed, and gain reduction, where a multiplicative suppression occurs for all features, could be associated with decreased population activity, even though the amount of information carried by the representations will be markedly different. Recently introduced multivariate pattern analytic approaches to human neuroimaging – specifically forward encoding modelling – allow for the quantification of stimulus-selective information contained within patterns of neural activity in human observers (Brouwer & Heeger, 2009; Garcia, Srinivasan, & Serences, 2013; King, Pescetelli, & Dehaene, 2016; Myers et al., 2015; Salti et al., 2015; Wolff, Jochim, Akyürek, & Stokes, 2017). This goes beyond typical multivariate pattern analyses, which normally produce only accuracy scores, by quantifying neural representations evoked by sensory stimuli to reveal both the accuracy and *tuning fidelity* for the specific feature dimension of interest.

Here we used these multivariate forward encoding methods to determine how repetition suppression and expectation affect the way the brain represents visual orientation information. To anticipate the results, we found that repetition suppression had no effect on visual orientation selectivity, but that violated expectations were associated with a significantly increased orientation-selective response through gain modulation soon after stimulus onset. This representation appeared to be transiently re-activated at around 200 ms, suggesting that feedback influences the initial sensory representation of the unexpected stimulus which allows for the updating of the sensory prior.

Results

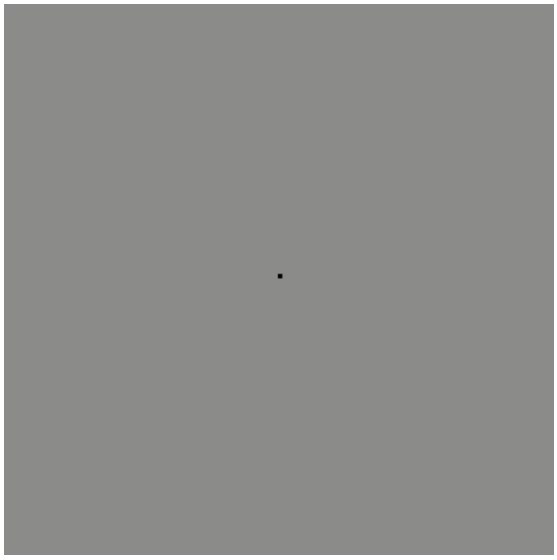
We used a modified version of the paradigm introduced by Summerfield and

colleagues (2008), replacing the face stimuli used in that study with oriented Gabors. These low-level stimuli allowed us to quantify the degree of orientation selectivity in EEG activity to determine how the representations of orientation is affected by prediction error and repetition suppression. Each of fifteen observers participated in two EEG sessions. On each trial, two Gabors were presented sequentially (100 ms presentation, 600 ms stimulus onset asynchrony), and these stimulus pairs either repeated or alternated in their orientation (Figure 1A, Movie 1). The predictability of the repeated and alternating pairs was varied in a block-wise manner to manipulate expectation. In a *repeating* block, the orientations of the two Gabors in a pair repeated in 80% of trials, and alternated for the remaining 20%. These contingencies were reversed in the *alternating* block (Figure 1B). The orientations of successive stimuli across a block were randomized to limit any accumulated effects of adaptation and prediction. As repetition suppression and expectation form orthogonal dimensions of the task, the design allowed us to isolate their respective contributions to neural responses. Participants completed an unrelated task of discriminating (red vs blue) rare coloured Gabors (which occurred on 10% of trials).



	Repeating block	Alternating block
Expected	Repeat (80%)	Alternating (80%)
Unexpected	Alternate (20%)	Repeat (20%)

Figure 1. Example stimulus displays and task design. **(A)** Schematic of the stimuli and timing used in the experiment. Participants viewed a rapid stream of pairs of Gabors and monitored for an infrequent coloured target (10% of trials). The stimulus orientations varied pseudorandomly across trials between 0° and 160° (in 20° steps), allowing estimation of orientation-selective information contained within patterns of EEG activity. **(B)** The orientation of the pairs of Gabors could either repeat or alternate. In one type of block, 80% of trials were orientation repeats and the remaining 20% alternated (Repeating blocks); in the other type of block these contingencies were reversed (Alternating blocks).



Movie 1. Example of a stimulus sequence of Gabors in a typical repetition block.

Repetition suppression and prediction error affect the overall level of neural activity

The Gabors elicited a large response over occipital-parietal areas (Figure 2A). Consistent with previous work (Cui, Wang, Park, Demb, & Butts, 2016; Keller et al., 2017; Rentzeperis, Nikolaev, Kiper, & Van Leeuwen, 2012; Summerfield et al., 2011; Todorovic et al., 2011; Todorovic & de Lange, 2012; Tootell et al., 1998), there was a significant repetition suppression effect (Repeat < Alternating), such that the response to repeated stimuli was significantly reduced compared with the response to alternating stimuli (Figure 2A). The repetition suppression effect was evident over a large cluster of occipital-parietal electrodes at two time intervals: an early effect from 79 to 230 ms, and a later effect at 250 to 540 ms after the onset of the second stimulus (cluster $p < .025$; Figure 2B and caption). A large cluster of frontal electrodes mirrored the repetition suppression effect with a similar time course: the ERP over these frontal sites had the same pattern, but was reversed in sign, suggesting it originated from the same dipole as the occipital response.

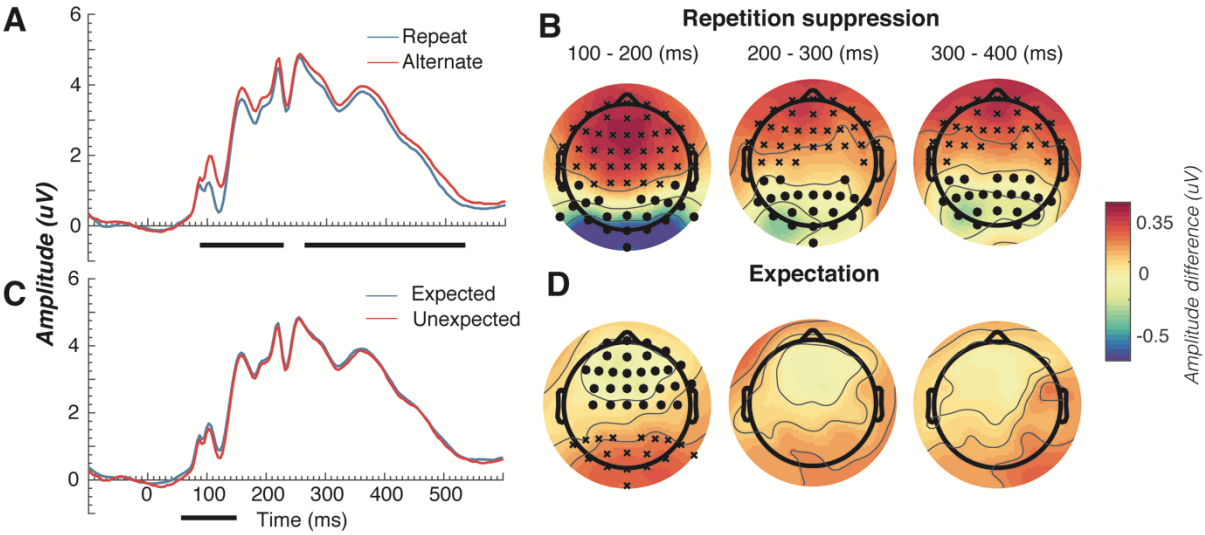


Figure 2. Univariate EEG results for the effect of repetition suppression and prediction error on the second stimulus in a pair (T2). Grand average ERP time courses for repetition suppression **(A)** and expectation **(C)** averaged over occipital-parietal electrodes (O1, O2, Oz, POz, PO7, PO3, PO8, PO4). The shaded region indicates significant differences between the conditions (two-tailed cluster-permutation, alpha $p < .05$, cluster alpha $p < .05$, N permutations = 20,000). The main effects of repetition suppression **(B)** and expectation **(D)** over three time periods and across all electrodes. The main effect of repetition suppression is displayed as Repeating minus Alternating trials. The main effect of expectation is displayed as Expected minus Unexpected trials. Circles indicate clusters of electrodes with significantly reduced activity, and crosses indicate clusters of electrodes with significantly increased activity (alpha $p < .05$, cluster $p < .025$, N permutations = 1500).

Also consistent with previous results (Garrido et al., 2009; Summerfield et al., 2011; Todorovic et al., 2011; Todorovic & de Lange, 2012), there was a significant prediction error effect (Expected < Unexpected), with a larger neural response over a cluster of frontal electrodes 75-150 ms after stimulus presentation (Figure 2C). As with the repetition suppression result described above, there was a prediction error effect of opposite polarity over occipital-parietal electrodes. This effect was significant at an early time point post-stimulus (79-130 ms), but not at later time points (320-390 ms; Figure 2D). Finally, there was no interaction between repetition suppression and prediction error (i.e., no significant positive or negative clusters, all

$p > .05$). Taken together, these results reveal both repetition suppression and prediction error effects in the neural data, which were indexed separately as shown in Figure 2.

Prediction errors increase the amount of orientation-selective information contained within patterns of EEG activity

We next examined the key question of whether repetition suppression and expectation differentially affect the neural representation of orientation information. To do this, we used a forward encoding approach to reconstruct orientation-selective information contained within the multivariate pattern of EEG activity distributed across the scalp (see Methods for details). Briefly, this technique transforms sensor-level responses into tuned ‘feature’ channels (Brouwer & Heeger, 2009; Garcia et al., 2013; Kay, Naselaris, Prenger, & Gallant, 2008; Myers et al., 2015), in this case, orientation-selective features. For each trial, the presented orientation was convolved with a canonical, orientation-selective tuning function and regressed against the pattern of EEG activity across all sensors at each time point. This created a spatial filter of the multivariate EEG activity that differentiated orientations (Figure 3D). These weights were then inverted to reconstruct the model, and multiplied against an independent set of test trials to produce responses in the modelled orientation channels. These sets of responses were then used to evaluate the degree of orientation selectivity in those trials. The procedure was repeated for all time points in the trial, and a cross-validated approach was used until all trials had been used for both training and testing.

As shown in Figure 3, the forward encoding revealed a strong, orientation-

selective response derived from the multivariate pattern of EEG activity. This orientation-tuned response was evident from ~50 to 470 ms after stimulus onset, and peaked between ~120-250 ms (Figure 3C). Examination of the regression weights revealed that this response was largely driven by activity centred over occipital-parietal areas (Figure 3D).

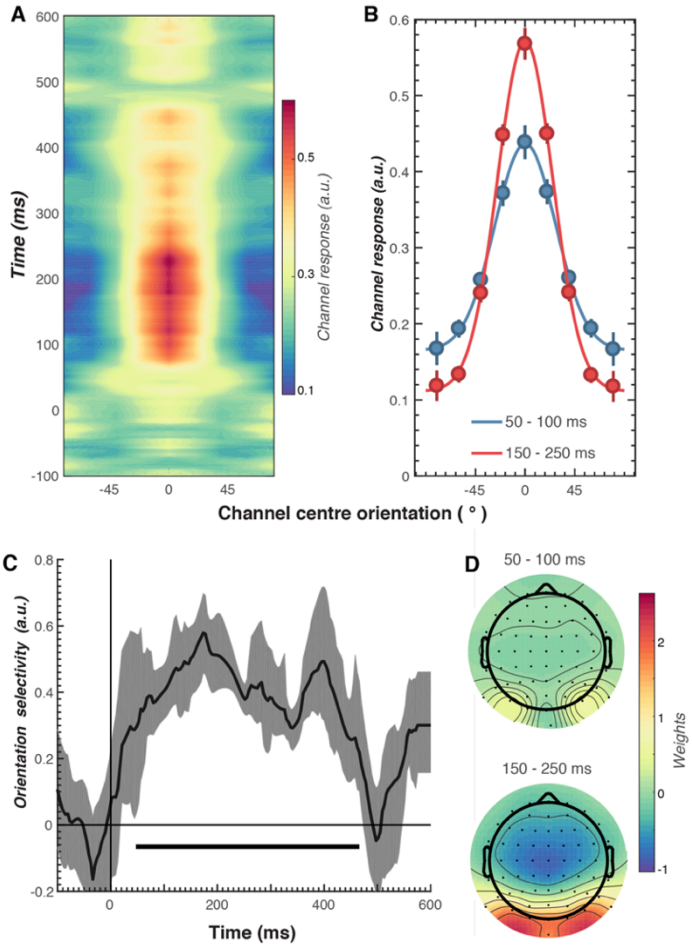


Figure 3. Results of the forward encoding modelling for orientation-selectivity. **(A)** Time-resolved orientation tuning curve across all participants and conditions in response to the second Gabor in the pair. The forward encoding approach resulted in a tuning curve for each of the nine presented orientations. These tuning curves were then centred at each presented orientation (here labelled as 0°) to combine across all orientations. The orientation-selective response is contained within the overall pattern of EEG; activity begins soon after stimulus onset and peaks at around 250 ms before declining. **(B)** Population tuning curve of the stimulus reconstruction across participants, averaged between 50-100 ms and 150-250 ms after stimulus presentation. Each line is a fitted Gaussian response with a variable offset used to quantify orientation selectivity. Error bars indicate ±1 standard error of mean across participants. **(C)** Amplitude of the channel response over time, averaged across all conditions (black line). The thick black line indicates significant encoding of orientation based on a cluster-permutation test across participants (cluster $p < .05$, N permutations = 20,000). Encoding accuracy was reliable from 52 to 470 ms post-stimulus onset. The error shading (in grey) indicates bootstrapped 95% confidence intervals of the mean. **(D)** Topographic plots of the weights (averaged across the 9 orientation channels across all participants) derived from forward encoding at the corresponding time points shown in panel B. (a.u. = arbitrary units).

Repetition suppression and prediction error

To examine our central question of whether repetition suppression and prediction error have differential effects on neural representations of orientation, we split and averaged the results of the forward encoding by trial type, and fitted these with Gaussians (see Methods) to quantify orientation selectivity (Figure 4). Repetition suppression did not affect the amount of orientation selectivity contained within the EEG data, with similar selectivity for repeated and alternating trials. This was the case even though the repeated trials had a markedly smaller EEG response over occipital and parietal electrodes, where the forward encoding model was maximally sensitive (Figure 2A). This result is consistent with the 'efficient representation' hypothesis of repetition suppression, which argues that the overall neural response is smaller with repetition suppression due to more efficient coding of stimulus information (Gotts, Chow, & Martin, 2012).

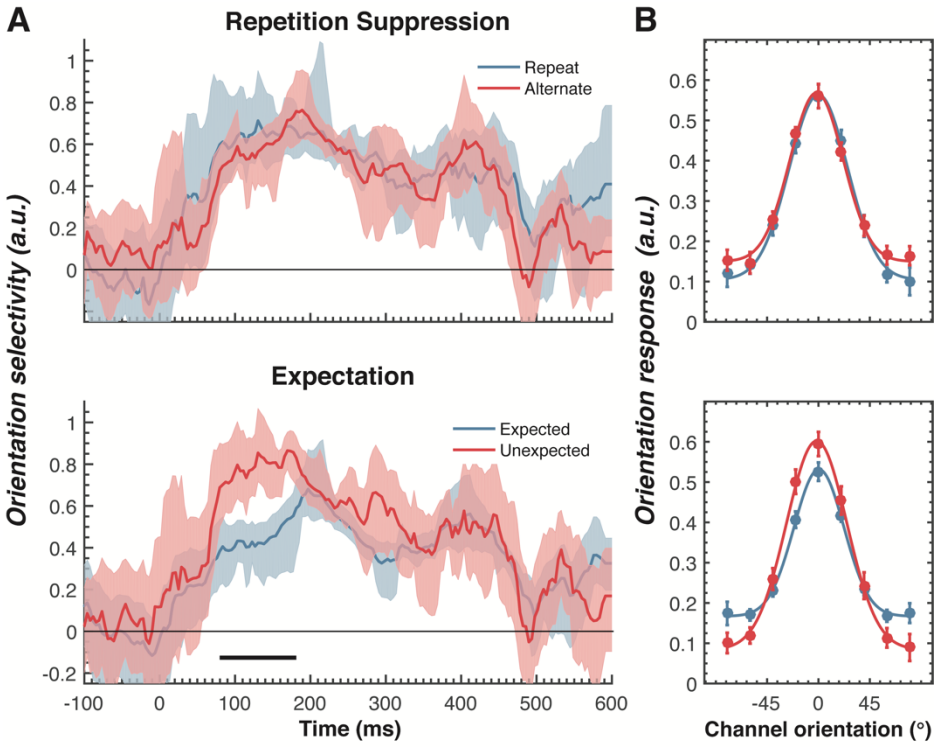


Figure 4. The effect of repetition suppression and expectation on orientation selectivity measured using forward encoding modelling. **(A).** Amount of orientation-selective information (given by the amplitude of the fitted gaussian) from the EEG signal in response to the second Gabor in a pair, shown separately for repetition suppression (upper panel) and prediction error (lower panel). The thick black line indicates significant differences between the conditions (two-tailed cluster-permutation, alpha $p < .05$, cluster alpha $p < .05$, N permutations = 20,000). Error shading indicates bootstrapped 95% confidence intervals of the mean. **(B)** The population tuning curve averaged over the significant time period (79 – 185 ms) to show how the overall stimulus representation is affected with a fitted Gaussian (line). For prediction error, the amplitude increased and the baseline decreased. There was no effect for the repetition suppression condition. Error bars indicate ±1 standard error.

Examining the effect of expectation revealed a markedly different pattern of results. At 79 - 185 ms after the onset of the second stimulus in the pair, orientation-selectivity increased significantly ($p < .0001$) when the stimulus was unexpected relative to when it was expected, i.e., the prediction error was associated with an increased representation of the stimulus at the earliest stages of the brain's response to that stimulus. Moreover, the prediction error signal

contained enhanced information about the specific features of the stimulus that violated the expectation, in this case the orientation of the second grating. We conducted the same statistical tests on the three other parameters defining the Gaussian function (namely, the width, centre orientation and baseline) to determine how repetition suppression and expectation might affect other properties of the neural representation. There was no reliable influence of repetition suppression on any of these Gaussian parameters (all $p > .32$). For prediction error, there was a significant decrease in baseline activity over the same time period as observed for the increase in amplitude (79-185 ms, $p = .001$), but there were no significant effects for the other parameters (all *ps* > .30).

We followed up this initial analysis to ensure we did not miss any small effects of repetition suppression or expectation on all aspects of stimulus representation. We increased the signal-to-noise by averaging the stimulus reconstruction over this early time period (79-185 ms after stimulus presentation), and fitted Gaussians to each participant's data individually (Figure 4B). This again showed that the amplitude of the response was significantly ($t(14) = 3.34$, $p = .005$) higher for unexpected ($M = 0.67$, $SE = 0.06$) than for expected ($M = 0.41$, $SE = 0.03$) stimuli. By contrast, the width of the representations was similar for unexpected ($M = 29.62°$, $SE = 4.72°$) and expected ($M = 26.72°$, $SE = 2.74°$) stimuli, ($t(14) = 0.78$, $p = .45$). There was also a small, but non-significant ($t(14) = 1.94$, $p = .06$) trend for a smaller baseline response (i.e., non-orientation tuned activity) in the unexpected ($M = -0.01$, $SE = 0.07$) than in the expected ($M = 0.13$, $SE = 0.02$) condition. For comparison, we also averaged the same time period for the repetition suppression

conditions, and found similar curves for the repeated and alternating trials (*all ps* > .18). This analysis confirms the previous result, which employed more conservative nonparametric cluster-based testing.

We used a number of approaches to determine whether repetition suppression and expectation interacted to affect orientation selectivity. First, we took the difference scores between the combination of factors (e.g., expected repetition minus unexpected repetition, and expected alternation minus unexpected alternation) and compared these using the same cluster-based permutation testing outlined above. This analysis revealed no significant interactions between the factors for any parameter (all *ps* > .10). Second, we found the largest orientation-selectivity, defined by the amplitude of the fitted Gaussian, across the 600 ms following stimulus presentation. For each participant, this resulted in a single value for the four conditions. Each of these values was subjected to a two-way repeated-measures ANOVA, which again revealed no significant interaction between the factors (all *ps* > .30)

To further examine whether orientation-selectivity contained within the overall pattern of EEG activity differed for unexpected and expected stimuli, we used multivariate discriminant analysis to determine whether more traditional backward decoding produces the same pattern of results as that yielded by the forward encoding approach described above. The same cross-validation procedure was used as in the forward encoding approach, but accuracy was now defined as the proportion of trials labelled with the correct orientation. To facilitate comparison with the results with Kok et al., (2013), we took the peak classification accuracy

within a 600 ms window after presentation of the second grating of the pair. This analysis confirmed the results of the forward encoding: orientations shown in unexpected trials were classified better than orientations shown in expected trials ($F(1,14)$ 76.42, p <. 00001). Again, there was no effect of repetition on classification accuracy ($F(1,14) = 0.027$, p = .87) nor a significant interaction ($F(1,14) = 2.52$, p = .13). This suggests the finding is not specific to the analysis method but rather reflects how expectation affects the representation of sensory information in general.

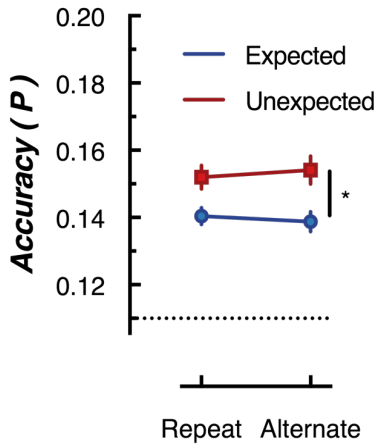


Figure 5. Peak (naive Bayes) classification accuracy of the presented grating orientation for expected and unexpected conditions. The dotted line indicates chance performance (1/9 orientations). The error bars indicate ±1 standard error of the mean.

Prediction error affects the temporal stability of stimulus representations

Next, we examined whether repetition suppression and prediction error affected dynamic, ongoing stimulus representations by using cross-temporal generalization (King & Dehaene, 2014; King, Gramfort, Schurger, Naccache, & Dehaene, 2014; Myers et al., 2015; Spaak, Watanabe, Funahashi, & Stokes, 2017; Stokes et al., 2013). To do this, we used the same forward encoding approach as in

the previous analysis, but now the weights were derived from one time-point on one set of trials, and then applied at every time point in the test trials. Again, a cross-validation approach was used, with all trials serving as both training and test. This analysis examined whether the same spatial pattern of EEG activity that allows for orientation selectivity generalizes to other time points, thus revealing whether the pattern of orientation-selective activity is stable or dynamic over time.

As shown in Figure 6, the best orientation selectivity was on-axis (training time equals test time) between 100 ms and 300 ms after stimulus presentation, suggesting that the stimulus representation changed dynamically over time (King & Dehaene, 2014). There was also significant off-axis orientation-selectivity from 100-500 ms after stimulus presentation, suggesting that some aspects of the neural representation of orientation were stable over time.

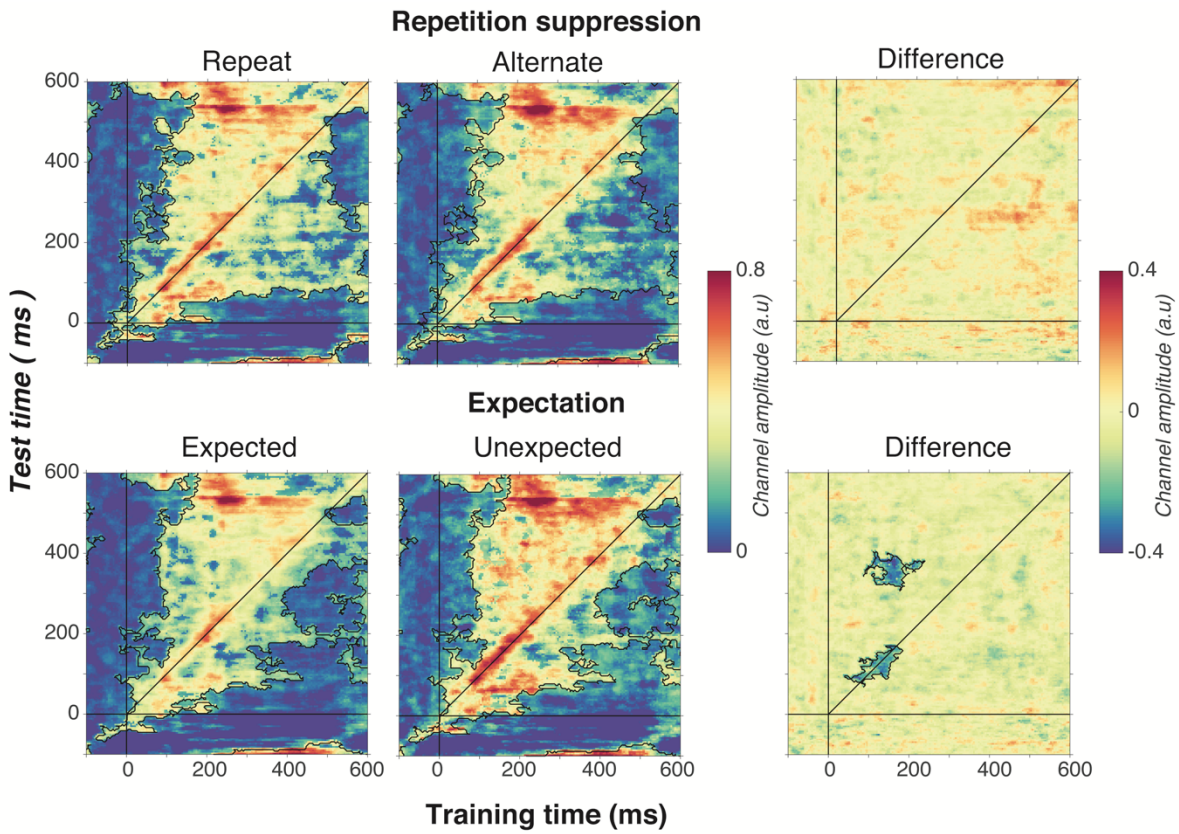


Figure 6. Cross-temporal generalization of the forward encoding model based on grating orientations for the main effects of repetition suppression (upper panels) and prediction error (lower panels). The maps have been thresholded (indicated by opacity) to show clusters (black outlines) of significant orientation selectivity (permutation testing, cluster threshold p < .05, corrected cluster statistic *p* < .05, 5,000 permutations). The difference between the conditions is shown in the right-hand column (permutation testing, cluster threshold *p* < .05, corrected cluster statistic *p* < .05). Opacity and outlines indicate significant differences.

There was no effect of repetition suppression on temporal generalization of orientation information (upper panels of Figure 6), suggesting that repetition suppression did not affect the temporal stability of neural representations of orientation. Examining the effect of expectation on cross-temporal generalization confirmed that there was significantly more on-axis orientation selectivity when the stimulus was unexpected than when it was expected (cluster *p* = .02). This increased on-axis orientation selectivity generalized off-axis at around 300-400 ms after stimulus onset (cluster *p* = .01). This finding suggests that the same

representation that is activated to process the prediction error is reactivated later as the stimulus continues to be processed. Such a signal could constitute the prior of the prediction, as this should be updated on the basis of incoming sensory evidence, which in turn would likely require reactivation of the unexpected stimulus.

Representation of the expected feature is activated even when it is not presented

In a final set of analyses, we investigated how expectation affects the interpretation of incoming sensory information, one of the key questions of predictive coding theory (Friston, 2005; Rao & Ballard, 1999). To do this, we focused on trials in which the orientation of the first Gabor within a pair did not match the orientation of the second Gabor (i.e., the alternating trials), and examined EEG responses to the second stimulus (see Figure 7). Here, we trained the forward encoding model on the orientation of the first Gabor in the pair to determine how the interaction between the expectation and the incoming sensory information was compared. We anticipated a lower expectation that the orientation would repeat for an expected alternating trial, relative to trials in the unexpected alternation condition, where the repeat should have been expected. We did not combine data across the repeat and alternating conditions because the orientations were now different upon presentation of the second Gabor in the pair.

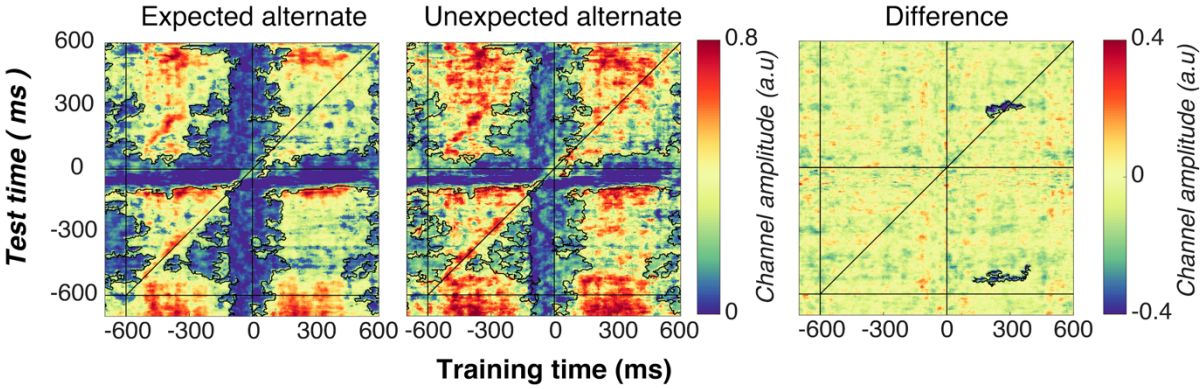


Figure 7. Cross-temporal generalization of the forward encoding model for the orientations presented in first grating relative to the onset of second grating (first grating was presented at -600 ms). The orientation in the second grating period is randomized but there are still periods for the unexpected alternation in which there is significant orientation selectivity. These maps have been thresholded (indicated by opacity) to show clusters (black outlines) of significant orientation selectivity (permutation testing, cluster threshold $p < .05$, corrected cluster statistic $p < .05$, 5,000 permutations). The opacity in the right-hand column indicates areas of statistically significant difference between the two conditions (permutation testing (n = 5,000), cluster threshold p < .05, corrected cluster statistic $p < .05$).

Inspecting the plot for the difference between expected and unexpected trials provides clues as to how this process might occur. As expected, for the time that first grating was presented there was significant on- and off-axis encoding for first grating orientation, with little difference between expectation conditions. More interestingly, in the unexpected alternation condition there was significantly better on-axis orientation selectivity for the first grating orientation between 150 and 300 ms after the onset of second grating, relative to the expected alternation condition (cluster p = .03; upper right quadrant of right panel in Figure 7). Furthermore, over the same time period, there was significantly better off-axis generalization of the first grating representation to the second grating epoch (cluster p = .03; lower right quadrant of right panel in Figure 7). This is confirmed by the occurrence of significantly increased off-axis orientation selectivity between training time of -500 ms and +300 ms. The time course of this effect is consistent with our earlier

findings, as the unexpected condition yielded better orientation selectivity for the second grating orientation between 79 and 150 ms after the second Gabor presentation. An intriguing possibility is that the relevant neural circuits re-activate the representation of the expected stimulus when faced with unexpected sensory signals.

Discussion

Our findings demonstrate that repetition suppression and expectation have distinct effects on neural representations of simple visual stimuli. We found that repetition suppression had no effect on orientation selectivity, even though the neural response to repeated stimuli was significantly decreased over occipital-parietal areas. Unexpected stimuli, on the other hand, showed significantly increased orientation-selectivity at an earlier time-point after stimulus onset than expected stimuli. This same early representation of the unexpected stimulus appeared to be reactivated at 200-300 ms after the initial neural response, supporting the idea that sensory expectations may be updated through comparison with incoming sensory evidence. These results suggest that repetition suppression and prediction error are separable and independent neural computations.

The present work provides a significant advance in our understanding of how predictive coding allows the brain to process incoming sensory information by comparing what is expected with what actually occurs. How expectations affect neural responses has been extensively investigated using mismatch negativity paradigms in which an unexpected stimulus causes a significantly larger neural response than an expected stimulus (Bekinschtein et al., 2009; Garrido et al., 2009;

Näätänen et al., 2007). This increased response to an unexpected stimulus has often been attributed to the generation of a prediction error that updates the expectation based on a conflict between the sensory evidence and the prior (Garrido et al., 2009). By contrast, our results suggest that the increased response may reflect two distinct processes that occur at separate times. Incoming sensory information is first evaluated against the prior (which occurs very early after stimulus presentation). When an unexpected stimulus is detected and generates a prediction error, the representation is amplified through gain enhancement. Later, around 300 ms after stimulus presentation, this same representation is reactivated to update the expectation against the initially predicted representation. Our findings thus provide a novel insight into how predictive coding might change neural representations of sensory information.

Multivariate pattern classification methods have been widely used to examine BOLD activity (Haxby, Connolly, & Guntupalli, 2014; Kamitani & Tong, 2005; 2006; Kriegeskorte et al., 2008), but have only recently been applied to EEG and MEG data (Cichy, Ramirez, & Pantazis, 2015; Garcia et al., 2013; King & Dehaene, 2014; King et al., 2013; 2014; 2016; Myers et al., 2015; Wolff et al., 2017). These kinds of analyses permit non-invasive imaging approaches to move beyond the measurement of overall levels of neural activity, and instead to uncover subtle patterns of neural activity associated with specific stimuli or task demands, which are more directly comparable with those obtained from invasive neurophysiological recordings of neuronal activity (Constantinescu, O'Reilly, & Behrens, 2016; Doeller, Barry, & Burgess, 2010; Kamitani & Tong, 2006; Sprague & Serences, 2013). Such

model-based approaches are especially appealing since changes in overall activity levels could be consistent with many different variations in the nature of the underlying neural representations. For instance, sharpening and gain reduction might both reduce the magnitude of an overall neural response but yield very different effects on the amount of information conveyed by those representations.

The current work applied these model-based approaches to EEG data to determine how prediction and repetition suppression affect neural representations of perceptual information. We chose to use EEG so we could recover the temporal dynamics of these effects – something that would not be possible with the BOLD signal used in fMRI – and because EEG is the most widely-used tool for measuring expectation effects in human participants (see Garrido:2009dp and Paavilainen, 2013 for review), thus facilitating comparison of our findings with those of other studies. We estimated orientation-selectivity using all EEG electrodes distributed across the scalp for two principal reasons. First, we wanted to limit experimenter degrees of freedom (Simmons, Nelson, & Simonsohn, 2011) potentially introduced through the post-hoc selection of subsets of electrodes. Second, given the broad spatial resolution of EEG, we reasoned that activity recorded from electrodes at any given scalp location could potentially carry important feature-selective information from a number of neural sources. The results revealed that orientation-selective information appears largely driven by electrodes over occipital-parietal regions (Figure 3D), consistent with a number of previous studies that employed visual decoding of M/EEG data (Cichy et al., 2015; Cichy, Pantazis, & Oliva, 2014; Stokes, Wolff, & Spaak, 2015). Having said this, and as noted above, it is entirely possible

that the effects we observed here arose from sources well beyond the occipital and parietal regions, potentially even including outside the visual cortical hierarchy. Future studies using invasive methods will be needed to directly measure neuronal responses across the brain to pinpoint the precise locations of these effects.

Surprisingly few studies have used invasive recording methods to examine how predictive coding affects stimulus representations at the neuronal level. One study in macaques (Kaliukhovich & Vogels, 2010) used the same design as the one introduced by Summerfield and colleagues, but using high-level objects (fractals and real-world objects) as stimuli and found that expectation did not attenuate repetition suppression in either spiking activity or local field potentials within the inferior temporal cortex. A later fMRI study in humans (Kovács, Kaiser, Kaliukhovich, Vidnyánszky, & Vogels, 2013) used a similar stimulus set, and also found no attenuation of repetition suppression by expectation in the same cortical region. A follow-up study provided a potential explanation for these findings by showing that the attenuation of neural responses associated with repetition suppression are found with familiar stimuli, but not with unfamiliar stimuli (Grotheer & Kovács, 2014). Viewed in this light, the stimulus sets used by Kaliukhovich and Vogels (2010) might not have been sufficiently familiar to yield effects of expectation in their animals. In any case, further work will need to be undertaken to determine whether expectation effects dependent on stimulus familiarity.

Other work has shown that context plays an important role in determining the magnitude of neuronal responses to sensory events. Thus, for example, (Ulanovsky, Las, & Nelken, 2003) found that rare auditory stimuli generate

significantly larger responses in primary auditory cortical neurons than more commonly occurring stimuli. This result has been interpreted as a single-neuron analogue of the mismatch negativity, but the design used in the study did not control for adaptation effects, thus making it difficult to draw an unambiguous comparison with the current work. In the visual domain, oddball stimuli have also been found to modulate neuronal activity in rats, characterised by an enhancement of responses in the higher-order latero-intermediate area (Vinken, Vogels, & Op de Beeck, 2017). Morever, Fiser et al. (2016) found that neurons in mouse primary visual cortex show a greater calcium response when task-irrelevant visual stimuli that had been presented during training were omitted, suggesting that an established expectation had been violated. This result is consistent with the literature on the mismatch negativity, in which the omission of an expected stimulus results in a large prediction error (Garrido et al., 2009; Wacongne et al., 2011). In non-human primates, neurons in the inferior temporal cortex show an enhanced response to unexpected relative to expected stimuli (Kaposvari, Kumar, & Vogels, 2018), and population decoding accuracy is higher for unexpected compared with expected stimuli (Kumar, Kaposvari, & Vogels, 2017).

Unlike the effects of expectation, there is a large body of electrophysiological work showing that sensory adaptation influences neuronal activity (Adibi et al., 2013b; Adibi, Clifford, & Arabzadeh, 2013a; Felsen et al., 2002; Kohn & Movshon, 2004; Patterson, Wissig, & Kohn, 2013) . For instance, there is a sharpening of stimulus selectivity in MT neurons following 40 s of adaptation to a drifting grating (Kohn & Movshon, 2004). As we have highlighted, however, prolonged adaptation is

likely also associated with a significant prediction that the next stimulus will be the same as the previous one. Perhaps more relevant to the current results, Patterson et al. (2013) found that the width of orientation tuning in V1 is only marginally sharpened following brief (400 ms) periods of adaptation. Again, however, their study did not control for expectation, so it is impossible to determine the role of predictive coding in their observations. Our finding that repetition suppression did not affect the bandwidth of orientation selectivity measured using EEG is also consistent with models of orientation adaptation based on human psychophysical data, which suggest that adaptation does not affect the tuning width of the adapted neural populations (Clifford, 2002; 2014; Dickinson, Almeida, Bell, & Badcock, 2010; Dickinson, Morgan, Tang, & Badcock, 2017; Tang, Dickinson, Visser, & Badcock, 2015).

We found that prediction errors were associated with increased gain in early stimulus representations, but did not affect the sharpness of their tuning. This lack of sharpening is in contrast to the findings of a previous study (Kok et al., 2012), in which a high-level prediction error led to ‘sharper’ multivariate decoding for expected versus unexpected visual stimuli. In their study, Kok et al. (2012) used an auditory tone to cue the orientation of a subsequent visual stimulus, and found significantly less off-label classification accuracy for predicted than for unpredicted stimuli. They concluded that predictions cause sharpening of stimulus representations. More recently, using the same task combined with a forward encoding approach, Kok et al. (2017) showed that response gain is increased for a predicted stimulus (Kok, Mostert, & de Lange, 2017).

Repetition suppression and prediction error

It is natural to ask why the results of the current study differ from those of Kok and colleagues outlined above. One possible explanation lies in the different approaches used to generate expectations across the studies. Specifically, whereas Kok et al. manipulated expectations by pairing an auditory cue with a visual stimulus, we exploited the properties of the visual stimuli themselves (i.e., their orientation) to generate expectations within blocks of trials. An intriguing possibility is that predictions requiring integration of sensory events from two or more modalities lead to increased gain, whereas predictions made within a single sensory modality lead to decreased gain. This might in turn relate to the noted differences between simple 'local' and higher-order 'global' type predictions (Bekinschtein et al., 2009; King et al., 2014), which lead to distinct patterns of stimulus-selective decoding. A similar discrepancy relating to the effects of attention on sensory representations has been widely discussed, with some studies finding sharpening of stimulus representations with attention, and others showing gain enhancement (Liu, Larsson, & Carrasco, 2007; Maunsell, 2015; Maunsell & Treue, 2006; Treue & Trujillo, 1999). *The differences between these results may potentially have arisen because the tasks relied upon different types of attention (e.g., spatial versus feature-based).* Future studies could determine whether this same divergence occurs for prediction effects.

In summary, we have shown that repetition suppression and expectation differentially affect the neural representation of simple, but fundamental, sensory features. Our results further highlight how the context in which a stimulus occurs, not just its features, affect the way it is represented by the brain. Our findings

suggest encoding priority through increased gain may be given to unexpected events which could potentially speed responses. This prioritized representation is then re-activated at a later time period, supporting the idea that feedback from higher cortical areas reactivates an initial sensory representation in early cortical areas.

Method

Participants

A group of 15 healthy adult volunteers (9 females, median age = 20.5 yr, range = 18 to 37 yr) participated in exchange for partial course credit or financial reimbursement (AUD$20/hr). We based our sample size on work that investigated the interaction between repetition suppression and prediction error (Summerfield et al., 2008), and that used forward encoding modelling to investigate orientation selectivity using MEG with a comparable number of trials as the current study (Myers et al., 2015). Each person provided written informed consent prior to participation, and had normal or corrected-to-normal vision. The study was approved by The University of Queensland Human Research Ethics Committee and was in accordance with the Declaration of Helsinki.

Experimental setup

The experiment was conducted inside a dimly illuminated room with the participants seated in a comfortable chair. The stimuli were displayed on a 22-inch LED monitor (resolution 1920 x 1080 pixels, refresh rate 120 Hz) using the PsychToolbox presentation software (Brainard, 1997; Pelli, 1997) for MATLAB (v7.3). Viewing distance was maintained at 45 cm using a chinrest, meaning the

screen subtended 61.18° x 36.87° (each pixel 2.4’ x 2.4’).

Task

The stimuli were Gabors (diameter: 5°, spatial frequency: 2 c/°, 100% contrast) presented centrally in pairs for 100 ms, separated by 500 ms (600 ms stimulus onset asynchrony) with a variable (650 to 750 ms) inter-stimulus interval between trials. Across the trials, the orientations of the Gabors were evenly spaced between 0° and 160° (in 20° steps) so we could reconstruct orientation selectivity contained within the EEG response using forward encoding modelling. The relationship of the orientations of the pairs Gabors was also used to construct the different repetition suppression and prediction conditions. The orientation presented in the second Gabor in the pair could either repeat or alternate with respect to the orientation of the first Gabor. In the alternation trials, the orientation of the first Gabor was drawn randomly, without replacement, from an even distribution of orientations that was different to the orientation of the second Gabor. To vary the degree of prediction, in half of the blocks 80% of the trials had repeated orientations and 20% of the trials had alternating orientations, whereas in the other half of the blocks these contingencies were reversed. This design allowed us to separately examine the effects of repetition suppression and prediction because of the orthogonal nature of the blocked design. The blocks of 135 trials (~3 mins) switched between the expectation of a repeating or alternating pattern, with the starting condition counterbalanced across participants.

The participants’ task was to monitor the visual streams for rare, faintly coloured (red or green) Gabors and to discriminate the colour as quickly and

accurately as possible. Any trial with a coloured target was excluded from analysis. The orientation match between the pairs was made to be consistent with the dominant contingency (i.e., repeated or alternating) within that block. Pilot testing was used prior to the main experiment to set the task at approximately threshold, to ensure that participants focused exclusively on the colour-discrimination task rather than the orientation contingencies associated with prediction and repetition. Only one participant reported being aware of the changing stimulus contingencies across the blocks when asked at the end of the experiment, and excluding this participant's data had no effect on the key results reported here. Self-paced breaks were provided between each of the 20 blocks within a session, at which time feedback was provided on performance in the preceding block. Each participant completed two sessions of 2700 trials each (5400 trials in total), with each session lasting around 70 mins of experimental time and 45 mins of EEG setup.

EEG acquisition and pre-processing

Continuous EEG data were recorded using a BioSemi Active Two system (BioSemi, Amsterdam, Netherlands). The signal was digitised at 1024 Hz sampling rate with a 24-bit A/D conversion. The 64 active scalp Ag/AgCl electrodes were arranged according to the international standard 10–20 system for electrode placement (Oostenveld & Praamstra, 2001) using a nylon head cap. As per BioSemi system design, the common mode sense and driven right leg electrodes served as the ground, and all scalp electrodes were referenced to the common mode sense during recording.

Offline EEG pre-processing was performed using EEGLAB in accordance

with best practice procedures (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015; Keil et al., 2014). The data were initially down-sampled to 256 Hz and subjected to a 0.5 Hz high-pass filter to remove slow baseline drifts. Electrical line noise was removed using *clean_line.m,* and *clean_rawdata.m* in EEGLAB (Delorme & Makeig, 2004) was used to remove bad channels (identified using Artifact Subspace Reconstruction), which were then interpolated from the neighbouring electrodes. Data were then re-referenced to the common average before being epoched into segments around each stimulus pair (-0.5 s to 1.25 s from the first stimulus in the pair). Systematic artefacts from eye blinks, movements and muscle activity were identified using semi-automated procedures in the SASICA toolbox (Chaumon, Bishop, & Busch, 2015) and regressed out of the signal. After this stage, any trial with a peak voltage exceeding ±100 *uV* was excluded from the analysis. The data were then baseline corrected to the mean EEG activity from -100 to 0 ms before the presentation of the second Gabor in the pair.

Design

We used a modified version of a factorial design that has previously been used to separately examine the effects of repetition suppression and prediction error (Kaliukhovich & Vogels, 2010; Kovács et al., 2013; Summerfield et al., 2008; 2011; Todorovic et al., 2011; Todorovic & de Lange, 2012). By comparing the two repeat conditions with the two alternating conditions, we could examine repetition suppression while controlling for different levels of expectation. Conversely, by comparing across the expected and unexpected trials, we could examine prediction error while controlling for repetition suppression.

Forward encoding modelling

We used a forward encoding approach to estimate the amount of orientation-selective information contained in the EEG data at each time point of the trial. This approach differs from standard decoding approaches by modelling each presented orientation as a continuous variable of a set of tuned orientation-selective channels. The forward-encoding technique has been successfully used to reconstruct colour (Brouwer & Heeger, 2009), spatial (Sprague & Serences, 2013) and orientation (Ester, Sutterer, Serences, & Awh, 2016) selectivity in fMRI data. More recently it has been extended to orientation (Garcia et al., 2013; Kok et al., 2017; Myers et al., 2015; Wolff et al., 2017) and space (Foster, Sutterer, Serences, Vogel, & Awh, 2016) encoding of MEG and EEG data, which allows for far greater temporal resolution than fMRI.

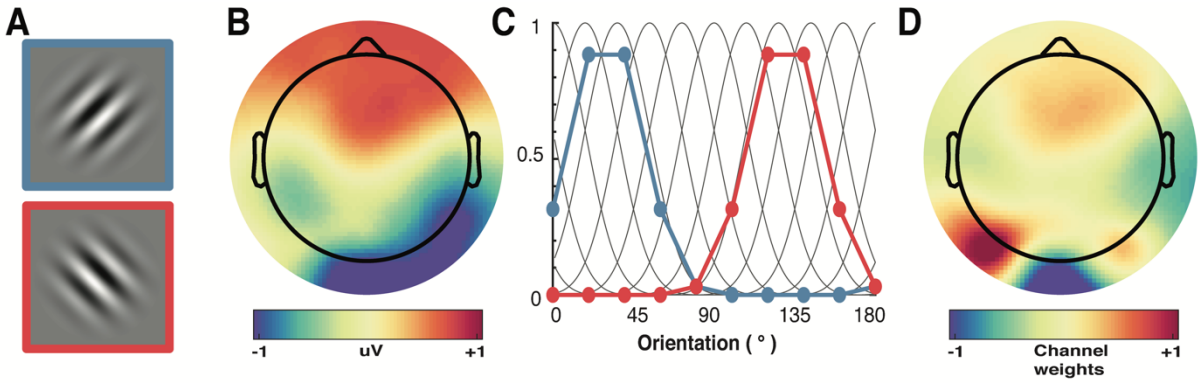


Figure 8. A schematic of the forward-encoding approach applied to EEG activity. **(A)** Participants viewed individual gratings at fixation, each with a specific orientation. **(B)** Neural activity evoked by each grating was measured over the entire scalp. **(C)** Evoked neural responses were convolved with canonical orientation-selective functions (grey lines in C) to determine coefficients for the different orientations (coloured dots and lines, which match the colours of the outlined gratings in A). These coefficients were then use to generate a regression matrix. **(D)** General linear modelling was used on a subset of training trials to generate weights for each channel. These weights were inverted and simultaneously applied to an independent test set of data to recover orientation selectivity in the EEG activity. As EEG activity has high temporal resolution we can apply the procedure to many epochs following stimulus presentation to determine the temporal dynamics of orientation processing (see Figure 3).

We applied forward encoding modelling to determine how repetition suppression and prediction error affected orientation selectivity. To do this, the second orientation (Figure 8A) in the Gabor pair in each trial was used to construct a regression matrix, with separate regressors for the 9 orientations used across the experiment. This regression matrix was convolved with a set of basis functions (half cosines raised to the 8th power (Figure 8C), which allowed complete and unbiased coverage of orientation space) to allow us to pool similar information patterns across nearby orientations (Brouwer & Heeger, 2009). We used this tuned regression matrix to estimate time-resolved orientation selectivity contained within the EEG activity in a 16 ms sliding window, in 4 ms steps (Figure 8B; Myers et al., 2015). To avoid overfitting, we used a leave-one-out cross-validation procedure

where the regression weights were estimated for a training set and applied to an independent test set (Figure 8D). All trial types (including target trials) were used in training and test sets. This was done by solving the linear equation:

$$B_1 = WC_1 \quad (1)$$

Where B_1 (64 sensors x N training trials) is the electrode data for the training set, C_1 (9 channels x N training trials) is the tuned channel response across the training trials, and W is the weight matrix for the sensors we want to estimate (64 sensors x 9 channels). W can be estimated using least square regression to solve equation (2):

$$W = (C_1\, C_1^T)^{-1}\, C_1^T\, B_1 \quad (2)$$

The channel response in the test set C_2 (9 channels x N test trials) was estimated using the weights in (2) and applied to activity in B_2 (64 sensors x N test trials).

$$C^2 = (W\, W^T)\, W^T\, B^2 \quad (3)$$

We repeated this process by holding one trial out as test, and training on the remaining trials until all trials had been used in test and training. The procedure was repeated for each trial within the trial epoch. We then shifted all trials to a common orientation, meaning that 0° corresponded to the orientation presented on each trial. The reconstructed channel activations were separated into the four conditions, and averaged over trials. These responses were then smoothed with a Gaussian kernel with a 16 ms window, and fitted with a Gaussian function (4) using non-linear least square regression to quantify the amount of orientation selective activity.

$$G(x) = A\, exp(-\frac{(x-\varphi)^2}{2\sigma^2}) + C \quad (4)$$

Where *A* is the amplitude representing the amount of orientation selective activity, is the orientation the function is centred on (in degrees), is the width (degrees) and

C is a constant used to account for non-orientation selective baseline shifts.

Multivariate pattern analysis

We conducted a multivariate pattern analysis to build upon the initial forward encoding results which showed that unexpected stimuli elicit greater orientation selectivity than expected stimuli. This analysis used the same data as the forward encoding analysis. We used the *classify* function from *Matlab 2017a* with the 'diaglinear' option to implement a Naive Bayes classifier. For each time point, we used the same cross-validation procedure as the forward encoding modelling with the same averaging procedure to select train and test sets of data. The classifier was given the orientations of the training data and predicted the orientation of the test data. A trial was labelled correct if the presented orientation was produced. To facilitate comparison of the results with those of (Kok et al., 2012), we found the peak classification accuracy for each participant in the 600 ms following stimulus presentation.

Statistical testing

A non-parametric sign permutation test was used to determine the null distribution for testing (Wolff et al., 2017). This method makes no assumptions about the underlying shape of the null distribution**.** This was done by randomly flipping the sign of the data for the participants with equal probability. Fifty thousand (50,000) permutations were used for the time-series data, whereas only 5000 were used for the temporal generalization plots because of the significantly greater computational demands.

Cluster-based non-parametric correction (50,000 permutations for timeseries

and 5,000 for temporal generalization) was used to account for multiple comparisons, and determined whether there were statistical differences between the contrasting conditions. Paired-samples t-tests were used to follow up the analysis in Figure 4 within a specified time window, and no correction was applied. A two-way repeated measures ANOVA (implemented using GraphPad Prism 7.0c, La Jolla California, USA) was used to analyse the multivariate pattern analysis results shown in Figure 5.

Acknowledgements

This work was supported by the Australian Research Council (ARC) Centre of Excellence for Integrative Brain Function (ARC Centre Grant CE140100007) to JBM and EA, and by an ARC Discovery Project (DP170100908) to EA. JBM was supported by ARC Australian Laureate Fellowship (FL110100103).

References

Adibi, M., McDonald, J. S., Clifford, C. W. G., & Arabzadeh, E. (2013). Adaptation improves neural coding efficiency despite increasing correlations in variability. *The Journal of Neuroscience*, *33*(5), 2108–2120. http://doi.org/10.1523/JNEUROSCI.3449-12.2013

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

Barron, H. C., Garvert, M. M., & Behrens, T. E. J. (2016). Repetition suppression: a means to index neural representations using BOLD? *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, *371*(1705), 20150355. http://doi.org/10.1098/rstb.2015.0355

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677. http://doi.org/10.1073/pnas.0809667106

Benucci, A., Saleem, A. B., & Carandini, M. (2013). Adaptation maintains population homeostasis in primary visual cortex. *Nature Neuroscience*, *16*(6), 724–729. http://doi.org/10.1038/nn.3382

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*, B153. http://doi.org/10.3389/fninf.2015.00016

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. http://doi.org/10.1163/156856897X00357

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and reconstructing color from responses in human visual cortex. *The Journal of Neuroscience*, *29*(44), 13992–14003. http://doi.org/10.1523/JNEUROSCI.3577-09.2009

Buckner, R. L., Goodman, J., Burock, M., Rotte, M., Koutstaal, W., Schacter, D., et al. (1998). Functional-Anatomic Correlates of Object Priming in Humans Revealed by Rapid Presentation Event-Related fMRI. *Neuron*, *20*(2), 285–296. http://doi.org/10.1016/S0896-6273(00)80456-0

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods*, *250*, 47–63. http://doi.org/10.1016/j.jneumeth.2015.02.025

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Publishing Group*. http://doi.org/10.1038/nn.3635

Cichy, R. M., Ramirez, F. M., & Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *NeuroImage*, *121*, 193–204. http://doi.org/10.1016/j.neuroimage.2015.07.011

Clifford, C. W. G. (2002). Perceptual adaptation: motion parallels orientation. *Trends in Cognitive Sciences*, *6*(3), 136–143.

Clifford, C. W. G. (2014). The tilt illusion: phenomenology and functional implications. *Vision Research*, *104*, 3–11. http://doi.org/10.1016/j.visres.2014.06.009

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468. http://doi.org/10.1126/science.aaf0941

Cui, Y., Wang, Y. V., Park, S. J. H., Demb, J. B., & Butts, D. A. (2016). Divisive suppression explains high-precision firing and contrast adaptation in retinal ganglion cells. *eLife*, *5*. http://doi.org/10.7554/eLife.19460

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. http://doi.org/10.1016/j.jneumeth.2003.10.009

Dickinson, J. E., Almeida, R. A., Bell, J., & Badcock, D. R. (2010). Global shape aftereffects have a local substrate: A tilt aftereffect field. *Journal of Vision*, *10*(13), 5–5. http://doi.org/10.1167/10.13.5

Dickinson, J. E., Morgan, S. K., Tang, M. F., & Badcock, D. R. (2017). Separate banks of information channels encode size and aspect ratio. *Journal of Vision*, *17*(3), 27. http://doi.org/10.1167/17.3.27

Diederen, K. M. J., Spencer, T., Vestergaard, M. D., Fletcher, P. C., & Schultz, W. (2016). Adaptive Prediction Error Coding in the Human Midbrain and Striatum Facilitates Behavioral Adaptation and Learning Efficiency. *Neuron*, *90*(5), 1127–1138. http://doi.org/10.1016/j.neuron.2016.04.019

Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature*, *463*(7281), 657–661. http://doi.org/10.1038/nature08704

Ester, E. F., Sutterer, D. W., Serences, J. T., & Awh, E. (2016). Feature-Selective Attentional Modulations in Human Frontoparietal Cortex. *The Journal of Neuroscience*, *36*(31), 8188–8199. http://doi.org/10.1523/JNEUROSCI.3935-15.2016

Fairhall, A. L., Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, *412*(6849), 787–792. http://doi.org/10.1038/35090500

Foster, J. J., Sutterer, D. W., Serences, J. T., Vogel, E. K., & Awh, E. (2016). The topography of alpha-band activity tracks the content of spatial working memory. *Journal of Neurophysiology*, *115*(1), 168–177. http://doi.org/10.1152/jn.00860.2015

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 815–836. http://doi.org/10.1098/rstb.2005.1622

Garcia, J. O., Srinivasan, R., & Serences, J. T. (2013). Near-real-time feature-selective modulations in human cortex. *Current Biology*, *23*(6), 515–522. http://doi.org/10.1016/j.cub.2013.02.013

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. http://doi.org/10.1016/j.clinph.2008.11.029

Gotts, S. J., Chow, C. C., & Martin, A. (2012). Repetition priming and repetition suppression: A case for enhanced efficiency through neural synchronization. *Cognitive Neuroscience*. http://doi.org/10.1080/17588928.2012.691277

Gross, C. G., Schiller, P. H., Wells, C., & Gerstein, G. L. (1967). Single-unit activity in temporal association cortex of the monkey. *Journal of Neurophysiology*, *30*(4), 833–43.

Grotheer, M., & Kovács, G. (2014). Repetition probability effects depend on prior experiences. *The Journal of Neuroscience*, *34*(19), 6640–6646. http://doi.org/10.1523/JNEUROSCI.5326-13.2014

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, *37*(1), 435–456. http://doi.org/10.1146/annurev-neuro-062012-170325

Henson, R. N. A., & Rugg, M. D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologia*, *41*(3), 263–270. http://doi.org/10.1016/S0028-3932(02)00159-8

Kaliukhovich, D. A., & Vogels, R. (2010). Stimulus Repetition Probability Does Not Affect Repetition Suppression in Macaque Inferior Temporal Cortex. *Cerebral Cortex*, *21*(7), 1547–1558. http://doi.org/10.1093/cercor/bhq207

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, *8*(5), 679–685. http://doi.org/10.1038/nn1444

Kamitani, Y., & Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Current Biology*, *16*(11), 1096–1102. http://doi.org/10.1016/j.cub.2006.04.003

Kaposvari, P., Kumar, S., & Vogels, R. (2018). Statistical Learning Signals in Macaque Inferior Temporal Cortex. *Cerebral Cortex*, *28*(1), 250–266. http://doi.org/10.1093/cercor/bhw374

Kay, K. N., Naselaris, T., Prenger, R. J., & Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature*, *452*(7185), 352–355. http://doi.org/10.1038/nature06713

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*(1), 1–21. http://doi.org/10.1111/psyp.12147

Keller, A. J., Houlton, R., Kampa, B. M., Lesica, N. A., Mrsic-Flogel, T. D., Keller, G. B., & Helmchen, F. (2017). Stimulus relevance modulates contrast adaptation in visual cortex. *eLife*, *6*, e21589. http://doi.org/10.7554/eLife.21589

King, J. R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. http://doi.org/10.1016/j.tics.2014.01.002

King, J. R., Faugeras, F., Gramfort, A., Schurger, A., Karoui, El, I., Sitt, J. D., et al. (2013). Single-trial decoding of auditory novelty responses facilitates the detection of residual consciousness. *NeuroImage*, *83*, 726–738. http://doi.org/10.1016/j.neuroimage.2013.07.013

King, J.-R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two Distinct Dynamic Modes Subtend the Detection of Unexpected Sounds. *PloS One*, *9*(1), e85791. http://doi.org/10.1371/journal.pone.0085791

King, J.-R., Pescetelli, N., & Dehaene, S. (2016). Brain Mechanisms Underlying the Brief Maintenance of Seen and Unseen Sensory Information. *Neuron*, *92*(5), 1122–1134. http://doi.org/10.1016/j.neuron.2016.10.051

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience*, *7*(7), 764–772. http://doi.org/10.1038/nn1267

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, *75*(2), 265–270. http://doi.org/10.1016/j.neuron.2012.04.034

Kok, P., Mostert, P., & de Lange, F. P. (2017). Prior expectations induce prestimulus sensory templates. *Proceedings of the National Academy of Sciences of the United States of America*, *26*, 201705652. http://doi.org/10.1073/pnas.1705652114

Kourtzi, Z., & Kanwisher, N. (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. *Science*, *293*(5534), 1506–1509. http://doi.org/10.1126/science.1061133

Kovács, G., Kaiser, D., Kaliukhovich, D. A., Vidnyánszky, Z., & Vogels, R. (2013). Repetition probability does not affect fMRI repetition suppression for objects. *The Journal of Neuroscience*, *33*(23), 9805–9812. http://doi.org/10.1523/JNEUROSCI.3423-12.2013

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. http://doi.org/10.1016/j.neuron.2008.10.043

Kumar, S., Kaposvari, P., & Vogels, R. (2017). Encoding of Predictable and Unpredictable Stimuli by Inferior Temporal Cortical Neurons. *Journal of Cognitive Neuroscience*, *29*(8), 1445–1454. http://doi.org/10.1162/jocn_a_01135

Kvale, M. N., & Schreiner, C. E. (2004). Short-Term Adaptation of Auditory Receptive Fields to Dynamic Stimuli. *Journal of Neurophysiology*, *91*(2), 604–612. http://doi.org/10.1152/jn.00484.2003

Larsson, J., & Smith, A. T. (2012). fMRI repetition suppression: neuronal adaptation or stimulus expectation? *Cerebral Cortex*, *22*(3), 567–576. http://doi.org/10.1093/cercor/bhr119

Liu, T., Larsson, J., & Carrasco, M. (2007). Feature-Based Attention Modulates Orientation-Selective Responses in Human Visual Cortex. *Neuron*, *55*(2), 313–323. http://doi.org/10.1016/j.neuron.2007.06.030

Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E., & Diamond, M. E. (2007). Shifts in Coding Properties and Maintenance of Information Transmission during Adaptation in Barrel Cortex. *PLoS Biology*, *5*(2), e19. http://doi.org/10.1371/journal.pbio.0050019

Maunsell, J. H. R. (2015). Neuronal Mechanisms of Visual Attention. *Annual Review of Vision Science*, *1*(1), 373–391. http://doi.org/10.1146/annurev-vision-082114-035431

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences*, *29*(6), 317–322. http://doi.org/10.1016/j.tins.2006.04.001

Movshon, J. A., & Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature*, *278*(5707), 850–852.

Myers, N. E., Rohenkohl, G., Wyart, V., Woolrich, M. W., Nobre, A. C., Stokes, M. G., & Frank, M. J. (2015). Testing sensory evidence against mnemonic templates. *eLife*, *4*, e09000. http://doi.org/10.7554/eLife.09000

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*(4), 713–719.

Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, *88*(2), 109–123. http://doi.org/10.1016/j.ijpsycho.2013.03.015

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. http://doi.org/10.1038/4580

Rasmussen, R. G., Schwartz, A., & Chase, S. M. (2017). Dynamic range adaptation in primary motor cortical populations. *eLife*, *6*, e21409. http://doi.org/10.7554/eLife.21409

Rentzeperis, I., Nikolaev, A. R., Kiper, D. C., & Van Leeuwen, C. (2012). Relationship between neural response and adaptation selectivity to form and color: an ERP study. *Frontiers in Human Neuroscience*, *6*. http://doi.org/10.3389/fnhum.2012.00089

Salti, M., Monto, S., Charles, L., King, J.-R., Parkkonen, L., Dehaene, S., & Johansen-Berg, H. (2015). Distinct cortical codes and temporal dynamics for conscious and unconscious percepts. *eLife*, *4*, e05652. http://doi.org/10.7554/eLife.05652

Schacter, D. L., & Buckner, R. L. (1998). Priming and the Brain. *Neuron*, *20*(2), 185–195.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. http://doi.org/10.1177/0956797611417632

Smirnakis, S. M., Berry, M. J., Warland, D. K., Bialek, W., & Meister, M. (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, *386*(6620), 69–73. http://doi.org/10.1038/386069a0

Spaak, E., Watanabe, K., Funahashi, S., & Stokes, M. G. (2017). Stable and Dynamic Coding for Working Memory in Primate Prefrontal Cortex. *The Journal of Neuroscience*, *37*(27), 6503–6516. http://doi.org/10.1523/JNEUROSCI.3364-16.2017

Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, *16*(12), 1879–1887. http://doi.org/10.1038/nn.3574

Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., & Duncan, J. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, *78*(2), 364–375. http://doi.org/10.1016/j.neuron.2013.01.039

Stokes, M. G., Wolff, M. J., & Spaak, E. (2015). Decoding Rich Spatial Information with High Temporal Resolution. *Trends in Cognitive Sciences*, *19*(11), 636–638. http://doi.org/10.1016/j.tics.2015.08.016

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M. M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. http://doi.org/10.1038/nn.2163

Summerfield, C., Wyart, V., Mareike Johnen, V., & de Gardelle, V. (2011). Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with Expectation. *Frontiers in Human Neuroscience*, *5*. http://doi.org/10.3389/fnhum.2011.00067

Tang, M. F., Dickinson, J. E., Visser, T. A. W., & Badcock, D. R. (2015). The broad orientation dependence of the motion streak aftereffect reveals interactions between form and motion neurons. *Journal of Vision*, *15*(13), 4. http://doi.org/10.1167/15.13.4

Todorovic, A., & de Lange, F. P. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *The Journal of Neuroscience*, *32*(39), 13389–13395. http://doi.org/10.1523/JNEUROSCI.2227-12.2012

Todorovic, A., van Ede, F., Maris, E., & de Lange, F. P. (2011). Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *The Journal of Neuroscience*, *31*(25), 9118–9123. http://doi.org/10.1523/JNEUROSCI.1425-11.2011

Tootell, R. B., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., & Dale, A. M. (1998). Functional analysis of primary visual cortex (V1) in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 811–817. http://doi.org/10.1038/340386a0

Tootell, R., Reppas, J. B., Dale, A. M., & Look, R. B. (1995). Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature*.

Treue, S., & Trujillo, J. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*. http://doi.org/10.1038/21176

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*(4), 391–398. http://doi.org/10.1038/nn1032

Vinken, K., Vogels, R., & Op de Beeck, H. (2017). Recent Visual Experience Shapes Visual Processing in Rats through Stimulus-Specific Adaptation and Response Enhancement. *Current Biology*, *27*(6), 914–919. http://doi.org/10.1016/j.cub.2017.02.024

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors

in human cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(51), 20754–20759. http://doi.org/10.1073/pnas.1117807108

Wolff, M. J., Jochim, J., Akyürek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience*, *20*(6), 864–871. http://doi.org/10.1038/nn.4546