Prior expectations induce pre-stimulus sensory templates

Peter Kok[1,2,3]*, Pim Mostert[1] and Floris P. de Lange[1]

1. Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands
2. Princeton University, Princeton Neuroscience Institute, 301 Peretsman-Scully Hall, Princeton, NJ 08544
3. Yale University, Department of Psychology, 1 Prospect Street, New Haven, CT 06511

***Corresponding author**

Yale University
Department of Psychology
305 SSS Hall
1 Prospect Street, New Haven, CT 06511
Phone: 1 203 432 9268
E-mail: peter.kok@yale.edu

Manuscript information: This manuscript contains 163 (Abstract) + 108 (Significance Statement) + 3442 (Main text) words, 4 Figures and 6 Supplementary Figures.

Classification: Biological sciences – Neuroscience / Psychological and Cognitive Sciences

Keywords: prediction; perceptual inference; predictive coding; feature-based expectation; feature-based attention

Abstract

Perception can be described as a process of inference, integrating bottom-up sensory inputs and top-down expectations. However, it is unclear how this process is neurally implemented. It has been proposed that expectations lead to pre-stimulus baseline increases in sensory neurons tuned to the expected stimulus, which in turn affects the processing of subsequent stimuli. Recent fMRI studies have revealed stimulus-specific patterns of activation in sensory cortex as a result of expectation, but this method lacks the temporal resolution necessary to distinguish pre- from post-stimulus processes. Here, we combined human MEG with multivariate decoding techniques to probe the representational content of neural signals in a time-resolved manner. We observed a representation of expected stimuli in the neural signal shortly before they were presented, demonstrating that expectations indeed induce a pre-activation of stimulus templates. The strength of these pre-stimulus expectation templates correlated with participants' behavioural improvement when the expected feature was task-relevant. These results suggest a mechanism for how predictive perception can be neurally implemented.

Significance Statement

The way we perceive the world is partly shaped by what we expect to see at any given moment. However, it is unclear how this process is neurally implemented. Recently, it has been proposed that the brain generates stimulus templates in sensory cortex to pre-empt expected inputs. Here, we provide evidence that a representation of the expected stimulus is present in the neural signal shortly before it is presented, demonstrating that expectations can indeed induce the pre-activation of stimulus templates. Importantly, these expectation signals resembled the neural signal evoked by an actually presented stimulus, suggesting that expectations induce similar patterns of activations in visual cortex as sensory stimuli.

\body

Introduction

Perception is heavily influenced by prior knowledge (1–3). Accordingly, many theories cast perception as a process of inference, integrating bottom-up sensory inputs and top-down expectations (4–6). However, it is unclear how this integration is neurally implemented. It has been proposed that prior expectations lead to baseline increases in sensory neurons tuned to the expected stimulus (7–9), which in turn leads to improved neural processing of matching stimuli (10, 11). In other words, expectations may induce stimulus templates in sensory cortex, prior to the actual presentation of the stimulus. Alternatively, top-down influences in sensory cortex may exert their influence only after the bottom-up stimulus has been initially processed, and the integration of the two sources of information may become apparent only during later stages of sensory processing (12).

The evidence necessary to distinguish between these hypotheses has been lacking. fMRI studies have revealed stimulus-specific patterns of activation in sensory cortex as a result of expectation (9, 13), but this method lacks the temporal resolution necessary to distinguish pre- from post-stimulus periods. Here, we combined MEG with multivariate decoding techniques to probe the representational content of neural signals in a time-resolved manner (14–17). The experimental paradigm was virtually identical to the ones employed in our previous fMRI studies that studied how expectations modulate stimulus-specific patterns of activity in the primary visual cortex (9, 11). We trained a forward model to decode the orientation of task-irrelevant gratings from the MEG signal (18, 19), and applied this decoder to trials in which participants expected a grating of a particular orientation to be presented. This analysis revealed a neural representation of the expected grating that resembled the neural signal evoked by an actually presented grating. This representation was present already shortly before stimulus presentation, demonstrating that expectations can indeed induce the pre-activation of stimulus templates.

Results

Participants (n=23) were exposed to auditory cues that predicted the likely orientation (45° or 135°) of an upcoming grating stimulus (Fig. 1*A*-*B*). This grating was followed by a second grating that differed slightly from the first in terms of orientation and contrast. In separate runs of the MEG session, participants performed either an orientation or contrast discrimination task on the two gratings (see Materials and Methods for details).

Behavioural results. Participants were able to discriminate small differences in orientation (3.9° ± 0.5°, accuracy = 74.0% ± 1.6%, mean ± sem) and contrast (4.6% ± 0.3%, accuracy = 76.6% ± 1.5%) of the cued gratings. There was no significant difference between the two tasks in terms of either accuracy ($F_{1,22}$ = 3.38, p = 0.080) or reaction time (mean RT = 633 ms vs. 608 ms, $F_{1,22}$ = 2.89, p = 0.10). Overall, accuracy and reaction times were not influenced by whether the cued grating had the expected or the unexpected orientation (accuracy: $F_{1,22}$ = 0.21, p = 0.65; RT: $F_{1,22} < 0.01$, p = 0.93), nor was there an interaction between task and expectation (accuracy: $F_{1,22}$ = 0.96, p = 0.34; RT: $F_{1,22}$ = 0.09, p = 0.77). Note that these discrimination tasks were orthogonal to the expectation manipulation, in the sense that the expectation cue provided no information about the likely correct choice.

During the grating localiser (Fig. 1*C*, see Materials and Methods for details), participants correctly detected 91.2% ± 1.6% (mean ± sem) of fixation flickers, and incorrectly pressed the button on 0.2% ± 0.1% of trials, suggesting that participants were successfully engaged by the fixation task.

MEG results – Localiser orientation decoding. As mentioned, participants were exposed to auditory cues that predicted the likely orientation of an upcoming grating stimulus. The question we wanted to answer was whether the expectations induced by these auditory cues would evoke templates of the visual stimuli prior to the presentation of the gratings. To be able to uncover such sensory templates, we trained a decoding model to reconstruct the orientation of (task-irrelevant) visual gratings (Fig. 1*C*) from the MEG signal, in a time-resolved manner. First, we found that this model was highly accurate at reconstructing the orientation of such gratings from the MEG signal (Fig. 2). Grating orientation could be decoded across an extended period of time (from 40 to 655 ms post-stimulus, $p < 0.001$, and from 685 to 730 ms, $p = 0.018$), peaking around 120-160 ms post-stimulus (Fig. 2*C*). Furthermore, in the period around 100 to 330 ms post-stimulus, orientation decoding generalised across time, meaning that a decoder trained on the evoked response at, for example, 120 ms post-stimulus could reconstruct the grating orientation represented in the evoked response around 300 ms, and vice versa (Fig. 2*D*). In other words, certain aspects of the representation of grating orientation were sustained over time.

MEG results – Expectation induces stimulus templates. Our main question pertained to the presence of visual grating templates induced by the auditory expectation cues during the main experiment. Therefore, we applied our model trained on task-irrelevant gratings to trials containing gratings that were either validly or invalidly predicted, respectively (Fig. 3*A*). In both conditions, the decoding model trained on task-irrelevant gratings succeeded in accurately reconstructing the orientation of the gratings presented in the main experiment (valid expectation: cluster from training time 60 to 410 ms and decoding time 60 to 400 ms, $p < 0.001$, and from training time 205 to 325 ms and decoding time 400 to 495 ms, $p = 0.045$; invalid expectation: cluster from training time 75 to 225 ms and decoding time 75 to 330 ms, $p = 0.0012$, and from training time 250 to 360 ms and decoding time 195 to 355 ms, $p = 0.027$).

If the cues induced sensory templates of the expected grating, one would expect these to be revealed in the difference in decoding between valid and invalidly predicted gratings (see Material and Methods for details of the subtraction logic). Indeed, this analysis demonstrated that the auditory expectation cues induce orientation-specific neural signals (Fig. 3*A*, bottom panel). These signals were present already 40 ms before grating presentation, and extended into the post-stimulus period (from decoding time -40 to 230 ms, $p = 0.0092$, and from 300 to 530 ms, $p = 0.016$). Furthermore, these signals were uncovered when the decoder was trained on around 120 to 160 ms post-stimulus during the grating localiser (Fig. 3*B*), suggesting that these cue-induced signals were similar to those evoked by task-irrelevant gratings. In other words, the auditory expectation cues evoked orientation-specific signals that were similar to sensory signals evoked by the corresponding actual grating stimuli (Fig. S1*A*).

In sum, expectations induced pre-stimulus sensory templates that influenced post-stimulus representations as well; invalidly expected gratings had to 'overcome' a pre-stimulus activation of the opposite orientation, while validly expected gratings were facilitated by a compatible pre-stimulus activation (Fig. S1*B*). The post-stimulus carryover of these expectation signals lasted throughout the trial (Fig. S1*C*).

As in previous studies using a similar paradigm (11, 20), there was no interaction between the effects of the expectation cue and the task (orientation vs. contrast discrimination) participants performed (no clusters with $p < 0.05$; Fig. S2*A*). In other words, expectations evoked pre-stimulus orientation signals to a similar degree in both tasks (Fig. S2*B*). This suggests that influences of expectation on neural representations are relatively independent of the task-relevance of the expected feature, in line with our previous fMRI study (11). Note though that, unlike in that study, there was no significant modulation of the orientation signal by task-relevance (no clusters with $p < 0.05$, Fig. S2*A*). The reason for this lack of difference is unclear, although it should be noted that there was a trend towards participants having higher accuracy and faster reaction times (see above) on the contrast task than on the orientation task. This may suggest the two tasks were not optimally balanced in terms of difficulty, precluding a proper comparison of the effect of task set in the current study.

In our previous fMRI study, we found a relationship between the effects of expectation on neural stimulus representations and performance on the orientation discrimination task. Specifically, participants for whom valid expectations led to the largest improvement in neural stimulus representations, also showed the strongest benefit of valid expectations on behavioural performance during the orientation discrimination task (11). This relationship was absent for the contrast discrimination task, when grating orientation was task-irrelevant. The current study allowed us to test for a similar relationship, with an important extension: here, we could test whether neural *pre-stimulus* expectation signals are related to behavioural performance improvements. We quantified the decoding of the expected orientation just before grating presentation (-50 to 0 ms, training window 120 to 160 ms) and correlated this with the difference in task accuracy for valid and invalid expectation trials, across participants. This analysis revealed that participants with a stronger pre-stimulus reflection of the expected orientation in their neural signal also had a greater benefit from valid expectations on performance on the orientation task (r = 0.44, p = 0.035; Fig. 4, left panel). No such relationship was found for the contrast task, where the orientation of the gratings was not task-relevant (r = -0.13, p = 0.55; Fig. 4, right panel). This is exactly the pattern of results we found in our previous fMRI study, but with the important extension that it is the *pre-stimulus* expectation effect that is correlated with behavioural performance, whereas the previous study did not have the temporal resolution to distinguish pre- from post-stimulus signals.

In the current study, neural orientation signals were probed by applying a forward model that takes the noise covariance between MEG sensors into account (see SI Materials and Methods for details). This model was superior to a forward model that did not correct for the noise covariance (Fig. S3), suggesting that feature covariance is an important factor to take into account when applying multivariate methods to MEG data. Corroborating this notion, a two-class decoder that corrected for noise covariance (16) was able to reproduce our effects of interest (Fig. S4), demonstrating that the expectation effects do not depend on a specific analysis technique, as long as the covariance between MEG sensors is taken into account.

Finally, there was no difference in the overall amplitude of the neural response evoked between validly and invalidly expected gratings (no clusters with $p < 0.4$, Fig. S5).

Discussion

Here, we show that expectations can induce sensory templates of the expected stimulus already before the stimulus appears. These results extend previous fMRI studies demonstrating stimulus-specific patterns of activation in sensory cortex induced by expectations, but which could not resolve whether these templates indeed reflected pre-stimulus expectations, or instead stimulus specific error signals induced by the unexpected omission of a stimulus (9, 13). Furthermore, the strength of these pre-stimulus expectation signals correlated with the behavioural benefit of a valid expectation, when the expected feature (i.e., orientation) was task-relevant (11). These results suggest that valid expectations facilitate perception by allowing sensory cortex to prepare for upcoming sensory signals. As in a previous fMRI study using a very similar experimental paradigm (11), the neural effects of orientation expectations reported here were independent of the task-relevance of the orientation of the gratings, suggesting that the generation of expectation templates may be an automatic phenomenon.

The fact that expectation signals were revealed by a decoder trained on physically presented (but task-irrelevant) gratings suggests that these expectation signals resemble activity patterns induced by actual stimuli. The expectation signal remained present throughout the trial, extending into the post-stimulus period, suggesting the tonic activation of a stimulus template. These results are in line with a recent monkey electrophysiology study (10), which showed that neurons in the face patch of IT cortex encode the prior expectation of a face appearing, both prior to and following actual stimulus presentation. When the subsequently presented stimulus is noisy or ambiguous, such a pre-stimulus template could conceivably bias perception towards the expected stimulus (21–24).

What is the source of these cue-induced expectation signals? One candidate region is the hippocampus, which is known to be involved in encoding associations between previously unrelated, discontiguous stimuli (25), such as the auditory tones and visual gratings used in the present study.

Furthermore, fMRI studies have revealed predictive signals in the hippocampus (13, 26, 27), and Reddy and colleagues (28) reported anticipatory firing to expected stimuli in the medial temporal lobe, including the hippocampus. One intriguing possibility is that predictive signals from the hippocampus are fed back to sensory cortex (13, 29, 30).

Previous studies have suggested, both on theoretical (31) and empirical (32, 33) grounds, that top-down (prediction) and bottom-up (stimulus-driven, or prediction error) signals are subserved by distinct frequency bands. Therefore, one highly interesting direction for future research would be to determine whether the expectation templates revealed here are specifically manifested in certain frequency bands (i.e., the alpha or beta band).

In addition to expectation, several other cognitive phenomena have been shown to induce stimulus templates in sensory cortex, such as preparatory attention (17, 34), mental imagery (35–37), and working memory (38, 39). In fact, explicit task preparation can also induce pre-stimulus sensory templates that last into the post-stimulus period (17). Note that in the current study the task did not require explicit use of the expectation cues, the task response was in fact orthogonal to the expectation. Furthermore, there was no difference in the expectation signal between runs in which grating orientation was task-relevant (orientation discrimination task) and when it was irrelevant (contrast discrimination task), suggestion expectation may be a relatively automatic phenomenon (11, 40). In fact, neural modulations by expectation have even been observed during states of inattention (41), sleep (42) and in patients experiencing disorders of consciousness (43). One important question for future research will be to establish whether the same neural mechanism underlies the different cognitive phenomena that are capable of inducing stimulus templates in sensory cortex, or whether different top-down mechanisms are at work. Indeed, it has been suggested that expectation and attention, or task preparation, may have different underlying neural mechanisms (20, 44, 45). For instance, predictive coding theories suggest that attention may modulate sensory signals in the superficial layers of sensory cortex, while predictions modulate the response in deep layers (5, 46).

One may wonder why the current study does not report a modulation of the overall neural response by expectation, while previous studies have found an increased neural response to unexpected stimuli (40, 47–51), including some using an almost identical paradigm as the current study (11, 20). Of course, the current study reports a null effect, from which it is hard to draw firm conclusions. However, it is possible that the type of measurement of neural activity plays a role in the absence of the effect. Most previous studies reporting expectation suppression in visual cortex used fMRI, while the current study used MEG. It is possible that the BOLD signal, a mass-action signal that integrates synaptic and neural activity, as well as integrating over time, is sensitive to certain neural effects that MEG, which is predominantly sensitive to synchronised activity in pyramidal neurons oriented perpendicular to the cortical surface, is not. It is even possible that within MEG, different types of sensors (i.e. magnetometers, planar and axial gradiometers) differ in their sensitivity to expectation suppression (52).

Recent theories of sensory processing state that perception reflects the integration of bottom-up inputs and top-down expectations, but ideas diverge on whether the brain continuously generates stimulus templates in sensory cortex to pre-empt expected inputs (10, 23, 53, 54), or rather engages in perceptual inference only after receiving sensory inputs (55, 56). Our results are in line with the brain being proactive, constantly forming predictions about future sensory inputs. These findings bring us closer to uncovering the neural mechanisms by which we integrate prior knowledge with sensory inputs to optimise perception.

Materials and Methods

Participants. Twenty-three (15 female, age 26 ± 9, mean ± SD) healthy individuals participated in the MEG experiment. All participants were right-handed and had normal or corrected-to-normal vision. The study was approved by the local ethics committee (CMO Arnhem-Nijmegen, The Netherlands) under the general ethics approval (“Imaging Human Cognition”, CMO 2014/288), and the experiment was conducted in accordance with these guidelines. All participants gave written informed consent according to the declaration of Helsinki.

Experimental design. Each trial consisted of an auditory cue, followed by two consecutive grating stimuli (750 ms SOA between auditory and first visual stimulus) (Fig. 1*A*). The two grating stimuli were presented for 250 ms each, separated by a blank screen (500 ms). A central fixation bull’s eye (0.7°) was presented throughout the trial, as well as during the intertrial interval (ITI, 2250 ms). The auditory cue consisted of either a low- (500 Hz) or high-frequency (1000 Hz) tone, which predicted the orientation of the first grating stimulus (45° or 135°) with 75% validity (Fig. 1*B*). In the other 25% of trials, the first grating had the orthogonal orientation. Thus, the first grating had an orientation of either exactly 45° or 135°, and a luminance contrast of 80%. The second grating differed slightly from the first in terms of both orientation and contrast (see below), as well as being in antiphase to the first grating (which had a random spatial phase). The contingencies between the auditory cues and grating orientations were flipped halfway through the experiment (i.e., after four runs), and the order was counterbalanced over subjects.

In separate runs (64 trials each, ~4.5 minutes), subjects performed either an orientation or a contrast discrimination task on the two gratings. When performing the orientation task, subjects had to judge whether the second grating was rotated clockwise or anticlockwise with respect to the first grating. In the contrast task, a judgment had to be made on whether the second grating had lower or higher contrast than the first one. These tasks were explicitly designed to avoid a direct relationship between the perceptual expectation and the task response. Furthermore, as in a previous fMRI study (11), these two different tasks were designed to manipulate the task-relevance of the grating orientations, to investigate whether the effects of orientation expectations depend on the task-relevance of the expected feature.

Interleaved with the main task runs, subjects performed eight runs of a grating localiser task (Fig. 1*C*). Each run (~2 min) consisted of 80 grating presentations (ITI uniformly jittered between 1000 and 1200 ms). The grating annuli were identical to those presented during the main task (80% contrast, 250 ms duration, 1.0 cycles/°, random spatial phase). Each grating had one of eight orientations (spanning the 180° space, starting at 0°, in steps of 22.5°), each of which was presented ten times per run in pseudorandom order. A black fixation bull's eye (4 cd/m^2, 0.7° diameter, identical to the one presented during the main task runs) was presented throughout the run. On 10% of trials (counterbalanced across orientations), the black fixation point in the centre of the bull's eye (0.2°, 4 cd/m^2) briefly turned gray (324 cd/m^2) during the first 50 ms of grating presentation. Participants task was to press a button (response deadline: 500 ms) when they perceived this fixation flicker. This simple task was meant to ensure central fixation, while rendering the gratings task-irrelevant. Trials containing fixation flickers were excluded from further analyses.

Orientation decoding analysis. To probe sensory representations in the visual cortex, we used a forward modelling approach to reconstruct the orientation of the grating stimuli from the MEG signal (17–19, 57). This method has been shown to be highly successful at reconstructing circular stimulus features, such as colour (18), orientation (17, 19, 57), and motion direction (22), from neural signals. Neural representations in MEG signals have also been successfully investigated using binomial classifiers (58), however, when it comes to a continuous stimulus feature such as orientation, forward model reconstructions provide a richer decoding signal than binomial classifier accuracy (59). We made certain changes to the forward model proposed by Brouwer and Heeger (18) (most notably, taking the noise covariance into account; see SI Materials and Methods for details) in order to optimise it for MEG data, given the high correlations between neighbouring sensors, based on (16)). In sum, this previously published and theoretically motivated decoding model was optimally suited for recovering a continuous feature from MEG data. For our main analyses, the forward model was trained on the data from the localiser runs, in which the gratings were task-irrelevant, and then applied to the main task data, in order to uncover sensory templates induced by pre-stimulus expectations (see SI Materials and Methods for details). Our effects of interest (see Fig. 3) were reproduced using a two-class decoder (Fig. S4).

The full methods can be found in the *Supporting Information*.

Data and code are available upon request.

Acknowledgements

This work was supported by The Netherlands Organisation for Scientific Research (NWO) to P.K. (Rubicon grant 446-15-004) and F.d.L. (VIDI grant 452-13-016) and the James S McDonnell Foundation to F.d.L. (Understanding Human Cognition 220020373). The authors would like to thank Mariya Manahova for data collection.

References

1. Helmholtz H von (1866) *Treatise on physiological optics* (The Optical Society of America, Menasha, WI). Translated from the third German edition, 1925.

2. Gregory RL (1997) Knowledge in perception and illusion. *Philos Trans R Soc Lond B Biol Sci* 352(1358):1121–1127.

3. Kersten D, Mamassian P, Yuille A (2004) Object Perception as Bayesian Inference. *Annu Rev Psychol* 55(1):271–304.

4. Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A* 20(7):1434.

5. Friston K (2005) A theory of cortical responses. *Philos Trans R Soc B Biol Sci* 360(1456):815–836.

6. Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15(11):745–756.

7. Wyart V, Nobre AC, Summerfield C (2012) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci* 109(9):3593–3598.

8. SanMiguel I, Widmann A, Bendixen A, Trujillo-Barreto N, Schroger E (2013) Hearing Silences: Human Auditory Processing Relies on Preactivation of Sound-Specific Brain Activity Patterns. *J Neurosci* 33(20):8633–8639.

9. Kok P, Failing MF, De Lange FP (2014) Prior Expectations Evoke Stimulus Templates in the Primary Visual Cortex. *J Cogn Neurosci* 26(7):1546–1554.

10. Bell AH, Summerfield C, Morin EL, Malecek NJ, Ungerleider LG (2016) Encoding of Stimulus Probability in Macaque Inferior Temporal Cortex. *Curr Biol* 26(17):2280–2290.

11. Kok P, Jehee JFM, De Lange FP (2012) Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron* 75(2):265–270.

12. Rao V, DeAngelis GC, Snyder LH (2012) Neural Correlates of Prior Expectations of Motion in the Lateral Intraparietal and Middle Temporal Areas. *J Neurosci* 32(29):10063–10074.

13. Hindy NC, Ng FY, Turk-Browne NB (2016) Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nat Neurosci* 19(5):665–667.

14. Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. *Nat Neurosci* 17(3):455–462.

15. King J-R, Dehaene S (2014) Characterizing the dynamics of mental representations: the temporal generalization method. *Trends Cogn Sci* 18(4):203–210.

16. Mostert P, Kok P, De Lange FP (2015) Dissociating sensory from decision processes in human perceptual decision making. *Sci Rep* 5:18253.

17. Myers NE, et al. (2015) Testing sensory evidence against mnemonic templates. *eLife* 4:e09000.

18. Brouwer GJ, Heeger DJ (2009) Decoding and Reconstructing Color from Responses in Human Visual Cortex. *J Neurosci* 29(44):13992–14003.

19. Brouwer GJ, Heeger DJ (2011) Cross-orientation suppression in human visual cortex. *J Neurophysiol* 106(5):2108–2119.

20. Kok P, Van Lieshout LLF, De Lange FP (2016) Local expectation violations result in global activity gain in primary visual cortex. *Sci Rep* 6:37706.

21. Chalk M, Seitz AR, Series P (2010) Rapidly learned stimulus expectations alter perception of motion. *J Vis* 10(8):2–2.

22. Kok P, Brouwer GJ, Van Gerven MAJ, De Lange FP (2013) Prior Expectations Bias Sensory Representations in Visual Cortex. *J Neurosci* 33(41):16275–16284.

23. Pajani A, Kok P, Kouider S, De Lange FP (2015) Spontaneous Activity Patterns in Primary Visual Cortex Predispose to Visual Hallucinations. *J Neurosci* 35(37):12947–12953.

24. St. John-Saaltink E, Kok P, Lau HC, De Lange FP (2016) Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *J Neurosci* 36(23):6186–6192.

25. Wallenstein GV, Hasselmo ME, Eichenbaum H (1998) The hippocampus as an associator of discontiguous events. *Trends Neurosci* 21(8):317–323.

26. Schapiro AC, Kustner LV, Turk-Browne NB (2012) Shaping of Object Representations in the Human Medial Temporal Lobe Based on Temporal Regularities. *Curr Biol* 22(17):1622–1627.

27. Davachi L, DuBrow S (2015) How the hippocampus preserves order: the role of prediction and context. *Trends Cogn Sci* 19(2):92–99.

28. Reddy L, et al. (2015) Learning of anticipatory responses in single neurons of the human medial temporal lobe. *Nat Commun* 6:8556.

29. Lavenex P, Amaral DG (2000) Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus* 10(4):420–430.

30. Bosch SE, Jehee JFM, Fernandez G, Doeller CF (2014) Reinstatement of Associative Memories in Early Visual Cortex Is Signaled by the Hippocampus. *J Neurosci* 34(22):7493–7500.

31. Bastos AM, et al. (2012) Canonical Microcircuits for Predictive Coding. *Neuron* 76(4):695–711.

32. Bastos AM, et al. (2015) Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron* 85(2):390–401.

33. Bauer M, Stenner M-P, Friston KJ, Dolan RJ (2014) Attentional Modulation of Alpha/Beta and Gamma Oscillations Reflect Functionally Distinct Processes. *J Neurosci* 34(48):16117–16125.

34. Stokes M, Thompson R, Nobre AC, Duncan J (2009) Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proc Natl Acad Sci* 106(46):19569–19574.

35. Stokes M, Thompson R, Cusack R, Duncan J (2009) Top-Down Activation of Shape-Specific Population Codes in Visual Cortex during Mental Imagery. *J Neurosci* 29(5):1565–1572.

36. Lee S-H, Kravitz DJ, Baker CI (2012) Disentangling visual imagery and perception of real-world objects. *NeuroImage* 59(4):4064–4073.

37. Albers AM, Kok P, Toni I, Dijkerman HC, De Lange FP (2013) Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. *Curr Biol* 23(15):1427–1431.

38. Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. *Nature* 458(7238):632–635.

39. Serences JT, Ester EF, Vogel EK, Awh E (2009) Stimulus-Specific Delay Activity in Human Primary Visual Cortex. *Psychol Sci* 20(2):207–214.

40. Den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE (2009) A Dual Role for Prediction Error in Associative Learning. *Cereb Cortex* 19(5):1175–1185.

41. Näätänen R (1990) The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behav Brain Sci* 13(02):201–233.

42. Nakano T, Homae F, Watanabe H, Taga G (2008) Anticipatory Cortical Activation Precedes Auditory Events in Sleeping Infants. *PLoS ONE* 3(12):e3912.

43. Bekinschtein TA, et al. (2009) Neural signature of the conscious processing of auditory regularities. *Proc Natl Acad Sci* 106(5):1672–1677.

44. Summerfield C, Egner T (2009) Expectation (and attention) in visual cognition. *Trends Cogn Sci* 13(9):403–409.

45. Summerfield C, Egner T (2016) Feature-Based Attention and Feature-Based Expectation. *Trends Cogn Sci* 20(6):401–404.

46. Kok P, Bains LJ, van Mourik T, Norris DG, De Lange FP (2016) Selective Activation of the Deep Layers of the Human Primary Visual Cortex by Top-Down Feedback. *Curr Biol* 26(3):371–376.

47. Summerfield C, Trittschuh EH, Monti JM, Mesulam M-M, Egner T (2008) Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 11(9):1004–1006.

48. Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus Predictability Reduces Responses in Primary Visual Cortex. *J Neurosci* 30(8):2960–2966.

49. Meyer T, Olson CR (2011) Statistical learning of visual transitions in monkey inferotemporal cortex. *Proc Natl Acad Sci* 108(48):19401–19406.

50. Todorovic A, van Ede F, Maris E, de Lange FP (2011) Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *J Neurosci* 31(25):9118–9123.

51. Wacongne C, et al. (2011) Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proc Natl Acad Sci* 108(51):20754–20759.

52. Cashdollar N, Ruhnau P, Weisz N, Hasson U (2016) The Role of Working Memory in the Probabilistic Inference of Future Sensory Events. *Cereb Cortex*:bhw138.

53. Berkes P, Orban G, Lengyel M, Fiser J (2011) Spontaneous Cortical Activity Reveals Hallmarks of an Optimal Internal Model of the Environment. *Science* 331(6013):83–87.

54. Fiser A, et al. (2016) Experience-dependent spatial expectations in mouse visual cortex. *Nat Neurosci* 19(12):1658–1664.

55. Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2(1):79–87.

56. Bar M, et al. (2006) Top-down facilitation of visual recognition. *Proc Natl Acad Sci U S A* 103(2):449–454.

57. Garcia JO, Srinivasan R, Serences JT (2013) Near-Real-Time Feature-Selective Modulations in Human Cortex. *Curr Biol* 23(6):515–522.

58. Cichy RM, Ramirez FM, Pantazis D (2015) Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *NeuroImage* 121:193–204.

59. Ester EF, Sprague TC, Serences JT (2015) Parietal and Frontal Cortex Encode Stimulus-Specific Mnemonic Representations during Visual Working Memory. *Neuron* 87(4):893–905.

60. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10(4):433–436.

61. Watson AB, Pelli DG (1983) Quest: A Bayesian adaptive psychometric method. *Percept Psychophys* 33(2):113–120.

62. Stolk A, Todorovic A, Schoffelen J-M, Oostenveld R (2013) Online and offline tools for head movement compensation in MEG. *NeuroImage* 68:39–48.

63. 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* 2011:1:1–1:9.

64. Bastiaansen MCM, Knösche TR (2000) Tangential derivative mapping of axial MEG applied to event-related desynchronization research. *Clin Neurophysiol* 111(7):1300–1305.

65. Blankertz B, Lemm S, Treder M, Haufe S, Müller K-R (2011) Single-trial analysis and classification of ERP components — A tutorial. *NeuroImage* 56(2):814–825.

66. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.

Figures

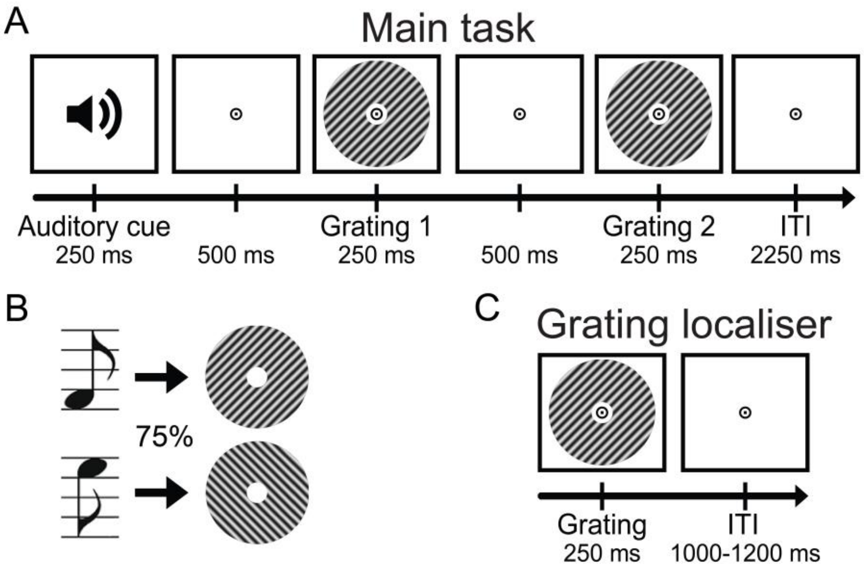

Fig. 1. Experimental paradigm. (*A*) Each trial started with an auditory cue that predicted the orientation of the subsequent grating stimulus. This first grating was followed by a second one, which differed slightly from the first in terms of orientation and contrast. In separate runs, participants performed either an orientation or contrast discrimination task on the two gratings. (*B*) Throughout the experiment, two different tones were used as cues, each one predicting one of the two possible orientations (45° or 135°) with 75% validity. These contingencies were flipped halfway through the experiment. (*C*) In separate grating localiser runs, participants were exposed to task-irrelevant gratings while they performed a fixation dot dimming task.

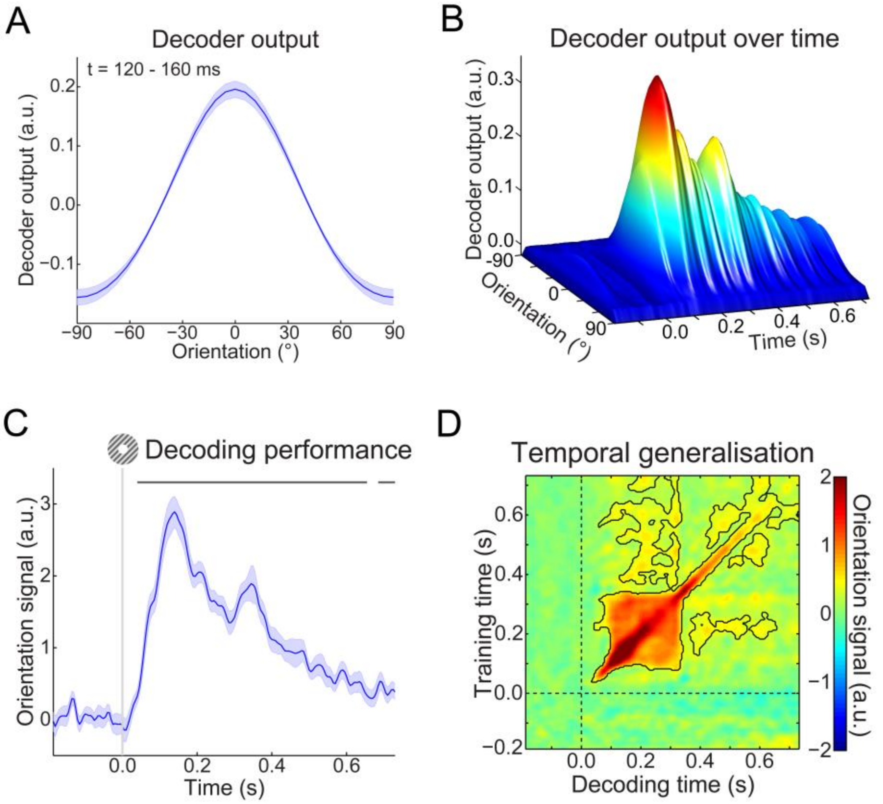

Fig. 2. Localiser orientation decoding. (*A*) The output of the decoder consisted of the responses of 32 hypothetical orientation channels, shown here decoders trained and tested on the MEG signal 120-160 ms post-stimulus during the grating localiser (cross-validated). Shaded region represent SEM. (*B*) Decoder output over time, trained and tested in 5 ms steps (sliding window of 29.2 ms), showing the temporal evolution of the orientation signal. (*C*) The response of the 32 orientation channels collapsed into a single metric of decoding performance (see SI Materials and Methods), over time. Shaded region represent SEM, horizontal lines indicate significant clusters ($p < 0.05$). (*D*) Temporal generalisation matrix of orientation decoding performance, obtained by training decoders on each time point, and testing all decoders on all time points (as above, steps of 5 ms and a sliding window of 29.2 ms). This method provides insight into the sustained versus dynamical nature of orientation representations (15). Solid black lines indicate significant clusters ($p < 0.05$), dashed lines indicate grating onset (t = 0s).

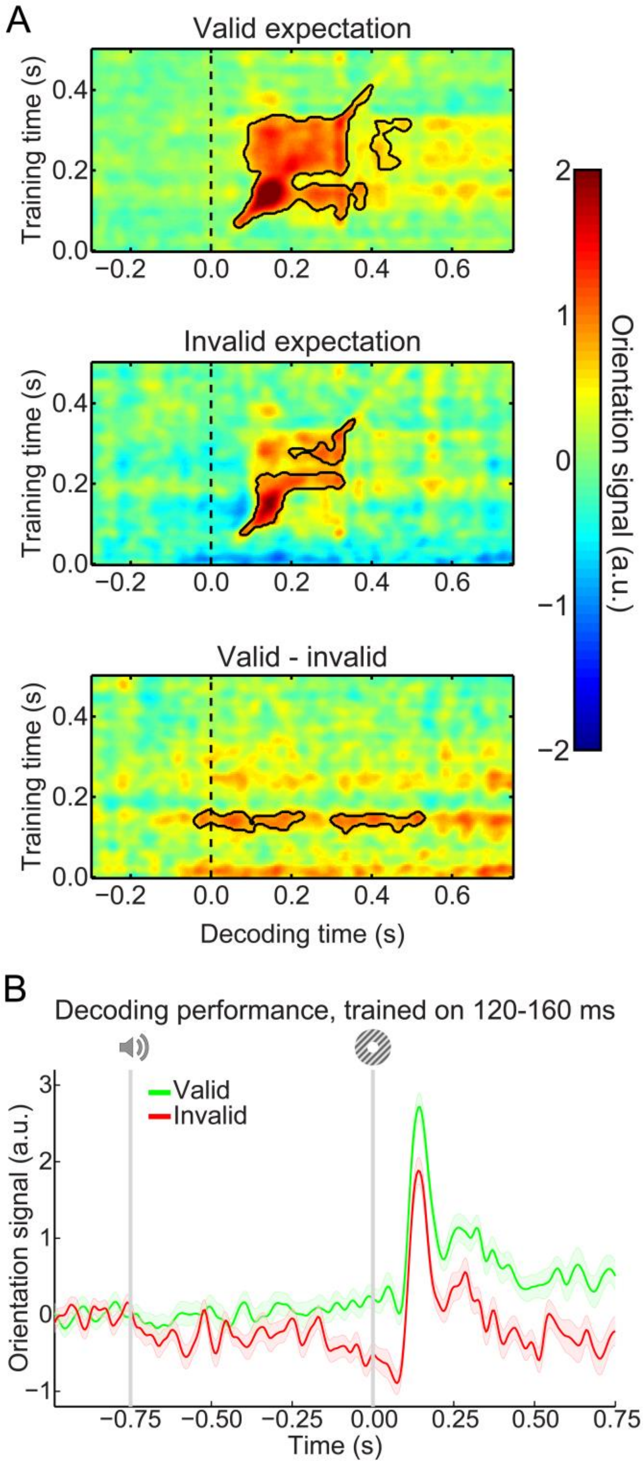

Fig. 3. Expectation induces stimulus templates. (*A*) Temporal generalisation matrices of orientation decoding during the main experiment. Decoders were trained on the grating localiser (training time on the y-axis) and tested on the main experiment (time on the x-axis; dashed vertical line indicates t = 0s, onset of the first grating). Decoding shown separately for gratings preceded by a valid expectation (top row), invalid expectation (middle row), and the subtraction of the two conditions (i.e., the expectation cue effect, bottom row). Solid black lines indicate significant clusters ($p < 0.05$). (*B*) Orientation decoding during the main task, averaged over training time 120 – 160 ms post-stimulus during the grating localiser. That is, a horizontal slice through the temporal generalisation matrices above at the training time for which we see a significant cluster of expected orientation decoding, for visualisation. Shaded regions indicate SEM.

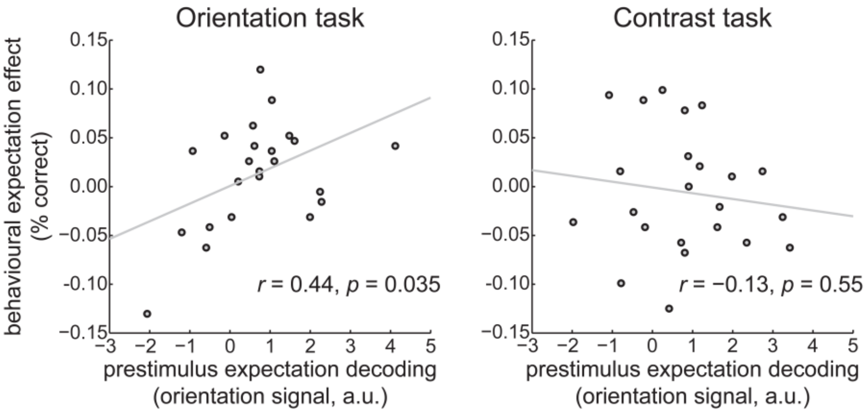

Fig. 4. Correlation between neural expectation signals and behavioural improvement by expectation. Neural pre-stimulus expectation decoding (on the x axis) correlated with behavioural improvement induced by valid expectations (on the y axis) during the orientation discrimination task (left panel). This correlation was absent during the contrast discrimination task (right panel).