

# Spontaneous Activity Patterns in Primary Visual Cortex Predispose to Visual Hallucinations

Auréliane Pajani,<sup>1</sup> Peter Kok,<sup>2</sup> Sid Kouider,<sup>1</sup> and Floris P. de Lange<sup>2</sup>

<sup>1</sup>Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, École Normale Supérieure, PSL Research University, 75005 Paris, France, and <sup>2</sup>Radboud University, Donders Institute for Brain, Cognition and Behaviour, 6525 EN Nijmegen, the Netherlands

According to theoretical frameworks casting perception as inference, vision results from the integration of bottom-up visual input with top-down expectations. Under conditions of strongly degraded sensory input, this may occasionally result in false perceptions in the absence of a sensory signal, also termed “hallucinations.” Here, we investigated whether spontaneous prestimulus activity patterns in sensory circuits, which may embody a participant’s prior expectations, predispose the observer toward false perceptions. Specifically, we used fMRI to investigate whether the representational content of prestimulus activity in early visual cortex is linked to subsequent perception during a challenging detection task. Human participants were asked to detect oriented gratings of a particular orientation that were embedded in noise. We found two characteristics of prestimulus activity that predisposed participants to hallucinations: overall lower prestimulus activity and a bias in the prestimulus activity patterns toward the to-be-detected (expected) grating. These results suggest that perceptual hallucinations may be due to an imprecise and biased state of sensory circuits preceding sensory evidence collection.

**Key words:** expectation; fMRI; hallucination; perception; spontaneous fluctuations

## Significance Statement

When sensory stimulation is strongly degraded, we occasionally misperceive a stimulus when only noise is present: a perceptual hallucination. Using fMRI in healthy participants, we investigated whether the state of early visual cortex preceding stimulus onset predisposes an observer to hallucinations. We found two characteristics of prestimulus activity that predisposed participants to hallucinations: overall lower prestimulus activity and a bias in the prestimulus activity patterns toward the expected grating. These results suggest that perceptual hallucinations are due to an imprecise and biased state of sensory circuits preceding sensation.

## Introduction

Dating back to Helmholtz’s theory of perception as “unconscious inference,” perception has been cast as a process of inference, of which the outcome is not simply determined by the bottom-up sensory input, but also by the top-down expectations that we have about our visual environment (Kersten et al., 2004; Sum-

merfield and de Lange, 2014). Under this framework, perceptual errors can arise, both when the bottom-up input is weak and imprecise and when top-down expectations are strong (Friston, 2005). For instance, one may mistake a spoon for a pen when looking for a pen in a dimly lit office and having the strong belief that the pen is there on the desk. This false percept would not occur if either the light were on (i.e., if the visual input were more precise) or if one would not have a strong expectation about the presence of a pen (i.e., under a weaker prior).

This example highlights two ingredients that may predispose a participant for erroneous perception: a state of imprecise sensory processing and strong prior expectations. The first ingredient, sensory precision, may be directly related to the overall activity levels in sensory regions through synaptic gain modulation (Feldman and Friston, 2010). Overall low levels of prestimulus activity in sensory circuits are associated with more perceptual errors (Hesselmann et al., 2010), both in terms of not perceiving stimuli that are present and falsely perceiving stimuli that are not present (i.e., “hallucinations”). Conversely, prior expectations about sen-

Received April 20, 2015; revised Aug. 13, 2015; accepted Aug. 14, 2015.

Author contributions: A.P., P.K., S.K., and F.P.d.L. designed research; A.P. and P.K. performed research; A.P., P.K., and F.P.d.L. analyzed data; A.P., P.K., S.K., and F.P.d.L. wrote the paper.

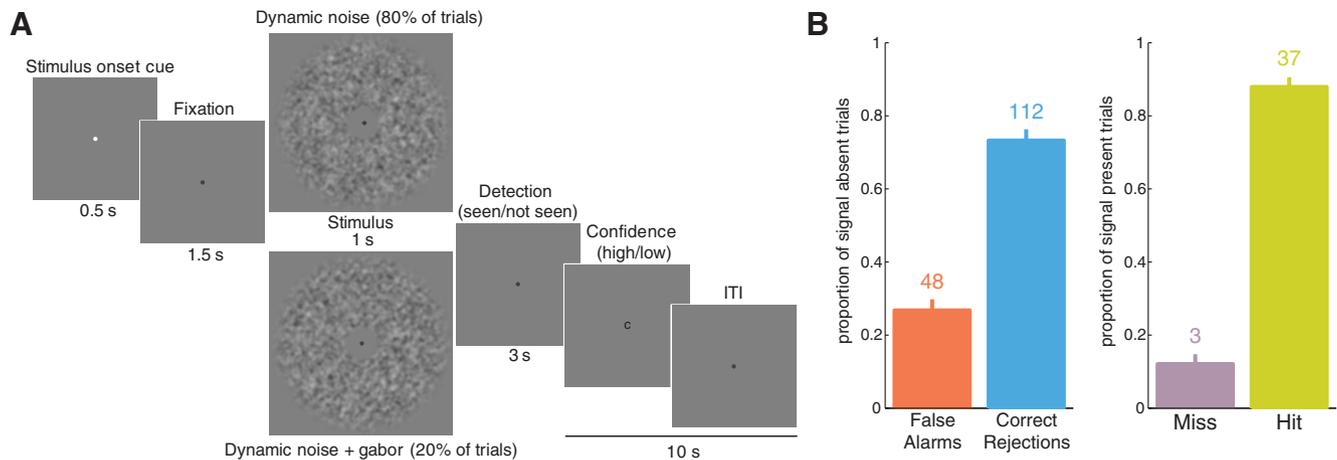
This work was supported by the James S. McDonnell Foundation (JSMF Scholar Award for Understanding Human Cognition to F.P.d.L.), the Ministère de l’Enseignement Supérieur et de la Recherche (A.P.), the European Research Council (A.P. and ERC Starting Grant DynaMind Project to S.K.), and Agence Nationale Recherche (Grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL to the Institut d’Études Cognitives). We thank Valentin Wyart and Tobias Donner for helpful discussions.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Floris P. de Lange, Donders Institute (Centre for Cognitive Neuroimaging), Room 0.26, Kapittelweg 29, P.O. Box 9101, NL 6500 HB Nijmegen, the Netherlands. E-mail: floris.delange@donders.ru.nl.

DOI:10.1523/JNEUROSCI.1520-15.2015

Copyright © 2015 the authors 0270-6474/15/3512947-07\$15.00/0



**Figure 1.** **A**, Experimental paradigm. Each trial started with a cue indicating the beginning of the trial. On 80% of the trials, participants were presented with dynamic noise alone; on the remaining 20%, a Gabor pattern at detection threshold was embedded in the noise. Participants performed a detection task and were then asked to provide a confidence judgment about their answer. Crucially, participants were informed that Gabor orientation was constant within a block, promoting the formation of orientation-specific stimulus expectation. **B**, Mean proportion of correct (CRs and hits) and incorrect (FAs and misses) responses within signal absent (left) and signal present (right) trials. Error bars indicate SEM. Numbers above bars indicate the median number of trials across participants.

sory events may be embodied in the specific patterns of prestimulus spontaneous activity in the relevant sensory regions (Berkes et al., 2011). Indeed, sensory expectations evoke patterns of activity that resemble those evoked by actual stimuli in early sensory cortices (SanMiguel et al., 2013; Kok et al., 2014) and such stimulus-like patterns have also been observed in spontaneous cortical activity (Kenet et al., 2003). Spontaneous fluctuations in neural patterns of activity may thus embody fluctuations in the content and strength of perceptual expectations. When strong, the latter may bias perception toward the detection of an expected stimulus even when it is absent. Whether these fluctuations predispose to visual hallucinations, however, is an open question.

In this study, we examined the influence of spontaneous prestimulus activity in the primary visual cortex (V1) on subsequent perceptual decision making. We used fMRI to measure overall neural activity and representational content in early visual cortex during a challenging detection task. Human participants were asked to decide whether an oriented grating was present in a dynamic noise patch, with perceptual difficulty titrated at perception threshold. Crucially, participants knew that only one grating orientation could be present in each block (either 45° or 135°), promoting the formation of a specific perceptual expectation.

In brief, we found two characteristics of prestimulus activity in the V1 that predisposed participants to false percepts: overall lower activity and a bias in the activity patterns toward the to-be-detected (i.e., expected) grating. These results suggest that perceptual hallucinations arise when the baseline state of sensory cortex is marked by an imprecise and biased state at the onset of sensory evidence collection, implying a sensory cause of decision biases in perception.

## Materials and Methods

### Participants

Twenty-five healthy right-handed individuals (14 female, age  $23.6 \pm 3.0$ , mean  $\pm$  SD) with normal or corrected-to-normal vision gave written informed consent to participate in this study, in accordance with the institutional guidelines of the local ethics committee (CMO region Arnhem-Nijmegen, The Netherlands). Data from six participants were excluded due to an insufficient number of false percepts (see “Behavioral results”).

### Stimuli

Greyscale luminance-defined sinusoidal Gabor grating stimuli and Gaussian noise patches were generated using MATLAB (The MathWorks) and the Psychophysics Toolbox (Brainard, 1997). In the behavioral session, stimuli were presented on a Samsung Syncmaster 940BF screen (1024  $\times$  786 resolution, 60 Hz refresh rate). In the fMRI session, they were displayed on a rear-projection screen using a luminance-calibrated Sony VPL FX40 projector (1024  $\times$  768 resolution, 60 Hz). The stimuli consisted of a set of 60 Gaussian noise patches presented successively in a random sequence, each displayed for 16.7 ms (total stimulus duration: 1.0 s) to which a Gabor pattern was added in 20% of the trials. They were displayed in an annulus (outer diameter: 20° of visual angle, inner diameter: 4°) surrounding a central fixation point.

The orientation of the Gabor patterns was constant within a block (either 45° or 135° clockwise), whereas their spatial phase was random (either 0 or 0.25, randomized) and their spatial frequency was fixed (1.0 cpd). Contrast detection threshold was determined for each participant and each orientation by a Bayesian adaptive staircase procedure (Watson and Pelli, 1983), which was set to an overall correct response percentage of 70%.

The noise patches were created by smoothing pixel-by-pixel Gaussian noise through a 2D Gaussian smoothing filter so that the spatial frequency of the noise equated that of the Gabor patterns (smoothing dimension: 0.167° of visual angle, noise contrast: SD of 20%). This allowed us to generate noise with similar low-level properties as the Gabor patterns that participants had to detect, which was meant to favor false detections. However, we wanted to exclude noise patches that actually resembled the Gabor patterns; that is, noise patches that contained a certain level of “signal energy” (Wyart et al., 2012). The signal energy of each noise patch was computed by processing it through a pool of Gabor energy filters with varying preferred orientations (1–180°) and 60 noise patches with low (<2%) signal energy for all orientations were selected. Those 60 noise patches were used for all participants and all trials. This ensured that the noise stimulation was kept constant across trials and participants (except for the order in which the patches were presented), so that there were no trial-to-trial fluctuations in noise. Therefore, false detections could not be driven by visual input, but could only be triggered by internal fluctuations.

### Experimental design

**Experiment.** The time course of a trial is displayed in Figure 1A. A gray central fixation point was presented throughout the trial and intertrial interval (ITI). At the beginning of a trial, the fixation point turned white for 0.5 s before turning back to gray for 1.5 s. Then the stimulus was displayed for 1 s: on 80% of the trials, participants were presented with

dynamic Gaussian noise only (“signal absent trials”) and, on the remaining 20%, a Gabor pattern was embedded in the noise (“signal present trials”). On signal present trials, the grating always had the same orientation within a block (either 45° or 135°), which subjects were explicitly instructed about at the beginning of each block. Signal absent and signal present trials were randomly interleaved. After the stimulus was displayed, participants pressed one of two buttons to indicate whether they thought the Gabor was present or absent using a two-button response box placed in one hand. They had to provide their answer within 2 s, otherwise the fixation point would turn blue for 0.2 s to signal that they had been too slow and their answer was not recorded. After this 2 s response period, participants kept fixation for 1 s and were then prompted to provide a 2-point confidence judgment when the letter “C” was centrally displayed. They pressed one of two buttons to indicate whether they had low or high confidence about their answer using a response box placed in the other hand than the one that they used for the detection response. Response hands for detection and confidence judgment were counterbalanced across participants. The trial was followed by a 10 s ITI (including response time for the confidence judgment).

All participants completed four blocks of 50 trials each (~13 min): two blocks with one orientation, followed by two blocks with the other orientation (order counterbalanced across participants). This yielded a total of 200 trials for each participant. Trials were split into the four signal-detection theoretical categories: hits (a grating was displayed and perceived), false alarms (FAs, no grating was displayed but a grating was perceived), correct rejections (CRs, no grating was presented nor perceived), and misses (a grating was presented but it was not perceived). Note that, given our interest in FAs, we used a low base rate of stimulus-present trials (20%), so misses were therefore relatively rare; therefore, we chose not to report the corresponding neural data. Because participants predominantly rated their FAs as low confidence, precluding a reliable estimation of neural activity separately for the two confidence levels, we chose not to further analyze the confidence judgment data.

Before each set of two blocks with a particular orientation, participants were given instructions and performed a short training block consisting of 10 trials with high-contrast Gabor patterns that were present on 50% of the trials. After that, they performed an adaptive staircase procedure to determine their detection threshold, which was used for the following two blocks. Participants thus performed the staircase twice, once at the beginning of the experiment for the first orientation and once halfway through the experiment for the second orientation. The purpose of this procedure was to keep task difficulty constant across orientations.

To familiarize participants with the task and to ensure stability of the detection threshold, the fMRI session was preceded by a behavioral session on the day before, which was identical to the imaging experiment except that ITI was reduced to 2 s and Gabor patterns were present on 50% of the trials.

**Functional localizer.** After the main experiment, participants performed a functional localizer task in the MR scanner. This consisted of flickering gratings (2 Hz) displayed in an annulus of the same size as during the experiment and presented at full contrast. Participants were presented with eight blocks of 144 s each; the third and sixth blocks were both followed by a 24 s fixation period. This yielded 20 min of scanning in total. A 144 s block consisted of 8 16 s subblocks corresponding to eight orientations (0–157.5° in steps of 22.5°) presented in pseudorandomized order, followed by a 16 s blank screen containing only a fixation bull’s-eye. Each subblock contained gratings with a fixed orientation, fixed spatial frequency (1.0 cpd), and random spatial phase. To ensure central fixation, participants were required to monitor a stream of letters displayed at 1 Hz in the bull’s-eye and to press a button whenever an “X” or a “Z” was displayed (18–36 times per block at unpredictable times).

**fMRI acquisition.** fMRIs were acquired using a 3T Skyra MRI system (Siemens) with a T2\*-weighted gradient-echo EPI sequence (TR/TE = 2000/30 ms, 29 transversal slices, voxel size 2 × 2 × 2 mm, 80° flip angle). Anatomical images were acquired with a T1-weighted MP-RAGE sequence (TR/TE = 2300/3.03 ms, voxel size 1 × 1 × 1 mm, 8° flip angle).

**Preprocessing.** SPM8 (<http://www.fil.ion.ucl.ac.uk/spm>, Wellcome Trust Centre for Neuroimaging, London, UK) was used for image preprocessing. The first six volumes of each block were discarded to allow T1

equilibration. All functional images were spatially realigned to the mean image, yielding six head movement parameters, and the structural image was coregistered with the functional volumes. Using custom MATLAB code, the data were high-pass filtered (cutoff 128 s) to remove low-frequency signal drifts and the time series of each voxel was z-scored.

**Voxel population selection.** We used *Freesurfer* (<http://surfer.nmr.mgh.harvard.edu/>) to identify the anatomical boundaries of the V1 (Benson et al., 2012) (V1: size = 2457 ± 294 voxels, mean ± SD).

Within V1, we first identified the 500 most visually responsive voxels using the independent localizer data. For this, we used SPM8 to construct a general linear model (GLM) with two regressors, representing grating blocks and blank screen blocks, respectively, which were constructed by convolving the onsets and durations of those blocks with a canonical hemodynamic response function. The six head movement parameters computed during preprocessing were added as nuisance regressors (Lund et al., 2005). We then selected the 500 voxels that showed the strongest preference for gratings versus a blank screen (highest T-value for the contrast grating > blank screen).

These voxels were defined based on their visual responsiveness to the stimulus during the localizer and were predominantly located close to the occipital pole, which corresponds to the retinotopic representation of the fovea. This is consistent with the idea that this 500-voxel population overlaps with the retinotopic representation of our stimuli.

Within the 500 most visually responsive voxels, we then selected two orientation-selective subpopulations that responded most strongly to 45° gratings and 135° gratings, respectively, using the independent localizer data. For this purpose, a GLM was constructed using SPM8 with separate regressors for the two grating orientations used in the experiment (45° and 135°) plus the six head movement regressors. Specifically, we selected the 100 voxels that showed the strongest preference for 45° (highest T-value for the contrast 45° > 135° gratings) and the 100 voxels that showed the strongest preference for 135° (highest T-value for the contrast 135° > 45° gratings).

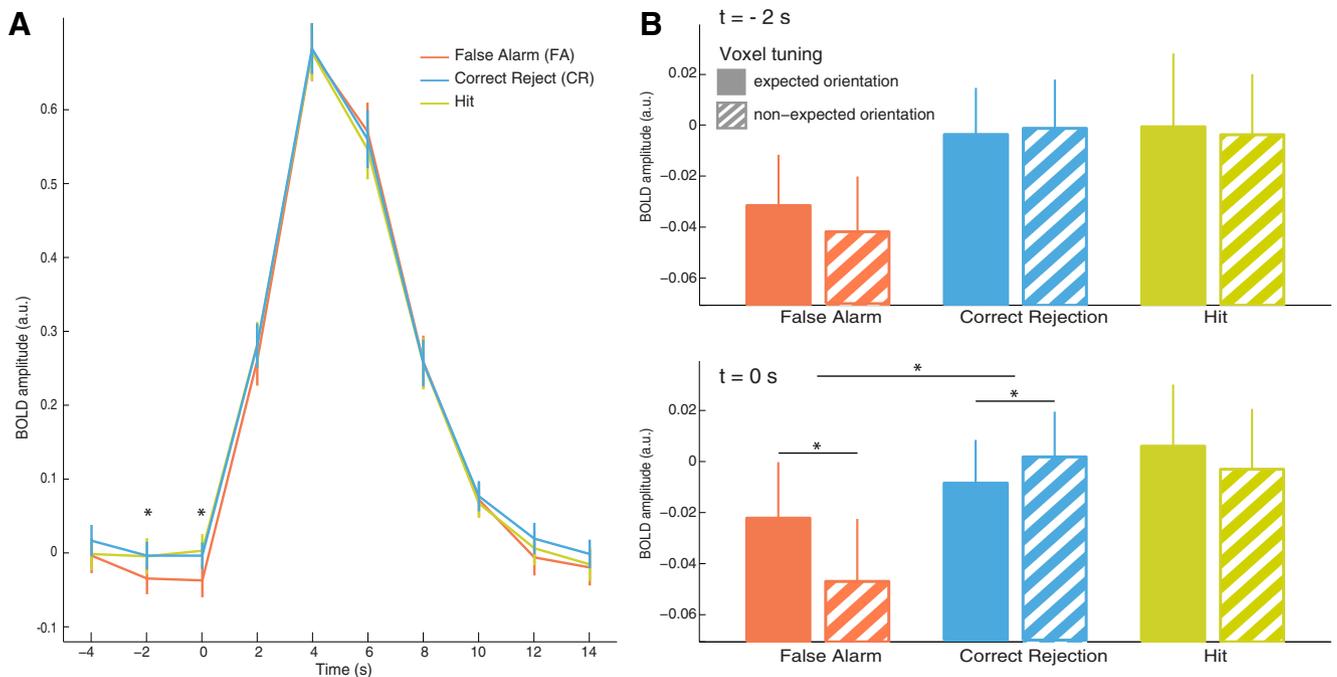
We probed the repartition of both voxel populations within V1 to assess its alignment with known properties of the visual cortex. The 45°-tuned voxels were predominantly located in the dorsal part of the right hemisphere and in the ventral part of the left hemisphere and, conversely, the 135°-tuned voxels were predominantly located in the ventral part of the right hemisphere and in the dorsal part of the left hemisphere. These results are consistent with the radial orientation bias in V1 (Sasaki et al., 2006), which means that there is a specific relationship between the retinotopic organization and orientation tuning within early visual cortex.

**Task data analysis.** Our slow event-related design (ITI = 10 s, trial duration = 16 s) allowed us to analyze the preprocessed functional images without needing to disentangle BOLD responses between the different conditions using a GLM. Consistently with previous work reporting a link between perception and prestimulus activity in sensory circuits (Hesselmann et al., 2008a), we restricted our analyses to the two prestimulus scans,  $t = -2$  s and  $t = 0$  s, corresponding to the functional volumes acquired between  $-2$  and 0 s and between 0 and 2 s, respectively. We considered the hemodynamic signal in these volumes to reflect prestimulus activity in view of the sluggishness of the hemodynamic response. Statistical significance testing was performed using R statistical software (R 3.0.3, GUI Rstudio 0.98.501); statistical tests were two-tailed.

To investigate whether spontaneous fluctuations in V1 biased perception in the absence of a stimulus, we focused our analyses on CRs and FAs. We also report the data for hits, whereas misses were not analyzed due to their low number (see “Behavioral results”).

To assess the mean prestimulus BOLD signal in V1, the functional images were averaged over the 500 most visually responsive V1 voxels. Differences in prestimulus overall activity levels between trial types were assessed by a two-way repeated-measures ANOVA with factors “trial type” (levels: CRs and FAs) and “time” (levels:  $t = -2$  s and  $t = 0$  s).

To probe the orientation specificity of these signals, the functional images were averaged separately in both 100-voxel orientation-selective subpopulations. We aggregated the data according to voxel preference and block orientation: activity in 45° voxels during 45° blocks and activity in 135° voxels during 135° blocks were averaged together as “voxels tuned



**Figure 2.** *A*, Time course of BOLD signal in V1 time locked to onset of visual stimulus. Data were averaged across 19 participants. Error bars indicate SEM. *B*, Amplitude of BOLD activity in V1 voxels preferring the orientation displayed in the current block (filled bars) or the orthogonal orientation (hatched bars) separately for FAs, CRs, and hits during the first ( $t = -2$  s, top panel) and second ( $t = 0$  s, bottom panel) time points before stimulus onset (see *x*-axis on *A*). Error bars indicate SEM.

to expected orientation” and activity in  $45^\circ$  voxels during  $135^\circ$  blocks and activity in  $135^\circ$  voxels during  $45^\circ$  blocks were averaged together as “voxels tuned to nonexpected orientation” (Kok et al., 2014). Differences in orientation-specific activity between trial types were assessed by a three-way repeated-measures ANOVA with the factors “voxel tuning” (levels: expected orientation and nonexpected orientation), “trial type” (levels: CRs and FAs), and “time” (levels:  $t = -2$  s and  $t = 0$  s).

Finally, to examine the relationship between overall and orientation-specific prestimulus activity, we computed the correlation between the strength of both neural effects across participants. Using Pearson’s product-moment correlation, we computed the correlation between the difference between CRs and FAs in prestimulus BOLD signal in the visually responsive V1 voxels and the difference between prestimulus activity in voxels tuned to expected and nonexpected orientations for FAs minus the same difference for CRs.

## Results

### Behavioral results

We estimated sensitivity,  $d'$ , and bias,  $c$ , using signal-detection theoretical methods (Green and Swets, 1974; Macmillan and Douglas, 2005). Perceptual sensitivity was high ( $d' = 2.03 \pm 0.61$ , mean  $\pm$  SD), indicating that participants were well able to perform the detection task. Because there were only 20% signal-present trials, misses were very infrequent ( $<4$  misses per orientation for 22/25 participants), so we did not further analyze this category. On average, participants’ criterion showed a slightly liberal bias ( $c = -0.17 \pm 0.56$ , mean  $\pm$  SD), resulting in, on average, 48 FAs. However, criterion values were variable, with 6/25 participants exhibiting  $<4$  FAs for 1 or both of the orientations during the experiment. Because this low number of trials precluded a reliable estimation of the BOLD activity patterns associated with FAs in these six participants, their fMRI data were not analyzed. All data analyses were done on the remaining 19 participants. Relative proportions and median number of each trial type for these participants are summarized in Figure 1*B*.

### fMRI results

#### Prestimulus activity

The aim of the experiment was to determine whether, in the context of identical bottom-up input, prestimulus neural fluctuations could result in false perception. To this end, we first compared overall prestimulus neural activity in V1 for FAs and CRs, two conditions in which stimulus input was identical (stimulus-absent) but perceptual choice diverged. We observed significantly lower overall activity preceding FAs than CRs (main effect of trial type:  $F_{(1,18)} = 8.35$ ,  $p = 0.01$ ; Fig. 2*A*) for both prestimulus scans (no trial type  $\times$  time interaction,  $F_{(1,18)} = 0.07$ ,  $p = 0.79$ ). The finding of overall lower prestimulus activity during FAs compared with CRs is consistent with previous research (Hesselmann et al., 2010). There were no significant differences in overall prestimulus activity between hits and CRs ( $F_{(1,18)} = 0.10$ ,  $p = 0.76$ ).

Next, we investigated whether FAs were associated with larger prestimulus activity in the voxels representing the expected percept, that is, the specific oriented stimulus that could be presented during a particular block. For this, we estimated neural activity in V1 voxels preferring  $45^\circ$  and  $135^\circ$  based on an independent localizer dataset (see Materials and Methods) during  $45^\circ$  and  $135^\circ$  blocks for both prestimulus time points. We aggregated these data into “voxels tuned to the expected orientation” and “voxels tuned to the nonexpected orientation” and again compared prestimulus neural activity patterns for FAs and CRs, two conditions in which stimulus input was identical (stimulus-absent) but perceptual choice diverged. The 3-way (voxel tuning  $\times$  trial type  $\times$  time) ANOVA yielded a significant 2-way interaction between voxel tuning and trial type ( $F_{(1,18)} = 5.90$ ,  $p = 0.026$ ; Fig. 2*B*), reflecting the presence of different patterns of activity before FAs and CRs. Prompted by a trend toward a 3-way interaction among voxel tuning, trial type, and time ( $F_{(1,18)} = 4.28$ ,  $p = 0.053$ ), we analyzed both prestimulus time points sep-

arately. At  $t = 0$  s, there was significantly stronger activity in the voxel population tuned to the expected than to the nonexpected orientation for FAs ( $t_{(18)} = 2.28$ ,  $p = 0.035$ ) and the opposite pattern was visible for CRs ( $t_{(18)} = -2.20$ ,  $p = 0.041$ ). These effects were not significant at  $t = -2$  s ( $t_{(18)} = 1.27$ ,  $p = 0.22$  and  $t_{(18)} = -0.55$ ,  $p = 0.59$ ; respectively for FAs and CRs). This 2-way interaction between voxel tuning and trial type for the prestimulus period was independent of the number of voxels selected in each orientation-selective population ( $p < 0.05$  for all voxel selection sizes between 75 and 250). There was no difference in prestimulus activity between populations for hits ( $t_{(18)} = 0.39$ ,  $p = 0.70$  and  $t_{(18)} = 0.92$ ,  $p = 0.37$ ; respectively for  $t = 0$  s and  $t = -2$  s).

Interestingly, there was a significant correlation across participants between the overall reduction of prestimulus neural activity and the increase in orientation-specific activity in the expected orientation voxel population before FAs ( $r = 0.48$ ,  $t_{(17)} = 2.25$ ,  $p = 0.038$ ). This positive correlation suggests that participants with an overall strong prestimulus decrease in V1 activity for FAs also exhibited strong orientation-specific patterns of activity before FAs compared with CRs.

#### Stimulus-evoked activity

We did not observe any poststimulus difference between conditions; the BOLD amplitude at the peak of the hemodynamic response (time points  $t = 4$  s and  $t = 6$  s; Fig. 2A) was similar for CRs, FAs, and hits ( $F_{(2,36)} = 0.73$ ,  $p = 0.49$ ) and there were no differences in orientation specificity ( $F_{(2,36)} = 1.22$ ,  $p = 0.31$ ). Perhaps surprisingly, we did not find evidence for stronger activity in voxels tuned to the expected orientation than in voxels tuned to the orthogonal orientation during hits ( $t_{(18)} = 1.36$ ,  $p = 0.19$ ) even though the expected grating was actually displayed; this was also true during FAs ( $t_{(18)} = 0.63$ ,  $p = 0.54$ ). However, given that the grating stimulus was extremely faint and embedded in high-contrast dynamic noise (60 successive noise patches during 1 s), it appears plausible that the stimulus-evoked fMRI activity is particularly driven by the noise, obscuring any possibly subtle differences between conditions. Note that our prestimulus measurements were not affected by the strong visual noise stimulation because there was no bottom-up input during these time points except for the fixation point, so they reflect only internal ongoing activity fluctuations.

#### Influence of the previous trial on prestimulus neural states

Although the prestimulus fluctuations in V1 that we observed may reflect spontaneous activity fluctuations, they could also result from the activity evoked by the stimulus or perceptual choice on the previous trial (Gao et al., 2009; de Lange et al., 2013; Fischer and Whitney, 2014). The proportion of FAs was indeed lower on stimulus-absent trials after CRs than after hits ( $t_{(18)} = 1.47$ ,  $p = 0.039$ ) or FAs ( $t_{(18)} = 3.94$ ,  $p = 0.001$ ), which means that participants were more likely to falsely detect a grating on a stimulus-absent trial when they perceived a grating on the previous trial. To investigate whether the reported prestimulus BOLD differences between FAs and other trial types (hits, CRs) could reflect differences in stimulus-evoked or choice-evoked signals on the previous trial, we probed prestimulus activity and representational content in V1 as a function of trial type during the previous trial. We observed no systematic differences in prestimulus orientation-specific patterns of activity between trials after CRs, FAs, or hits ( $F_{(2,36)} = 1.33$ ,  $p = 0.28$ ), nor were there any differences in the amplitude of prestimulus activity ( $F_{(2,36)} = 0.21$ ,  $p = 0.81$ ). Therefore, our results cannot be explained as

resulting from residual activity from the stimulus or choice that was made on the previous trial.

## Discussion

According to inferential accounts, perception is the result of the integration of bottom-up inputs with top-down expectations. Therefore, the same noisy stimulus may or may not trigger a false percept depending on the state of sensory circuits before its presentation (Hesselmann et al., 2008b, 2010). Here, we found that false percepts (hallucinations) arise due to a conjunction of two processes in the V1: low overall prestimulus activity, potentially reflecting low sensory precision, and a bias in the prestimulus activity patterns toward the to-be-detected (expected) percept.

Our results are consistent with recent studies showing that lower prestimulus activity in sensory cortices is associated with false perception (Hesselmann et al., 2010; Schölvinck et al., 2012). However, in those studies, overall activity in sensory circuits is lower not only before FAs, but also before misses. This suggests that a low prestimulus baseline predisposes to perceptual errors in general, so it may embody the precision of forthcoming perceptual inference. From a normative perspective, sensory inputs should be weighted depending on how reliable they are. The predictive coding framework posits that the brain represents an internal model of its environment with which incoming sensory input is constantly compared: when there is a mismatch, the internal model is updated to better fit the external inputs received through the sensory receptors. However, in a noisy environment, a mismatch should not always lead to model revision because it is likely to be a random upshot of noise, in which case sensory inputs should be attributed little weight (Hohwy, 2012). Sensory precision pertains to the estimation the brain makes about how noisy its sensations are as follows: a noisy environment, in which sensory inputs are unreliable, is associated with low sensory precision, which entails little weighting of sensory inputs. Precision could be neurally implemented through modulations of the synaptic gain of sensory neurons, which would be reflected in ongoing BOLD signal. In this framework, attention and precision are computationally equivalent: attention leads to increased synaptic gain in sensory areas and therefore increased weighting of sensory inputs (Friston, 2009; Feldman and Friston, 2010). Empirical evidence shows that visuospatial attention is associated with increased prestimulus BOLD signal (Munneke et al., 2008; Sylvester et al., 2009), which in turn is associated with more accurate perceptual inference (Hesselmann et al., 2010). Altogether, this is consistent with the view that prestimulus BOLD signal reflects the weight attributed to sensory inputs during perceptual inference, possibly through attentional modulation. Here, we show that low sensory precision (as indexed by low prestimulus BOLD activity in sensory cortex) is one ingredient that may result in the erroneous detection of the expected stimulus in the absence of a grating. Further support in favor of this hypothesis comes from the M/EEG literature, which establishes a similar link among attention, detection accuracy, and prestimulus occipital power in the alpha band. High occipital alpha power is associated with low BOLD signal in the occipital cortex (Laufs et al., 2003a, 2003b; Moosmann et al., 2003; Scheeringa et al., 2011), as well as with low levels of visuospatial attention (Worden et al., 2000; Thut et al., 2006). Interestingly, high occipital alpha power before stimulus onset is associated with decreased visual discrimination ability (Ergenoglu et al., 2004; van Dijk et al., 2008), as is the case for low BOLD signal in early sensory cortex. Strong posterior alpha activity is thought to reflect low gain of the visual stream (Jokisch and Jensen, 2007), which, in the predictive coding framework, is

equivalent to low precision of the bottom-up sensory input (Friston, 2008). Altogether, these results suggest that overall prestimulus activity in early visual cortex reflects attentional levels and affects the precision of subsequent sensory input during perceptual inference, favoring erroneous detections in the absence of a grating.

Interestingly, visual hallucinations were marked by the presence of prestimulus activity patterns in V1 that contained orientation information specific to the to-be-detected stimulus. Specifically, when early visual cortex contained orientation-specific activity in voxels tuned to the currently expected grating, a noise stimulus was more likely to be perceived as containing the expected grating (FA) than as mere noise (CR). Previous research has shown that these orientation-specific patterns may embody prior expectations. For example, in the absence of a stimulus, expectations evoke patterns of activity in early visual cortex similar to those evoked by actual stimuli (SanMiguel et al., 2013; Kok et al., 2014). Such complex and regular spatiotemporal patterns have also been characterized in spontaneous brain fluctuations (Fox et al., 2005), where they are thought to embody prior expectations about the immediate, directly relevant future (Schacter et al., 2007). This suggests that the visual cortex forms an internal model of previously encountered stimuli that is reflected in spontaneous neuronal activity. Moreover, such expectations influence subsequent perception and bias sensory processing toward the perception of the expected stimulus (Kok et al., 2013). In our study, the to-be-detected grating had the same orientation over the course of an entire block, allowing participants to form a stable expectation template of the grating they had to detect. Fluctuations in the activation of this template may have led to fluctuations in perception: when the template was strongly activated and sensory precision was low, a false perception of a grating occurred.

One source of prestimulus activity could be residual activity from the previous trial due to “choice stickiness” or priming by the previous choice (Bode et al., 2012; de Lange et al., 2013; Fischer and Whitney, 2014). However, our results do not show any systematic bias in prestimulus patterns of activity for trials following hits (when a Gabor was presented) or for trials following FAs (when a Gabor was perceived). Therefore, the orientation-specific patterns of activity that we observed before FAs are more likely to reflect spontaneous fluctuations in the sensory circuits than lingering activity from the previous trial, although they may influence perception through similar mechanisms.

Another alternative interpretation of our results relates to feature-based attention because this phenomenon is known to trigger stimulus-like patterns of activity in the visual cortex (Kamitani and Tong, 2005; Stokes et al., 2009). However, the orientation-specific patterns that characterize V1 activity before FAs arise in conjunction with overall low activity and are therefore more consistent with attenuated levels of attention (Supèr et al., 2003; Hesselmann et al., 2010). Therefore, it seems unlikely that FAs are due to increased feature-based attention.

Furthermore, the current task design makes it unlikely that eye movements contributed to the presence of an orientation-specific signal in prestimulus patterns. We selected our orientation-selective voxel populations on the basis of presentation of unattended gratings while participants performed a task at fixation with no incentive for eye movements.

In sum, our data suggest that a combination of spontaneous fluctuations of sensory precision and sensory expectations can lead to false perceptions (i.e., hallucinations). The magnitude of these factors were correlated, suggesting that these phenomena

may be linked rather than fluctuate independently. This could result from the existence of two stable states of sensory systems, between which spontaneous activity would fluctuate in the absence of a stimulus (Deco and Corbetta, 2011). “Inner-oriented” states would be characterized by low weighting of sensory inputs and strong expectations, favoring percepts that reflect inner templates rather than the external world, as is the case during FAs. Conversely, “outer-oriented” states would be associated with low sensory noise and relatively flat priors, altogether leading to visual representations mostly driven by sensory stimulation. In this case, the absence of a grating would lead to a correct rejection. In conclusion, our data support the view that false alarms are not pure strategic guesses, but rather genuine perceptual errors during which participants falsely perceive the stimulus that they are expecting due to a strong prior expectation and low sensory precision.

## References

- Benson NC, Butt OH, Datta R, Radoeva PD, Brainard DH, Aguirre GK (2012) The retinotopic organization of striate cortex is well predicted by surface topology. *Curr Biol* 22:2081–2085. [CrossRef Medline](#)
- Berkes P, Orbán G, Lengyel M, Fiser J (2011) Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science* 331:83–87. [CrossRef Medline](#)
- Bode S, Sewell DK, Lilburn S, Forte JD, Smith PL, Stahl J (2012) Predicting perceptual decision biases from early brain activity. *J Neurosci* 32:12488–12498. [CrossRef Medline](#)
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436. [CrossRef Medline](#)
- Deco G, Corbetta M (2011) The dynamical balance of the brain at rest. *Neuroscientist* 17:107–123. [CrossRef Medline](#)
- de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci* 33:1400–1410. [CrossRef Medline](#)
- Ergenoglu T, Demiralp T, Bayraktaroglu Z, Ergen M, Beydagi H, Uresin Y (2004) Alpha rhythm of the EEG modulates visual detection performance in humans. *Cogn Brain Res* 20:376–383. [CrossRef Medline](#)
- Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. *Front Hum Neurosci* 4:215. [Medline](#)
- Fischer J, Whitney D (2014) Serial dependence in visual perception. *Nat Neurosci* 17:738–743. [CrossRef Medline](#)
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci U S A* 102:9673–9678. [CrossRef Medline](#)
- Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–836. [CrossRef Medline](#)
- Friston K (2008) Hierarchical models in the brain. *PLoS Comput Biol* 4:e1000211. [CrossRef Medline](#)
- Friston K (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn Sci* 13:293–301. [CrossRef Medline](#)
- Gao J, Wong-Lin K, Holmes P, Simen P, Cohen JD (2009) Sequential effects in two-choice reaction time tasks: decomposition and synthesis of mechanisms. *Neural Comput* 21:2407–2436. [CrossRef Medline](#)
- Green DM, Swets JA (1974) Signal detection theory and psychophysics. Oxford: Krieger.
- Hesselmann G, Kell CA, Eger E, Kleinschmidt A (2008a) Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proc Natl Acad Sci U S A* 105:10984–10989. [CrossRef Medline](#)
- Hesselmann G, Kell CA, Kleinschmidt A (2008b) Ongoing activity fluctuations in hMT+ bias the perception of coherent visual motion. *J Neurosci* 28:14481–14485. [CrossRef Medline](#)
- Hesselmann G, Sadaghiani S, Friston KJ, Kleinschmidt A (2010) Predictive coding or evidence accumulation? False inference and neuronal fluctuations. *PLoS One* 5:e9926. [CrossRef Medline](#)
- Hohwy J (2012) Attention and conscious perception in the hypothesis testing brain. *Front Psychol* 3:96. [Medline](#)
- Jokisch D, Jensen O (2007) Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J Neurosci* 27:3244–3251. [CrossRef Medline](#)

- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685. [CrossRef Medline](#)
- Kenet T, Bibitchkov D, Tsodyks M, Grinvald A, Arieli A (2003) Spontaneously emerging cortical representations of visual attributes. *Nature* 425:954–956. [CrossRef Medline](#)
- Kersten D, Mamassian P, Yuille A (2004) Object perception as Bayesian inference. *Annu Rev Psychol* 55:271–304. [CrossRef Medline](#)
- Kok P, Failing MF, de Lange FP (2014) Prior expectations evoke stimulus templates in the primary visual cortex. *J Cogn Neurosci* 26:1546–1554. [CrossRef Medline](#)
- Kok P, Brouwer GJ, van Gerven MA, de Lange FP (2013) Prior expectations bias sensory representations in visual cortex. *J Neurosci* 33:16275–16284. [CrossRef Medline](#)
- Laufs H, Kleinschmidt A, Beyerle A, Eger E, Salek-Haddadi A, Preibisch C, Krakow K (2003a) EEG-correlated fMRI of human alpha activity. *Neuroimage* 19:1463–1476. [CrossRef Medline](#)
- Laufs H, Krakow K, Sterzer P, Eger E, Beyerle A, Salek-Haddadi A, Kleinschmidt A (2003b) Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proc Natl Acad Sci U S A* 100:11053–11058. [CrossRef Medline](#)
- Lund TE, Nørgaard MD, Rostrup E, Rowe JB, Paulson OB (2005) Motion or activity: their role in intra- and inter-subject variation in fMRI. *Neuroimage* 26:960–964. [CrossRef Medline](#)
- Macmillan NA, Douglas C (2005) *Detection theory: a user's guide*, Ed 2. Mahwah, NJ: Lawrence Erlbaum Associates.
- Moosmann M, Ritter P, Krastel I, Brink A, Thees S, Blankenburg F, Taskin B, Obrig H, Villringer A (2003) Correlates of alpha rhythm in functional magnetic resonance imaging and near infrared spectroscopy. *Neuroimage* 20:145–158. [CrossRef Medline](#)
- Munneke J, Heslenfeld DJ, Theeuwes J (2008) Directing attention to a location in space results in retinotopic activation in primary visual cortex. *Brain Res* 1222:184–191. [CrossRef Medline](#)
- SanMiguel I, Widmann A, Bendixen A, Trujillo-Barreto N, Schröger E (2013) Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *J Neurosci* 33:8633–8639. [CrossRef Medline](#)
- Sasaki Y, Rajimehr R, Kim BW, Ekstrom LB, Vanduffel W, Tootell RB (2006) The radial bias: a different slant on visual orientation sensitivity in human and nonhuman primates. *Neuron* 51:661–670. [CrossRef Medline](#)
- Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci* 8:657–661. [Medline](#)
- Scheeringa R, Fries P, Petersson KM, Oostenveld R, Grothe I, Norris DG, Hagoort P, Bastiaansen MC (2011) Neuronal dynamics underlying high- and low-frequency EEG oscillations contribute independently to the human BOLD signal. *Neuron* 69:572–583. [CrossRef Medline](#)
- Schölvinck ML, Friston KJ, Rees G (2012) The influence of spontaneous activity on stimulus processing in primary visual cortex. *Neuroimage* 59:2700–2708. [CrossRef Medline](#)
- 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 U S A* 106:19569–19574. [CrossRef Medline](#)
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756. [CrossRef Medline](#)
- Supèr H, van der Togt C, Spekreijse H, Lamme VA (2003) Internal state of monkey primary visual cortex (V1) predicts figure-ground perception. *J Neurosci* 23:3407–3414. [Medline](#)
- Sylvester CM, Shulman GL, Jack AI, Corbetta M (2009) Anticipatory and stimulus-evoked blood oxygenation level-dependent modulations related to spatial attention reflect a common additive signal. *J Neurosci* 29:10671–10682. [CrossRef Medline](#)
- Thut G, Nietzel A, Brandt SA, Pascual-Leone A (2006)  $\alpha$ -Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci* 26:9494–9502. [CrossRef Medline](#)
- van Dijk H, Schoffelen JM, Oostenveld R, Jensen O (2008) Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci* 28:1816–1823. [CrossRef Medline](#)
- Watson AB, Pelli DG (1983) QUEST: a Bayesian adaptive psychometric method. *Percept Psychophys* 33:113–120. [CrossRef Medline](#)
- Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci* 20:RC63. [Medline](#)
- Wyart V, Nobre AC, Summerfield C (2012) Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proc Natl Acad Sci U S A* 109:3593–3598. [CrossRef Medline](#)