Behavioral/Cognitive

# Apparent Motion Induces Activity Suppression in Early Visual Cortex and Impairs Visual Detection

Lu Shen,[1,2,3] Biao Han,[1,2,3]* and Floris P. de Lange[3]

[1]School of Psychology, South China Normal University, Guangzhou, 510631, China, [2]Key Laboratory of Brain, Cognition and Education Sciences, South China Normal University, Guangzhou, 510631, China, and [3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, 6500 HB, The Netherlands

Apparent motion (AM) is induced when two stationary visual stimuli are presented in alternating sequence. Intriguingly, AM leads to an impaired detectability of stimuli along the AM path (i.e., AM-induced masking). It has been hypothesized that AM triggers an internal representation of a moving object in early visual cortex, which competes with stimulus-evoked representations of visual stimuli on the motion path in early visual cortex of 25 human adults (16 female). We tested this hypothesis by measuring BOLD responses in early visual cortex during the process of AM-induced masking, using fMRI and population receptive field methods. Surprisingly, and counter to our hypothesis, we showed that AM suppressed, rather than increased, BOLD responses along early visual (V1 and V2) representations of the AM path, including regions that were not directly activated by the AM inducer stimuli. This activity suppression of the visual response predicted the subsequent reduction in detectability of the target that appeared in the middle of the AM path. Our data thereby provide direct empirical evidence for suppressive neural mechanisms underlying AM and suggest that illusory motion can render us blind to objects on the motion path by suppressing neural activity at the earliest cortical stages of visual perception.

*Key words:* apparent motion; early visual cortex; fMRI; masking; population receptive field

## Significance Statement

When two spatially distinct visual objects are presented in alternating sequence, apparent motion (AM) occurs and impairs detectability of stimuli along its path. The underlying mechanism is thought to be that increased activation in human early visual cortex evoked by AM interferes with the representation of the stimulus. Strikingly, however, we show that AM suppresses neural activity along the motion path, and the strength of activity suppression predicts the subsequent behavioral performance decrement in terms of detecting a stimulus along the AM path. Our findings provide empirical evidence for a suppressive, rather than faciliatory, mechanism underlying AM.

## Introduction

When two stationary visual stimuli are presented in alternating sequence, they will often induce the illusory perception of a moving stimulus, that is, apparent motion (AM) (Wertheimer, 1912). The experience of AM has been considered as a case of perceptual filling-in in early visual cortex, during which a percept is internally generated at a location that is not physically stimulated

(Pessoa and De Weerd, 2003). The neural mechanisms underlying AM are not fully understood. Some neuroimaging studies reported that AM evokes activation in primary visual cortex (V1) along the illusory path of AM (Muckli et al., 2005; Sterzer et al., 2006), which may be driven by feedback from higher visual areas involved in motion (MT/V5) (Sterzer et al., 2006) or form processing (visual ventral regions) (Ferrera et al., 1994; Zhuo et al., 2003). However, other studies only observed AM-related activity in higher motion processing areas, but no stronger activation in early visual areas (Mikami et al., 1986; Goebel et al., 1998; Muckli et al., 2002; Liu et al., 2004).

Intriguingly, the detection and identification of a simple visual form in the path of AM are impaired by AM, which has been taken as indirect evidence of AM-related activation in V1 (Yantis and Nakama, 1998; Hidaka et al., 2011). For instance, Hidaka et al. (2011) observed a reduced detectability of a target stimuli in the AM path; this masking effect of the target was maximal when the target orientation was the same as the inducers. The authors argued that the AM inducers may evoke responses in V1 neurons

that are tuned to the orientation of the inducers and to the locations along the illusory path. This AM-evoked activation may then compete with the neural response to the target on the path, thus impairing the visibility of the target.

However, AM masking could potentially be understood in a different manner. When elements are grouped into coherent shapes, this results in higher activity in higher visual areas (representing the shape) and concurrently reduces activity in early visual areas (Murray et al., 2002). This reduction of activity in early visual areas may also result in reduced visual sensitivity, as well as impaired detection performance (Ress et al., 2000; Jacobs et al., 2012). Therefore, if AM induces a suppression of activity in early visual cortex along the motion path, this reduced excitability may subsequently impair the visibility of stimuli presented along the whole path. A recent neurocomputational model that used a V1–like population code model of early visual processing (Van Humbeeck et al., 2016) indeed predicts strong suppression, rather than activation, of early sensory responses during AM.

To empirically test these hypotheses, we created an AM percept by repeatedly alternating two oriented Gabor gratings. Following this, a target grating appeared in the middle of the AM path. During this process, we measured BOLD response in early visual cortex along the path using fMRI and population receptive field (pRF) methods. If AM masking is indeed the result of AM-induced activation, we expected to observe increased activity during AM compared with a control condition in which no AM was induced. Strikingly, however, our results showed that, instead of increased activation, AM induced suppressed activity in early sensory areas along the whole illusory motion path before the target was presented. This suppression further predicted the impaired detectability of the subsequent target.

## Materials and Methods

*Data availability.* All data and code used for stimulus presentation and analysis are available from the Donders Institute for Brain, Cognition, and Behavior repository at https://data.donders.ru.nl.

*Participants.* Twenty-seven right-handed participants were recruited in the present study. Sample size was decided *a priori*, and ensured at least 80% power to detect experimental effects with at least moderate to large effect size (Cohen's $d > 0.6$). All participants gave their written informed consent in accordance with the institutional guidelines of the local ethical committee (CMO region Arnhem-Nijmegen, The Netherlands) and received monetary compensation for their participation. All participants had normal or corrected-to-normal visual acuity. All participants were invited to participate in two separate scanning sessions within maximally 2 weeks' time. One participant completed only one of the two sessions and 1 participant was excluded due to excessive head motion. Only the remaining 25 participants (16 female, mean age 25.5 years) were included in all analyses.

*Stimuli.* All stimuli used in the experiment were Gabor patches, created by multiplying a cosine grating with a 2D Gaussian envelope. The spatial frequency of all gratings was 1.5 cycles per degree. Stimuli were displayed on a gray background (Michelson contrast of 50%). Two types of visual display were used in the main experiment: AM and flicker (FL). Both the AM and FL inducer stimuli had a Michelson contrast of 100%, while the target stimulus had a 30% Michelson contrast. The orientation of the inducer gratings is 45° or 135°. Target orientation is the same or different from the inducers. The target grating was presented at 4° eccentricity right from a fixation point (0.05°). The inducers were vertically separated by 10°, and the target stimulus was presented exactly in between the two inducers, at a distance of 5° from each inducer.

For the AM condition, the inducers were presented for 100 ms alternately at the top and bottom position at the right side of the screen with an interstimulus interval of 150 ms, corresponding to a motion frequency at 2 Hz. This specific frequency was chosen based on previous

results showing that the perception of long-range AM and motion masking is optimal for presentation rates between 1.5 and 3 Hz (Finlay and von Grünau, 1987; Selmes et al., 1997; Yantis and Nakama, 1998). This AM sequence was repeated 16 times to induce a strong percept of stimuli moving back and forth. The target was flashed briefly for 20 ms during the last AM sequence, 65 ms after the presentation of the inducer at the top position, and at an intermediate position in the interstimulus interval. In the end of each trial, the observers were asked to respond whether the target was present or not. For the FL condition, the two inducers were presented in synchrony. The presentation frequency of the FL stimuli could be the slow (i.e., same as the motion frequency of AM, 2 Hz), or fast (i.e., same as the stimulus presentation frequency in AM sequence, 4 Hz). The two different FL sequences were set to control the influence of bottom-up stimulation.

*Experimental procedures.* The experiment consisted of two fMRI sessions. Each of the two fMRI sessions lasted ∼2 h. During the two sessions, participants were asked to perform 12 or 14 experimental runs, 1 functional localizer run, and 4 functional runs with the moving bar sequence to estimate pRFs (Dumoulin and Wandell, 2008).

Each experimental run contained 24 trials, consisting of 12 AM sequence trials (4 with same target, 4 with different target, 4 with target-absent), and 12 FL sequence trials (4 with same target, 4 with different target, 4 with target-absent). The order of stimulus conditions was randomized. Each trial lasted 19.8 s (corresponding to 22 fMRI volumes), consisting of a 9.8 s stimulus sequence, 1.5 s response, and 8.5 s intertrial interval. Each run lasted 8 min and started with 4.5 s of fixation that was discarded from the analysis. Each run contained either fast or slow FL sequence trials. The order of fast and slow FL sequences was pseudo-randomized over runs, with the restriction that consecutive runs containing the same FL sequence should not exceed three times.

The functional localizer was used for a functional definition of three stimuli locations represented in early visual cortex for each participant. The grating stimuli in each of three locations was presented 15 times, each time flashing at 2 Hz (250 ms on, 250 ms off) for 10 s. The order of three stimulus locations was pseudorandomized with the restriction that the two consecutively presented trials were different. It also contained six null events of 10 s in which only a fixation was displayed. The null events occurred at random positions throughout the localizer run. Participants were instructed to fixate the fixation dot and respond by button press whenever the fixation changed color. To ensure stable fixation, an infrared eye tracker (SensoMotoric Instruments) was used to monitor eye movements online. If the participant's gaze moved away from the central fixation (e.g., toward the inducer), the experiment would be paused and participants would be reminded to maintain fixation.

At the end of the experiment, participants performed four functional runs with moving bar stimuli to map the pRFs of voxels in early visual cortex. During these runs, bars with full contrast contrast-reversing checkerboards (2 Hz) moved across the screen in a circular aperture with a diameter of 11°. The bars moved in eight different directions (four cardinal and four diagonal directions) in 11 steps of 1°, one step per TR (900 ms); a colored fixation dot was presented in the center of the screen. Participants were instructed to fixate the fixation dot and respond by button press whenever the fixation changed color.

*fMRI data acquisition.* Functional and anatomic images were acquired using a 3T Prisma MRI system (Siemens) equipped with a 64-channel head coil. Functional activity was measured using a T2*-weighted multiband-4 sequence (60° flip angle, voxels size $2 \times 2 \times 2$ mm, TR/TE = 900/39.8 ms, 30 transversal slices). Structural images were acquired using a T1-weighted MP-RAGE sequence (GRAPPA acceleration factor = 2, 8° flip angle, voxel size $1 \times 1 \times 1$ mm, TR/TE = 2300/3.03 ms).

*pRF estimation.* The data from the moving bar runs were used to estimate the pRF of each voxel in the functional volumes we obtained, using MrVista (http://white.stanford.edu/software/). Before estimation, the BOLD time courses per voxel from the four runs were averaged together. During estimation, a predicted BOLD signal is calculated from the known stimulus parameters and a model of the underlying neuronal population. The model of the neuronal population consisted of a 2D

Gaussian pRF, with parameters x0, y0, and s, where x0 and y0 are the coordinates of the center of the receptive field, and s indicates its spread (SD), or size. All parameters were stimulus-referred, and their units were degrees of visual angle. These parameters were adjusted to obtain the best possible fit of the predicted to the actual BOLD signal. The goodness of this fit was expressed as proportion of each voxel's variance explained by its pRF model. For all subsequent analyses, voxels were considered "visually active" when at least 40% of their variance was explained by the pRF model. With increasing receptive field size, voxels will respond to multiple locations. In other words, the smaller the pRF size, the less overlap in response profiles and thus the more accurate the responses of voxels, but also the smaller the number of the selected voxels, leading to noisier estimates. In order to prevent overlap in response profiles while keeping a sufficient number of selected voxels, the following analysis was further restricted to voxels with a pRF size $\leq 4°$, where the number of selected voxels started to asymptote. For details of this procedure, see Dumoulin and Wandell (2008). This method has been shown to reconstruct the cortical visual field map more accurately than conventional retinotopic mapping methods, as well as produce pRF size estimates that agree well with electrophysiological receptive field measurements in monkey and human visual cortex.

*Preprocessing of fMRI data.* Functional images were preprocessed using FSL (Oxford) including motion correction (six-parameter affine transform), temporal high-pass filtering (128 s), and spatial smoothing using a Gaussian kernel of 5 mm FWHM for each run separately. All analyses were conducted in the native subject space.

*Specification of ROIs.* For the functional localizer, onsets and durations of the stimuli were convolved with a double-$\gamma$ HRF and fitted using a GLM. For each subject, activation associated with a particular stimulus location was revealed by a *t* test between one location and the other two locations (e.g., for the middle location, the contrast was "middle – 0.5 * (top + bottom)"). Next, V1 and V2 were determined using the automatic cortical parcellation provided by FreeSurfer based on individual T1 images. The combined V1 and V2 mask was further restricted to voxels with receptive fields along the AM path. Within the combined V1 and V2 regions, the 100 most active voxels based on *t* values were selected as the ROI for the specific location. To note, all reported results are based on these 100 most active voxels in V1 and V2, unless specified otherwise.

*Behavioral analyses.* For behavioral data, to measure the detection sensitivity in the context of AM masking, it is important to consider both the proportion of hits/misses and false alarms/correct rejections. Therefore, we computed $d'$, an index of detection sensitivity for the target, on the basis of signal detection theory (Macmillan and Creelman, 2005). The responses of target-present were regarded as "hits" in the trials with a target and as "false alarms" in the trials without a target. The proportions of hits and false alarms with 0% or 100% were corrected as $1/n$ or $(n – 1)/n$, respectively, where $n$ was the total number of presented trials (Anscombe, 1956; Sorkin, 1999). The $d'$ values were submitted to a 2 (inducer condition: AM vs FL) × 2 (target condition: same vs different relative to the inducer stimulus) repeated-measures ANOVA.

*BOLD signal analyses.* For each voxel in the ROIs, the BOLD time course was extracted separately for each condition. The activity at time 0 (stimulus onset) was used as the baseline to calculate percent signal change. Time courses for each condition were then averaged over trials and runs for each ROI. The mean activity within 0.9-9 s window relative to the inducer stimulus was used to index the magnitude of the response during the inducing period. The mean activity within 10.8-17.1 s was used to index the magnitude of the response during the post-target period. Two-tailed paired *t* test was used to assess the significance of time course differences in magnitude between different conditions for inducing period and post-target period separately.

*Stimulus reconstruction.* The estimated pRF parameters allowed a straightforward and intuitive reconstruction of the BOLD effects from cortical space to visual space. Each voxel's receptive field can be represented by a 2D Gaussian, with peak coordinates (x0, y0) and SD s. The reconstruction in visual space consisted of the sum of the 2D Gaussians of all voxels in a given visual area, weighted by the voxels' BOLD response as follows:

$$\sum_{i=1}^{n} (a_i - b_i) \cdot g(x0_i, y0_i, s_i)$$

where $n$ is the number of voxels in a given area, $a_i$ and $b_i$ are responses to certain conditions, and $g(x0_i, y0_i, s_i)$ is the 2D Gaussian defining the voxel's receptive field. The rationale is the following: if a voxel in V1–V2 is highly activated, then this reflects activity in V1–V2 neurons corresponding to the region of visual space modeled by the 2D Gaussian. In order to represent this activity in visual space, we multiplied the voxel's 2D Gaussian receptive field with its activity (i.e., BOLD signal change) and projected the result on a 2D map of visual space. By doing this for all V1–V2 voxels, we obtained a reconstruction of the BOLD signal in V1–V2 in visual space.

To show the spatial specificity of the activity spread, a pRF-based stimulus construction was conducted based on voxels in V1–V2 with a pRF size $\leq 4°$, covering the visual space from $x = 2°$ to $6°$ and from $y = -7°$ to $7°$. For example, to reconstruct the response to the middle location versus the other two locations in functional localizers, the contrast between the parameter estimates of the middle location and the other two locations was calculated for each voxel, and these values were used as voxel weights multiplied with voxels' 2D Gaussian defined by pRF estimates, and then averaged over the voxel dimension, resulting in a stimulus reconstruction. Similarly, to reconstruct the BOLD difference induced by different experimental conditions (e.g., AM vs FL; target-present vs target absent), the average BOLD difference between the two conditions of each voxel was used for the reconstruction.

*Calculation of relative suppression.* To further characterize the magnitude of AM-induced suppression, we calculated the relative suppression index as follows:

$$Relative\ suppression\ index = \frac{BOLD_{AM} - BOLD_{FL}}{|BOLD_{AM}|}$$

where $BOLD_{AM}$ and $BOLD_{FL}$ are the BOLD amplitude averaged over inducing period (0.9–9 s) for AM and FL conditions, respectively, for each of the three ROIs representing the three stimuli locations. One-sample *t* test was used to assess the significance of the suppression index compared with 0. A one-way ANOVA was further used to test the difference in the relative suppression index between these three locations.

## Results

In the AM condition, two repeatedly alternated grating stimuli induced a strong percept of AM along a vertical path at the right side of the screen. A target grating with low contrast (30% Michelson contrast) was presented in the middle of the path in 66.6% of all trials. In the control FL condition, the two inducers were presented in synchrony, which abolished the motion percept. To control for differences in bottom-up visual stimulation, we used two presentation rates during FL, either fast or slow. The fast FL sequence had the same stimulus presentation frequency as the AM sequence (and thus double the amount of visual stimulation), while the slow FL had the same amount of physical stimulation (but half the rate) at each presentation location as the AM. The orientation difference between the target and inducer was either 0° (i.e., Same) or 90° (i.e., Different) (Fig. 1A; see also Materials and Methods).

### AM-induced masking effect

For each condition, we computed $d'$, an index of detection sensitivity for the target, on the basis of signal detection theory (Macmillan and Creelman, 2005). As expected, we observed
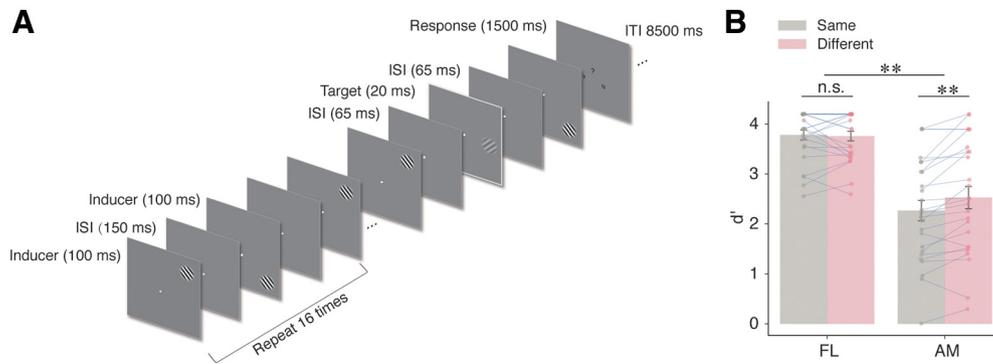
**Figure 1.** Paradigm and behavioral results. **A**, Illustration of one trial sequence in the AM condition. Participants were instructed to detect whether the target presented or not. **B**, AM resulted in a reduced detection performance ($d'$) compared with FL. $d'$ was lower for targets with the same orientation as the inducers compared with targets with different orientation as the inducers in AM condition. Error bars indicate $\pm$ standard error of the mean (SEM). n.s., $p > 0.05$. **$p < 0.001$.

strong masking by AM (AM: $d' = 2.40 \pm 0.21$, mean $\pm$ SEM; FL: $d' = 3.77 \pm 0.09$; $F_{(1,24)} = 58.59$, $p = 6.82 \times 10^{-8}$). The impairment in sensitivity induced by AM, compared with FL, was orientation-specific (inducer $\times$ stimulus interaction: $F_{(1,24)} = 9.64$, $p = 0.005$). Specifically, only in the AM condition, detection performance for targets with the same orientation as the inducers was significantly lower compared with targets with the orthogonal orientation as the inducers ($t_{(24)} = -4.01$, $p = 5.17 \times 10^{-4}$), while there was no significant difference in the FL condition ($t_{(24)} = 0.37$, $p = 0.35$) (Fig. 1B), indicating orientation tuning in AM masking. This orientation tuning effect in AM replicates previous findings (Hidaka et al., 2011; Van Humbeeck et al., 2016).

**AM-induced suppression in early visual cortex**
We defined cortical ROIs within the early visual cortex (combined V1 and V2 areas) representing the three stimulus positions for each participant using independent functional localizers (see Materials and Methods). BOLD signals were extracted from these three ROIs (representing the upper, middle, and lower right visual field) for the AM and FL conditions, respectively. There was significantly lower BOLD activity in the ROIs representing top (Fig. 2A, top) and bottom right locations (Fig. 2A, bottom), where the inducer stimuli were presented, during the AM condition compared with the FL condition, during both the inducing period (0.9-9 s relative to the inducer stimulus onset; top location: $t_{(24)} = -5.04$, $p = 3.68 \times 10^{-5}$; bottom location: $t_{(24)} = -3.96$, $p = 5.7 \times 10^{-4}$) and post-target period (10.8-17.1 s relative to the inducer stimulus onset; top location: $t_{(24)} = -2.50$, $p = 0.01$; bottom location: $t_{(24)} = -4.88$, $p = 5.53 \times 10^{-5}$). Crucially, for the ROI representing the middle position, there was no physical stimulus presented before the target onset; however, there was a reliable reduction of activity during AM compared with FL in this location in the period before the onset of the target ($t_{(24)} = -2.10$, $p = 0.046$), indicating AM-related suppression in the early visual cortex. After the target was presented, there were strong evoked responses in both AM and FL conditions, and no significant differences between the two conditions ($t_{(24)} = -1.07$, $p = 0.29$) (Fig. 2A, middle). As the response patterns in V1 and V2 were similar (Fig. 3), and no significant differences in the suppression effect between V1 and V2 were found for all three ROIs (two-sample $t$ test: top location: $t_{(48)} = 1.55$, $p = 0.12$; middle location: $t_{(48)} = 0.39$, $p = 0.69$; bottom location: $t_{(48)} = 1.11$, $p = 0.27$), we did not distinguish between V1 and V2 in the analysis.

Because AM led to reduced activity at both the bottom and top locations (where the inducer stimuli appeared) and the

middle location (where no stimulus was presented), one might wonder whether the difference in the middle location might be due to the spreading of differences in feedforward input between AM and FL in the top and bottom locations. First, to illustrate the spatial specificity of stimulus-evoked activity, we performed a pRF-based reconstruction of the functional localizer during which the three stimuli in the bottom, middle, and top locations were presented, based on all voxels in V1 and V2 (see Materials and Methods). As can be seen from Figure 2B, physically presenting the stimuli at each of the three positions triggered higher activity that was strictly limited to the corresponding retinotopic location (Fig. 2B). This renders it less likely that activity difference in the middle location is simply due to a spatial spreading of the activity difference for the top/bottom locations. Moreover, when comparing the BOLD response to the fast and slow FL trials, which have different amount of feedforward input in the inducer locations, we found that the fast FL evoked higher activity than the slow FL during the inducing period in both the top (Fig. 4A; FL_fast vs FL_slow: $t_{(24)} = 3.11$, $p = 4.69 \times 10^{-3}$) and bottom locations (Fig. 4C; FL_fast vs FL_slow: $t_{(24)} = 4.03$, $p = 4.86 \times 10^{-4}$) but not in the middle (Fig. 4B; FL_fast vs FL_slow: $t_{(24)} = 0.52$, $p = 0.60$), indicating that the feedforward input induced activity differences were constrained to the stimulus location and did not spread spatially. Furthermore, when comparing the AM condition with the slow FL condition, which had the exactly same amount feedforward input at the inducers, the suppression effects were still presented during the inducing period in the three locations (Fig. 4; FL_slow vs AM; top location: $t_{(24)} = 4.83$, $p = 6.24 \times 10^{-5}$; middle location: $t_{(24)} = 2.19$, $p = 0.03$; bottom location: $t_{(24)} = 3.76$, $p = 9.58 \times 10^{-4}$). Therefore, the AM-induced suppression in the three locations appears to be the result of the illusory motion and cannot be simply explained by bottom-up stimulus differences. Our results were largely independent of the voxel selection procedure that we employed (Fig. 5).

**AM-induced suppression predicts subsequent masking**
Next, we examined whether the observed AM-related suppression was behaviorally relevant, in terms of predicting subsequent masking of the target. For this, we capitalized on the intersubject variability in terms of the efficacy of masking induced by AM. We collapsed across the same and different target conditions to calculate the behavioral $d'$ difference between AM and FL and average BOLD difference between AM and FL during the inducing period, respectively. Then, Pearson correlations were
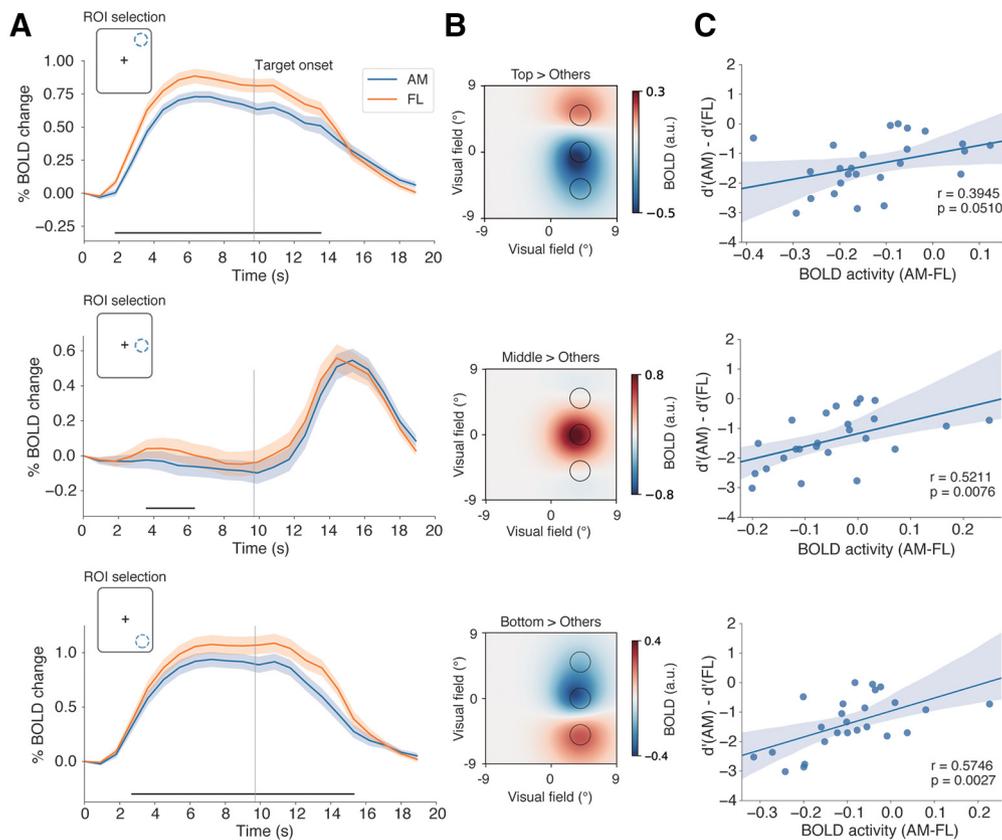
**Figure 2.** AM-induced suppression in early visual cortex (V1/V2) and the strength of suppression predicted the impaired detectability of the target. *A*, Average BOLD signal change for each of three ROIs in combined V1 and V2 areas, representing three stimuli locations, is plotted for the AM and FL conditions, respectively. To avoid double dipping, horizontal black bar represents significant time points ($p < 0.05$, uncorrected) just for visualization purpose. Shaded areas represent ± SEM. *B*, pRF-based reconstruction of the stimuli presented at the three locations in functional localizer. Images were obtained by weighting all voxels' Gaussian receptive fields in combined V1 and V2 by the respective activity in terms of $z$ values in each condition and then averaging these responses over all pRFs. Black circles represent the spatial position of the stimuli. *C*, Correlation between the average BOLD difference in the inducing period (0.9-9 s relative to the inducer stimulus onset) in different ROIs and the behavioral $d'$ difference between AM and FL conditions across subjects.

calculated between the $d'$ difference (indexing the amount of motion masking) and BOLD differences in V1/V2 at the three aforementioned retinotopic locations across subjects. For the middle location, AM-induced suppression was strongly correlated with AM-induced masking (Fig. 2C, middle; $r = 0.52$, $p = 0.0076$). In other words, the more AM suppressed early visual cortex, the more the detectability of the target was decreased. Interestingly, similar trends could also be found in the top (Fig. 2C, top; $r = 0.39$, $p = 0.051$) and bottom positions (Fig. 2C, bottom; $r = 0.57$, $p = 0.0027$). These results suggest that AM generates a suppression along the entire motion path that may impair the visibility of a visual target presented at any location on the AM path.

**AM suppresses the whole AM path**
We further characterized the relative suppression effect of AM during the inducing period (0.9-9 s relative to the inducer stimulus onset) by calculating the relative suppression index (see Materials and Methods). The lower the AM induced activity in early visual cortex compared with the FL condition, the more negative the relative suppression index will be. Consistent with Figure 2A, all regions of V1–V2 representing the three locations showed a significant suppression effect (Fig. 6A; top location: $t_{(24)} = -3.24$, $p = 1.74 \times 10^{-3}$; middle location: $t_{(24)} = -2.19$, $p = 0.019$; bottom location: $t_{(24)} = -4.84$, $p = 3.07 \times 10^{-5}$). Moreover, the strength of suppression between these three locations was significantly different ($F_{(2,74)} = 3.89$, $p = 0.025$).

Specifically, the relative suppression in the middle location was significantly larger than both the top and bottom locations (*post hoc* LSD test, $p$ values < 0.05).

Since the AM suppressed the three locations of the AM path, we reasoned that AM might suppress the whole illusory path. To verify this, we further reconstructed the average BOLD difference between AM and FL during the inducing period (0.9-9 s relative to the inducer stimulus onset), based on all voxels in V1 and V2. The visualization clearly showed that the suppression was distributed over the whole AM path (Fig. 6B), indicating the suppressive influence of AM on the whole illusory path.

**Target evokes a focal activation pattern**
We further examined the target evoked activity during post-target period (10.8-17.1 s relative to the inducer stimulus onset). The BOLD signal in the middle location was extracted from target-present and target-absent trials (collapsed over all AM and FL trials), respectively. Unsurprisingly, after the target onset, the target-present evoked higher response than target-absent in the middle location (Fig. 7A; $t_{(24)} = 4.60$, $p = 1.14 \times 10^{-4}$). Both the AM and FL conditions showed a similar pattern (Fig. 7C, AM: $t_{(24)} = 4.33$, $p = 2.26 \times 10^{-4}$; Fig. 7D, FL: $t_{(24)} = 3.76$, $p = 9.58 \times 10^{-4}$). Of note, for both the AM and FL conditions, there was no significant BOLD difference between same and different target during the post-target period (all $p$ values > 0.05). We further reconstructed the average BOLD difference between the target-present and target-absent conditions during post-
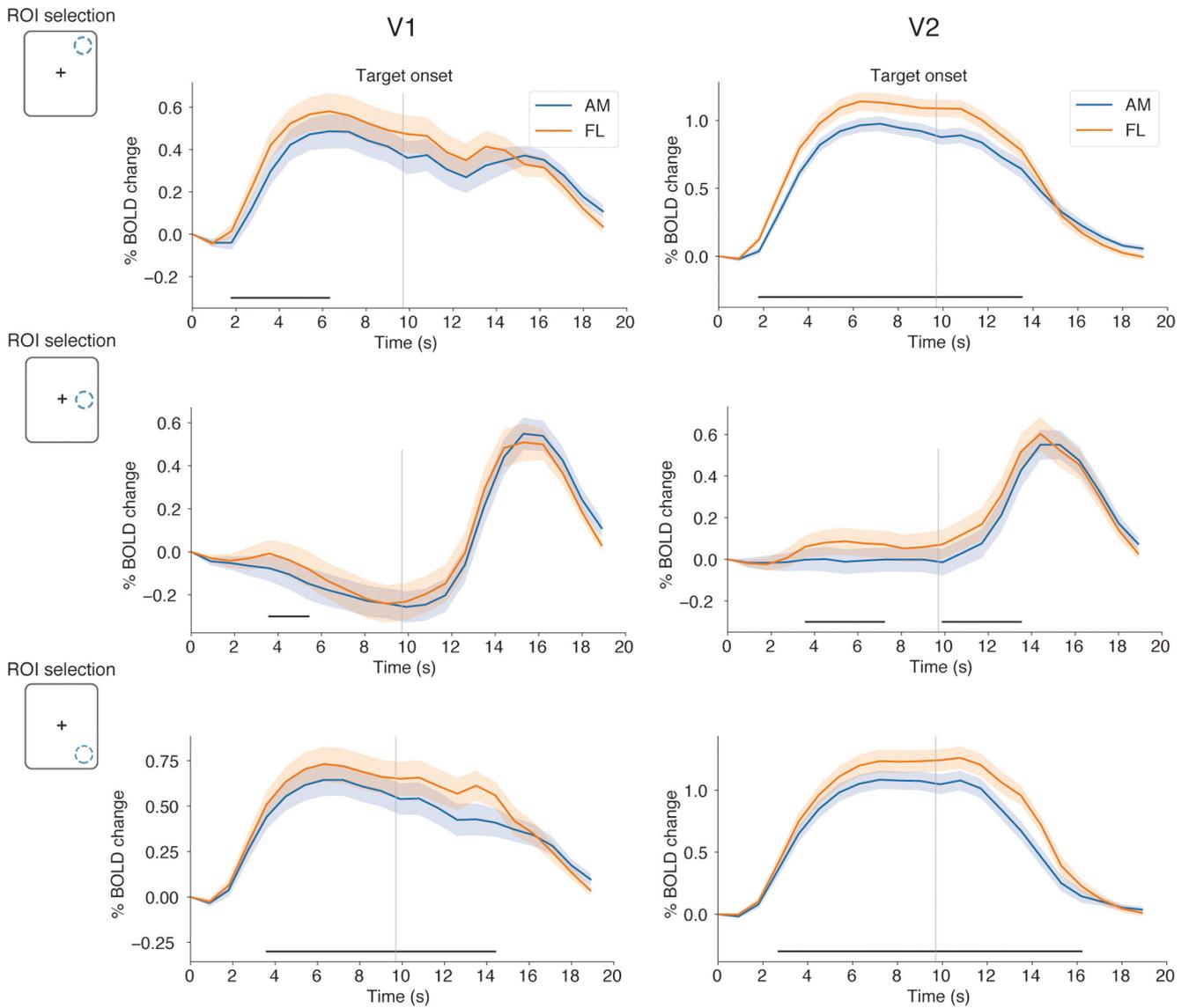
**Figure 3.** AM-induced suppression in V1 and V2, respectively. Average BOLD signal change of AM and FL conditions is extracted from each of three locations represented in V1 (left column) and V2 (right column), separately. Horizontal black bar represents significant time points ($p < 0.05$). Shaded areas represent $\pm$ SEM.
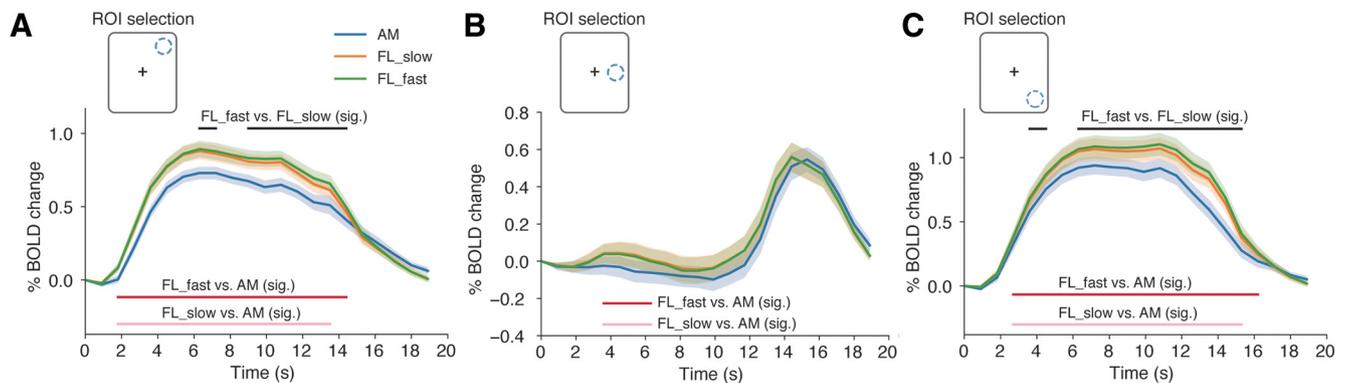


**Figure 4.** Effects of difference in feedforward input on AM-induced suppression. **A**, Average BOLD signal change for the ROI in combined V1 and V2 areas representing the top location, are plotted for the AM, fast FL, and slow FL conditions, respectively. Horizontal black bar represents significant time points indicated by the horizontal bar ($p < 0.05$), testing the difference between fast and slow FL conditions (black), fast FL and AM conditions (red), slow FL and AM conditions (pink), respectively. Shaded areas represent $\pm$ SEM. **B**, Same as **A**, but for the middle location. **C**, Same as **A**, but for the bottom location.
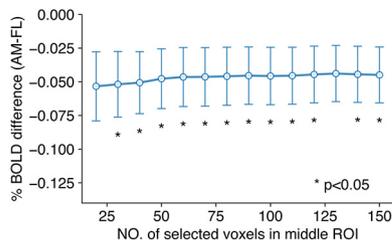
**Figure 5.** AM-induced suppression is independent of the number of voxels selected. Average BOLD difference between AM and FL during inducing period (0.9-9 s relative to the inducer stimulus onset) is plotted as a function of the number of selected voxels from ROI representing the middle location.
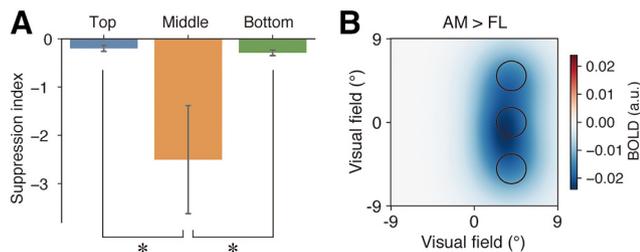


**Figure 6.** AM-induced suppression along the whole AM path. *A*, Relative suppression effects are presented in all three ROIs. The strength of relative suppression is significantly stronger in the middle location compared with the top and the bottom locations. *p < 0.05. *B*, pRF-based reconstruction of the average BOLD difference between AM and FL during inducing period (from 0.9 to 9 s relative to the inducer stimulus onset). Black circles represent the spatial position of the stimuli.

consistent with previous studies demonstrating increased activity in visual cortex in the absence of visual stimulation when subjects covertly directed their attention to a peripheral location expecting the onset of visual stimuli (Kastner et al., 1999; Murray, 2008).

## Discussion

In the present study, we examined how AM reduces the ability to detect stimuli appearing along the AM path. Our behavioral data indicate that AM indeed strongly impairs the detection of a visual target presented on the AM path, especially when the stimulus orientation matches the inducer's orientation. fMRI results further show that AM leads to a suppressed BOLD response along early visual representations of the whole AM path, including regions that are not directly activated by the AM inducer stimuli. This suppression of the visual response predicts the subsequent reduction in detectability of the visual target appearing in the middle of the AM path.

Contrary to the perceptual filling-in hypothesis, which suggests that AM masking is the result of AM-induced activation in early visual areas that competes with the response to the target (Yantis and Nakama, 1998; Pessoa and De Weerd, 2003; Hidaka et al., 2011), our results show that AM elicits activity suppression that subsequently reduces the detectability of a target. Specifically, greater suppression in V1/V2 resulted in worse detectability. This is in line with earlier neuroimaging studies in humans that observed a relationship between BOLD activity in early visual cortex and participants' detection (Ress et al., 2000) and discrimination ability (Boynton et al., 1999; Ress and Heeger, 2003). Consistent with these studies, AM masking may result from a reduced excitability of the early visual cortex caused by AM-induced suppression.

It may seem surprising that AM induces activity suppression in early visual cortex, instead of activation. The evidence for AM-induced activation is however mixed: whereas some studies found that AM led to activation in primary visual cortex along the AM path (Muckli et al., 2005; Sterzer et al., 2006), others only observed AM-related activity in higher areas, but no stronger activation in early visual areas (Mikami et al., 1986; Goebel et al., 1998; Muckli et al., 2002; Liu et al., 2004). Furthermore, optical imaging studies involving cats and monkeys have demonstrated suppressive effects in early visual cortex during various types of illusory motion percepts, for example, line-motion percepts (Jancke et al., 2004), temporal sequence of dark and bright stimuli eliciting motion percepts (Rekauzke et al., 2016), and AM (Chemla et al., 2019). A recent neurocomputational modeling study that used a highly similar experimental paradigm (Van Humbeeck et al., 2016) used a V1–like population code model of early visual processing, based on a standard contrast normalization model, to examine the cause of AM masking. In their model, masking of the visual target only occurs when V1 is suppressed by AM. Our results empirically confirm the



**Figure 7.** Target evoked response. *A*, Average BOLD signal change for target-present and target-absent conditions collapsed across AM and FL conditions. Target-present condition evokes increased activity compared with the target-absent condition. For visualization purpose, horizontal black bar represents significant time points (p < 0.05). *B*, pRF-based reconstruction of average BOLD difference between target-present and target-absent conditions during the post-target period (from 10.8 to 17.1 s relative to the inducer stimulus onset). *C*, Average BOLD signal change for target-present and target-absent in AM condition. *D*, Average BOLD signal change for target-present and target-absent in FL condition.

target period in the whole V1 and V2. The physical stimuli induced activation difference was focal, constrained to the middle location (Fig. 7B). It is worth noting that there was strong BOLD response in V1–V2 also when no target was presented. This is

their model, masking of the visual target only occurs when V1 is suppressed by AM. Our results empirically confirm the

predictions made by this computational model and strongly suggest that suppression of early sensory responses induced by AM causes masking of subsequent matching input.

How can AM lead to a suppression of activity in early visual cortex? One possibility is that this suppression is the result of feedback signal from higher-level areas to early visual cortex. These higher-level visual areas have larger receptive fields, allowing to determine the trajectory of long-range AM (Angelucci and Bullier, 2003; Angelucci and Bressloff, 2006). It has been suggested that feedback from MT to V1 plays a role in processing AM (Wibral et al., 2009; Vetter et al., 2015), as well as the involvement of ventral visual areas (Ferrera et al., 1994; Zhuo et al., 2003; Roe et al., 2012). Many studies have observed an inhibitory role of feedback signals from higher-level areas to low-level areas. Several studies have found decreased activity in early visual areas for more predictable stimuli (de Lange et al., 2018): for example, it has been shown that early visual areas respond less to coherent motion than incoherent motion (McKeefry et al., 1997; Harrison et al., 2007; Bartels et al., 2008), and to coherent shapes than randomly arranged lines (Murray et al., 2002). Therefore, AM-induced suppression in early visual cortex may be the result of inhibitory feedback from higher area MT or ventral visual areas. However, there is another possibility that AM-induced suppression can result from intracortical processing within early visual cortex. Some studies have demonstrated that local processing within V1 plays an important role in long-range AM (Jancke et al., 2004; Gerard-Mercier et al., 2016; Muller et al., 2018). The precise retinotopic map in V1 allows for more sophisticated neural computations for representing the trajectory of AM (Mumford, 1991; Lee et al., 1998). A recent monkey study demonstrated that a gain control mechanism within V1 can generate AM-related suppressive activity (Chemla et al., 2019), which helps higher areas read out motion information (Adelson and Bergen, 1985; Mumford, 1991, 1992). Because of the low temporal resolution of fMRI, we could not test these two possibilities directly in our study. Therefore, it still remains to be determined whether the AM-induced suppression in early visual cortex results from feedback in higher areas or local intracortical processing within early visual cortex.

The present results may appear to be at odds with some previous studies that also found reduced activity related to expectation (predictive feedback) but linked this to an increased (i.e., sharpened) sensory representation (Alink et al., 2010; Kok et al., 2012; Edwards et al., 2017). Several relevant points should be noted here. First, compared with the previous studies (Alink et al., 2010; Kok et al., 2012; Edwards et al., 2017; Richter et al., 2018; Han et al., 2019), the activity suppression that we observed is not occurring during the presentation of the stimulus, but precedes it. Furthermore, the AM-related suppression was observed on the whole AM path rather than only in the target area. These results suggest that the activity suppression induced by AM is not the result of an interaction between predictive feedback and a target stimulus, but rather precedes it and the cause of the reduced detectability. Second, in previous studies that observed sharpening effect (Alink et al., 2010; Kok et al., 2012; Edwards et al., 2017), expected and unexpected conditions were directly compared and predictive feedback existed in both conditions. In the current study, the suppressive feedback signal only existed in the AM condition but not in the FL condition. Third, how expectation affects the sensory representation is still under debate: some studies suggest that expectation may indeed dampen the sensory representation (Richter et al., 2018; Han et al., 2019) rather than sharpen it. As suggested by these studies, whether the predictive feedback is directly related to the task may be critical for whether sharpening or dampening of the sensory representation is observed. For example, in previous studies, which demonstrated that temporally expected targets were detected more often than temporally unexpected targets in the spatiotemporal dynamic context provided by AM (Schwiedrzik et al., 2007; Vetter et al., 2012; Edwards et al., 2017), the target detection task may focus participants' temporal attention on the AM-provided spatiotemporal contingencies, which has been found to sharpen stimulus representations, and which is associated with increased detectability (Rohenkohl et al., 2012). In the current study, spatiotemporal attention is identical for both the same and different targets, which were presented at the same spatiotemporal positions of the AM trace with equal probability. In other words, the information carried by the predictive feedback does not provide any information that aids the detection task, akin to several studies that observed dampening of representations (Meyer and Olson, 2011; Kumar et al., 2017; Richter et al., 2018; Han et al., 2019).

A speculative explanation for the observed sensory suppression in the present study is that it represents the interaction between the AM-provided predictive feedback and pretarget sensory noise in neuronal populations that have their receptive fields on the AM path (Faisal et al., 2008), leading to a form of "expectation suppression." Moreover, if the predictive feedback is feature-specific (Angelucci and Bressloff, 2006; Maunsell and Treue, 2006; Huh et al., 2018), sensory neurons representing the same orientation as present in the predictive feedback signal would be suppressed more than the neurons representing a different orientation, consistent with the predictions of a recent computational model (Van Humbeeck et al., 2016). Accordingly, the detectability for a "same" target would be lower than a "different" target.

In conclusion, our results demonstrate a suppressive mechanism underlying AM. Specifically, AM induces suppressed responses in V1/V2 along the entire illusory motion path, and the strength of this activity suppression predicts the strength of subsequent masking of visual stimuli. This suppression is in line with predictive coding models of cortical processing, which propose that higher-level predictions try to explain away lower-level responses to expected input.

## References

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2:284–299.

Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. J Neurosci 30:2960–2966.

Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog Brain Res 154:93–120.

Angelucci A, Bullier J (2003) Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? J Physiol Paris 97:141–154.

Anscombe FJ (1956) On estimating binomial response relations. Biometrika 43:461–464.

Bartels A, Zeki S, Logothetis NK (2008) Natural vision reveals regional specialization to local motion and to contrast-invariant, global flow in the human brain. Cereb Cortex 18:705–717.

Boynton GM, Demb JB, Glover GH, Heeger DJ (1999) Neuronal basis of contrast discrimination. Vision Res 39:257–269.

Chemla S, Reynaud A, di Volo M, Zerlaut Y, Perrinet L, Destexhe A, Chavane F (2019) Suppressive traveling waves shape representations of illusory motion in primary visual cortex of awake primate. J Neurosci 39:4282–4298.

de Lange FP, Heilbron M, Kok P (2018) How do expectations shape perception? Trends Cogn Sci 22:764–779.

Dumoulin SO, Wandell BA (2008) Population receptive field estimates in human visual cortex. Neuroimage 39:647–660.

Edwards G, Vetter P, McGruer F, Petro LS, Muckli L (2017) Predictive feedback to V1 dynamically updates with sensory input. Sci Rep 7:16538.

Faisal AA, Selen LP, Wolpert DM (2008) Noise in the nervous system. Nat Rev Neurosci 9:292–303.

Ferrera VP, Rudolph KK, Maunsell JH (1994) Responses of neurons in the parietal and temporal visual pathways during a motion task. J Neurosci 14:6171–6186.

Finlay D, von Grünau M (1987) Some experiments on the breakdown effect in apparent motion. Percept Psychophys 42:526–534.

Gerard-Mercier F, Carelli PV, Pananceau M, Troncoso XG, Frégnac Y (2016) Synaptic correlates of low-level perception in V1. J Neurosci 36:3925–3942.

Goebel R, Khorram-Sefat D, Muckli L, Hacker H, Singer W (1998) The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. Eur J Neurosci 10:1563–1573.

Han B, Mostert P, de Lange FP (2019) Predictable tones elicit stimulus-specific suppression of evoked activity in auditory cortex. Neuroimage 200:242–249.

Harrison LM, Stephan KE, Rees G, Friston KJ (2007) Extra-classical receptive field effects measured in striate cortex with fMRI. Neuroimage 34:1199–1208.

Hidaka S, Nagai M, Sekuler AB, Bennett PJ, Gyoba J (2011) Inhibition of target detection in apparent motion trajectory. J Vis 11:2.

Huh CY, Peach JP, Bennett C, Vega RM, Hestrin S (2018) Feature-specific organization of feedback pathways in mouse visual cortex. Curr Biol 28:114–120.e5.

Jacobs C, Goebel R, Sack AT (2012) Visual awareness suppression by prestimulus brain stimulation: a neural effect. Neuroimage 59:616–624.

Jancke D, Chavane F, Naaman S, Grinvald A (2004) Imaging cortical correlates of illusion in early visual cortex. Nature 428:423–426.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761.

Kok P, Jehee JF, de Lange FP (2012) Less is more: expectation sharpens representations in the primary visual cortex. Neuron 75:265–270.

Kumar S, Kaposvari P, Vogels R (2017) Encoding of predictable and unpredictable stimuli by inferior temporal cortical neurons. J Cogn Neurosci 29:1445–1454.

Lee TS, Mumford D, Romero R, Lamme VA (1998) The role of the primary visual cortex in higher level vision. Vision Res 38:2429–2454.

Liu T, Slotnick SD, Yantis S (2004) Human MT$^+$ mediates perceptual filling-in during apparent motion. Neuroimage 21:1772–1780.

Macmillan N, Creelman C (2005) Detection theory: a user's guide. Mahwah, NJ: Erlbaum.

Maunsell JH, Treue S (2006) Feature-based attention in visual cortex. Trends Neurosci 29:317–322.

McKeefry DJ, Watson JD, Frackowiak RS, Fong K, Zeki S (1997) The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. Neuroimage 5:1–12.

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

Mikami A, Newsome WT, Wurtz RH (1986) Motion selectivity in macaque visual cortex: II. Spatiotemporal range of directional interactions in MT and V1. J Neurophysiol 55:1328–1339.

Muckli L, Kriegeskorte N, Lanfermann H, Zanella FE, Singer W, Goebel R (2002) Apparent motion: event-related functional magnetic resonance imaging of perceptual switches and States. J Neurosci 22:RC219.

Muckli L, Kohler A, Kriegeskorte N, Singer W (2005) Primary visual cortex activity along the apparent-motion trace reflects illusory perception. PLoS Biol 3:e265.

Muller L, Chavane F, Reynolds J, Sejnowski TJ (2018) Cortical travelling waves: mechanisms and computational principles. Nat Rev Neurosci 19:255–268.

Mumford D (1991) On the computational architecture of the neocortex. Biol Cybern 65:135–145.

Mumford D (1992) On the computational architecture of the neocortex: II. The role of cortico-cortical loops. Biol Cybern 66:241–251.

Murray SO (2008) The effects of spatial attention in early human visual cortex are stimulus independent. J Vis 8:2.1–11.

Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. Proc Natl Acad Sci USA 99:15164–15169.

Pessoa L, De Weerd P (2003) Filling-in: from perceptual completion to cortical reorganization. New York: Oxford UP.

Rekauzke S, Nortmann N, Staadt R, Hock HS, Schöner G, Jancke D (2016) Temporal asymmetry in dark-bright processing initiates propagating activity across primary visual cortex. J Neurosci 36:1902–1913.

Ress D, Heeger DJ (2003) Neuronal correlates of perception in early visual cortex. Nat Neurosci 6:414–420.

Ress D, Backus BT, Heeger DJ (2000) Activity in primary visual cortex predicts performance in a visual detection task. Nat Neurosci 3:940–945.

Richter D, Ekman M, de Lange FP (2018) Suppressed sensory response to predictable object stimuli throughout the ventral visual stream. J Neurosci 38:7452–7461.

Roe AW, Chelazzi L, Connor CE, Conway BR, Fujita I, Gallant JL, Lu H, Vanduffel W (2012) Toward a unified theory of visual area V4. Neuron 74:12–29.

Rohenkohl G, Cravo AM, Wyart V, Nobre AC (2012) Temporal expectation improves the quality of sensory information. J Neurosci 32:8424–8428.

Schwiedrzik CM, Alink A, Kohler A, Singer W, Muckli L (2007) A spatio-temporal interaction on the apparent motion trace. Vision Res 47:3424–3433.

Selmes CM, Fulham WR, Finlay DC, Chorlton MC, Manning ML (1997) Time-till-breakdown and scalp electrical potential maps of long-range apparent motion. Percept Psychophys 59:489–499.

Sorkin RD (1999) Spreadsheet signal detection. Behav Res Methods Instrum Comput 31:46–54.

Sterzer P, Haynes JD, Rees G (2006) Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT$^+$/V5. Neuroimage 32:1308–1316.

Van Humbeeck N, Putzeys T, Wagemans J (2016) Apparent motion suppresses responses in early visual cortex: a population code model. PLoS Comput Biol 12:e1005155.

Vetter P, Edwards G, Muckli L (2012) Transfer of predictive signals across saccades. Front Psychol 3:176.

Vetter P, Grosbras MH, Muckli L (2015) TMS over V5 disrupts motion prediction. Cereb Cortex 25:1052–1059.

Wertheimer M (1912) Experimentelle studium uber das sehen von bewegung. Z Psychol Z Angew Psychol 61:161–265.

Wibral M, Bledowski C, Kohler A, Singer W, Muckli L (2009) The timing of feedback to early visual cortex in the perception of long-range apparent motion. Cereb Cortex 19:1567–1582.

Yantis S, Nakama T (1998) Visual interactions in the path of apparent motion. Nat Neurosci 1:508–512.

Zhuo Y, Zhou TG, Rao HY, Wang JJ, Meng M, Chen M, Zhou C, Chen L (2003) Contributions of the visual ventral pathway to long-range apparent motion. Science 299:417–420.