A neural basis for inference in perceptual ambiguity

Philipp Sterzer, and Andreas Kleinschmidt

PNAS published online Dec 26, 2006;
doi:10.1073/pnas.0609006104

A neural basis for inference in perceptual ambiguity

Philipp Sterzer[†‡§] and Andreas Kleinschmidt[¶‖]**

[†]Department of Neurology, Johann Wolfgang Goethe University, Theodor-Stern-Kai 7, D-60590 Frankfurt am Main, Germany; [‡]Department of Psychiatry, Charité University Hospital/St Hedwig Hospital, Turmstrasse 21, D-10559 Berlin, Germany; [¶]Institut National de la Santé et de la Recherche Médicale Unité 562, Cognitive Neuroimaging, 91401 Orsay Cedex, France, [‖]Commissariat à l'Energie Atomique, Direction des Sciences du Vivant, Département de Recherche Médicale, Service Hospitalier Frédéric Joliot, Orsay, 91401 Cedex, France; and **Université Paris-Sud, Orsay, 91405, France

Edited by Marcus E. Raichle, Washington University School of Medicine, St. Louis, MO, and approved November 8, 2006 (received for review October 11, 2006)

When looking at ambiguous visual stimuli, the observer experiences frequent spontaneous transitions between two competing percepts while physical stimulation remains unchanged. Despite recent advances in understanding the neural processes underlying such perceptual rivalry, a key question has remained unresolved: Does perceptual rivalry result merely from local bistability of neural activity patterns in sensory stimulus representations, or do higher-order areas play a causal role by shifting inference and, thus, initiating perceptual changes? We used functional MRI to measure brain activity while human observers reported successive spontaneous changes in perceived direction for an ambiguous apparent motion stimulus. In a control condition, the individual sequences of spontaneous perceptual switches during bistability were replayed by using a disambiguated version of the stimulus. Greater activations during spontaneous compared with stimulus-driven switches were observed in inferior frontal cortex bilaterally. Subsequent chronometric analyses of event-related signal time courses showed that, relative to activations in motion-sensitive extrastriate visual cortex, right inferior frontal cortex activation occurred earlier during spontaneous than during stimulus-driven perceptual changes. The temporal precedence of right inferior frontal activations suggests that this region participates in initiating spontaneous switches in perception during constant physical stimulation. Our findings can thus be seen as a signature of when and where the brain ''makes up its mind'' about competing perceptual interpretations of a given sensory input pattern.

apparent motion | bistable perception | functional MRI

We usually experience our perception of the visual world as unitary and stable even though the information available to our brains is often noisy or ambiguous. How does the brain nonetheless ''make sense'' of this information? From a theoretical perspective, translating the images cast onto the retina into meaningful percepts can be addressed as an inference problem, but related neurophysiological evidence is sparse (1). This need for inference is best illustrated by ambiguous stimuli as the Necker cube or the apparent motion quartet (2, 3) (Fig. 1), where perception typically fluctuates spontaneously between two mutually exclusive interpretations of the same sensory input. Such bistable perception is paralleled by activity changes in separate neuronal populations that represent one or the other perceptual interpretation (4), but what causes the change to occur in the first place remains unclear. In binocular rivalry, where discrepant information is presented to the two eyes, neural activity at the earliest visual processing stages correlates with spontaneous perceptual fluctuations (5–9), suggesting a ''gate-keeping'' function of these structures in visual awareness. Alternatively, perceptual changes could be initiated by a higher-level process that, by way of inference, can stabilize, bias, or topple the current interpretation of the sensory input (10, 11).

These two scenarios differ in the causal chain assumed to underlie changes in visual awareness, but it remains difficult to infer causality from correlative neurophysiological measures. Still, temporal precedence is generally considered good evidence in favor of a putative causal role (12). Invasive neurophysiological recordings would therefore appear ideally suited to resolve this question but suffer from the uncertainty of where exactly to place recording electrodes, for instance, in the frontal lobe. Functional neuroimaging provides whole-brain coverage and has indeed shown activations related to perceptual change not only in visual but also frontal and parietal regions (10, 13–16). However, the conventional amplitude-based analytical techniques used in these studies could not clarify whether frontoparietal regions respond to a feed-forward signal from the sensory cortex that is driven by perceptual change or whether these regions generate a feedback signal to the sensory cortex before perceptual change. Recent novel analytical approaches have shown that the temporally dispersed neurovascular response is reliable enough to resolve latency differences in the range of a couple of hundreds of milliseconds (17–22). Here, we used such techniques to analyze the onset latencies of functional MRI (fMRI) activiations during spontaneous perceptual changes in bistable apparent motion (rivalry) compared with stimulus-driven changes in a matched control condition (replay) where the sequence of reversals as reported in the preceding rivalry condition was replayed by using a disambiguated version of the apparent motion quartet (Fig. 1). We hypothesized that a top-down process, where frontal or parietal regions play a causal role in initiating perceptual change, should be associated with an earlier onset of activations in these regions during spontaneous compared with stimulus-driven perceptual changes, relative to activations in visual cortex [supporting information (SI) Fig. 5].

Results

We first mapped transient activations that occurred whenever perception changed, irrespective of whether this happened spontaneously during rivalry or in response to a stimulus change during replay. In our paradigm, change in perception refers to a change in direction of apparent motion (from vertical to horizontal or vice versa, see Fig. 1), and activation refers to a signal increase over and above constant ongoing sensory processing. As reported with similar paradigms (16, 23–25), switch-related activations occurred at the lateral occipitotemporal junction, corresponding to the human motion complex (hMT+/V5) (Fig. 2*A*). Additionally and again in accord with previous findings (13, 14, 16), we observed a distributed activation pattern associated with perceptual switches that included bilateral inferior frontal lobe regions and the right inferior parietal lobule but also the calcarine cortex, consistent with early visual areas V1/V2 (see SI

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

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: fMRI, functional magnetic resonance imaging; hMT+/V5, human motion complex; BOLD, blood oxygenation level dependent; HRF, hemodynamic response function; ROI, region of interest.

[§]To whom correspondence should be addressed. E-mail: philipp.sterzer@charite.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0609006104/DC1.

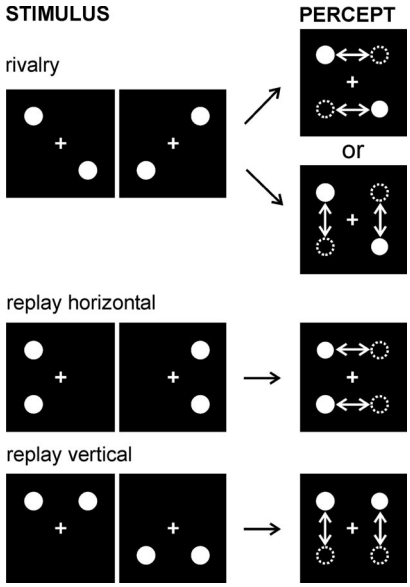

Fig. 1. Stimulus display. Ambiguous and disambiguated versions of the apparent motion quartet used in the rivalry and replay conditions, respectively, are shown. The single frames alternated at 4 Hz. When looking at the rivalry stimulus, perception is bistable and fluctuates spontaneously between periods of horizontal and vertical apparent motion perception. Disambiguated versions of the stimulus were used to change participants' perception of apparent motion with the same temporal sequence as during the rivalry condition.

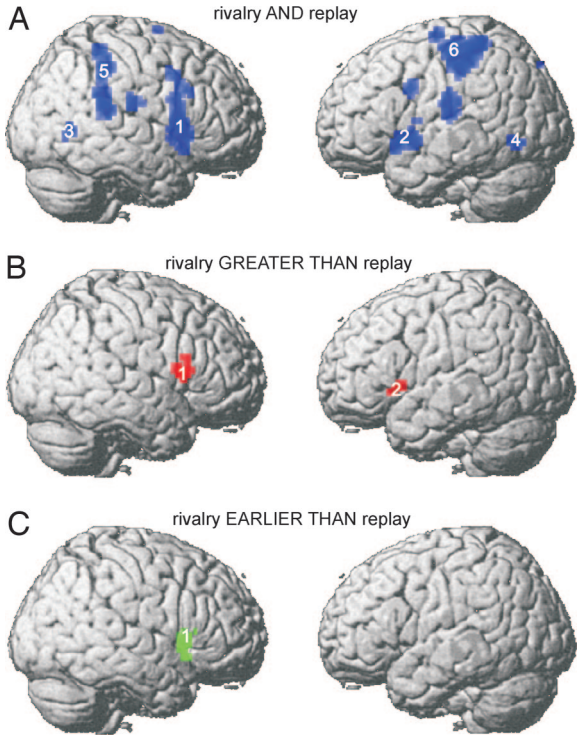

Fig. 2. Transient activation during perceptual switches. (*A*) Regions commonly activated in response to both spontaneous and stimulus-driven perceptual switches are rendered in blue onto a standard anatomical template image ($n = 12$, random effects, for visualization thresholded at $P < 10^{-5}$, uncorrected, cluster threshold $k > 12$ voxels). Numbers 1–6 indicate the regions that were subsequently used for detailed analyses of signal time courses (see Figs. 3 and 4): 1, 2 = right and left inferior frontal cortex; 3, 4 = right and left hMT+/V5; 5 = right inferior parietal lobule; 6 = left sensorimotor cortex. (*B*) Greater response amplitudes during spontaneous as opposed to stimulus-driven switches were observed in bilateral inferior frontal regions and are shown in red (x, y, z = −45, 18, and 0, $P < 0.001$; and x, y, z = 66, 18, 12, $P < 0.016$, small-volume corrected; thresholded for visualization at $P < 0.005$, uncorrected, cluster threshold $k > 12$ voxels). (*C*) Earlier responses during spontaneous as opposed to stimulus-driven switches were observed in the right inferior frontal gyrus (x, y, z = 63, 18, and 15; $P < 0.017$, small-volume corrected; thresholded for visualization at $P < 0.005$, uncorrected, cluster threshold $k > 12$ voxels). Latency differences were assessed by using a published voxel-wise method (21).

Table 1). As in these studies, we also found several further activations related to the preparation and execution of motor reports, e.g., in the hand representations of the left primary sensorimotor cortex and the cerebellum as well as in the supplementary motor area (14).

We next tested whether those regions that responded to perceptual changes in both conditions, rivalry as well as replay, showed differences in response amplitude between these two conditions. Greater response amplitudes during spontaneous vs. stimulus-driven perceptual switches were found bilaterally in inferior frontal regions (Fig. 2*B*; see also SI Table 2). This result is in general agreement with findings from a previous study (14) that used a similar paradigm with a rivalry and a replay condition but a different stimulus (binocular rivalry). In contrast to this earlier study, no significant response amplitude differences were detected in parietal or occipital regions (14).

Finally, and thus addressing the key issue of this study, we asked whether there was evidence for temporal precedence of inferior frontal lobe activations over those in other brain regions, especially those in hMT+/V5. There is ample evidence to suggest that perceived changes in motion direction, no matter whether they are spontaneous or stimulus-driven, are reflected by event-related hMT+/V5 activations (16, 23–28). We hypothesized that if the inferior frontal cortex was causally involved in inducing switches during bistable perception, then frontal activations in association with spontaneous perceptual switches should precede those in hMT+/V5. Regional variations in the shape of the hemodynamic response (18, 22) might invalidate any direct comparison of response onsets across regions. We overcame this problem by testing our hypothesis as the region-by-condition interaction that probed whether an earlier onset of activations during spontaneous as opposed to stimulus-driven switches occurred in frontal regions over and above any potential differences in hMT+/V5. A finding of earlier frontal activations during spontaneous compared with stimulus-driven switches relative to hMT+/V5 activations would hence argue for a role of frontal regions in inducing spontaneous perceptual switches (SI Fig. 5).

In a first step, we used a previously described voxel-wise method that allowed us to capture slight timing variations in blood oxygenation level-dependent (BOLD) responses by using a canonical hemodynamic response function (HRF) and its temporal derivative as model functions (21). Testing for earlier response onsets in spontaneous vs. stimulus-driven switches, we found a significant difference in the right inferior frontal cortex at virtually the same location where switch-related activations were greater in amplitude during rivalry than replay (Fig. 2*C*, see also SI Table 3). No indication of significant onset-latency differences was detected in any of the other regions that activated during switches.

To validate these results independently with a different approach, we then analyzed our data using a Fourier set of basis functions, which avoids assumptions about the shape of the hemodynamic response (29) (Fig. 3). From the resulting response fits, we measured onset latencies for both conditions in each participant at the exact individual frontal lobe location that responded more strongly during rivalry than replay switches. For

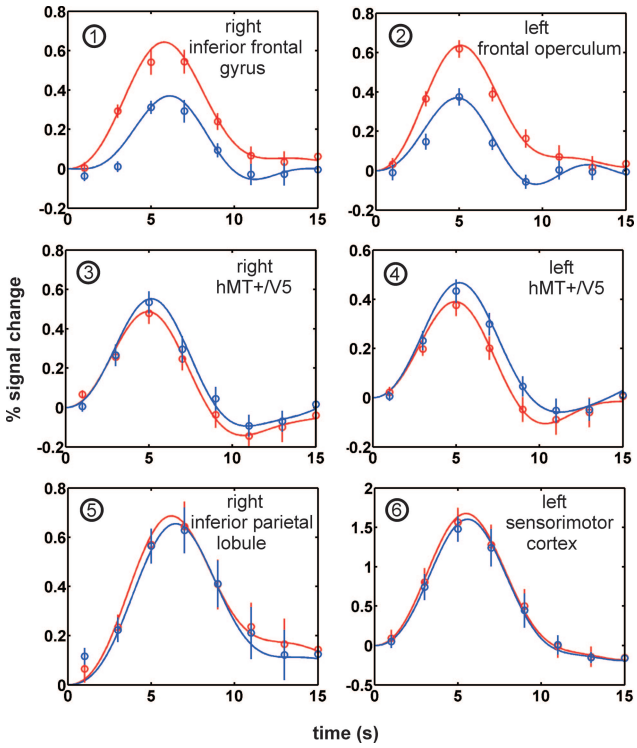

Fig. 3. Event-related signal time courses. Circles represent the signal change per 2-s time bin averaged across participants (error bars indicate standard error), and smooth lines represent the average fitted response from a fourth-order Fourier model (29). Postswitch time (s) is plotted on the *x* axes and percentage signal change on the *y* axes. Responses to spontaneous perceptual switches (rivalry condition) are plotted in red and responses to stimulus-driven switches (replay condition) are plotted in blue. The data and fitted responses were extracted from individual participants' response maxima in *a prior*-defined ROIs. The numbering of the panels 1–6 corresponds to the numbering of the regions in Fig. 2. The individual fitted responses were used for the assessment of response-onset latencies (see Fig. 4).

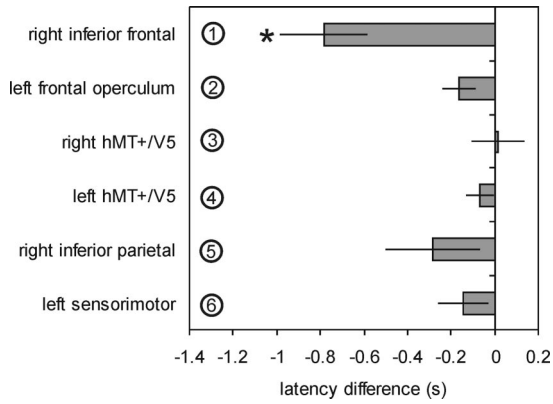

Fig. 4. Onset-latency differences. Onset-latency differences between regional brain activations during perceptual switches in the rivalry and those in the replay condition averaged across participants (0 on the *x* axis = response onset in the replay condition). The numbering of the bars 1–6 corresponds to the numbering of the regions in Fig. 2. Right inferior frontal response onsets during rivalry occurred on average 784 ms ±200 SEM earlier, whereas no such difference was detectable in other frontal or parietal regions and hMT+/V5. Error bars indicate standard errors. *, $P < 0.002$ (two-tailed paired *t* test for onset-latencies of responses to spontaneous compared with stimulus-driven switches).

comparison, we sampled data from the individual response maxima common to rivalry and replay switches, i.e., right inferior parietal cortex, hMT+/V5 bilaterally, and left sensorimotor cortex (see SI Fig. 6). Strikingly, we found the response onset in right inferior frontal cortex to occur, on average, ≈800 ms earlier during rivalry compared with replay (Fig. 4), confirming the result of the voxel-wise analysis. No significant latency difference was detectable in any other region, resulting in a significant region-by-condition interaction ($P < 0.003$, repeated-measures ANOVA).

Discussion

In this study, we compared the amplitude and timing of brain activations during spontaneous changes in visual perception with those during stimulus-driven changes. Although these two types of perceptual change recruited largely overlapping brain structures, greater response amplitudes during spontaneous perceptual changes were observed in inferior frontal cortical regions bilaterally. Crucially, we could also demonstrate that, relative to brain responses during stimulus-driven switches, right inferior frontal activations occurred earlier than those in other brain regions, most notably those in hMT+/V5.

According to evidence from neurophysiological studies in nonhuman primates (30, 31) and human fMRI studies (32), subjective fluctuations in perception are reflected by relative activity levels of neuronal populations representing either perceptual content. The finding of switch-related hMT+/V5 activations in bistable motion perception (16, 24–26, 28) has led to the suggestion that separate direction-selective neuronal populations coding for different directions of motion might be in rivalry and that, at perceptual switches, one of these populations displays a transient rise in activity. This interpretation is supported by the electrophysiological observation in monkeys that responses of directionally selective cells in MT signal the perceived direction of motion in ambiguous motion stimuli (27). Spontaneous fluctuations and adaptation of percept-related neural activity would then be two mechanisms contributing to perceptual dominance of either percept during ongoing rivalry.

In this study, event-related hMT+/V5 activations were indistinguishable between spontaneous and stimulus-driven perceptual changes. This suggests that hMT+/V5 activity reflects perceived changes in motion direction irrespective of whether they are "real" or generated spontaneously in the absence of any stimulus change. We also found extremely stable relationships without evidence of latency shifts between the activations in hMT+/V5 and the key presses reporting perceptual changes as well as the activations in the primary sensorimotor hand representation. These two observations indicate that the sensory and perceptual differences between the two conditions were so subtle that they were reflected neither in the fMRI responses recorded at the level of visual cortex nor in the delay of reporting the changes. Reaction-time differences between reporting spontaneous as opposed to stimulus-driven perceptual changes should have translated into condition-dependent latency differences between the motor cortex and hMT+/V5, which is at odds with our chronometric findings. Moreover, our approach was not sensitive to putative reaction-time differences between rivalry and replay, because our critical result, a region-by-condition interaction for latencies, was based on chronometric analyses of frontal lobe activity referenced against the timing of a neural signal, the response in hMT+/V5, instead of a behavioral report (SI Fig. 5).

Although responses in sensory and motor areas were insensitive to the actual condition, activity in frontal lobe regions differed between the two conditions not only in amplitude but, more importantly, also in timing. Simply put, we used the replay condition to measure the minimal latency of neural processes that might occur just because perception has changed. The

latency of the frontal responses in the replay condition therefore determines the earliest possible time point for feed-forward effects caused by processes occurring at lower visual processing stages. Conversely, any earlier activation in the rivalry condition inevitably, and irrespective of its absolute latency, expresses that perception will change.

One concern could be that the earlier response onset in the right inferior frontal lobe merely reflected the fact that this region also showed greater response amplitudes during rivalry than during replay. However, this scenario is unlikely for several reasons. First, we did not find a difference in response latency in left inferior frontal cortex, where the response amplitude differed as much as in the right inferior frontal cortex (Fig. 3). This result shows that greater BOLD responses are not systematically associated with earlier response onsets. Furthermore, we used two different and independent analytical approaches to determine latency, both of which gave mutually confirmatory results. One of our approaches used a bias-free model that captures variations in amplitude or duration without translating them into an estimated difference in onset (29). Finally, there is no indication from previous studies that greater neural responses *per se* can lead to an earlier onset of the BOLD response (22, 33). It is also noteworthy that frontal and occipital switch-related activations were comparable in strength (Fig. 3), which makes it unlikely that potential neural latency differences in the occipital cortex might have been undetectable because of substantially weaker BOLD responses.

We found no significant latency effects in other regions that might also have been expected to show temporal precedence in our setting as, for instance, in the left inferior frontal and right inferior parietal lobes. Of course, one explanation may be that fMRI chronometry simply lacks the temporal resolution to detect what could be a whole chain of incremental latency differences at the various stages of the visual-processing hierarchy. We thus cannot exclude the possibility of temporal precedence at the neural level in other regions (34), but our results do suggest that the latency difference in the right inferior frontal cortex is on a larger scale than elsewhere. Interestingly, recent work using noninvasive MRI-based fiber-tracking has revealed direct connections between the inferior frontal and lateral occipital cortex in humans (35). Such a frontooccipital "fast-track" may provide the anatomical routing for the physiological responses we studied. We observed a latency difference of ≈800 ms between right inferior frontal activations during spontaneous and stimulus-driven switches, which (in the absence of a latency difference in hMT+/V5) indicates that the neural process that initiates a spontaneous perceptual reorganization may start several hundred milliseconds before perception actually switches. This interval seems surprisingly long, no matter whether one assumes interactions by direct connections or a cascade involving intermediate processing levels at other sites such as the parietal lobe. Conceivably, however, the process of perceptual reinterpretation may require recursive interactions between high-level and sensory areas that could last for a few hundred milliseconds before eventually resulting in a perceptual switch. Such a mechanism would not need to be associated with earlier fMRI signal increases in hMT+/V5, because the driving force in the sensory area before a perceptual reversal could be adaptation to the still-dominant percept.

The notion that adaptation of percept-related neural activity may also play an important role in perceptual bistability (4) is not irreconcilable with a causal role of higher-order processes in initiating a perceptual switch as previously suggested (11, 36). The actual alternations of perception could be determined by the joint effect of local processes embedded into a more global process. That is, whenever local processes (e.g., adaptation) act to destabilize activity that underpins the currently dominating percept, higher-order evaluative processes can take effect and initiate a perceptual reorganization that translates into and stabilizes a new "balance of power" in specialized visual areas. This interpretation is further supported by behavioral alteration of perceptual bistability that has recently been described in patients with prefrontal cortex lesions (37).

Our results cannot necessarily be extended to studies using binocular rivalry, where perceptual ambiguity is generated by interocular sensory conflict. This latter setting may involve additional mechanisms specifically operating at the monocular levels of stimulus processing (5, 8, 9). Conversely, our findings can very well be extended to "real-life" visual perception, because the ambiguity in interpreting natural visual scenes falls within the range between the two extremes probed by our experiment. On one end, the rivalry condition maximizes the need for perceptual inference; perception cannot be informed by physical stimulus properties when sensory input is perfectly ambiguous, and thus remains "arbitrary." On the other end, the replay condition is almost perfectly unambiguous and thus requires little if any inference. By using these two extremes, our experiment showed arbitrating neural activity by virtue of latency differences, but we propose that the same inferential mechanisms are also called upon after any sensory input change and that their activity level depends on the degree of perceptual ambiguity in the sensory input.

Although apparent motion paths and object identities across frames are confounded in our rivalry stimulus, it appears intuitive to suspect a spatial inference to act on percept selection. This speculation is compatible with earlier observations that central symbolic (endogenous) cues activate right inferior frontal cortex more than salient peripheral (exogenous) cues, particularly when cues are invalid, and this mismatch of expectation and sensory input yields reorienting (38). In the rivalry case, expectation would correspond to the currently dominant percept. With progressive adaptation of the related hMT+/V5 signal, the mismatch or error would increase and might drive reorienting. In the replay case, this error signal would also arise but only briefly once sensory input changes. This interpretation, which draws on predictive coding theory (39), would thus also account for the smaller and later switch-related activations that we observed in the inferior frontal cortex during replay. Together, our results provide empirical neurofunctional evidence for theoretical models of visual perception that are grounded in Bayesian and related frameworks (1, 40) instead of relying purely on feed-forward architectures (41).

Materials and Methods

Experimental Paradigm and fMRI Data Acquisition. Twelve healthy right-handed participants aged 28–39 years (8 males; written informed consent) with normal or corrected-to-normal vision participated in a fMRI experiment on a 3T whole-body scanner with gradient booster and standard head coil (Trio; Siemens, Erlangen, Germany). We obtained blood oxygenation-sensitive (T2* weighted) echo-planar image volumes every 2 s (echo time 30 ms; flip angle 60°; 34 contiguous transverse slices, voxel size = 3.3 × 3.3 × 3 mm).

During scanning, participants were presented with alternating blocks of the rivalry and replay conditions (Fig. 1). The stimuli were projected onto a screen located 25 cm from their eyes and viewed by means of a mirror attached to the head coil. During rivalry, a bistable apparent motion quartet consisting of a white central fixation cross and two white dots was presented. The latter were flashed simultaneously in diagonally opposite corners of an implicit rectangle in 4-Hz alternation with two dots in the other two corners (Fig. 1). This yields spontaneously alternating perception of horizontal or vertical apparent motion, both perceived directions being mutually exclusive (2, 3); for a demonstration of this phenomenon see http://psy.ucsd.edu/chip/illu_ambig_apprnt_mot.html from the laboratory of V. S. Ram-

achandran). Background luminance was 12 cd/m^2, dot and fixation cross luminance 76 cd/m^2, dot diameter 25′ of visual angle, and center-to-center horizontal dot distance 2°. The vertical dot distance was adjusted individually for each participant in preceding behavioral sessions to obtain approximately equal average durations of stable horizontal and vertical motion perception. It thus varied across individuals between 2°32′ and 3°13′. Participants were chosen for imaging if they reported clear-cut transitions between stable states of horizontal and vertical apparent motion perception without experiencing further percepts (e.g., circular motion or flicker) and if the percepts lasted on average at least 8 s, thus optimizing power for the event-related analysis. The participants were instructed to maintain fixation on the central cross and to indicate changes of motion direction by key presses, by using the right index finger for perceptual switches to vertical and the right middle finger for switches to horizontal motion perception. All participants confirmed after scanning that they could clearly determine and report the transitions between vertical and horizontal apparent motion in both conditions. Interreversal times between perceptual switches followed a γ distribution as generally found for bistable stimuli (13, 14, 16, 25, 42). Mean percept duration across all participants was 13.08 ± 0.37 s (SEM) and thus well suited for an event-related analysis of fMRI signals. There was no significant difference in duration of stable horizontal and vertical motion perception (12.65 ± 0.36 s (SEM) vs. 13.50 ± 0.42 s, $P = 0.071$, Wilcoxon test).

For replay, a disambiguated version of the apparent motion quartet was used to repeat the sequence of horizontal and vertical percepts as reported by the participant in the preceding rivalry condition. The stimulus also consisted of a central fixation cross and two dots flashing simultaneously at 4 Hz (see Fig. 1), dot sizes and distances being identical to the rivalry condition. For horizontal apparent motion, two dots in the upper left and lower left corners of an implicit rectangle were shown in alternation with two dots in the upper right and lower right corners. Accordingly, two dots in the upper left and right corners alternating with two dots in the lower left and right corners yielded vertical apparent motion. Scanning comprised 3 rivalry and 3 replay runs of 245 scans each. Stimuli were displayed and key presses registered by using ERTS software (Experimental Run Time System, Version 3.18, 1996, J. Beringer, Darmstadt, Germany).

Imaging Data Analysis. fMRI data were analyzed by using SPM2 (www.fil.ion.ucl.ac.uk/spm). After removal of the first five scans, all functional image volumes were realigned, unwarped, slice-timing corrected, spatially normalized into Montreal Neurological Institute (MNI) neuroanatomical space (www.bic.mni.mcgill.ca/brainweb), and smoothed by using a 10-mm full-width-at-half-maximum Gaussian kernel. We removed low-frequency fluctuations by a high-pass filter with a cut-off at 128 s and used an autoregressive model of order one (AR (1) + white noise) to correct for temporal autocorrelations.

A mixed-effects analysis was adopted, using a two-stage procedure. First, a fixed-effects analysis was applied to the preprocessed data of each participant by using the general linear model implemented in SPM2. This model fits the data with a linear combination of regressors in a design matrix to produce 3D maps of parameter estimates (β weights), which represent the contribution of a particular regressor to the data. The design matrix consisted of six regressors. There were two experimental conditions, spontaneous switches (occurring during rivalry) and stimulus-driven-switches (during replay), both of which were modeled by using three regressors in the following way: each participant's individual average reaction time from stimulus change to key press during replay (average across participants 843 ± 14 ms) was subtracted from the times at which the key presses occurred to create a sequence of estimated time points of the actual perceptual switches in both conditions. The evoked hemodynamic responses to these perceptual switches were modeled as δ ("stick") functions convolved with a canonical HRF as implemented in SPM2 and both its temporal and dispersion derivatives. For each participant, t maps of the effects for both switch-types as well as separately for both event types relative to a baseline of stable motion perception were computed by using the HRF regressors and then submitted to second-level analyses (random-effects).

The main effect for perceptual switches, irrespective of whether they occurred during the rivalry or the replay condition, was assessed by performing a one-sample t test on all t maps for both event types. This contrast was inclusively masked with all voxels that were activated by both single-switch types at $P < 0.001$, uncorrected, to limit the analysis to regions where both conditions contributed positively. Differential effects (spontaneous > stimulus-driven switches and vice versa) were assessed by using paired t tests. For the main effect of perceptual switches, activations were considered significant at $P < 0.05$, by using a family-wise error correction for multiple comparisons across the whole brain and, alternatively at $P < 0.001$, uncorrected, if predicted by previous findings, or in the case of a response in the homologous contralateral area at $P < 0.05$, corrected. The analysis of differential effects was constrained to candidate regions that were defined *a priori* on the basis of previous findings (13–16). These regions of interest (ROI) included the inferior frontal cortices, right inferior parietal cortex, and extrastriate occipital regions showing switch-related responses, i.e., in hMT+/V5 bilaterally. We used spherical search volumes (diameter, 20 mm) centered on the maxima of the main effect for both switch types in these regions and corrected for multiple comparisons across the voxels within these search volumes (small-volume correction).

Latency differences between conditions were assessed in two steps, by using two separate and independent analyses. First, we used a previously described method (21) that is based on the voxel-wise computation of ratios of the parameter estimates for the temporal derivative and the HRF. In brief, latency maps for each participant and each condition were created by transforming the derivative/HRF ratio for each voxel with a sigmoidal logistic function. A paired t test was then performed on the second level to test for between-condition latency differences across participants. To restrict this analysis to voxels with a positive response in both conditions and where the HRF provided a reasonable fit to the data, the second-level maps were masked with t maps of voxels activated by both single-event types at $P < 0.001$, uncorrected. For statistical inference, we again used a small-volume correction for the same ROIs as in the original analysis, with a significance threshold of $P < 0.05$, corrected.

To validate the results further from this voxel-wise analysis (21), we assessed the onset latencies in a more detailed analysis based on event-related signal time courses in each individual participant. A common approach in previous fMRI studies investigating response latencies was the use of HRF models (usually based on γ functions) parameterized with respect to onset latency (17, 19, 20, 22). To account fully for the variability in response shapes due to regional differences in vascularization and the duration of neural responses, the model used also has to be parameterized for other variables, such as peak latency and response width. Because these parameters may be interdependent and the resulting parameter estimates hence influenced by each other, we took a different approach and used a model that is free of *a priori* assumptions with regard to response shape. Importantly, the selection of voxels was guided by our original analysis using the HRF, but the fine-grained information regarding the exact response onset was assessed by using a Fourier

model that can, in principle, fit any response shape as a linear combination of its basis functions (29). We first determined in each participant the individual peak voxel for regions that had shown a group effect for spontaneous > stimulus-driven switches at a lowered statistical threshold ($P < 0.05$, uncorrected; in 4 of 12 participants, the threshold had to be lowered to $P < 0.1$ to identify response maxima for this contrast; see SI Fig. 6). Inclusive masking with voxels surviving $P < 0.05$, uncorrected, in both respective single contrasts was used in each participant individually to restrict the onset-latency analysis to voxels showing a positive response and where the HRF provided a reasonable fit to the data in both conditions. Individual maxima for ROIs not showing significant activity differences were analogously determined from the main effect for all perceptual switches. We extracted data from individual peak voxels rather than averaging across entire ROIs to maximize the sensitivity and specificity of our analysis. However, because of spatial smoothing (see above) the signal in each single voxel nevertheless represents a weighted average of nearby voxels in the range of the smoothing kernel used. The data were then reanalyzed with a fourth-order Fourier set windowed with a Hanning function, which has proven useful for characterizing the temporal aspects of evoked hemodynamic responses (29). Condition-specific onset latencies were then calculated for each participant from the slope of the resulting fitted responses in each ROI. The onset was operationally defined as the time at which the slope exceeded 10% of the maximum slope in the ascending part of the BOLD response. This measure accounts for, but is not confounded by differences in response amplitude. Again, onset latencies were subjected to second-level analyses testing across participants for significant differences between conditions within each region (one-sample t tests) and for a condition-by-region interaction (repeated-measures ANOVA).

We thank Karl Friston for expert advice on chronometric analyses and Stanislas Dehaene and Geraint Rees for helpful comments on an earlier version of the manuscript. This work was supported by the Volkswagen Foundation (P.S. and A.K.) and by the German Research Society (P.S.).

1. Kersten D, Yuille A (2003) *Curr Opin Neurobiol* 13:150–158.
2. von Schiller P (1933) *Psychologische Forschung* 17:179–214.
3. Ramachandran VS, Anstis SM (1983) *Nature* 304:529–531.
4. Blake R, Logothetis NK (2002) *Nat Rev Neurosci* 3:13–21.
5. Haynes JD, Deichmann R, Rees G (2005) *Nature* 438:496–499.
6. Haynes JD, Rees G (2005) *Curr Biol* 15:1301–1307.
7. Polonsky A, Blake R, Braun J, Heeger DJ (2000) *Nat Neurosci* 3:1153–1159.
8. Tong F, Engel SA (2001) *Nature* 411:195–199.
9. Wunderlich K, Schneider KA, Kastner S (2005) *Nat Neurosci* 8:1595–1602.
10. Kleinschmidt A, Buchel C, Hutton C, Friston KJ, Frackowiak RS (2002) *Neuron* 34:659–666.
11. Leopold DA, Logothetis NK (1999) *Trends Cogn Sci* 3:254–264.
12. Parker AJ, Krug K (2003) *Curr Opin Neurobiol* 13:433–439.
13. Kleinschmidt A, Büchel C, Zeki S, Frackowiak RSJ (1998) *Proc R Soc London B* 265:2427–2433.
14. Lumer ED, Friston KJ, Rees G (1998) *Science* 280:1930–1933.
15. Portas CM, Strange BA, Friston KJ, Dolan RJ, Frith CD (2000) *Proc R Soc London B* 267:845–850.
16. Sterzer P, Russ MO, Preibisch C, Kleinschmidt A (2002) *NeuroImage* 15:908–916.
17. Bellgowan PS, Saad ZS, Bandettini PA (2003) *Proc Natl Acad Sci USA* 100:1415–1419.
18. Buckner RL, Bandettini PA, O'Craven KM, Savoy RL, Petersen SE, Raichle ME, Rosen BR (1996) *Proc Natl Acad Sci USA* 93:14878–14883.
19. Calhoun V, Adali T, Kraut M, Pearlson G (2000) *Magn Reson Med* 44:947–954.
20. Formisano E, Linden DE, Di Salle F, Trojano L, Esposito F, Sack AT, Grossi D, Zanella FE, Goebel R (2002) *Neuron* 35:185–194.
21. Henson RN, Price CJ, Rugg MD, Turner R, Friston KJ (2002) *NeuroImage* 15:83–97.
22. Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000) *NeuroImage* 11:735–759.
23. Muckli L, Kriegeskorte N, Lanfermann H, Zanella FE, Singer W, Goebel R (2002) *J Neurosci* 22:RC219.
24. Sterzer P, Eger E, Kleinschmidt A (2003) *NeuroReport* 14:2337–2441.
25. Sterzer P, Kleinschmidt A (2005) *Eur J Neurosci* 21:3097–3106.
26. Castelo-Branco M, Formisano E, Backes W, Zanella F, Neuenschwander S, Singer W, Goebel R (2002) *Proc Natl Acad Sci USA* 99:13914–13919.
27. Dodd JV, Krug K, Cumming BG, Parker AJ (2001) *J Neurosci* 21:4809–4821.
28. Kleinschmidt A, Thilo KV, Buchel C, Gresty MA, Bronstein AM, Frackowiak RS (2002) *NeuroImage* 16:873–882.
29. Friston KJ, Frith CD, Frackowiak RS, Turner R (1995) *NeuroImage* 2:166–172.
30. Leopold DA, Logothetis NK (1996) *Nature* 379:549–553.
31. Sheinberg DL, Logothetis NK (1997) *Proc Natl Acad Sci USA* 94:3408–3413.
32. Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) *Neuron* 21:753–759.
33. Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) *J Neurosci* 16:4207–4221.
34. Williams ZM, Elfar JC, Eskandar EN, Toth LJ, Assad JA (2003) *Nat Neurosci* 6:616–623.
35. Ffytche DH, Catani M (2005) *Philos Trans R Soc London B* 360:767–779.
36. Rees G, Kreiman G, Koch C (2002) *Nat Rev Neurosci* 3:261–270.
37. Windmann S, Wehrmann M, Calabrese P, Gunturkun O (2006) *J Cogn Neurosci* 18:456–471.
38. Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M (2005) *J Neurosci* 25:4593–4604.
39. Rao RP, Ballard DH (1999) *Nat Neurosci* 2:79–87.
40. Pouget A, Dayan P, Zemel RS (2003) *Annu Rev Neurosci* 26:381–410.
41. Riesenhuber M, Poggio T (2002) *Curr Opin Neurobiol* 12:162–168.
42. Borsellino A, De Marco A, Allazetta A, Rinesi S, Bartolini B (1972) *Kybernetik* 10:139–144.

Supplementary Material

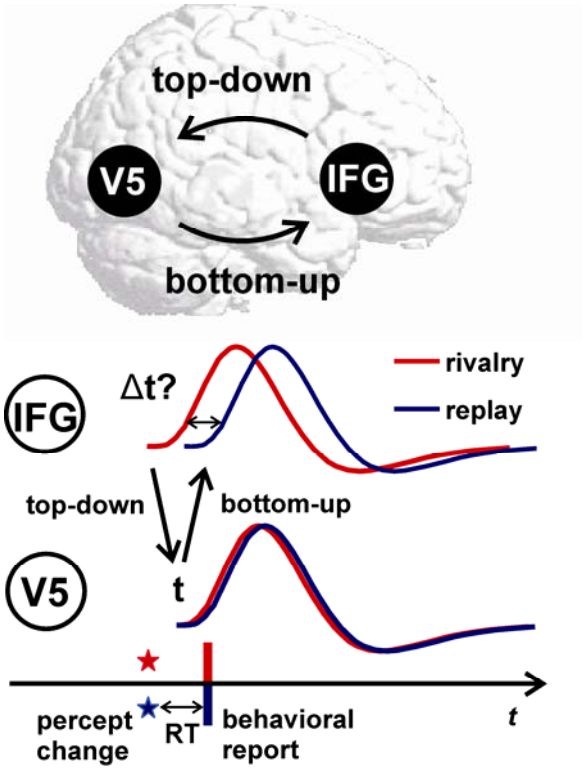

Figure 5. Analytical approach. The events of interest are neural activations associated with perceptual changes (asterisks) during rivalry (red) and replay (blue). Precise timing (and thus reaction time) is only known in the case of replay where perceptual changes are stimulus-driven. In a first step, the analysis is based on the timing of the key presses (short vertical lines) reporting perceptual changes in either condition. A second step addresses local onset latency differences (Δt) between conditions (as graphically illustrated here for IFG, the inferior frontal gyrus, and reported for the experimental data in Fig. 2C). This analysis is sensitive to reaction time differences between conditions and would yield apparently earlier response onsets in case of longer reaction times. As this effect should be systematic across all regions it can be addressed by testing whether latency differences in IFG occur over and above those in a neural instead of behavioral reference signal. This approach is graphically illustrated by aligning responses for rivalry and replay in the motion-sensitive area hMT+/V5 (V5), and corresponds to a region-by-condition interaction (reported in Fig. 4). Although replay might also involve top-down effects, it defines those onset latency differences between hMT+/V5 and IFG that can be explained with a mere bottom-up account. Significant temporal precedence of IFG over hMT+/V5 responses in rivalry compared to replay would provide evidence for a top-down mechanism.

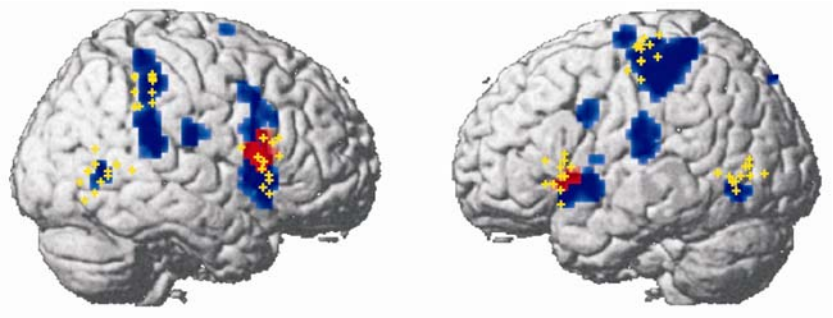

Figure 6. Individual response maxima. Regions commonly activated in response to both spontaneous and stimulus-driven perceptual switches across the whole group of participants are shown in blue (for visualization thresholded at $p < 10^{-5}$, uncorrected, cluster threshold $k > 12$ voxels) and greater response amplitudes during spontaneous as opposed to stimulus-driven switches are superimposed in red (for visualization thresholded at $p < 0.005$, uncorrected, cluster threshold $k > 12$ voxels; see also Fig. 2A and B). Yellow crosses indicate the individual participants' response maxima in regions of interest that were used for the extraction of the event-related signal time courses as shown in Fig. 3 (for full details, see Materials and Methods). These regions included the right inferior frontal gyrus and the left frontal operculum as determined by the contrast spontaneous > stimulus-driven switches. Right and left hMT+/V5, right inferior parietal lobule, and left sensorimotor cortex as determined from the main effect for both spontaneous and stimulus-driven perceptual switches were used as control regions.

Table 1. Regions commonly activated by spontaneous and stimulus-driven switches at p < 0.05 (corrected across whole brain)

Region	Coordinates			corrected p-value	t-value
	x	y	z		
Frontal					
Inferior frontal gyrus/ frontal	*-48*	*12*	*-6*	*< 0.0005*	*15.94*
operculum	*60*	*12*	*9*	*0.001*	*13.72*
Anterior cingulate gyrus	0	9	42	0.002	12.65
Middle frontal gyrus	-57	6	30	0.021	9.94
	57	12	39	0.054	9.05*
Left precentral gyrus (MI)	*-48*	*-21*	*54*	*< 0.0005*	*16.53*
Parietal					
Left postcentral gyrus (SI)	-45	-39	63	< 0.0005	17.94
Parietal operculum (SII)	-54	-21	12	0.003	12.13
	51	-21	27	0.003	12.26
Right temporo-parietal junction	60	-39	24	0.001	13.45
Right inferior parietal lobule	*54*	*-39*	*45*	*0.001*	*13.69*
Right intraparietal sulcus	39	-60	45	0.038	9.38
Occipital					
Calcarine cortex (V1/V2)	0	-75	-3	0.003	11.89
	12	-75	3	< 0.0005	15.98
Occipito-temporal junction	*-57*	*-69*	*-6*	*0.10*	*8.49**
(hMT+/V5)	*54*	*-63*	*3*	*0.029*	*9.65*
Subcortical					
Thalamus	-12	-21	6	0.001	13.72
	15	-12	3	0.007	11.26
Superior colliculus	-6	-30	-12	0.002	12.38
	9	-21	-9	0.006	11.25
Right anterior cerebellar lobe	9	-54	-21	0.007	11.14

* significant at p < 0.001, uncorrected;

Italics indicate regions used for subsequent region-of-interest analyses.

Table 2. Regions of interest for the comparison of spontaneous vs. stimulus-driven perceptual switches

Region	Coordinates			corrected p-value (SVC)	t-value
	x	y	z		
Left frontal operculum	*-45*	*18*	*0*	*0.001*	*7.21*
Right inferior frontal gyrus	*66*	*18*	*12*	*0.016*	*5.45*
Right inferior parietal lobule	57	-30	42	0.45	2.54
Left hMT+/V5*	-	-	-	-	-
Right hMT+/V5*	-	-	-	-	-

* no positive t-values; SVC = small-volume correction;

Italics indicate significant effects at 0.05, corrected across a sphere of 20 mm diameter.

Table 3. Regions of interest for the voxel-wise analysis of latency differences using the ratio of parameter estimates for the temporal derivative and the hemodyamic response function

Region	Coordinates			corrected p-value (SVC)	t-value
	x	y	z		
Left frontal operculum	-48	3	-9	0.22	2.78
Right inferior frontal gyrus	*63*	*18*	*15*	*0.017*	*5.14*
Right inferior parietal lobule	54	-42	48	0.20	2.86
Left hMT+/V5	-57	-69	-3	0.63	1.04
Right hMT+/V5	60	-63	3	0.37	1.94
Left precentral gyrus	-42	-24	48	0.64	1.04

SVC = small-volume correction;

Italics indicate significant effects at 0.05, corrected across a sphere of 20 mm diameter.