Behavioral/Systems/Cognitive

A Hierarchy of Temporal Receptive Windows in Human Cortex

Uri Hasson,[1,2] **Eunice Yang,**[1] **Ignacio Vallines,**[3,4] **David J. Heeger,**[1,2] **and Nava Rubin**[1]

[1]Center for Neural Science and [2]Department of Psychology, New York University, New York, New York 10003, [3]Institute for Experimental Psychology, University of Regensburg, 93053 Regensburg, Germany, and [4]Department of Experimental Psychology, Ludwig-Maximilian University of Munich, 80539 Munich, Germany

Real-world events unfold at different time scales and, therefore, cognitive and neuronal processes must likewise occur at different time scales. We present a novel procedure that identifies brain regions responsive to sensory information accumulated over different time scales. We measured functional magnetic resonance imaging activity while observers viewed silent films presented forward, backward, or piecewise-scrambled in time. Early visual areas (e.g., primary visual cortex and the motion-sensitive area MT+) exhibited high response reliability regardless of disruptions in temporal structure. In contrast, the reliability of responses in several higher brain areas, including the superior temporal sulcus (STS), precuneus, posterior lateral sulcus (LS), temporal parietal junction (TPJ), and frontal eye field (FEF), was affected by information accumulated over longer time scales. These regions showed highly reproducible responses for repeated forward, but not for backward or piecewise-scrambled presentations. Moreover, these regions exhibited marked differences in temporal characteristics, with LS, TPJ, and FEF responses depending on information accumulated over longer durations (~36 s) than STS and precuneus (~12 s). We conclude that, similar to the known cortical hierarchy of spatial receptive fields, there is a hierarchy of progressively longer temporal receptive windows in the human brain.

Key words: temporal coding; fMRI; cortex; receptive fields; functional organization; time

Introduction

It is well established that neurons along the visual cortical pathways have increasingly larger spatial receptive fields (Hubel, 1988). This is a basic organizing principle of the visual system; neurons in higher-level visual areas receive inputs from many neurons with smaller receptive fields in early visual areas, accumulating information from the large portions of space occupied by the objects and scenes they process.

Real-world events occur not only over extended regions of space, but also over extended periods of time. We therefore hypothesized that a hierarchy analogous to that found for spatial receptive field sizes should also exist for the temporal response characteristics of different brain regions. Examples of perceptual and cognitive processes that unfold over time and must therefore rely on accumulation of information over long durations include inferences of cause and effect (Fonlupt, 2003), processing linguistic information at various scales (syllables, words, sentences), understanding narrative (Xu et al., 2005), event segmentation

(Zacks et al., 2001b), human social interaction, and "theory of mind" (Gallagher and Frith, 2003; Saxe et al., 2004). Although we are a long way from understanding how the brain performs these cognitive functions, we argue here that only brain areas that accumulate information over sufficiently long periods of time are candidates for being directly involved in such tasks.

Specifically, defining the temporal receptive window (TRW) of a neuron as the length of time before a response during which sensory information may affect that response, we hypothesized that there is a hierarchy of increasing TRWs as one moves from low level (sensory) to higher level (perceptual and cognitive) brain areas. Because the TRWs of neurons in a brain area determine the length of time into the past from which information is available for processing, we further hypothesized that the range of TRWs in each area must correspond to its functional role. TRWs in early sensory areas should be short, enabling rapid processing of the ever changing sensory input. In contrast, TRWs in some higher level areas should be longer, allowing them to process information from perceptual and cognitive events that unfold over time. (Note, however that the specific aim of this study was to assess the TRWs in each brain area independently of its functional role.)

Assessing TRWs of neurons that are continually processing information calls for using the entire profile of their time-dependent responses, not just its overall magnitude. Our approach is therefore based on measuring the reliability of response profiles over time, building on two disparate sets of previous findings. First, electrophysiological studies indicate that the reli-

ability of the response of a neuron can vary greatly depending on the stimulus. Upon repeated presentations, certain stimuli elicit spikes at predictable, repeatable times whereas others, although they may elicit a similar average firing rate, do not drive the neuron in a precisely predictable way (Mainen and Sejnowski, 1995; Mechler et al., 1998; Yao et al., 2007). Response reliability (reproducibility) therefore offers a measure of how effectively a stimulus is driving activity that is complementary to the more common measure of response amplitude. The second set of findings on which our approach is based comes from human functional magnetic resonance imaging (fMRI) studies, indicating that many brain regions are strongly correlated in their activity within and across individuals watching the same movie (Bartels and Zeki, 2004; Hasson et al., 2004; Golland et al., 2006; Wilson et al., 2008). Although the spatial and temporal resolutions of the fMRI blood oxygen level-dependent (BOLD) signal are much coarser than those of electrophysiology, we conjectured that the correlations between the responses over time to repeated presentations of the same movie similarly provide a measure of how reliably that movie drives each brain area.

We therefore measured cortical activity with fMRI while observers viewed repeated presentations of complex, naturalistic stimuli (silent films). We then compared the reliability with which an intact movie drove each area to that obtained when the temporal structure of the movie was disrupted. Lower and upper bounds on the TRWs of many cortical areas can be inferred from differences in reliability. Regions with short TRWs will exhibit high response reliability regardless of disruptions in temporal structure, whereas such disruptions will reduce the response reliability in regions with longer TRWs. Our results show that intact movies and time-scrambled movies indeed drive fMRI activity at consistently different levels of reliability in certain, but not all, brain areas. Specifically, response reliability in early visual areas [primary visual cortex (V1), V2, V4, and the motion-sensitive area MT+] was not affected by the temporal structure of the movie, indicating that neurons in those areas have short TRWs (a second or less). In contrast, responses in several higher brain areas were affected by information accumulated over longer time scales, revealing a hierarchy of TRWs spanning from short (~4 s) to intermediate (~12 s) and long (~36 s). Importantly, disrupting temporal order had no affect on response amplitudes even for brain areas in which it dramatically reduced response reliability, establishing a clear dissociation between these two measures.

Materials and Methods

Observers. Nine observers, ages 21–40, participated in one or more of the experiments. Eight observers participated in the time-reversal experiment, five of those participated also in the piecewise-scrambled experiment, and five (one new) participated in the block-alternation control experiment. Procedures were in compliance with the safety guidelines for MRI research and approved by the University Committee on Activities Involving Human Subjects at New York University. All observers had normal or corrected-to-normal vision and provided written informed consent.

MRI acquisition. We used functional magnetic resonance imaging at 3T (Allegra; Siemens, Erlangen, Germany) to measure BOLD changes in cortical activity. During each fMRI scan, a time series of volumes was acquired using a T2*-weighted echo-planar imaging pulse sequence (repetition time, 2000 ms; echo time, 30 ms; flip angle 80°; 32 slices; 3 × 3 × 3 mm voxels; field of view, 192 mm), and using custom radio frequency coils (NM-011 transmit head coil and NMSC-021 four-channel phased array receive coil; NOVA Medical, Wakefield, MA). T1-weighted high-resolution (1 × 1 × 1 mm) anatomical images were acquired for each observer with a magnetization-prepared rapid acquisition gradient

echo pulse sequence to allow accurate cortical segmentation and three-dimensional (3D) surface reconstruction.

Movie stimuli. Stimuli for the time-reversal experiment were compiled from three classic silent films [*The Adventurer* (1917); *The Navigator* (1924); *City Lights* (1931)]. Using silent films allowed us to drive activity simultaneously in many brain areas while side-stepping potential complications associated with temporal scrambling of an audio track and speech. The three clips were on average 3.39 min in duration (total duration was 12 min). The term "forward" (F) condition means simply playing the films from beginning to end. In the "backward" (B) condition, the films were presented reversed in time. For both conditions, 20 s of blank frames followed by 10 s of dynamic, symmetrical, texture patterns were shown before and at the end of the movie. The forward and backward movies were presented twice to each observer in a counter balanced order: either B1, F1, B2, F2 or F1, B1, F2, B2.

Two of the silent films (*City Lights* and *The Adventurer*) were used for the piecewise-scrambled experiment. The films were subdivided into segments and scrambled at each of three time scales: short (4 ± 1 s), intermediate (12 ± 3 s), and long (36 ± 4 s). Each original film was first divided into the segments defined by the director's cuts. For the short time scale, segments that were longer than 6 s was further divided into the minimal number of equal-length segments of duration 6 s or less (this led to only a few cuts beyond those of the directors, ~20%). The intermediate and long time scale movies did not require further subdivision of the directors' cut segments; instead, segments that were shorter than a preset minimum were joined with consecutive segments from the original movie. Each movie started with 30 s of blank frames followed by 10 s of dynamic textures and ended with 30 s of blank, and was presented twice.

Control experiment: forward versus backward block alternation. A block-design protocol was used to compare the level of activity evoked by forward versus backward presentations, consisting of 18 movie clips of diverse human biological action (e.g., kicking a ball). The entire duration was 10 min, 20 s. Each 10 s epoch contained a forward or backward clip followed by 6 s of blank frames, with a total of 36 epochs (18 forward and 18 backward), alternating randomly. Each movie clip was played once forward and once backward. The experiment started with 20 s of blank frames and 10 s of texture patterns, and ended with 12 s of blank screen.

Localizer for objects, faces, and places. Observers viewed 32 movie clips (15 s each), categorized into four distinct object types: medium shots of faces, urban buildings, natural outdoor scenes, and various objects. The experiment began with 30 s of blank frames, followed by 9 s of pattern stimuli, and ended with 21 s of blank frames (for additional information, see Hasson et al., 2004). The localizer was used to define several regions of interests (ROIs): the fusiform face area (FFA), parahippocampal place area (PPA), an area in superior temporal sulcus responsive to faces (STS-faces), and the object-sensitive lateral occipital (LO) complex.

MT+ localizer. Observers viewed 7 min of low-contrast (6%) rings surrounding a central fixation point (0.5 cycles/degree; thin light rings on a black background). Two experimental conditions of 18 s duration were presented in alternation: a stationary condition, in which each low-contrast visual ring was presented for 3 s and a moving condition in which the rings alternately expanded for 2 s and contracted for 2 s. Observers were instructed to fixate on a central cross-hair throughout the experiment.

Frontal eye field localizer. Six of the subjects also participated in an experiment that was used to localize the frontal eye fields (FEFs). Subjects performed a series of trials making saccadic eye movements, reaching hand movements, or holding the eyes and hand stationary. The FEF was localized by comparing the responses during saccade and fixation trials, ignoring the reaching hand movement trials (for additional details, see Levy et al., 2007). The average FEF regions localized in this way [Talairach coordinates: left hemisphere (LH), −32, −9, −51; right hemisphere (RH), 19, −8, 61] completely overlapped with the long temporal receptive window region found in the superior frontal sulcus (Talairach coordinates: LH, −33, −9, −51; RH, 19, −8, 61) (see Fig. 2). This anatomical agreement was used as an independent confirmation for the functional definition of this area. A more ventral area in the vicinity of the precentral sulcus and middle frontal gyrus, however, which exhibited

weak responses during saccades, may correspond to the area identified as human FEF using electrical cortical stimulation (Blanke et al., 2000).

Data analysis: preprocessing. fMRI data were analyzed with the Brain-Voyager software package (Brain Innovation, Masstricht, Netherlands) and with additional software written with Matlab (MathWorks Natick, MA). Preprocessing of functional scans included intersession and intra-session 3D motion correction, linear trend removal and high-pass filtering (up to six cycles per experiment). Spatial smoothing was applied using a Gaussian filter of 6 mm (full-width at half-maximum value). The functional maps were projected on a 3D reconstruction (inflated) or flattened representation of the cortical surface.

Correlation-based analysis. The correlation-based analysis provides a measure of the reliability of responses to time-dependent stimuli by quantifying the variability between responses to repeated presentations of the same stimulus. The correlation coefficient $C(r_1, r_2)$ is given as follows:

$$C(r_1, r_2) = \frac{r_1(t) \cdot r_2(t)}{\sqrt{(r_1(t) \cdot r_1(t))(r_2(t) \cdot r_2(t))}},$$

where $r_1(t)$ and $r_2(t)$ are the response time courses of a voxel (or brain area) to two movie presentations (e.g., two repeated presentations of the same forward movie).

In the time-reversal experiment, correlation coefficients were calculated between the responses to the following conditions: the first and second presentations of the original, forward movie ($C_{F1:F2}$); the first and second presentations of the backward movie ($C_{B1:B2}$); response to the forward movie and time-reversed response to the backward movie ($C_{rB:F}$; F and B time courses were averaged across the two presentations); finally, as a control measure, we calculated the responses to the forward movie and the backward movie ($C_{F,B}$). The highest value exhibited by any voxel in the $C_{F,B}$ analysis served as a conservative statistical criterion for thresholding the intersubject correlation maps ($r > 0.3$). The correlations were calculated separately for each voxel, or after averaging the time series across voxels within a predefined ROI. To compare the forward and time-reversed backward responses we corrected for the hemodynamic delay of the BOLD responses, by shifting the time courses by $\Delta t = 5$ s before calculation of the correlation coefficients ($C_{rB:F}$). This time-shift was chosen by calculating the correlation between the reversed-backward and forward time courses separately at different time lags in each of the predefined ROIs, and picking the value that yielded the peak correlations (see Fig. 1E and supplemental Fig. 2, available at www.jneurosci.org as supplemental material). In addition, we performed a different correction procedure: in the frequency domain, we shifted each frequency component by the phase shifts predicted by a standard model of the hemodynamic impulse response function (HRF) (Boynton et al., 1996). Similar results were obtained with the two methods. The time courses for each run (e.g., F1, F2, B1, B2) were averaged across observers before computing the pair wise correlations between the conditions. We also computed the correlations for each observer separately and then averaged the correlation coefficients across observers. Very similar results were obtained with this method, albeit with lower overall correlation values because of the higher noise in the individual time courses. Similar analysis was performed in the piecewise-scrambled experiment, but there we computed the correlation coefficients between the first and second presentations of the short-scale scrambled (S) movies, and similarly for the intermediate-scale (M) and the long-scale scrambled (L) movies ($C_{S1:S2}$, $C_{M1:M2}$, and $C_{L1:L2}$, respectively). The responses to the blank and texture pattern periods were not included in calculating the correlation coefficients.

Intersubject correlation analysis. As a complementary assessment of response reliability, we also computed the correlations across subjects within each run (e.g., the intersubject correlation for the first presentation of the short-scale scrambled movie) (supplemental Fig. 3, available at www.jneurosci.org as supplemental material). The intersubject correlation was computed for each pair of subjects on a voxel by voxel basis and within each region of interest (see below), for the forward film (F1: F2) and for each of the piecewise-scrambled films (L1:L2, M1:M2, S1:S2). We then calculated the average correlation coefficient (r) per voxel or

ROI, after applying Fisher transformation to individual coefficients. To ensure that these mean correlation values were not biased by outliers, we performed a second-order t test analysis on the pairwise values to confirm that the mean was significantly different from zero. These mean correlation values reflected the degree of similarity across subjects. The correlations across subjects between the forward and backward (not flipped) time courses in the time-reversal experiment were also computed and, as expected, were extremely low. The highest value exhibited by any voxel in this analysis served as a conservative statistical criterion for thresholding the intersubject correlation maps ($r > 0.07$).

ROI analysis. A single set of ROIs was used to analyze data from all three experiments (time-reversal, piecewise-scrambling, and block-alternation control). MT+ and object-related regions (LO, FFA, PPA, STS-faces) were defined functionally (see localizer sections above). V1 was defined anatomically by marking the calcarine sulcus (for simplicity, we refer to this ROI as V1 although it probably included part of the neighboring V2 visual cortical area). All other ROIs were defined using data from the first silent film (*The Adventurer*), identifying voxels for which the correlation values between the two presentations of the forward film were high ($C_{F1:F2} > 0.3$), and the correlations between the reversed-backward and forward time courses were small ($C_{rB:F} < ½C_{F1:F2}$). The response time courses were then sampled from the two remaining silent films for further analysis of the time-reversal experiment. This procedure ensured that the measured correlation coefficient levels were unbiased by the statistical test used to define each ROI.

To assess the variability of the results across different segments of the movies, we divided the time courses into nonoverlapping segments (30 time points each), calculated the correlation within each segment, and derived the SD for each ensemble. Response time courses from the time-reversal experiment, sampled from the two silent films that were not used to define the ROIs, were divided into six segments. Response time courses evoked by the scrambled films were divided into five segments. We then calculated the average correlation and SEM across segments, for each comparison (e.g., $C_{F1:F2}$). t tests were used to assess the statistical significance of differences in the z-transformed correlation values (supplemental Table 2, available at www.jneurosci.org as supplemental material).

Analysis of localizer responses. Visual cortical ROIs were defined from the localizers using standard methods. Our analysis consisted of a multiple regression with a regressor for each condition (e.g., faces, objects, houses, blank) in the experiment, using a box-car shape convolved with a standard model of the hemodynamic impulse response function (Boynton et al., 1996). The analysis was performed independently for the time course of each individual voxel. After computing the coefficients for all regressors, we performed a t test between coefficients of different conditions (e.g., faces vs blank). Each ROI was defined by combining the data across subjects, i.e., by aligning each subject's data to the Talairach coordinate system and treating differences between subjects as a random effect. This was done to facilitate a more direct comparison with the results from the intersubject correlation analysis which averaged across subjects in Talairach coordinates. Similar results were obtained, however, when the ROIs were defined based on the localizer responses separately for each individual subject.

Coherence. Coherence was calculated using standard methods (Mitra and Pesaran, 1999). The coherence c(f) is given as follows:

$$c(f) = \frac{\|\langle R_1(f) R_2^*(f) \rangle\|}{\langle \|R_1(f)\| \rangle \langle \|R_2(f)\| \rangle},$$

where $R(f)$ is the Fourier transform of a response time series $r(t)$, * denotes complex conjugate, $\|\ \|$ is magnitude, and $<\ >$ is expected value (i.e., mean). We subdivided each time series into 15 overlapping (but orthogonally windowed) segments, computed the Fourier transform of each segment, and then finally computed the expected values by averaging across segments.

Eye movements. Gaze direction was measured during the fMRI sessions with an infrared video camera (model 504LRO; Applied Science Laboratories, Bedford, MA) which acquired two sample points per video frame (60 Hz). Movie frames and eye-movement recordings were synchronized
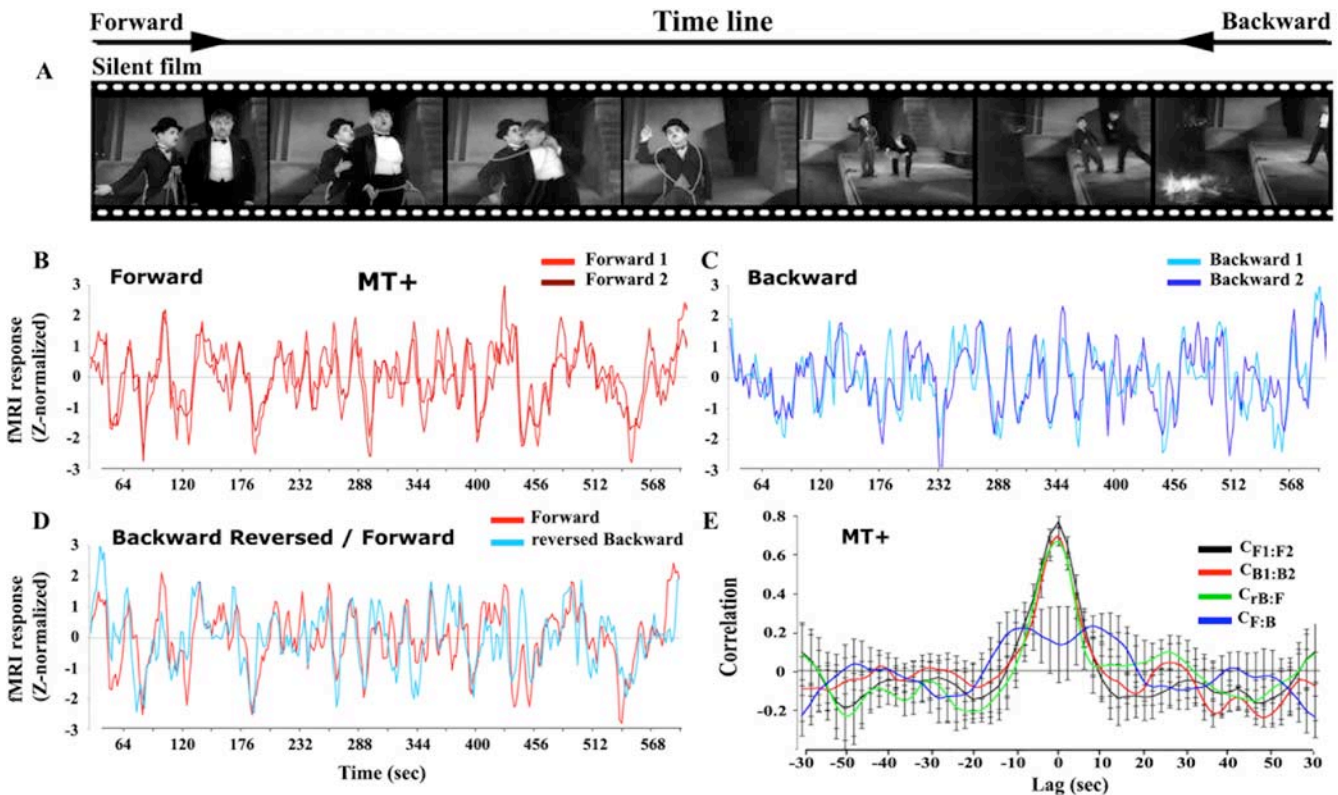
Figure 1. Correlation analysis reveals that responses in many visual cortical areas are insensitive to time reversals: illustration from area MT+. ***A***, Representative frames from the silent films (taken from *City Lights*). Observers viewed two presentations of each original movie (forward) and each time-reversed movie (backward). ***B***, Average time courses (8 observers) sampled from MT+ for the two forward presentations (F1, F2) of the silent films. ***C***, The average time courses in MT+ for the two backward presentations (B1, B2). ***D***, Superimposed traces of the time-reversed version of the backward time course (rB) and the forward time course (F). Both time courses were shifted ($\Delta t = 5$ s) to correct for hemodynamic delay in the fMRI responses. ***E***, Cross-correlations: forward 1 versus forward 2 ($C_{F1:F2}$, black); backward 1 versus backward 2 ($C_{B1:B2}$, red); forward versus backward ($C_{F:B}$, blue); reversed backward versus forward ($C_{rB:F}$, green). The correlation was computed at different lags (2 s intervals); the peak at lag 0 indicates reproducible responses that were time-locked to the movies. $C_{F:B}$, which is an estimate of the arbitrary correlation values that can be expected from such complex stimuli, is much lower than the other three cross-correlations. Error bars represent the SEM across nonoverlapping movie segments.

by adding to the movie a sound channel containing a pulse for every other video frame. These pulses were recorded together with the eye-position measurements, through an ancillary channel in the eye tracker. These data were synchronized with the MRI acquisition by also recording electrical pulses acquired from the MRI scanner that marked the onset of each volume acquisition. Nine-point calibrations were performed at the beginning and at the end of each fMRI run. Eye position was then calibrated, based on a third order polynomial-fit to the calibration data, and transformed to visual space (video frame coordinates). Saccades, blinks, and artifacts were eliminated by median filtering, but also checked manually for precision. Cross-correlation values (± 12 s) were calculated independently for horizontal and vertical eye position. Correlation coefficients were first computed within each observer and then averaged across observers.

Results

We performed two fMRI experiments, both of which involved manipulating the temporal order of classic silent films by Charlie Chaplin and Buster Keaton. We used silent films to drive activity simultaneously in as many brain areas as possible, while sidestepping potential complications associated with scrambling an audio track and speech. In one of these experiments, we parametrically varied the temporal structure of the movie sequence by scrambling the movie at three different temporal scales: short (4 ± 1 s), intermediate (12 ± 3 s), and long (36 ± 4 s). Each of the scrambled films was presented twice, and we measured the reproducibility of the responses across repeated presentations, separately for each time scale. In brain areas whose responses are

driven primarily by the instantaneous sensory input the responses should be similar across repeated presentations, regardless of temporal scrambling. In contrast, in brain regions where responses depend on sensory information accumulated over several seconds or more, the reliability of the responses should depend on the temporal scale of scrambling.

The other experiment used time-reversal as a complementary manipulation. For brain areas driven primarily by the momentary content of the movie (i.e., unaffected by preceding events), the response time course to the backward movie would be predicted from the time-reversal of the response to the forward movie. Therefore, the correlation between the time-reversed response to the backward movie and the response to the forward movie, denoted $C_{rB:F}$, would be comparable with the correlation between the response time courses to two repeats of the forward presentation ($C_{F1:F2}$). In contrast, in brain regions where responses depend on sensory information accumulated over several seconds or more, the responses to the backward movie would be unreliable (i.e., $C_{B1:B2}$ values would be low) because of the breakdown of temporal continuity caused by the time reversal, and they would also be very different from the responses to the forward movie (low $C_{rB:F}$ values).

Although a long TRW will give rise to low $C_{rB:F}$ values, the reverse is not always true: by itself, a low $C_{rB:F}$ does not necessarily imply that a region has long TRWs. For example, brain regions with short TRWs but temporally asymmetric responses (e.g.,
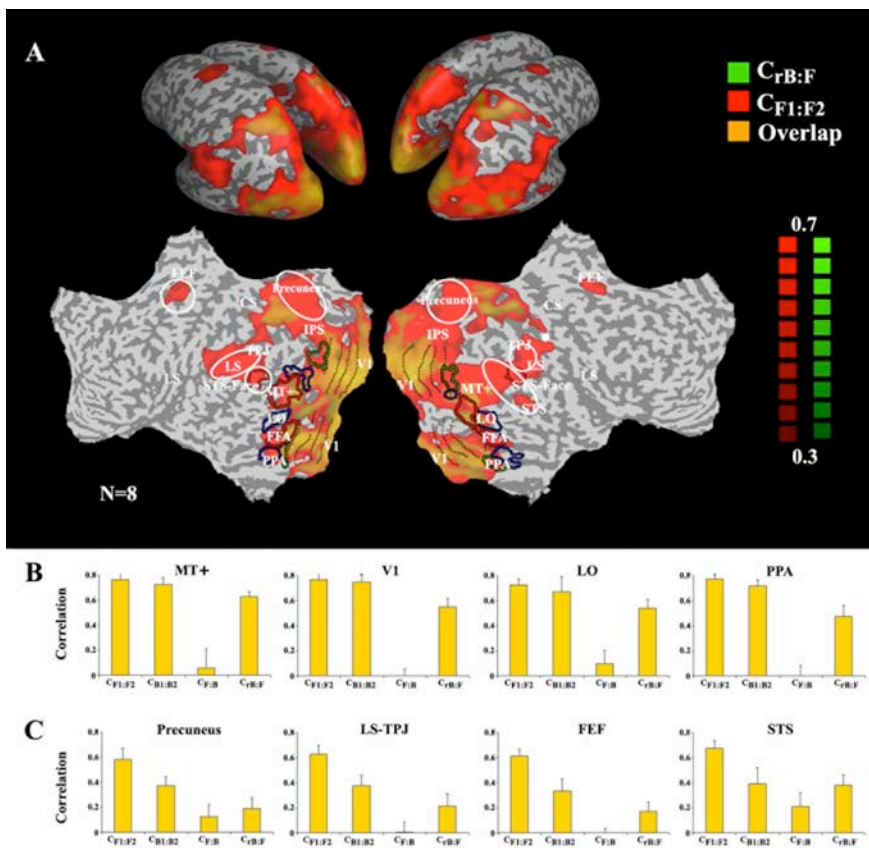
Figure 2. Effect of time reversal across the cortical surface. **A**, Maps of the correlations between the two forward time courses ($C_{F1:F2}$, red) and the reversed-backward and forward time courses ($C_{rB:F}$, green). Regions in which the responses were time-reversible exhibit both high $C_{F1:F2}$ and high $C_{rB:F}$ (overlap, orange). Correlation maps are shown on inflated (top) and unfolded (bottom) left and right hemispheres. The maps show only voxels for which the correlation exceeded a threshold value (0.3, chosen because it was above the highest $C_{F:B}$ value exhibited by any voxel). $C_{rB:F}$ was high in a number of posterior cortical regions, indicating that the responses to the backward films were a simple time reversal of the responses to the forward films. In other brain regions, we observed high $C_{F1:F2}$, but low $C_{rB:F}$ (red). White outlines mark the main regions in which responses were not time reversible. Anatomical abbreviations: ITS, inferior temporal sulcus; LS, lateral sulcus; STS, superior temporal sulcus; TPJ, temporal parietal junction; CS, central sulcus; IPS, intraparietal sulcus. Several higher-order visual areas were functionally defined based on their responses to faces (red outlines), objects (blue outlines), and houses (green outlines). Functionally and anatomically defined cortical areas: V1, primary visual cortex; MT+, MT complex responsive to visual motion; PPA, parahippocampal place area; FFA, fusiform face area; LO, lateral occipital complex responsive to pictures of objects; STS-face, area in superior temporal sulcus responsive to faces. **B**, Reliability of response time courses for regions that exhibited time-reversible responses: MT+, V1, LO, and PPA. In each of these cortical regions, the values of $C_{rB:F}$ and $C_{B1:B2}$ were similar to the $C_{F1:F2}$ values. Error bars represent the SEM across nonoverlapping movie segments. **C**, Brain regions in which responses were not time reversible (white outlines in **A**): precuneus, posterior LS, TPJ, FEFs, and the posterior STS. In each of these cortical regions, the values of $C_{B1:B2}$ and $C_{rB:F}$ were much less than $C_{F1:F2}$.

describing the results from several ROIs in early visual cortex. The ROIs were functionally defined using independent fMRI measurements of cortical activity evoked by object categories and by visual motion (for Talairach coordinates of all ROIs, see supplemental Table 1, available at www.jneurosci.org as supplemental material) (see Materials and Methods). The purpose of this analysis was to establish the validity of the basic assumptions of our approach. We focus on the motion-sensitive cortical area MT+ because it is an illustrative example (Fig. 1), although similar results were also obtained in other independently defined ROIs (Fig. 2B).

The responses in cortical area MT+ were strikingly similar for repeated presentations of both the forward and the backward films, indicating that both types of stimuli produced highly reliable responses in this brain area. Figure 1B shows the response time courses, averaged across observers, for the two forward presentations. The correlation between the two curves ($C_{F1:F2}$), provides an upper-bound estimate of what can be expected from highly reproducible responses, given the inherent variability across repeated measurements resulting from a number of cognitive and instrumental factors. The reliability of MT+ responses was also maintained when the films were played backward in time (Fig. 1C, $C_{B1:B2}$), with the values of $C_{F1:F2}$ and $C_{B1:B2}$ not statistically distinguishable from one another (Fig. 1E and supplemental Table 2, available at www.jneurosci.org as supplemental material).
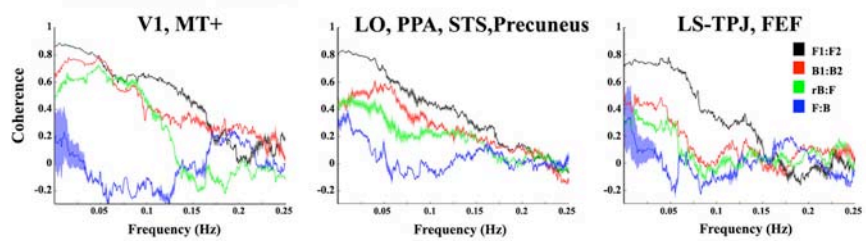
Cortical area MT+ exhibited time-reversible responses. Figure 1D shows the responses to the forward films superimposed on the time-reversed responses to the backward films (both time courses were first shifted by $\Delta t = 5$ s to correct for the hemodynamic delay) (see Materials and Methods). The reversed-backward time course (rB) closely matched the time course obtained for the forward films (Fig. 1D). The cross-correlation for $C_{rB:F}$ was nearly identical to those for $C_{F1:F2}$ and $C_{B1:B2}$. (Fig. 1E), with a peak at lag 0 indicating reproducible responses that were time-locked to the movies. Finally, given that at every moment the content of the films was different between the forward and backward presentations, the correlation between the responses to the forward and backward presentations (without time-reversing the responses) provides a lower bound estimate for the arbitrary correlation values that can be expected from such complex stimuli. As expected, $C_{F:B}$ was markedly lower in MT+ (Fig. 1E), with no peak at lag 0, confirming again that the responses that gave rise to the high $C_{F1:F2}$, $C_{B1:B2}$ and $C_{rB:F}$ values were time locked to, and hence driven by, the content of the films.

The independence of temporal order of the fMRI responses in MT+ may seem puzzling at first glance, because it is well documented that MT+ contains direction selective neurons (Huk et

stronger onset than offset responses) will yield responses that are not time reversible, because an onset in the forward movie (with a strong response) becomes an offset in the backward movie. Nevertheless, we used low $C_{rB:F}$ together with low $C_{B1:B2}$ compared with high $C_{F1:F2}$ values to identify candidate brain regions that might have long temporal receptive windows. We then used the piecewise-scrambled experiment to disambiguate the interpretation, by showing that response reliability in these candidate brain areas depended systematically on the temporal scale of scrambling.

High correlation between forward and reversed-backward time courses

Three 4 min excerpts from classic silent films (Fig. 1A) were viewed twice forward and twice backward (time reversed). We analyzed the data throughout the brain, but we begin by

al., 2001; Huk and Heeger, 2002). Indeed, a direction selective cell that responded to (say) rightward motion will not respond to the same event in the backward movie, because the motion will then be in the left ("null") direction. However, if for every cell tuned to one direction there is a nearby cell tuned to the opposite direction (Albright, 1984; Malonek et al., 1994), then it is to be expected that the fMRI signal, which sums over large numbers of neighboring cells, would yield on average equal responses to the forward and backward presentations of the movie. (Note that although this explanation is based on the relatively coarse spatial resolution of fMRI, it predicts the same results also for methods that can sample the population activity at much higher rates, such as local field potentials or EEG.)

The high $C_{rB:F}$ value found in MT+ imposes significant constraints on the temporal receptive widows of its neurons: neurons with TRWs longer than ~1 s could not yield such results. To see why, consider first the simple case of neurons that integrate (linearly sum) the incoming signal over time (possibly convolved with a kernel that may be temporally symmetric or not). Because the movie stimuli were temporally asymmetric, responses to the forward and backward movies would be shifted from each other by twice the duration of the TRW (even for neurons with temporally symmetric, or reversible kernels). Indeed, this is how the hemodynamics affected the underlying neural activity to yield the fMRI response time courses, an effect that we compensated for by shifting the time courses by 5 s (see Materials and Methods, Data analysis: preprocessing). The high $C_{rB:F}$ in MT+ was much smaller when the shift was different from this well established hemodynamic delay (Fig. 1*E* and supplemental Fig. 2, available at www.jneurosci.org as supplemental material). If the TRWs had been long in duration, then significantly longer shifts of the response time courses would have been required to maximize $C_{rB:F}$. This implies an upper bound for the characteristic TRWs of MT+ neurons of ~1 s (i.e., limited by the precision in estimating the peak of the HRF, not by the full duration of the HRF which is an order of magnitude longer). Although the analysis above is precise only in the simple case of linear temporal integration, asymmetric input to a nonlinear time-dependent process is likely to lead to even more divergent F and rB response time courses, unlikely to be brought to registration by a simple time shift. Thus, given the asymmetric nature of our stimuli, the high $C_{rB:F}$ provides strong evidence for short TRWs in MT+. Furthermore, it implies that responses to stimulus onsets and offsets were very nearly symmetrical (across the population), although we did note some evidence for adaptation effects (i.e., transient response peaks after movement onset) that were not time reversible.
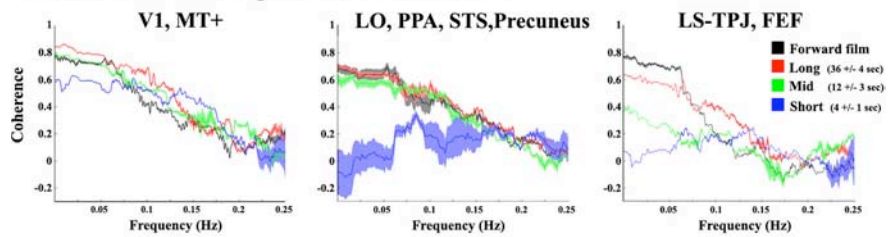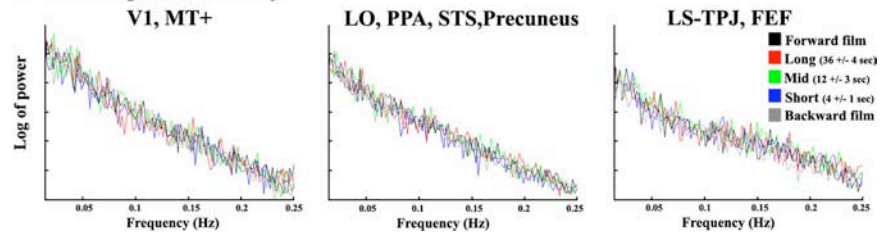

Figure 3. Coherence analysis and power spectra analysis. *A*, Coherence (correlation as a function of frequency) between the two presentations of the forward ($c_{F1:F2}$, black), backward ($c_{B1:B2}$, red), reversed-backward versus forward ($c_{rB:F}$, green), and forward versus backward ($c_{F:B}$, blue) silent films. The coherence was computed for each ROI, and averaged across regions that exhibited similar dependency on time scale in the correlation-based analyses: left, V1 and MT+; middle, LO, PPA, STS, and precuneus; right, LS, TPJ, and FEF. Error bars represent the SE across ROIs. The results of this analysis support the findings presented in Figure 2, *B* and *C*. Left, The coherence values of $c_{rB:F}$ and $c_{B1:B2}$ were similar to the $c_{F1:F2}$ values across all frequencies. Right, The coherence values of $c_{rB:F}$ and $c_{B1:B2}$ were much less than $c_{F1:F2}$ and closer to $c_{F:B}$ across all frequencies. Middle, The coherence values of $c_{rB:F}$ and $c_{B1:B2}$ were in between $c_{F1:F2}$ and $c_{F:B}$. *B*, Coherence between the two presentations of the uninterrupted forward films ($c_{F1:F2}$, black), and piecewise scrambled films at a long time scale (36 ± 4 s, red), intermediate time scale (12 ± 3 s, green), and short time scale (4 ± 1 s, blue). The results of this analysis support the findings presented in Figure 5. Left, No difference in the coherence values across conditions. Middle, Lower coherence values when the films clips were shuffled at a short time scale. Right, High coherence values only for the longest time scales. *C*, Power spectra of the responses were similar across all experimental conditions. Log of power spectrum for the forward (black), backward (gray), and randomly shuffled clips at a long time scale (36 ± 4 s, red), intermediate time scale (12 ± 3 s, green), and short time scale (4 ± 1 s, blue). Power spectra were computed separately for each subject and each ROI, and then averaged across subjects and ROIs. The results of this analysis support the findings presented in Figure 6 *A*. The power spectra were similar for all five conditions, demonstrating that observed differences in correlation values across regions were not confounded with a decrease in the response amplitudes in any frequency band. A variant of this analysis was also performed, with similar results, in which the response time courses were first averaged across subjects before computing the power spectra.

Large regions of posterior cortex exhibited time-reversible responses like cortical area MT+. These regions exhibited reproducible responses to repeated presentations of the forward films as well as the backward films (supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Critically, the correlations between the reversed-backward and forward time courses ($C_{rB:F}$) were also high (Fig. 2*A*, overlap between $C_{F1:F2}$ and $C_{rB:F}$, orange, and supplemental Table 2, available at www.jneurosci.org as supplemental material). These brain regions included retinotopic visual cortical areas, most of the higher-order visual areas (the only exception being the FFA), and a region of anterior intraparietal sulcus. Analyses of time-course data sampled from the independently defined ROIs revealed results similar to those observed in area MT+ in three additional ROIs in visual cortex (Fig. 2*B*): VI,
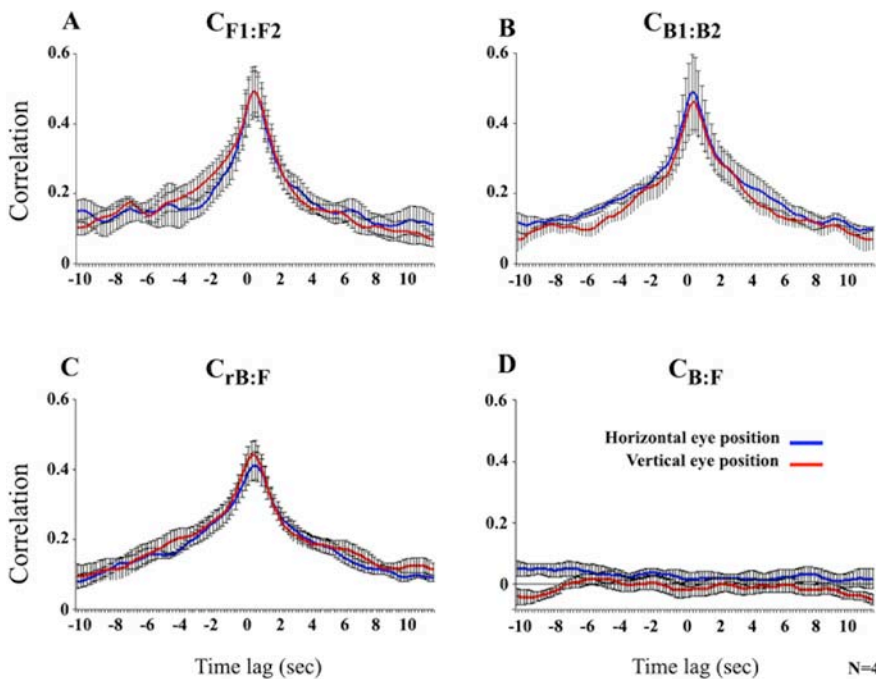
Figure 4. Reproducible eye movements regardless of time reversal. ***A***, Cross-correlations of eye positions over time for the two forward presentations of the silent films. Red curve, Vertical eye position. Blue curve, Horizontal eye position. The correlation was computed at different lags (0.2 s intervals); the peak at lag 0 indicates a reproducible sequence of eye positions in the two presentations, that was time-locked to the movies. Error bars indicate SEM across observers. ***B***, Cross-correlations for the two backward presentations. ***C***, Cross-correlations for the reversed-backward and forward presentations. The high correlations indicate that the sequence of eye positions was similar (but reversed in order) during the forward and backward viewings. Correlation coefficients at zero lag corresponding to forward:forward, backward:backward, and forward:reversed-backward were not statistically distinguishable from one another ($C_{F1:F2}$:$C_{B1:B2}$ horizontal eye position, $p = 0.54$; $C_{F1:F2}$:$C_{B1:B2}$ vertical eye position, $p = 0.43$; $C_{F1:F2}$:$C_{rB:F}$ horizontal eye position, $p = 0.22$; $C_{F1:F2}$:$C_{rB:F}$ vertical eye position, $p = 0.3$). ***D***, Cross-correlations for the backward (not flipped) and forward presentations.

bilateral ROIs (criterion used: $C_{rB:F} < \frac{1}{2}C_{F1:F2}$). We then extracted for each ROI the time courses of the responses to the other two silent films, and calculated correlation values for all movie pairs for each ROI. The cross-correlations of response time courses for the two forward presentations peaked at lag 0, with nearly the same cross-correlation peak widths in all ROIs (supplemental Fig. 2, available at www.j-neurosci.org as supplemental material). The correlations between the reversed-backward and forward time courses were, however, considerably smaller than the correlations between repeated forward presentations (Fig. 2*C* and supplemental Table 2, available at www.jneurosci.org as supplemental material). Importantly, the correlations between the two backward presentations ($C_{B1:B2}$) were also low, indicating that the reliability of responses in these regions was disrupted by time-reversing the movies. Note that these results also demonstrate that the lack of $C_{rB:F}$ in high-order brain areas cannot be attributed trivially to low-level differences between the forward and backward films (e.g., transients in luminance at scene cuts, contrast, motion, etc.), which were the same within repeated presentations of the backward ($C_{B1:B2}$) and forward ($C_{F1:F2}$) films. Additionally, the coherence values of $c_{rB:F}$ and $c_{B1:B2}$ in those ROIs were much less than $c_{F1:F2}$ and closer to $c_{F:B}$ across all frequencies (Fig. 3*A*).

LO, and PPA (see Materials and Methods for their localization and definitions).

Low correlation between forward and reversed-backward time courses

In contrast to the high $C_{rB:F}$ in posterior cortical regions, there were a number of brain regions that showed low $C_{rB:F}$ and low $C_{B1:B2}$ values. These regions included the posterior superior temporal sulcus (STS), posterior lateral sulcus (LS; also known as Wernicke's area), temporal parietal junction (TPJ), right intraparietal sulcus (IPS), the precuneus, and the superior frontal sulcus in the vicinity of the FEFs (for independent confirmation of the functional localization of the FEF, see Materials and Methods). These regions exhibited reproducible responses to repeated presentations of the forward films (Fig. 2*A*, red, high $C_{F1:F2}$), but the correlations between the reversed-backward and forward time courses were low (Fig. 2*C*), as were the correlations between the two backward presentations (Fig. 2*C* and supplemental Fig. 1, available at www.jneurosci.org as supplemental material). One possible cause for these results is that these brain areas have long TRWs because, as detailed previously, if the responses in a certain brain area depend on the past history of stimulation then it would show low $C_{rB:F}$ values. A second possibility is that these regions have short TRWs, but with temporally asymmetric responses (e.g., stronger onset than offset responses). These two alternatives were disambiguated by the piecewise-scrambled experiment (see below).

To further examine the response profile in these regions, we used data from one of the three silent films to define a number of

Reproducible eye movements regardless of time reversal

The dependence of the fMRI response correlations on temporal order could not be attributed to differences in eye movements between the forward and backward presentations (Fig. 4). The measured eye positions were independent of temporal order, exhibiting high values for the $C_{F1:F2}$, $C_{B1:B2}$, and $C_{rB:F}$ comparisons. In contrast, the correlations between the forward and the nonreversed backward ($C_{B:F}$) eye movements were low, as expected. These results indicate that observers fixated on similar image locations for similar durations, but in the opposite order, when the films were presented backwards. Moreover, the reproducibility of the eye movements suggests a comparable level of engagement while observers viewed the forward and backward films, removing potential concerns that the unreliable responses to the backward movies were because observers paid less attention to them (see also analysis of response amplitude below).

Response reliability across brain regions for short, intermediate, and long temporal scales

Time reversal of the forward movie provides one particular manipulation that disrupts the movie's temporal structure. The low $C_{rB:F}$ and $C_{B1:B2}$ values observed in high-order areas suggests that the reliability of response in these areas depends on the movie's temporal structure. To directly test this hypothesis, and to insure that the results are not specific to the time-reversal manipulation, we further manipulated the temporal structure of the movies by parametrically scrambling their temporal structure. Specifically,

we hypothesized that interfering with the temporal continuity of a forward movie at a given time scale would reduce response reliability only for neural processes that depend on longer time scales. To test this hypothesis, we measured the reproducibility of the activity evoked by scrambled versions of the forward films. This was done by first segmenting the forward films based on the director's cuts, and then randomly shuffling the order of the resulting film clips at three different time scales (see Materials and Methods): short (4 ± 1 s), intermediate (12 ± 3 s), and long (36 ± 4 s). Each of the scrambled films was presented forward two times, and we computed the correlations between responses to repeated presentations separately for each time scale.

The results revealed three profiles of response reliability as a function of time scale (Fig. 5). These can be seen both in the ROI analysis (Fig. 5A) and in the correlation map (Fig. 5B). Early visual areas (including V1 and MT+) exhibited high correlations for all three scrambled films, corroborating the results from the time-reversal experiment that they have short temporal receptive windows, regardless of the stimulus time scale. In contrast, correlation values in the STS, precuneus, FEF, LS, and TPJ were much lower for the short-time scale films than the uninterrupted forward films or the long-time scale films. The correlations were low in FEF, LS, and TPJ also for the intermediate-time scale films. These differences in reliability demonstrate that responses in LS, TPJ, and FEF depended on information presented over a longer time scale than responses in STS and precuneus, which in turn depended on a longer time scale than responses in visual cortex. Similar results were obtained also in the coherence analysis (Fig. 3B).

Given that the movies at the three levels of temporal scrambling were composed of the same short segments (4 ± 1 s) and had similar number of transient cuts, the low reliability observed in high-order areas for the scrambled movies could not have been affected by differences in the composition of adjacent events across movies or, more generally, by asymmetries in the neural response properties (e.g., stronger onset than offset responses). Note that the same logic should be applied to the backward–backward comparison ($C_{B1:B2}$) in the time-reversal experiment.

The only manipulated factor here was the "history," i.e., which segments preceded each point in time. That this history affected the response reliability of high-level cortical areas is strong evidence that they have long TRWs.

Strong response amplitudes regardless of stimulus temporal structure

Areas that responded with poor reliability to the temporally scrambled and time-reversed movies nevertheless showed high response amplitudes to those same stimuli (Fig. 6). Response
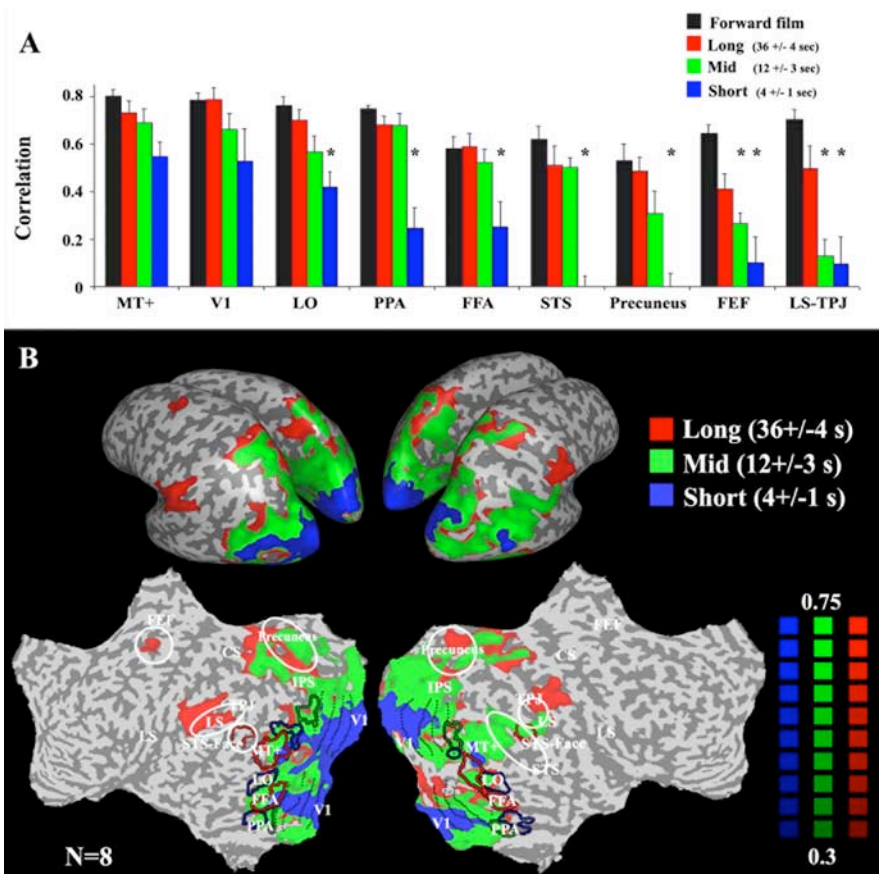

Figure 5. Effect of scrambling at different time scales. *A*, The reliability of response time courses in each of several brain regions. Time courses were sampled from the same regions of interest as in Figure 2, *B* and *C* (V1, MT+, LO, PPA, FFA, STS, precuneus, LS, TPJ, and FEF). Black, Correlations for repeated presentations of the uninterrupted forward films ($C_{F1:F2}$). Red, Piecewise scrambled at a long time scale (36 ± 4 s). Green, Intermediate time scale (12 ± 3 s). Blue, Short time scale (4 ± 1 s). Asterisks denote statistically significant differences between a given time scale and the uninterrupted forward film ($p < 0.05$, one-tailed t test). Early visual areas (V1, MT+) exhibited no difference across conditions. High-order visual areas (LO, FFA, PPA, STS, precuneus) exhibited smaller correlation values when the films clips were scrambled at a short time scale. LS, TPJ, and FEF responses were reliable only for the longest time scales. *B*, Regions in which the responses were reliable ($r > 0.3$) for all piecewise scrambled movies (at long, middle, and short time scales; e.g., MT and V1 above) are marked in blue. Regions in which the responses were reliable ($r > 0.3$) only for the long and intermediate time scales, but not when the film clips were scrambled at a short time scale (e.g., LO, PPA, FFA, STS above) are marked in green. Regions in which the responses were reliable ($r > 0.3$) only for the longest time scales (e.g., LS, TPJ, FEF above) are marked in red. Figure layout is identical to that in Figure 2 *A*.

amplitudes were quantified in two ways. First, we computed the SD of the ROIs' responses over time to each movie presentation to assess the "dynamic range" of the activity in each brain region (Nir et al., 2006). The SDs of the fMRI responses were indistinguishable for all five movie types used (continuous forward film, backward film, short-, intermediate-, and long-time scale scrambled films) in all brain areas examined, including those with long TRWs (Fig. 6A). Similarly, the power spectra were indistinguishable for all five conditions (Fig. 3C), demonstrating that observed differences in correlation values across regions were not a result of a decrease in the response amplitudes in any frequency band. These results provide support to the idea that the correlation based analysis, which provides indication of the reliability, or reproducibility of a brain region's response to a stimulus, is complementary to the more conventional analysis of response amplitudes. Indeed, we found a clear dissociation between response amplitudes and response correlations for certain stimuli.

Using a second way of quantifying response amplitudes, in a separate control experiment we confirmed that response amplitudes evoked by backward movies were the same as or larger than
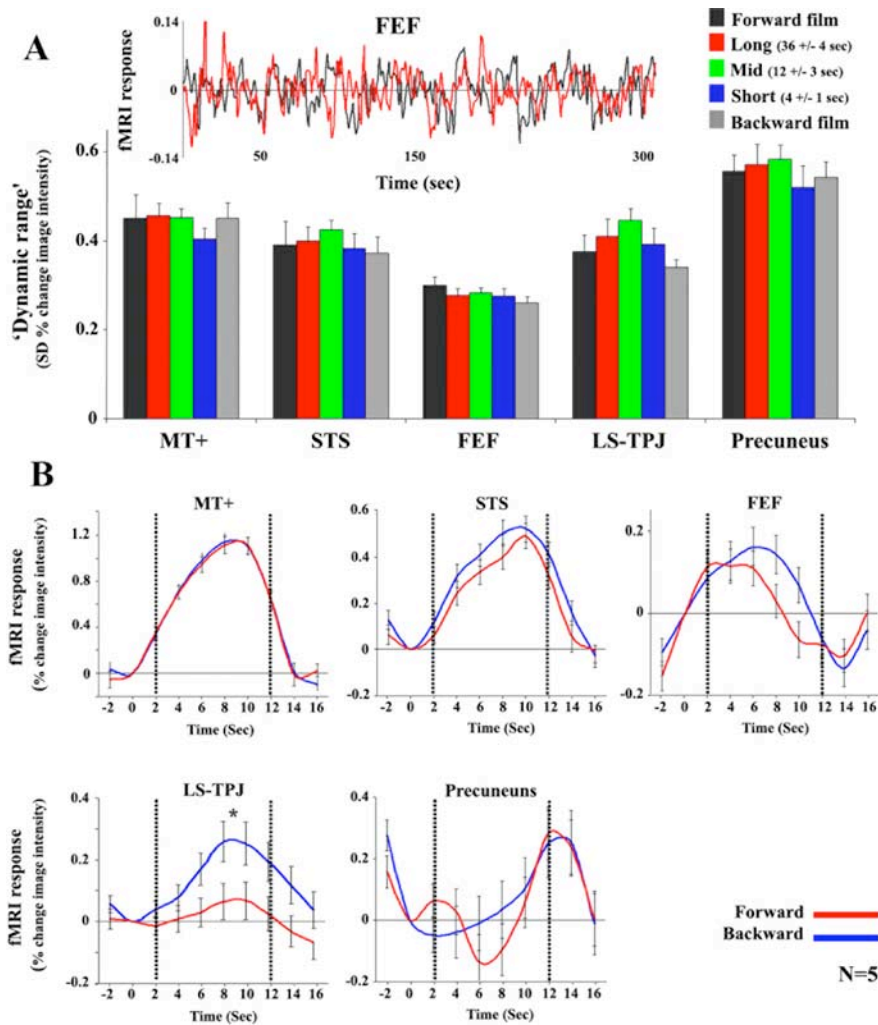
Figure 6. Dissociation between reliability and response amplitude. **A**, Dynamic range of responses. SD of the fMRI responses (dynamic range) for forward, time-reversed, and scrambled films. Time courses were sampled from a subset of the same regions of interest (MT+, STS, LS, TPJ, FEF and precuneus) used in Figure 2, *B* and *C*. Inset, Time course of responses for the forward and reversed-backward silent films from FEF in one representative subject. Bars and error bars indicate the mean and SD across observers (*n* = 8 for the forward and backward films; *n* = 5 for the scrambled films). Note that the dynamic range of the fMRI responses within each brain region was similar for all five films. **B**, fMRI responses to the block-alternation control experiment, in which short movie clips were presented forward (red) and backward (blue) relative to blank screen. Time courses were sampled from a subset of the same regions of interest (MT+, STS, FEF, LS, TPJ and precuneus) used in Figure 2. Dotted lines mark the beginning and end of the epoch. In all regions the response modulations evoked by the backward movie clips were the same as or larger than those evoked by forward clips.

those evoked by forward movie clips (Fig. 6*B*). We used a standard block-alternation protocol in which the responses to forward and backward movie clips were compared against a baseline blank condition. Each 10 s movie clip was presented once forward and once backward, in randomly shuffled order, with each clip separated by a 6 s blank screen (see Materials and Methods). Time courses were sampled from the same ROIs as those used for the analysis of the main experiments. In MT+, STS, and FEF, the level of activity was similar in the forward and backward conditions (*t* test, MT+, $p = 0.46$; STS, $p = 0.23$; FEF, $p = 0.53$). In LS and TPJ, the responses were larger for backward movies than forward movies ($p < 0.04$). In precuneus, although there was a decrease in activity during external stimulation (movie clips) relative to baseline, there was no significant amplitude difference between the forward and backward conditions ($p = 0.26$).

Intersubject correlation analysis

In a complementary analysis of the piecewise-scrambled experiment, the response time courses from individual observers were used to compute all pairwise intersubject correlations for each voxel and ROI, for each piecewise-scrambled film (supplemental Fig. 3, available at www.jneurosci.org as supplemental material). Large regions of posterior cortex exhibited reproducible responses across subjects (i.e., high intersubject correlations) for the scrambled films at all three time scales. Those regions showed also high intersubject correlations for the backward film (data not shown). In contrast, the intersubject correlations in the STS, precuneus, FEF, LS, and TPL were high only for the long time scale scrambled films, but not for the short time scale. Similar results were obtained for the time-reversal experiment (data not shown). These results indicate that in these high-level cortical areas, the history of stimulation also affected the reproducibility of responses across different individuals.

Discussion

To assess the TRWs across different brain areas, we varied parametrically the temporal structure of silent films, by either time reversing the movie sequence (playing it backward) or by scrambling the movie at different temporal scales. Using a novel correlation-based analysis we presented evidence for a hierarchy of different characteristic TRWs in different brain areas, analogous to the well established hierarchy of spatial receptive field sizes in visual cortex (Hubel, 1988). Our results indicate that, whereas neuronal responses in early visual cortex (V1, MT+) were determined primarily by the instantaneous sensory input, responses in STS and precuneus were affected by information accumulated over intermediate time scales (~12 s), and those in the posterior portion of the LS, TPJ, and FEFs were affected by yet longer time scales (~ 36 s) (Figs. 2, 3*A*,*B*, 5).

Furthermore, we found a clear dissociation between the reliability of the responses (Figs. 2, 3) and response amplitudes (Figs. 3*C*, 6). For example, in the LS and TPJ, we observed large response amplitudes for all movies, but the responses to the scrambled and time-reversed movies were much less reliable than the responses to the intact forward movies. We interpret the strong response amplitudes in these brain regions as reflecting incessant processing, presumably aimed to extract meaningful information from the stimuli. At the same time, the low reliability of the responses to temporally disrupted movies indicates a failure to attain a consistent sequence of neural states in those regions (and, correspondingly, of cognitive states) while viewing the nonecological stimuli. It is useful to draw an analogy with the multistable fluctuations in perception of highly ambiguous visual display,

such as the rotating snake motion illusion (Murakami et al., 2006), the Marroquin pattern (Wilson et al., 2000), or the perceptual organization effects explored by some of the op artists in the 1960s. In all of these cases, visual neurons presumably respond with large amplitudes while processing the stimuli, but the responses are unreliable, leading to a failure to "lock in" to a consistent and stable perceptual organization. Response reliability, therefore, provides information about neural processing that is complementary to that derived from more traditional measurements of response amplitude, and can uncover phenomena that response amplitudes alone do not reveal, such as the long temporal receptive windows found in this study.

Short temporal scale of processing in early visual areas
The short temporal receptive windows observed for early visual cortex are in agreement with the notion that the visual system is optimized for rapidly processing the instantaneous visuospatial properties of a stimulus (Potter, 1975; Rolls and Tovee, 1994; Thorpe et al., 1996). Using continuous rapid serial visual presentation of unrelated natural images, for example, it was shown that some neurons in the visual system can preserve their stimulus selectivity even at surprisingly fast presentation rates of 14 ms/ image (Keysers et al., 2001). Such remarkable selectivity to a single frame is an indication of nearly instantaneous, or time independent, responses. But other neurons in early visual areas, for example direction selective neurons in V1 and MT, necessarily accumulate information over temporal scales from 20 to 80 ms up to 200–400 for slow moving stimuli (Bair and Movshon, 2004). Differentiating between those time scales was not possible in our experiments; detailed characterization of the very fast spatiotemporal response properties of individual neurons in early visual cortex would require methods with better spatial as well as temporal resolutions (e.g., single-unit electrophysiology).

Long temporal scale of processing in high-order areas
On the other end of the spectrum, responses in LS, TPJ, and FEF were affected by information accumulated over very long time scales (>30 s) (Fig. 5A). Although such long temporal scales are consistent with the temporal durations of many real world events (see also below), the possibility of such long TRWs was largely ignored until now.

These results should not, however, be taken to imply the brain areas we found are the only ones with long temporal receptive windows. Although our complex, naturalistic visual stimuli (silent films) evoked activity simultaneously in many brain areas, we could not assess the temporal receptive windows in areas that were not driven reliably by those stimuli (e.g., auditory cortex and regions of prefrontal cortex other than FEF).

Nor should our results be taken to imply that each area has a single fixed temporal receptive window. Similar events can vary in their temporal pace. Speech processing provides a useful illustration of this. Because some people have faster or slower speech rates than others, listeners are exposed to the same word in utterances that can differ considerably in their duration. The number of constituent phonemes, however, is invariant of speech rate. Therefore, it makes sense for the temporal receptive windows in speech perception areas to depend on the number of incoming phonemes, and not just on time per se. To envision how such flexible TRWs may be achieved, it is again constructive to draw on findings about the organization of visual spatial receptive fields. Although the spatial receptive fields of early visual cortical neurons are relatively inflexible in terms of stimulus size, inferiotemporal neurons respond to the same shape over a range of

spatial scales, as if responding to a particular combination of building blocks that are relatively invariant to spatial scale. By analogy, TRWs in higher-order cortical areas may likewise operate over a range of time scales: for example, a neuron involved in speech perception may accumulate information from both fast and slow versions of its constituent phonemes.

In this context, it is important to note that real life events unfold with a natural pace that varies over a fairly restricted range. For example, it typically takes a fraction of a second to utter a word and a few seconds to utter a sentence. We conjecture, therefore, that each brain area is tuned to extract information at the range of time scales that is ecologically appropriate for the content it processes. Accordingly, it should be harder to extract information at a pace that is significantly faster or slower than this ecologically determined range. In this study, we scrambled the temporal structure of a movie at three very distinctive temporal scales (~4, ~15, and ~40 s), while keeping the rate of presentation constant. Additional studies, which will vary the presentation rate, will be needed to further characterize the range of time scales required for evoking reliable responses within each brain region.

Temporal receptive window and functional specialization
Although knowledge of the TRW does not specify the function of a brain area (indeed, functional specialization was very specifically not the goal of the current study), the TRW does place important limits on what the function(s) of an area might be. The short TRWs observed in early visual cortex are in agreement with the notion that these brain areas are optimized for rapidly processing the instantaneous visuospatial properties of a stimulus. However, many cognitive processes require accumulation of information over time. Several examples follow.

First, activity related to the processing of narrative has previously been reported in the posterior part of LS, also known as Wernicke's area (Ferstl et al., 2005; Xu et al., 2005; Schmithorst et al., 2006). In our experiment, the LS had a long temporal receptive window, which coincides with the idea that accumulation of information over time in high-order linguistic areas is needed for processing the movie's plot. Note, however, that our study used nonverbal stimulation whereas previous studies of narrative all used verbal stimuli. Second, LS, TPJ, FEF, and STS, together with the anterior paracingulate cortex (not seen in this study), have been reported to be involved in theory of mind: the capacity to discern the beliefs, desires, and goals of others and to predict their actions (Fletcher et al., 1995; Gallagher et al., 2000; Gallagher and Frith, 2003; Saxe and Kanwisher, 2003; Saxe et al., 2004; Vollm et al., 2006). In a separate behavioral study, we confirmed that viewers encountered severe difficulties in describing the characters' objectives, intentions, and motivations during the backward films (supplemental Fig. 4, available at www.jneurosci.org as supplemental material). However, the TPJ region that we identified (mean Talairach coordinates: 51, −37, 23) was ~1.5 cm anterior to that reported previously to be activated by theory of mind tasks [mean Talairach coordinates: 51, −54, 27 (Saxe and Kanwisher, 2003)], leaving open the possibility that these are two neighboring (perhaps related) functional areas. Third, activity in the posterior lateral sulcus and precuneus has been reported to be correlated with the level of predictability of a sequence of visual stimuli (Bischoff-Grethe et al., 2000; Han et al., 2006), another example of a cognitive process that requires accumulation of information over time. Our results extend these findings by showing that although reducing predictability by reversing the movies led to a dramatic reduction in the reliability of neural activity in

similar brain regions, it did not cause a decrease in the amplitude of activity. Fourth, activity in right middle temporal cortex, right posterior superior temporal cortex, and right inferior postcentral gyrus has been reported to be also correlated with the perceived coherence of visual events (Bischoff-Grethe et al., 2000; Han et al., 2006). Our study goes beyond those studies by parametrically measuring the window of temporal coherency needed to evoke reproducible activation in these regions. Moreover, our results suggest that measuring the reliability of neural activity is a more sensitive tool to look for such temporal differences than the more standard measurement of response amplitude. Fifth, MT+, STS, FEF, and precuneus have been reported to be involved in segmenting the incoming stream of sensory information into meaningful events (Zacks et al., 2001a,b), yet another cognitive process that requires long temporal receptive windows. Although we are a long way from understanding how the brain performs these complex cognitive functions, in this study we have delimited brain areas with long temporal receptive windows, an essential characteristic for being directly involved in these long time scale functions.

Perhaps the most surprising of our results was that FEF exhibited a relatively long temporal receptive window. One might hypothesize that the breakdown of reproducible activity in FEF was caused by different patterns of eye movements during the backward and scrambled films. However, we found reproducible eye movements while viewing the backward films (Fig. 4), dissociated from the lack of reproducible FEF activity (Fig. 2 and supplemental Figs. 1, 3, available at www.jneurosci.org as supplemental material). From this, we conclude that FEF activity was not solely tied to eye position as has often been assumed, but rather that it also depended on a longer time scale, perhaps serving a function other than driving instantaneous eye position. One explanation for these findings is that the area we localized as FEF (Levy et al., 2007), although it responded strongly during saccades, may not be the same area of the brain as that characterized in the macaque monkey. Neurons in several frontal areas, in addition to the FEF, are active during visually guided saccades such that FEF can be identified definitively only by using electrical stimulation of the cortex to evoke eye movements (Bruce et al., 1985; Blanke et al., 2000). A more ventral area in the vicinity of the precentral sulcus and middle frontal gyrus, which exhibited weak responses during saccades in our experiments, may correspond to the area identified as human FEF using electrical cortical stimulation (Blanke et al., 2000). Regardless, consistent with our conclusion, some of the known properties of FEF physiology exhibit long time scales, independent of eye position. Activity in FEF, as well as other prefrontal areas active during saccades, can reflect the plan or intention to perform a saccade well before the eyes are moved, even if the motor plan is later cancelled and no eye movements are made (Hanes et al., 1998; Schall and Thompson, 1999). Activity in these prefrontal areas is also believed to be involved in spatial attention (Moore and Fallah, 2001) and spatial working memory (Curtis, 2006), both of which can be maintained over time without moving the eyes.

A hierarchy of temporal receptive windows in cortex

In summary, our results provide support for the hypothesis that there is a hierarchy of progressively longer temporal receptive windows in the human brain, similar to the known cortical hierarchy of spatial receptive fields. Importantly, one should distinguish between TRWs and long term changes in the responses of neurons which are caused by adaptation and learning. Again by analogy, although past exposure and/or attentional and contextual constrains can affect the response properties of spatial receptive fields (Moran and Desimone, 1985; Sheinberg and Logothetis, 2001; Ahissar and Hochstein, 2004; Furmanski et al., 2004), the notion of a cortical hierarchy of spatial receptive fields is still useful for characterizing the relationships between neurons in each region. Similarly, although the TRW of a neuron may vary with learning, context, adaptation, and attentional state, we still propose that it is a useful concept. Despite contextual modulation and plasticity, there will always be a division of labor between regions with short TRWs that process the incoming sensory signals and regions with longer TRWs that accumulate information for processing complex real life events which unfold over time.

References

Ahissar M, Hochstein S (2004) The reverse hierarchy theory of visual perceptual learning. Trends Cogn Sci 8:457–464.

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52:1106–1130.

Bair W, Movshon JA (2004) Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. J Neurosci 24:7305–7323.

Bartels A, Zeki S (2004) Functional brain mapping during free viewing of natural scenes. Hum Brain Mapp 21:75–85.

Bischoff-Grethe A, Proper SM, Mao H, Daniels KA, Berns GS (2000) Conscious and unconscious processing of nonverbal predictability in Wernicke's area. J Neurosci 20:1975–1981.

Blanke O, Spinelli L, Thut G, Michel CM, Perrig S, Landis T, Seeck M (2000) Location of the human frontal eye field as defined by electrical cortical stimulation: anatomical, functional and electrophysiological characteristics. NeuroReport 11:1907–1913.

Boynton GA, Engel SA, Glover G, Heeger D (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16:4207–4221.

Bruce CJ, Goldberg ME, Bushnell MC, Stanton GB (1985) Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. J Neurophysiol 54:714–734.

Curtis CE (2006) Prefrontal and parietal contributions to spatial working memory. Neuroscience 139:173–180.

Ferstl EC, Rinck M, von Cramon DY (2005) Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. J Cogn Neurosci 17:724–739.

Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, Frith CD (1995) Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition 57:109–128.

Fonlupt P (2003) Perception and judgement of physical causality involve different brain structures. Brain Res Cogn Brain Res 17:248–254.

Furmanski CS, Schluppeck D, Engel SA (2004) Learning strengthens the response of primary visual cortex to simple patterns. Curr Biol 14:573–578.

Gallagher HL, Frith CD (2003) Functional imaging of "theory of mind." Trends Cogn Sci 7:77–83.

Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD (2000) Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. Neuropsychologia 38:11–21.

Golland Y, Bentin S, Gelbard H, Benjamini Y, Heller R, Nir Y, Hasson U, Malach R (2006) Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation. Cereb Cortex 17:766–777.

Han S, Jiang Y, Mao L (2006) Right hemisphere dominance in perceiving coherence of visual events. Neurosci Lett 398:18–21.

Hanes DP, Patterson II WF, Schall JD (1998) Role of frontal eye fields in countermanding saccades: visual, movement, and fixation activity. J Neurophysiol 79:817–834.

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. Science 303:1634–1640.

Hubel DH (1988) Eye, brain, and vision. New York: Scientific American Library.

Huk AC, Heeger DJ (2002) Pattern-motion responses in human visual cortex. Nat Neurosci 5:72–75.

Huk AC, Ress D, Heeger DJ (2001) Neuronal basis of the motion aftereffect reconsidered. Neuron 32:161–172.

Keysers C, Xiao DK, Foldiak P, Perrett DI (2001) The speed of sight. J Cogn Neurosci 13:90–101.

Levy I, Schluppeck D, Heeger DJ, Glimcher PW (2007) Specificity of human cortical areas for reaches and saccades. J Neurosci 27:4687–4696.

Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. Science 268:1503–1506.

Malonek D, Tootell RB, Grinvald A (1994) Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. Proc R Soc Lond B Biol Sci 258:109–119.

Mechler F, Victor JD, Purpura KP, Shapley R (1998) Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. J Neurosci 18:6583–6598.

Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. Biophys J 76:691–708.

Moore T, Fallah M (2001) Control of eye movements and spatial attention. Proc Natl Acad Sci USA 98:1273–1276.

Moran J, Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. Science 229:782–784.

Murakami I, Kitaoka A, Ashida H (2006) A positive correlation between fixation instability and the strength of illusory motion in a static display. Vision Res 46:2421–2431.

Nir Y, Hasson U, Levy I, Yeshurun Y, Malach R (2006) Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. NeuroImage 30:1313–1324.

Potter MC (1975) Meaning in visual search. Science 187:965–966.

Rolls ET, Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. Proc R Soc Lond B Biol Sci 257:9–15.

Saxe R, Kanwisher N (2003) People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind." NeuroImage 19:1835–1842.

Saxe R, Carey S, Kanwisher N (2004) Understanding other minds: linking developmental psychology and functional neuroimaging. Annu Rev Psychol 55:87–124.

Schall JD, Thompson KG (1999) Neural selection and control of visually guided eye movements. Annu Rev Neurosci 22:241–259.

Schmithorst VJ, Holland SK, Plante E (2006) Cognitive modules utilized for narrative comprehension in children: a functional magnetic resonance imaging study. NeuroImage 29:254–266.

Sheinberg DL, Logothetis NK (2001) Noticing familiar objects in real world scenes: the role of temporal cortical neurons in natural vision. J Neurosci 21:1340–1350.

Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381:520–522.

Vollm BA, Taylor AN, Richardson P, Corcoran R, Stirling J, McKie S, Deakin JF, Elliott R (2006) Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. NeuroImage 29:90–98.

Wilson HR, Krupa B, Wilkinson F (2000) Dynamics of perceptual oscillations in form vision. Nat Neurosci 3:170–176.

Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. Cereb Cortex 18:230–242.

Xu J, Kemeny S, Park G, Frattali C, Braun A (2005) Language in context: emergent features of word, sentence, and narrative comprehension. NeuroImage 25:1002–1015.

Yao H, Shi L, Han F, Gao H, Dan Y (2007) Rapid learning in cortical coding of visual scenes. Nat Neurosci 10:772–778.

Zacks JM, Tversky B, Iyer G (2001a) Perceiving, remembering, and communicating structure in events. J Exp Psychol Gen 130:29–58.

Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME (2001b) Human brain activity time-locked to perceptual event boundaries. Nat Neurosci 4:651–655.