A novel measure to determine viewing priority and its neural correlates in the human brain

Jan-Bernard C. Marsman — NeuroImaging Center, Laboratory for Experimental Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Frans W. Cornelissen — Laboratory for Experimental Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Michael Dorr — Institute for Human-Machine Communication, Technische Universität München, München, Germany

Eleonora Vig — Xerox Research Center Europe, Meylan, France

Erhardt Barth — Institute for Neuro- and Bioinformatics, University of Lübeck, Lübeck, Germany

Remco J. Renken — NeuroImaging Center, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

A key property of human visual behavior is the very frequent movement of our eyes to potentially relevant information in the environment. Observers thus continuously have to prioritize information for directing their eyes to. Research in this field has been hampered by a lack of appropriate measures and tools. Here, we propose and validate a novel measure of priority that takes advantage of the variability in the natural viewing behavior of individual observers. In short, our measure assumes that priority is low when observers' gaze behavior is inconsistent and high when it is very consistent. We calculated priority for gaze data obtained during an experiment in which participants viewed dynamic natural scenes while we simultaneously recorded their gaze position and brain activity using functional magnetic resonance imaging. Our priority measure shows only limited correlation with various saliency, surprise, and motion measures, indicating it is assessing a distinct property of visual behavior. Finally, we correlated our priority measure with the BOLD signal, thereby revealing activity in a select number of human occipital and parietal areas. This suggests the presence of a cortical network involved in computing and representing viewing priority. We conclude that our new analysis method allows for empirically establishing the priority of events in near-natural vision paradigms.

Introduction

During natural visual behavior such as visual search, human observers make eye movements toward locations that require further attention and scrutiny (Yarbus, 1967). Therefore, eye movements indicate what observers (or at least their visual systems) consider potentially important in a scene (Henderson, 2003). A key challenge in visual neuroscience has been to explain how the human brain prioritizes certain events or objects in a scene over others, with priority being defined as a combination of all bottom-up and top-down factors driving overt gaze behavior. Unfortunately, no measure of priority exists today. An

Citation: Marsman, J.-B. C., Cornelissen, F. W., Dorr, M., Vig, E., Barth, E., & Renken, R. J. (2016). A novel measure to determine viewing priority and its neural correlates in the human brain. *Journal of Vision*, *16*(6):3, 1–18, doi:10.1167/16.6.3.

doi: 10.1167/16.6.3 Received January 17, 2015; published April 8, 2016 ISSN 1534-7362

important reason is that a computational derivation of priority is challenging, if not impossible.

Assuming that eye movements form the observable output of the putative prioritization process, we hypothesized that we can take advantage of the variability of the viewing behavior across individuals in order to design a novel measure to empirically determine priority in naturalistic viewing paradigms. The second hypothesis that we tested is if this measure can be used to trace the neural correlates of priority in the human visual cortex of observers looking at dynamic natural stimuli.

Two aspects appear to determine where an observer will gaze at in a scene: first, locations that "stand out" in one way or the other from their surrounding may attract gaze. Second, top-down factors such as the task at hand or instructions also affect the guidance of our eyes (Buswell, 1935; Gitelman et al., 1999; Hayhoe & Ballard, 2005; Rothkopf, Ballard, & Hayhoe, 2007; van Beilen, Renken, Groenewold, & Cornelissen, 2011). A number of theories and computational models propose that the brain uses stimulus information, such as local contrast or orientation differences, to determine regions that stand out from their surroundings, thereby constructing a saliency map (a map highlighting the most prominent features). Such a map can be used to predict the gaze behavior of observers (Bruce & Tsotsos, 2009; Itti & Koch, 2001; Itti, Koch, & Niebur, 1998; Koch & Ullman, 1985). However, even though models based on the concept of saliency have become increasingly sophisticated and can handle natural dynamic stimuli, they still fail to account for a substantial portion of observers' fixations (Hayhoe & Ballard, 2005). Moreover, task demands can easily overrule stimulus saliency as the dominant factor determining viewing behavior (Henderson, Brockmole, Castelhano, & Mack, 2007). For this reason, it has been argued that task, rather than saliency, may be the primary determinant of viewing behavior (Hayhoe & Ballard, 2005). Furthermore, memory (Aivar, Hayhoe, Chizk, & Mruczek, 2005) and anticipatory behavior play an important role in guiding viewing behavior (Hayhoe & Ballard, 2005). This implies that contextual information can modulate the expression of saliency. Indeed, the neuronal expression of object saliency is higher when it is consistent with the behavioral goals of the observer (Fecteau & Munoz, 2006). This finding led Fecteau and Munoz (2006) to propose that priority—the integrated representation of saliency and task-related relevance—could describe the firing patterns of neurons in the human brain more effectively.

Here, we here propose an empirical and quantitative measure of viewing priority based on an observer's gaze position that can be used to quantify viewing priority during the free observation of dynamic natural scenes. Our measure exploits the variability in the natural viewing behavior of individual observers. Our intuition was the following: When the gaze behavior of an observer is consistent with that of other observers, the information gazed at has a high priority. If, on the other hand, the observer's gaze behavior is inconsistent with others, the information has a low priority. In the Methods section below, we provide a more formal description of our measure. We found that our new measure was highly consistent across participants, yet was clearly different from state-of-the-art image-based measures of saliency, surprise, and motion. This indicates that our measure of priority captures both the bottom-up and top-down influences that drive gaze behavior in natural dynamic conditions.

Next, we explored the use of our measure to trace the neural correlates of viewing priority using concurrent functional magnetic resonance imaging (fMRI) and eye tracking in participants viewing movies of natural scenes. Our results suggest viewing priority is represented in the temporal parietal junction (TPJ), the precuneus, and area V5/MT in the human brain.

Methods

Definition of viewing priority

We defined viewing priority as the combination of salience and relevance and assume it underlies the viewing pattern of an observer. We assumed that a participant gazes at aspects of a scene to which his brain assigned the highest viewing priority. Based on this assumption, we hypothesized that viewing priority can be derived from gaze behavior of a single participant as follows: When the viewing behavior of a participant is consistent with that of a reference group, it has a high viewing priority. In other words, if everyone is fixating at the same point in the same context, the fixation must be a prioritized one. Therefore, fixation events have a higher priority when the individual's viewing behavior is more consistent with that of a reference group of observers.

Mathematically, this behavior—the conformity of a fixation to a reference group of fixations—can be seen as a typical example of a clustering problem. In the general case, the clustering of items is based on calculating optimal class centers based on minimizing distances within the class and maximizing distances between classes. Here, if the distance between the fixation and the reference group is small, it has a high priority. The further away this distance falls from the center of the reference group, the lower the priority value should become. Therefore, our goal is to derive a value of priority based on the distances between the fixation at hand and the reference group. However, in standard clustering algorithms (e.g., k-nearest neighbor) the labeling is discrete and would only discrimi-

nate whether or not the fixation belongs to the reference group. An improvement upon this method is fuzzy c-means (FCM) clustering (Bezdek, 1981), allowing for probability of the final class labels (i.e., the degree of belongingness of the fixation to the reference group and therefore providing in a priority value).

The following definitions were used: (a) Fixation of interest: the individual fixation made by an observer in a certain context (e.g., watching a video sequence, or performing a cognitive task) at a specific time. When comparing an observers' fixation with that of a reference group, the context has to be the same. Both fixation location and duration are taken into account. (b) Reference set: the distribution of fixations made around the same time by other observers in the same context. (c) Random set: the distribution of fixations made around the same time by other observers in a different context (e.g., watching a different video sequence).

Priority was calculated by comparing the fixation of interest with regards to the reference set. The null distribution for priority is given by the distribution of the priorities for each fixation of the random set to the reference set (see the example in Figure 1). The priority associated with a fixation during a certain context is characterized by calculating the degree of belongingness to each reference fixation where the definition of belongingness is derived from the FCM methodology (Bezdek, 1981).

FCM is traditionally applied to divide a finite set of data into a number of classes. The strength of FCM compared to traditional k-means clustering is having continuous membership values instead of discrete ones. In FCM, these membership values are calculated for each element with respect to each class, indicating to what degree the element belongs to the class. The sum over all distances between each element x_i and each class center c_j is minimized (Equation 1). The formula for belongingness of element x_i to class c_j is given in Equation 2:

$$\operatorname{argmin} \sum_{i=1}^{n} \sum_{j=1}^{c} w_{ij}^{m} ||x_i - c_j||^2 \quad (1)$$

$$w_{i,j} = \frac{1}{\sum_{k=1}^{c} \left(\frac{||x_i - c_j||}{||x_i - c_k||} \right)^{\frac{2}{m-1}}} \quad (2)$$

We adapted this algorithm and state that each element (in our case a fixation location) defines its own class, and make use of the continuous membership scores for the current fixation location toward the location of reference fixations.

A step-by-step procedure for calculating the priority measure is given below.

Step 1

Given a fixation (f) of an observer, it has a location within a certain time interval τ (i.e., all time points falling inside the start and end of the fixation) in a given context (e.g., a movie). A reference set of fixation locations (X) is selected from all fixations of other observers that show a temporal overlap with τ during that same context. Similarly, a random set of fixation locations (R) is selected, again from the other observers, using the same time interval but made during a set of different contexts.

Step 2

Let x_k be the kth element of X; let x_l be an element of X with $l \neq k$; and let r_i be the ith element of R. First, belongingness μ_k of an arbitrary fixation q having the location f_q is defined as:

$$\mu_k(f_q) = \frac{1}{\left(\sum_{l \neq k} \frac{d(f_q, x_k)}{d(x_l, x_k)} \right)^{\frac{2}{m-1}}}$$

where $d(f_q, x_k)$ and $d(x_p, x_k)$ are the Euclidian distances between f_q and x_l to x_k, respectively and m is a fuzzy parameter. The ratio between $d(f_q, x_k)$ and $d(x_p, x_k)$ represents the weighted distance between f_q and x_k. In clustering algorithms, m is commonly defined such that the total belongingness is equal to 1, and the individual belongingness μ_k represents a fraction. As a result, m will adapt to the spatial distribution of X (see Figure 2 and Appendix A). To prevent the algorithm from becoming unstable, zero distances are not taken into account in the calculation of belongingness. In principle, each arbitrary fixation f_q will result in a different value for m. Therefore, m_i is calculated for each fixation in R, r_i (Figure 2).

Step 3

The parameter m is fixed to the median of all m_is for all contexts. In this way, the belongingness captures the effects of the spatial distribution of X. Note that typically, m is chosen a priori with a desired level of fuzziness. Due to possible spatial bias in both the random and reference set, m has to be determined on the data first, before this a priori value can be set.

Step 4

In order to calculate priority, the total belongingness ($\sum \mu_k$) is calculated using the fixed m as described above, for the fixation of interest (f_q) and all random fixations. Priority is now defined as the fraction of total belongingness of R below total belongingness of f_q.

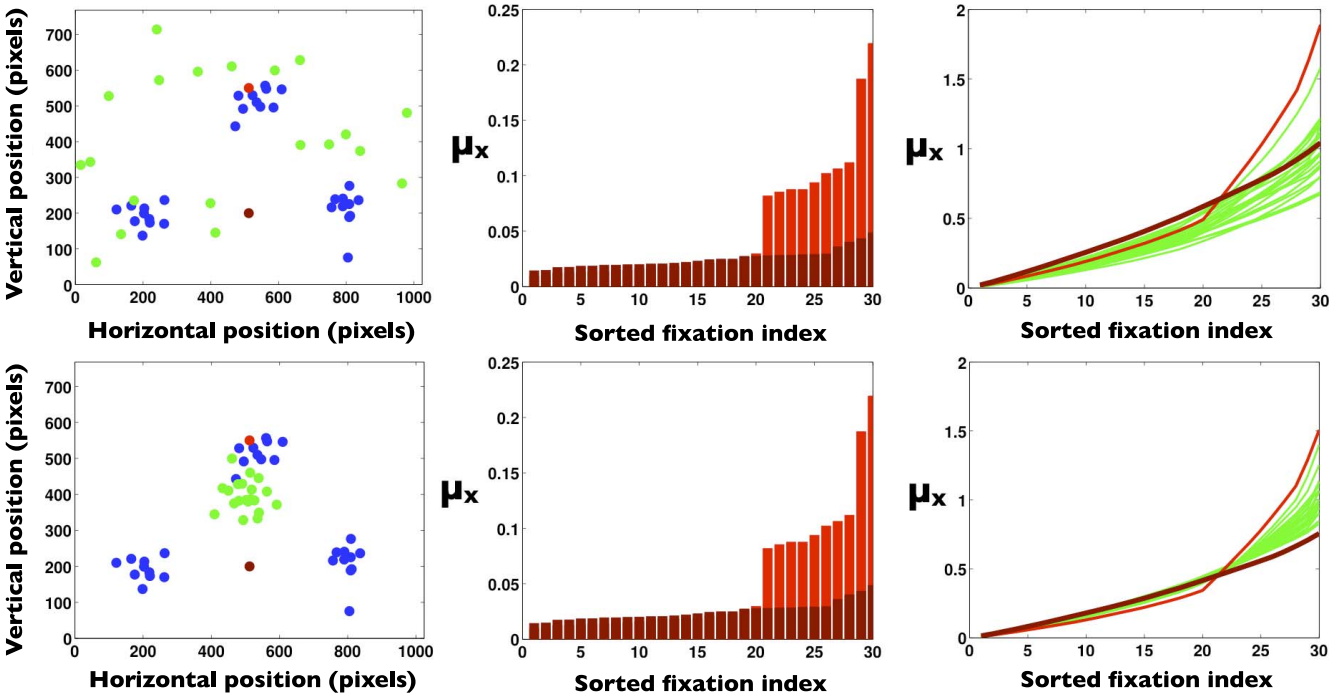

Figure 1. Simple example. The graph on the upper left shows simulated data: two fixations of interest (bright and dark red dots), reference fixations (blue dots), and random fixations (green dots). The graph on the upper middle shows the sorted distribution of belongingness μ_x for the two fixations of interest based on the differences in location to all reference fixations (*x*-axis denotes fixation index of the reference set, *y*-axis denotes μ_x). The graph on the upper right shows the cumulative histograms of belongingness μ_x for the two fixations of interest (red colors) and all random fixation locations (green). Priority values are 1 and 0.6 for the bright red and dark red fixations, respectively. Note that in this simple case, fuzzy parameter *m* has already been fixed (value: 3.486; see Figure 2 for more details). The lower graphs denote the case where a bias is present in the random set (green). This has no effect on the distribution of μ_x for the fixations of interest (lower middle graph). For this example, priority values are 1 and 0 for the bright red and dark red fixations, respectively.

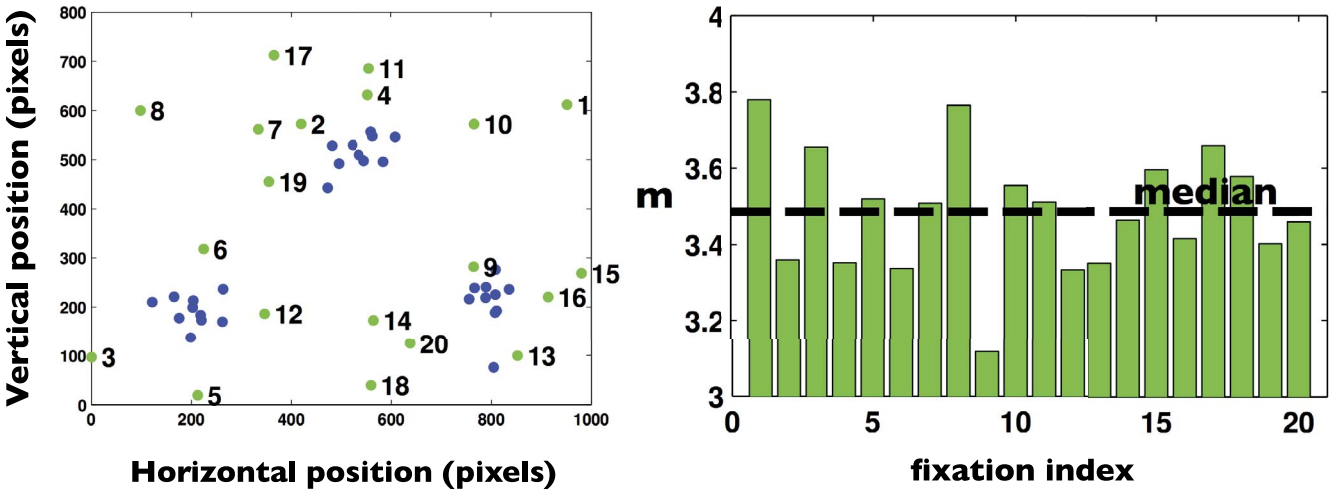

Figure 2. Illustration of fuzzy parameter *m*. The graph on the left shows simulated data as shown in Figure 1: reference fixations (blue dots) and random fixations (green dots). Indices are written next to each random fixation, and correspond to the *x*-axis of the graph on the right. The graph on the right shows values of *m* for each random fixation. The black dashed line indicates the median value of *m*, which is subsequently used to calculate viewing priority.

Functional magnetic resonance imaging

Participants

The study included 18 healthy participants (eight women, 10 men, aged 18–24 years) who reported normal or corrected-to-normal vision. All participants were right-handed as tested using the Edinburgh handedness inventory. Although we initially scanned 19 participants with the fMRI, one participant was excluded due to technical problems in the registration of eye movements during the experiment. All participants gave their informed consent prior to participation. The medical ethics committee of the University Medical Center Groningen approved this study.

Stimuli, task, and design

The stimuli used in the present study came from a previous eye-tracking study by our group (Böhme, Dorr, Martinetz, & Barth, 2006). A high-definition television video camera (JVC JY-HD10, JVC, Yokohama, Japan) was used to record 18 video sequences of a variety of real-world scenes in and around the city of Lübeck (videos are available online: http://www.inb.uni-luebeck.de/tools-demos/gaze). Video resolution was 1024×768 pixels. Each video sequence lasted approximately 20 s, where seven sequences depicted people (e.g., in a pedestrian area, on the beach, playing in a park), four sequences depicted traffic scenes (e.g., populated streets and roundabouts), four sequences depicted animals, and three sequences depicted miscellaneous scenes, such as a ship passing by at a distance. The same video sequences were used as stimuli in the fMRI experiment described in this paper. To prevent abrupt onset effects, video sequences faded in and faded out of a gray background (Tukey window, $\alpha = 0.5$ [i.e., the total fading transition lasted 4.5 s]).

Stimulus presentation

Video sequences were presented using custom software written using the Psychtoolbox (Cornelissen, Peters, & Palmer, 2002; Pelli, 1997) in Matlab (Version 7.8, MathWorks, Natick, MA) and displayed using a Barco LCD Projector G300 (Barco, Kortrijk, Belgium) on a translucent display at a resolution of 1024×768 pixels. The dimensions of the translucent display were 44×34 cm. This subtends a visual angle of $32° \times 25.5°$ for the entire screen. An Apple MacBook Pro (Apple, Cupertino, CA) was used to drive the stimulus display. Via a 45° tilted mirror, placed on top of the head coil, the subject was able to see the entire presentation display. The distance from the eyes to the screen (via the mirror) was 75 cm.

Eye-tracking

Gaze behavior was recorded using a 50-Hz magnetic resonance (MR) compatible eye tracker (SMI, Teltow, Germany), connected via fiber optics to a dedicated PC (Pentium 600 Mhz, 256 MB RAM) running IView X software (Version 1.0, SMI, Teltow, Germany). Communication between this PC and the stimulus PC took place using ethernet (user datagram protocol) and using various routines of the Eyelinktoolbox (Cornelissen et al., 2002). A second mirror relayed the image of the eye to the infrared camera of the eye tracker mounted at the foot of the scanner bed. Manual calibration (nine-point calibration technique) and validation of the eye position signal took place before the start of each session.

The fMRI experiment consisted of two sessions, with different instructions given in each session. During the first session (free viewing), participants were instructed to observe the video sequences as if they were watching television. This instruction was to prevent them from making excessive eye movements. During the second session (fixating), participants were asked to fixate on a 1° marker at the center of the screen. The marker was present during the entire fixation session. During each session, a series of 18 natural video sequences was shown twice. To minimize contextual effects across subjects, the order of presentation during a series was fixed. A 30-s blank grey screen (average luminance: 3262 cd/m^2) with a central fixation cross preceded the first series of video sequences. This was shown again between the two series and after the second series. Nine participants first performed the fixation session.

fMRI data acquisition

Scanning was performed using a 3.0 T MRI Scanner (Philips, Best, the Netherlands) with an eight-channel SENSE head coil. Functional recordings (axial slices recorded in a descending manner) were made using the following parameter settings: flip angle: 70°; echo time: 28 ms; repetition time: 1800 ms, field of view: $224 \times 156 \times 224$ mm; 35 slices were acquired in descending order (slice thickness of 4 mm, in-plane resolution of 3.5 mm). Two sessions were recorded, each consisting of 482 dynamic volumes (approximate duration 14.5 min). Each volume contained 35 slices recorded in a descending manner with an in-plane resolution of 3.5 mm and slice thickness of 4.0 mm. Between the two

sessions, a high-resolution anatomical scan was recorded. The anatomical T1 volume was made with an in-plane resolution of 1×1 mm and contained 160 slices, recorded transversal. The field of view was $232 \times 170 \times 256$ mm.

Eye-tracking data analysis

Fixations were calculated based on the recorded gaze behavior using IView X software (SMI, Teltow, Germany). A fixation duration threshold of 80 ms was used. Fixations having the same position and separated by a short blink (indicating noise) were concatenated. Fixations were exported for each participant using the IDF (Iview X DataFile) Event detector (IDFconverter) with default settings (Version 2.0.1.13, SMI, Teltow, Germany) with a minimal fixation period of 80 ms and a maximal dispersion of 100 pixels. All further eye movement analyses were performed in Matlab (Version 7.8, Mathworks, Natick, MA) using in-house written software.

In the remainder of the analysis, only fixations made during the presentation of the video sequences were taken into account; those recorded during resting periods were excluded. Additionally, we excluded fixations falling outside the display boundaries (approximately 0.5%).

Viewing priority measures

For each participant, viewing priority was calculated for each fixation made during the free-viewing session. The sequence of viewing priorities associated with all fixations made during a session by one participant is hereafter referred to as the viewing priority time series.

Baseline viewing priority time series

A baseline viewing priority time series was calculated for each subject based on eye movement data for a random subject other than the selected subject recorded during a different movie than the movie at hand. This baseline viewing priority will act as a null distribution capturing any systematic/structured effects (e.g., capturing gaze behavior related to the onset of a stimulus). In contrast to choosing completely random spatial data, we maintain the structure of consecutive viewing in similar settings, but for a different context. The expected value of the resulting baseline viewing priority time series will be 0.5, if fixations are made uniform across the screen (see Appendix A). The real viewing priority time series are tested against these baseline time series.

Group-average viewing priority time series

A group-average viewing priority time series was calculated to investigate the idiosyncrasy of our measure. To calculate a group-average viewing priority and variance of viewing priority across participants (red lines in Figure 4), individual viewing priority time series were linearly interpolated to obtain viewing priority values at each 100-ms time point. This interpolation step was necessary because fixation onsets and durations vary across participants. To explore the effect of presenting the same movie for a second time on viewing behavior, we compared the variances in the viewing priorities obtained for the first and second presentation of the series of movies (during free viewing); a Kolmogorov-Smirnov test was performed.

Correlations between viewing priority and eye movement parameters

To test for associations between viewing priority and eye movement parameters, correlations were calculated between viewing priority and (a) fixation duration, (b) amplitude of the previous saccade, (c) phase angle of the previous saccade, (d) amplitude of the subsequent saccade, (e) phase angle of the subsequent saccade, (f) distance vector (i.e., the distance of the fixation to the center of the screen), and (g) phase angle (i.e., the angle of the distance vector).

Correlations between viewing priority and a number of image-based saliency measures

We also assessed correlations between viewing priority and a number of image-based—or bottom-up—measures of saliency, surprise, and motion. Traditionally, saliency measures are represented in maps having the dimension of the image or movie. For our present method, a saliency value was computed for each fixation by placing it on a downsampled saliency map. A summary statistic for a small region centered on the position of the fixation and at the time of occurrence was computed using the following algorithms: (a) standard Itti Koch saliency measure (Itti & Koch, 2001; Itti et al., 1998); (b) a Bayesian measure of saliency and surprise, using an implementation of the Fast Saliency Using Natural statistics algorithm with default parameters (Butko, Zhang, Cotrell, & Movellan, 2008); (c) an analytical saliency measure (Vig, Dorr, & Barth, 2011), which is determined using a classification algorithm that learns visual representations on multiple scales at the location of fixations (geometrical invariants of the structure tensor);[1] and (d) a velocity or optic flow measure based on the Minors of the Structure Tensor algorithm (Barth, 2000).[2]

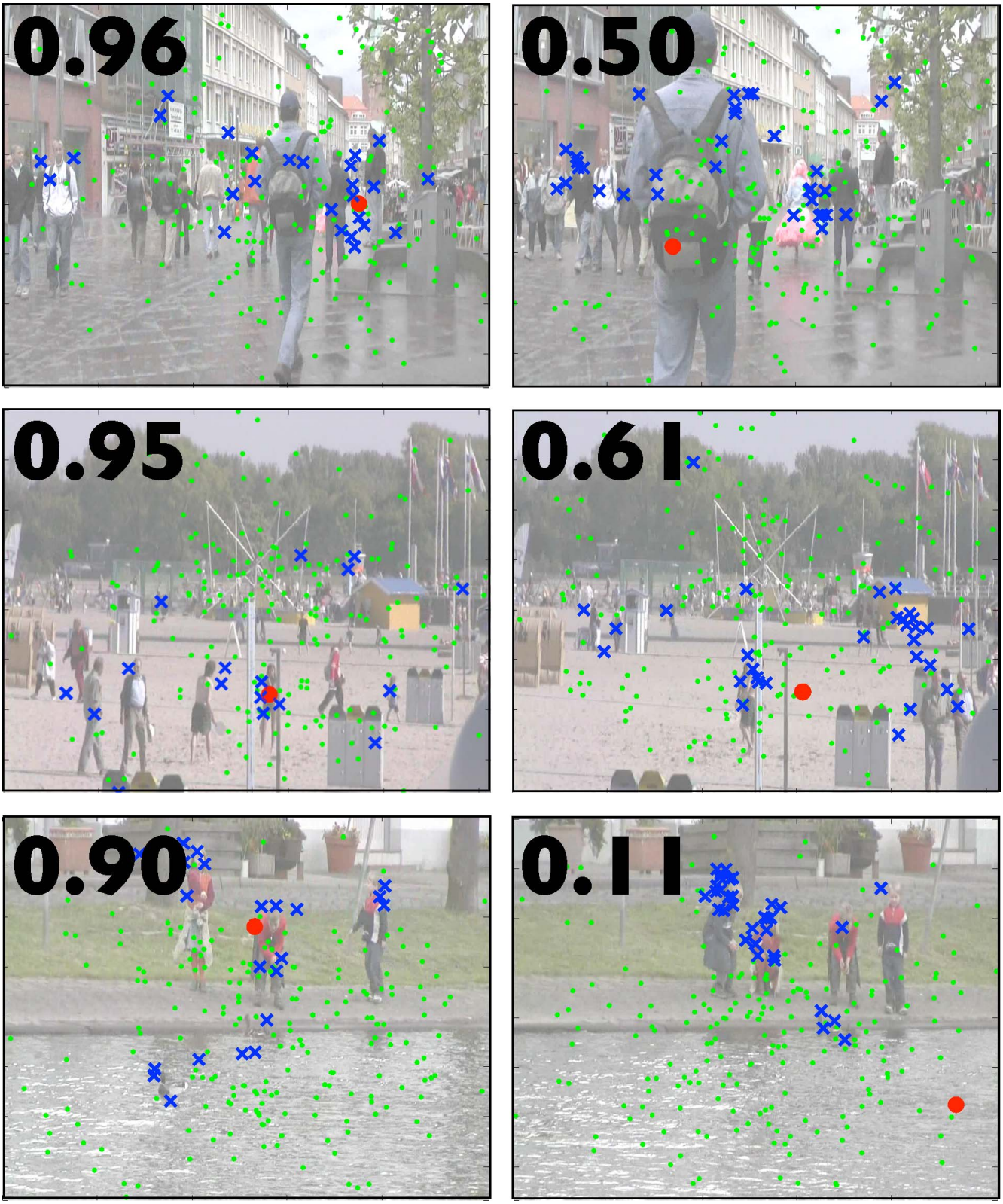

Figure 3. A graphical depiction of viewing priority. Movie stills overlaid with examples of single fixations of an individual participant (indicated by red dots) that were assigned high (left column) to low (right column) viewing priority values. The priority values are displayed in the upper left corner for each movie still. Blue crosses: fixations made by all other participants while viewing the same movie in the time window defined by the participant's fixation (reference set); green dots: fixations made within the same time window by all other participants while viewing all other movies (random set).

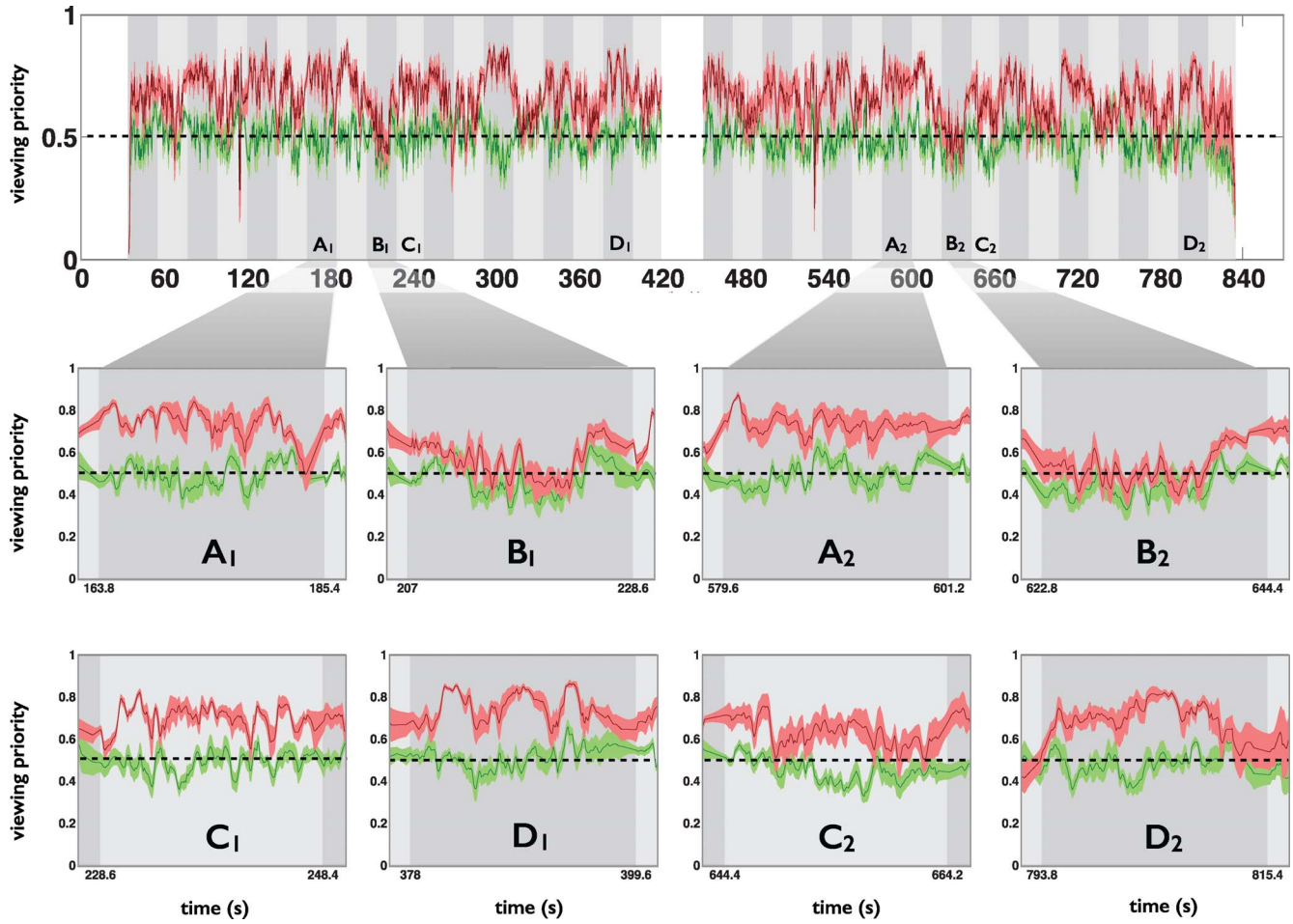

Figure 4. Viewing priority time series (red) and group-average random viewing priority time series (green) across 18 participants. Horizontal axes denote time in seconds. Panels A_1–D_1 detail viewing priority during four selected movies presented during the first series of presentations. Panels A_2–D_2 show viewing priority for the same movies during their second presentation in the same free viewing session. Background: Gray colorings (light and dark gray) denote the presentation and change of movies; white coloring denotes the presentation of a central fixation cross on a uniform gray background (0–20 s, 420–440 s, 840–860 s).

fMRI analysis

All fMRI analyses were performed using SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK) in Matlab 7.4 (Mathworks, Natick, MA). Preprocessing consisted of realignment to correct for subject movement, coregistration to align all functional data to the subject's anatomical volume, normalization to convert all images to Montreal Neurological Institute (MNI) space and spatial smoothing with a Gaussian kernel of 8 mm (full width at half maximum). No slice timing correction was applied.

Analysis of fMRI data based on individually determined viewing priority time series

In a previous study we showed that fixations are meaningful events in fMRI analysis, despite their nonrandom statistical properties, thus enabling Fixation Based Event-Related (FIBER) fMRI analysis (Marsman, Renken, Velichkovsky, Hooymans, & Cornelissen, 2012). We used the same approach in the present study. In this method, fixations are used as rapid events over time to construct a design matrix, which is consequently used in a general linear model. An additional measure can be added per fixation and used as a covariate (i.e., a parametric modulator). A (voxelwise) general linear test for this covariate results in a brain map containing the effect measure on top of the average response (the main effect) for fixations. This map of t statistics can be thresholded and interpreted as the modulatory effect of the covariate. Event-related analyses were conducted using onsets and durations of fixations recorded during the presentation of movies. Two linear parametric modulations were added in the following order: distance and viewing priority; viewing priority correlated with distance. By including the modulations in this order, brain activity

related to viewing priority modulation (residual or otherwise) can be uniquely assigned (van Rootselaar et al., 2008). We used the priority measure to model the activity prior to the eye movement (see also the macaque research study; Mirpour & Bisley, 2013).

To allow for variation in the width and onset of the estimated hemodynamic responses, an informed basis set was used (Henson, Rugg, & Friston, 2001). This set comprised the canonical hemodynamic response function and its temporal and dispersed derivatives. At the first level of the fMRI analysis, contrasts were calculated for (a) the main effect of a fixation event, (b) the viewing priority parameter, and (c) the distance parameter, all using the amplitude component only (i.e., the first basis function of the informed hemodynamic response function model). At the second level of the fMRI analysis, a one-sample t test of each of these contrasts was performed ($p < 0.05$ Family Wise Error [FWE] cluster corrected, with an initial threshold of $p = 0.001$). To visualize the results, functional maps were overlaid on a three-dimensional brain in Caret using average fiducial mapping with the metric-enclosing-voxel algorithm (Van Essen et al., 2001). The corresponding z-scores depicted in the fMRI activation maps were calculated using 3dcalc implemented in afni (http:// afni.nimh.nih.gov/afni/).

Analysis of fMRI data recorded during central fixation based on individually determined viewing priority time series

To verify the influence of covert attention and to assess the relevance of measuring overt eye movements, we also conducted an analysis of the fMRI data recorded while participants were fixating on the fixation marker at the center of the screen. In this analysis, we applied the same statistical parametric mapping (SPM) model to the functional imaging data for central fixation in the first level analyses (with events and priority time series for each subject used in the free viewing condition). Other aspects were identical as described in the section "Analysis based on individually determined viewing priority."

Analysis based on group-averaged viewing priority time series

To verify the relevance of using individually determined viewing priority time series—rather than group averages, the first level of the fMRI analysis was performed using the average viewing priority time series across participants (see above). In this analysis, zero-duration events were set every 100 ms and assigned the corresponding average viewing priority value. Distance was not included in the SPM model. Other aspects were similar to those described in the section "Analysis based on individually determined viewing priority."

Viewing priority versus analytical saliency

Since other measures of saliency showed negligible correlations with priority (See Results section "Correlations between Viewing Priority, Saliency, Surprise and Motion") only analytical saliency was taken into further inspection. To explore the differences between viewing priority and analytical saliency even further, we compared fMRI models of either regressor. Similar models to viewing priority models were estimated using analytical saliency as a regressor (informed basis set). In these models, eccentricity was also added as a regressor before analytical saliency.

Four bilateral regions of interest were created based on the main effect of fixation using Marsbar ($p < 0.05$, FWE corrected; Brett, Anton, Valabregue, & Poline, 2002). From this contrast, peak voxels within the remaining clusters were selected as centers for our regions of interest (ROIs). Spheroid ROIs were defined with a radius of 3 mm. We calculated the fractions of variance explained by the fMRI models (combining amplitude, time derivative, and dispersed information) with respect to the total explained variance of the model for both viewing priority and analytical saliency for all voxels within these ROIs for each subject using in-house written software. To test for significance of the difference in these explained variance fractions between viewing priority and saliency, we employed permutation testing on this difference comprising 5,000 permutations; an effect was considered significant at $p < 0.05$, FWE corrected for the search volumes comprising all ROIs).

Results and discussion

Evaluation of the priority measure

We propose a new method to quantify viewing priority in observers that freely observed dynamic natural scenes (i.e., movies). The method is based on fuzzy clustering, allowing for continuous class membership values inside the algorithm. This also serves as the core mechanism in determining priority in our algorithm, and would therefore not work with standard clustering techniques such as k-means clustering as a starting point. Some examples of viewing priority are shown in Figure 3. Our measure for viewing priority adapts itself to the spatial distribution of the fixations made at about the same time by other participants viewing the same stimulus material (see Figure 3; Appendix A). Biases present in both the reference fixation set (i.e., the blue crosses in Figure 3) and the random set (i.e., the green dots in Figure 3), such as a tendency to view the center of the screen more frequently, are therefore taken into account when

calculating viewing priority. Simulations were run to verify the robustness of our measure for different types and degrees of bias (see Appendix A). Furthermore, we found that viewing priority generally showed very low correlations with classical eye-movement parameters. An exception was distance of the fixation location from the center of the screen, which showed a moderate correlation of −0.55.

In earlier work, Borji, Sihite, and Itti (2013) have attempted to disentangle fixation behavior into a saliency component and a top-down component. Based on voting mechanisms they count the number of fixations on an object to determine its saliency agreement. The main difference between our methods is that their aim was to segregate saliency from top-down context, while we wanted to develop a rich measure encompassing all contributing factors to overt gaze behavior. Moreover, Borji et al. (2013) employed computationally expensive image processing methods that are based on visual analysis of the stimuli themselves. In contrast, our method only needs the positions of fixations and is therefore a relatively lightweight algorithm (once m is determined, that is), which could even be feasible for real-time usage.

Group-average viewing priority and baseline viewing priority time series

Figure 4 shows group-average viewing priority time series across observers (red) and random viewing priority time series (green). A number of observations can be made.

First, viewing priority varied relatively little across participants. This can be seen from the standard error around the red lines in Figure 4. This suggests that despite the presence of individual differences, participants gazed at the movies in a relatively similar manner. Moreover, during some time frames, there was a high viewing priority with hardly any variance, indicating highly coherent gaze behavior across all participants (e.g., in Figure 4, panel D_1 an abrupt movie effect was present in the beginning [after approximately 4 s] which resulted in almost no variation in the viewing priority).

Second, group-averaged viewing priorities during the first and second movie presentations were similar. Nevertheless, the variance was significantly larger for the second series of movie presentations (Kolmogorov–Smirnov test, $p < 0.0001$; see Figure 4). This is not entirely surprising. During the second presentation of a movie, participants may choose to gaze at different—perhaps less salient—events, may anticipate or ignore events, or may only direct covert attention to certain events. Note that these are some of the same reasons why image-based saliency measures have failed to predict a substantial portion of fixations (Birmingham, Bischof, & Kingstone, 2009; Henderson et al., 2007). A similar finding has been reported in Dorr, Martinez, Gegenfurter, and Barth (2010), who have characterized viewing behavior on the exact same stimuli for 54 subjects extensively (Dorr et al., 2010). Third, the group-averaged random viewing priority (green) fluctuated around 0.5, which was in accordance with our expectation for a random, uniform distribution of fixations (see "Baseline viewing priority time series" in the Methods section).

Analysis of eye movement parameters

We analyzed the fixation duration across participants across movies. No systematic trends in fixation duration were present over the time course of the experiment. Furthermore, the pace of eye movements was similar across all subjects throughout the experiment. However, similar to Böhme et al. (2006), we found a strong center bias present in the eye-tracking data. In Appendix A, we describe why this does not affect our priority measure, as it is adaptive for this bias, which is present in both the random and the reference set.

Correlations between viewing priority, saliency, surprise, and motion

A high viewing priority could be the result of a suddenly appearing conspicuous object in the movie that draws the gaze of all participants. In this case, our measure will not be very different from an image-based salience or surprise measure. However, our measure is based exclusively on gaze behavior and does not take image properties into account. Therefore, high values of viewing priority could result equally well from consistent behavior among participants that is driven by contextual information or dynamic events in a movie sequence. For example, the destination of a car could be anticipated by gazing at it, irrespective of the conspicuity of that location.

We found that image-based saliency, surprise, and velocity measures have a relatively low (and nonsignificant) correlation with viewing priority (median values of 0.1, 0.1, 0.12, and 0.08, respectively). These low correlations indicate that our measure of viewing priority does not simply reflect bottom-up saliency; it supports the assumption that saliency is but one aspect that determines participants' gaze and that it is only one of the components that determine viewing priority.

We have now characterized the behavior of viewing priority and have compared it with eye movement

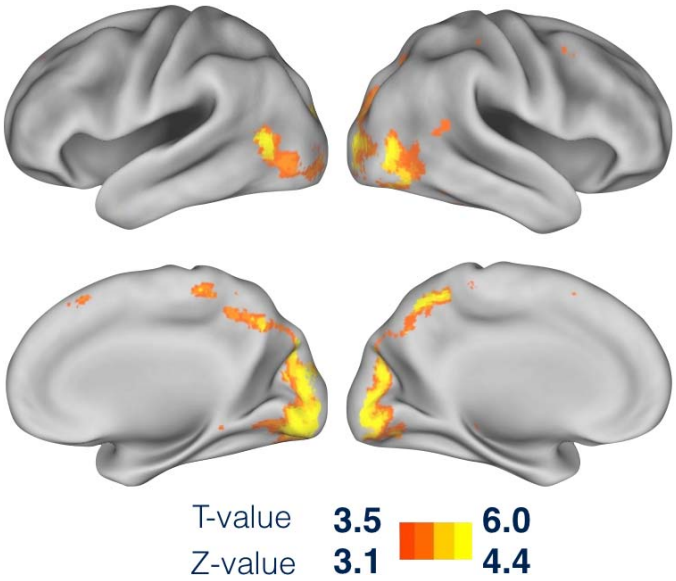

Figure 5. Maps of the brain activation in visual and parietal cortical regions associated with the occurrence of fixations (main effect of fixation event) Average results for 18 participants; maps thresholded at $p = 0.001$, uncorrected. This activation pattern shows that occipital and parietal regions are involved during natural viewing behavior.

properties as well as measures of saliency. The following sections deal with results from the functional imaging experiments and elaborate on the neural correlates of viewing priority, as determined by our measure.

fMRI results for viewing priority

To investigate the neural correlates of viewing priority, we conducted an exploratory parametric analysis of the free viewing sessions using individual fixations as events. Two parametric modulations were added to these events: distance of the fixation to the center of the screen (distance) and viewing priority. We investigated three effects: the main effect of fixation, the modulatory effect of distance, and the modulatory effect of viewing priority (the latter effect after correcting for the modulatory influence of distance). Moreover, we performed a number of control analyses. The priority measures derived from the free viewing condition were also applied to the session in which participants fixated the center of the screen while viewing the same movies. This can indicate to what extent the activations depend on the actual execution of eye-movements. Furthermore, we conducted parametric analyses, one based on the group-averaged viewing priority time series and one based on the random viewing priority time series.

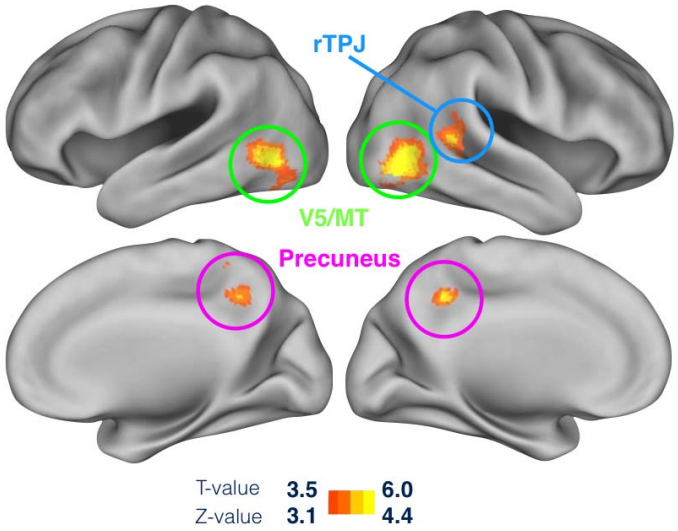

Figure 6. Maps illustrating that brain activation in a select number of visual and parietal cortical regions is associated with individually derived viewing priority during the presentation of natural movies. Average results for 18 participants; maps thresholded at $p < 0.05$ (FWE cluster corrected, initial threshold $p = 0.001$). This signifies that the priority measure can be specifically linked to certain visual processes in the following reported brain regions: bilateral precuneus (highlighted in pink), bilateral V5/MT (highlighted in green) and rTPJ (highlighted in blue).

fMRI results for central fixation viewing based on individually determined viewing priority time series

In this analysis (results shown in Figure 8) we found brain regions that correlate with our measure of viewing priority that strikingly resemble the original activation map for viewing priority (Figure 6). At lower thresholds, also V5/MT, the precuneus, and the TPJ appear (Figure can be found in Appendix B). This suggests that our priority measure not only accounts for brain activity that is tied to the execution of eye movements, but also accounts for attentionally modulated activity.

fMRI results for group averaged priority time series

Figure 9 shows the brain activity when a group-average viewing priority signal is used as a regressor instead of the individual one. In comparison with Figures 6 and 7, we see that most activity for the group-averaged viewing priority is not present, although V5/MT shows a very weak activation. Therefore, priority can only be considered to be meaningful at the level of the individual.

Viewing priority compared with analytical saliency

First, we investigated which regions were active during fixation events (shown in Figure 5). We found

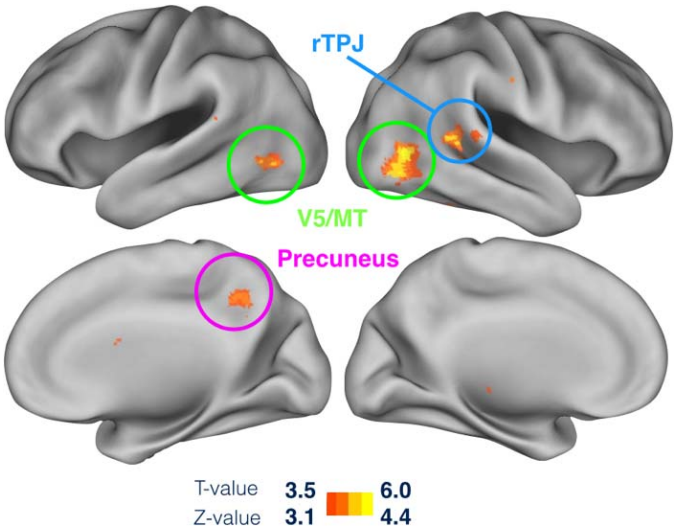

Figure 7. Maps for individually derived viewing priority during the presentation of natural movies once velocity has been accounted for. Average results for 18 participants; maps thresholded at $p < 0.05$ (FWE cluster corrected, initial threshold $p = 0.001$). Right precuneus (highlighted in pink), bilateral V5/MT (highlighted in green), and rTPJ (highlighted in blue) remain specifically active when the velocity component is taken out of the priority measure.

that early visual regions and regions along the dorsal and ventral streams were active when a fixation is made. No brain regions showed significant activation that correlated with distance. So, even though distance

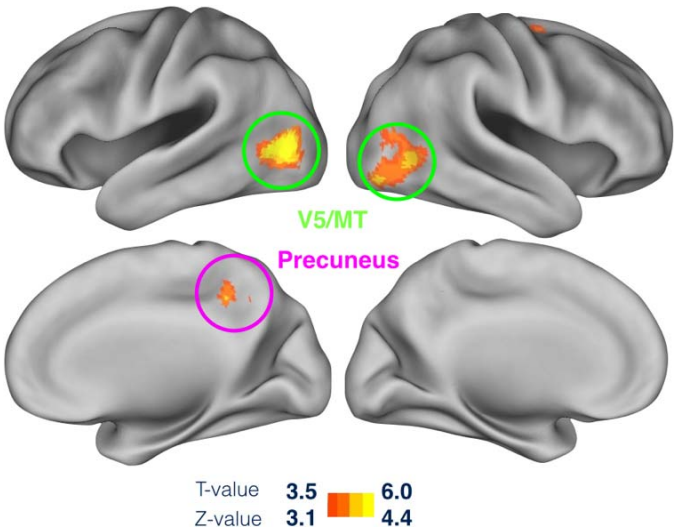

Figure 8. Maps illustrating neural correlates of viewing priority determined from functional data acquired while subjects were fixating centrally during the viewing of movies. The viewing priority time series used in the analysis came from conditions in which observers could freely view the same movies. Average results for 18 participants; maps thresholded at $p < 0.05$ (FWE cluster corrected, initial threshold of $p = 0.001$).

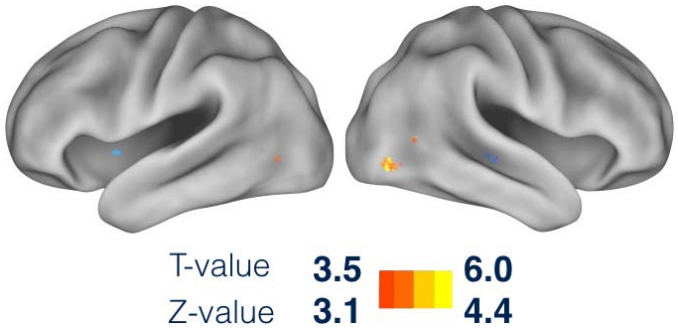

Figure 9. Maps illustrating brain activation associated with group-averaged viewing priority. Average results for 18 participants; maps thresholded at $p = 0.001$, uncorrected. Note that in the medial views of these maps no activations were present.

of the fixation location to the center of the screen correlated with viewing priority, it is not associated with brain activity.

Both Figure 6 and Table 1 show cortical regions whose response correlated with viewing priority. Because we modeled viewing priority orthogonally to distance, we can justify attributing these brain activations to the modulation of viewing priority. Figure 8 shows that priority signals also correlated—albeit less strongly—with data obtained when participants held central fixation. Figure 9 shows that group-averaged viewing priority explains only a minute fraction of the brain activity found with individually determined viewing priority. Despite their relatively higher variance compared to the average viewing priority time series, the use of individual priority time series in the fMRI analysis turned out to be crucial to establish the results. Note that a control analysis using random priority signals revealed no significant correlations in the brain at all. Therefore, no figure was included for these results.

The regions shown in Figure 6 are V5/MT, precuneus, and TPJ. Coordinates of cluster centers of V5/MT coordinates (Table 1) correspond with the

Region	*x* (mm)	*y* (mm)	*z* (mm)	*t* value	*Z* value
V5/MT (left)	−48	−70	0	6.79	4.59
V5/MT (right)	48	−66	0	7.49	4.84
Temporal parietal junction (right)	54	−44	8	5.38	4.01
Precuneus (left)	−10	−50	44	5.59	4.10
Precuneus (right)	−8	−52	46	5.13	3.89

Table 1. Locations of brain regions that correlate with viewing priority. *Notes*: MNI Coordinates refer to the voxels of maximum activation within significant clusters ($p < 0.05$, FWE corrected at cluster level, initial threshold $p < 0.001$).

coordinates reported by Dumoulin et al. (2000). Furthermore, TPJ coordinates (Table 1) overlap with previously reported centers for the TPJ in visuospatial attention paradigms (Downar, Crawley, Mikulis, & Davis, 2000; Natale, Marzi, & Macaluso, 2009; van Beilen et al., 2011). Finally, we found activity in the bilateral precuneus (Table 1).

V5/MT

We found strong correlations between viewing priority and activity in bilateral V5/MT (Figure 6). V5/MT activation is commonly associated with motion processing (Tootell et al., 1995). Previous research has suggested that V5 has a central role in the processing of change during natural stimulation (Bartels & Zeki, 2004). To evaluate whether the contribution of V5/MT can (solely) be attributed to motion processing, we discuss the outcomes of a number of our analyses. First, the correlations between image-based motion and surprise measures on the one hand and viewing priority on the other are rather low. Second, in the fMRI analysis based on the group-averaged priority, we find only a weak V5/MT response. In addition, we investigated post hoc the brain activation maps for priority when velocity was accounted for (Figure 7). Here, we found the same pattern of activation with the main difference that all regions show weaker activation. From these results, we conclude that velocity (i.e., the motion component in a movie) plays only a relatively minor role in determining viewing priority. Third, we regressed the individual viewing priority time series acquired during free viewing also to an fMRI data set acquired in a different experiment (Figure 8). In this experiment, subjects had to fixate in the center of the screen. Here, we found a similar pattern of brain activity only with weaker activations. Moreover, no activity in right TPJ showed up. These findings indicate that (a) the generation of eye movements is a component of the priority signal, and that (b) priority does also correlate with covert attention (when no eye movements are made). Overall, these results imply that the responses in the precuneus, right TPJ (rTPJ), and V5/MT not only represent motion information per se, but may integrally represent the priority this information has for guiding gaze behavior of an individual.

Precuneus

The second set of regions whose activity correlated with viewing priority is the bilateral precuneus. The precuneus lies along the dorsal stream of the visual system (Milner & Goodale, 1993) and has been associated with visuospatial attention and orienting (Cavanna & Trimble, 2006). Interpreted in the light of the premotor theory of attention (Rizolatti, Riggio, Dascola, & Umilta, 1987), previous fMRI studies suggested that the precuneus, together with the neighboring posterior parietal cortex, is involved in shifting attention without accompanying eye movements (Beauchamp, Petit, Ellmore, Ingeholm, & Haxby, 2001; Gitelman et al., 1999; Rizolatti et al., 1987). This is also in line with our current finding, as we find that the precuneus shows a strong response for individual viewing priority (Figure 6), which even remains present in the activation maps obtained from data in which participants held central fixation (Figure 8).

Temporal parietal junction

We also found strong correlations between viewing priority and activity in the rTPJ. This TPJ activation in human observers corroborates the finding in macaques that the lateral intraparietal area contains a priority map (Bisley & Goldberg, 2010). The TPJ has previously been identified as a circuit breaker for ongoing cognitive events. Corbetta et al. (2000) suggested that the rTPJ may underlie the process of spatially redirecting the focus of attention towards the location of unattended stimuli. Other studies also suggest that the rTPJ redirects attention to behaviorally relevant sensory stimuli that are outside the present focus of processing (Cavanna & Trimble, 2006; Corbetta & Shulman, 2002). Correlations between rTPJ activity and novelty have been reported in multimodal sensory processing by Downar et al. (2000, 2002). Additionally, Shulman, Astafiev, McAvoy, d'Avossa, and Corbetta (2007) have reported deactivation of the rTPJ during visual search. These authors suggested that the rTPJ filters out irrelevant input, which is in line with our present findings that the rTPJ shows a different activation for viewing priority compared to saliency (Figure 8). Furthermore, the rTPJ activity revealed in the activation maps for analytical saliency is much lower compared to that for the priority measure (and did not survive FWE correction). This difference in activation corroborates previous results reported in the literature and provides further evidence that the rTPJ is involved only when eye movements are made. For this reason, it should be considered part of the oculomotor component of the attentional system.

Our results suggest a unified interpretation for these findings and hypotheses in terms of priority: The presumed filter role of the TPJ is the result of prioritizing various information and selectively responding to high priority events only, while the

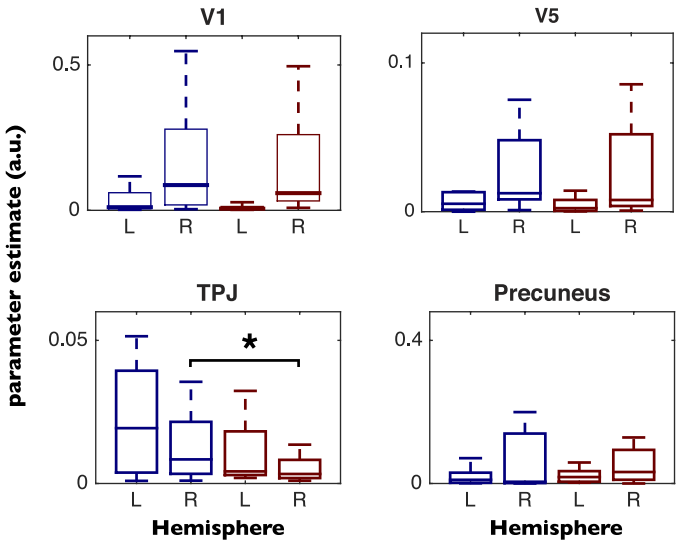

Figure 10. Fractions of explained variance of viewing priority and analytical saliency with respect to the total explained variance of the model. Four regions selected (bilaterally) based on the main effect for fixations. Red bars denote analytical saliency; blue bars denote viewing priority. A significant difference between analytical saliency and viewing priority was found for the rTPJ only ($p < 0.05$ FWE corrected), even though a similar trend can be observed in the left TPJ.

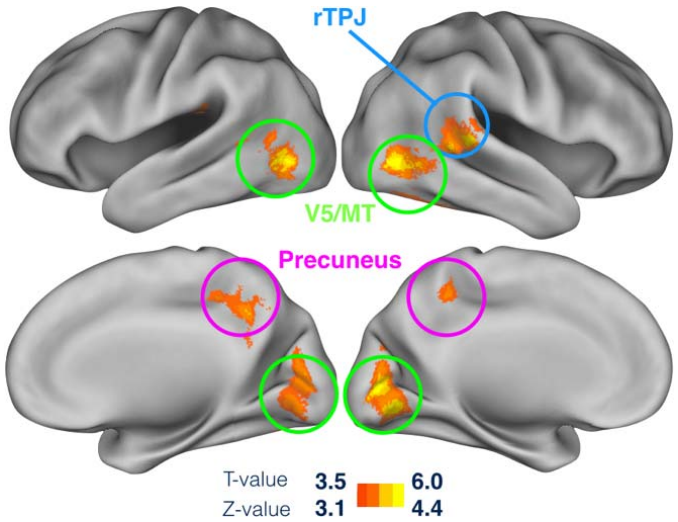

Figure 11. Maps illustrating brain activation associated with analytical saliency in free viewing. Average results for 18 participants; maps thresholded at $p = 0.001$, uncorrected. Bilateral precuneus (highlighted in pink), bilateral V5/MT (highlighted in green), and rTPJ (highlighted in blue) show a moderate association for analytical saliency.

novelty-associated responses of the TPJ are the result of novel events receiving a high priority.

fMRI results for analytical saliency

Overall, our present results indicate that activity in specific cortical regions correlates with our measure of viewing priority. This is most clear for data recorded while participants were overtly attending the movies (Figure 6; Table 1). This supports our assumption that priority can be inferred from eye movements. In addition, we believe that the regions activated play a role in the processing of visual input in assessing priority, in contrast to the steering of the eyes. Although this interplay must be inherently existent throughout the brain, we have restricted our interpretations in interpreting these as processing regions.

A comparison between viewing priority and analytical saliency in terms of explained variance of the parametric models reveal a substantial resemblance for V1, V5/MT, and precuneus (Figure 10). This also appears when comparing Figures 6 and 11. Although the same regions show activity in the maps for both analytical saliency and viewing priority, the effects are much lower for analytical saliency (Figure 11; Table 2), in Figure 11 the maps have not been corrected for multiple comparisons. After such correction, only the bilateral V5/MT remained significantly active. Moreover, the right TPJ shows a significant difference in explained variance (Figure 10). This indicates that priority can be considered distinct from analytical saliency, in support of the theory of Fecteau and Munoz (2006). According to their framework, both salience and relevance affect the same brain regions. However, while each yields a different neural signal the spiking patterns of oculomotor system neurons appear to reflect their combination.

Limitations and future extensions

Regarding the algorithm of viewing priority, one limitation is that reference data needs to be collected and that random data needs to be determined prior to

Region	*x* (mm)	*y* (mm)	*z* (mm)	*t* value	*Z* value
V1 (left)	−16	−96	6	10.66	5.22
V5/MT (right)	66	−34	4	9.01	4.88
Superior colliculi	−12	−22	−6	8.23	4.63
V5/MT (right)	46	−54	−30	7.99	4.55

Table 2. Locations of cluster centers that correlate with analytical saliency. *Notes*: MNI Coordinates in mm refer to the location of voxels having the maximum of activation within significant clusters ($p < 0.05$, FWE corrected at cluster level, initial threshold $p < 0.001$). Please note that activity in V5/MT is listed twice since two significant clusters are found in that area.

the calculation of priority. Moreover, it should be noted that our method—in its current form—does not predict gaze behavior (yet it could provide predictive power for a future observer). It provides a current statistic on the fixation at hand that is based on human behavior rather than that it guesses what the visual cortex finds important. As such, our method provides a new and unique approach.

In our present experiment, only passive viewing without a task was studied and, consequently, our fMRI analyses are primarily exploratory. As such, the neural correlates of viewing priority that we found need to be corroborated by future—more elaborate—experiments. For example, the addition of various tasks might help tease apart the contributions of the priority and saliency components to viewing priority and reveal the roles of specific brain areas in priority computations.

Our goal for the present method was to track and use priority variations over time. Technically, it is also possible to use our method to generate spatial priority maps similar to saliency maps, which might perhaps find employ in predicting viewing behavior in new observers. In addition, spatiotemporal varying priority maps could potentially be used in a fashion somewhat akin to retinotopic mapping to localize cortical maps representing priority.

Conclusions

We presented a new measure to determine viewing priority based on eye movements. It shows only limited correlation with various saliency, surprise, and motion measures, indicating it is assessing a distinct property of visual behavior. Moreover, we showed that it can be used to determine regions in the brain whose responses correlate with viewing priority. Neural correlates of viewing priority were observed in the precuneus, the rTPJ, and V5/MT. Our measure of viewing priority correlated most strongly with human brain activity when actual eye movements were made and when the individual participants' priority signals were taken into account. We conclude that our new analysis method allows for empirically establishing the priority of events in near-natural vision paradigms.

Keywords: eye movements, priority, attention, fMRI, saliency

Acknowledgments

The authors would like to thank Professor Dr. J. B. T. M. Roerdink for his very useful comments on an earlier draft of the manuscript, Mr. L. Nanetti for his helpful discussions, A. Sibeijn-Kuiper, and J. Streurman-Werdekker for their help during data acquisition, and the Center for High Performance Computing and Visualization of the University of Groningen for providing computing time on the cluster. FWC, EB and RR designed the experiments; JBM and RR developed the viewing priority algorithm; JBM performed the research and analyzed the fMRI data; MD and EV analyzed other saliency measures; JBM, RR and FWC prepared the manuscript. All authors have contributed to drafts of the manuscript. This research was supported by European Commission Grant 033816 (GazeCom) to EB and FWC and 043261 (Percept) to FWC. The work reflects only the authors' views.

Commercial relationships: None.
Corresponding author: Jan-Bernard C. Marsman.
Email: j.b.c.marsman@umcg.nl.
Address: NeuroImaging Center, Laboratory for Experimental Ophthalmology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

Footnotes

[1] Specific settings: third spatial scale pyramid level, and third temporal pyramid level 4 cyc/°, 7.5 fps.

[2] Specific settings: Input smoothing with a spatial five-tap and a temporal seven-tap binomial, an omega of seven-tap binomials in space and time, a final (spatial) blur with a 21-tap binomial, $T_m = 0.001$, $T_\Theta = 45°$, and the two relative and absolute thresholds for M_{ii} set to 0.001 and 10^{-7}, respectively. Velocity was estimated on each scale of a pyramid with five spatial and three temporal levels and then combined (1 pixel/frame on the second spatial, second temporal scale corresponds to 4 pixels/frame on the highest scale).

References

Aivar, M. P., Hayhoe, M. M., Chizk, C. L., & Mruczek, R. E. (2005). Spatial memory and saccadic targeting in a natural task. *Journal of Vision*, *5*(3):3, 177–193, doi:10.1167/5.3.3. [PubMed] [Article]

Bartels, A., & Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping*, *21*(2), 75–85, doi:10.1002/hbm.10153.

Barth, E. (2000). *The minors of the structure tensor* (pp. 221–228). Berlin, Heidelberg: Springer.

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *NeuroImage*, *14*(2), 310–321, doi:10.1006/nimg.2001.0788.

Bezdek, J. C. (1981). Models for pattern recognition. *Pattern recognition with fuzzy objective function*. New York: Plenum Press. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4757-0450-1_1

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research*, *49*(24), 2992–3000, doi:10.1016/j.visres.2009.09.014.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neurosciences*, *33*, 1–24.

Böhme, M., Dorr, M., Martinetz, T., & Barth, E. (2006). Eye movement predictions on natural videos. *Neurocomputing*, *69*, 1996–2004. Retrieved from http://www.inb.uni-luebeck.de/publications/pdfs/BoDoKrMaBa06.pdf

Borji, A., Sihite, D. N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research*, *91*, 62–77, doi:10.1016/j.visres.2013.07.016.

Brett, M., Anton, J.-L. Valabregue R., & Poline, J.-B. (2002). Region of interest analysis using an SPM toolbox [Abstract]. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2-6, 2002, Sendai, Japan. Available on CD-ROM in NeuroImage, Vol. 16, No. 2.

Bruce, N. D., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: an information theoretic approach. *Journal of Vision*, *9*(3):5, 1–24, doi:10.1167/9.3.5. [PubMed] [Article]

Buswell, T. G. (1935). *How people look at pictures*. Chicago: University of Chicago Press.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain: A Journal of Neurology*, *129*(Pt 3), 564–583, doi:10.1093/brain/awl004.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience*, *3*(3), 292–297, doi:10.1038/73009.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215, doi:10.1038/nrn755.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617.

Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision*, *10*(10):28, 1–17, doi:10.1167/10.10.28. [PubMed] [Article]

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience*, *3*(3), 277–283, doi:10.1038/72991.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2002). A cortical network sensitive to stimulus salience in a neutral behavioral context across multiple sensory modalities. *Journal of Neurophysiology*, *87*(1), 615–620.

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Pike, B. G., & Evans, A. C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cerebral Cortex*, *10*(5), 454–463.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382–390, doi:10.1016/j.tics.2006.06.011.

Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y. H., Meyer, J. R., & Mesulam, M. (1999). A large-scale distributed network for covert spatial attention: further anatomical delineation based on stringent behavioural and cognitive controls. *Brain: A Journal of Neurology*, *122*(Pt 6), 1093–1106.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences*, *9*(4), 188–194, doi:10.1016/j.tics.2005.02.009.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual Saliency does not account for eye movements during visual search in real-world scenes. In *Eye movements: A window on mind and brain*. (pp. 537–562). Oxford, UK: Elsevier.

Henson, R. N. A., Rugg, M. D., & Friston, K. J.

(2001). The choice of basis functions in event-related fMRI. *NeuroImage*, *13*, 127.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203, doi:10.1038/35058500.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(11), 1254–1259, doi:10.1109/34.730558.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

Marsman, J. B., Renken, R., Velichkovsky, B. M., Hooymans, J. M., & Cornelissen, F. W. (2012). Fixation based event-related fMRI analysis: Using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. *Human Brain Mapping*, *33*(2), 307–318, doi:10.1002/hbm.21211.

Milner, A., & Goodale, M. A. (1993). Visual pathways to perception and action. *Progress in Brain Research*, *95*, 317–337.

Mirpour, K., & Bisley, J. W. (2013). Evidence for differential top-down and bottom-up suppression in posterior parietal cortex. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *368*(1628), 20130069, doi:10.1098/rstb.2013.0069.

Natale, E., Marzi, C. A., & Macaluso, E. (2009). fMRI correlates of visuo-spatial reorienting investigated with an attention shifting double-cue paradigm. *Human Brain Mapping*, *30*(8), 2367–2381, doi:10.1002/hbm.20675.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442, doi:10.1163/156856897X00366.

Rizolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*, 31–40.

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision*, *7*(14):16, 1–20, doi:10.1167/7.14.16. [PubMed] [Article]

Shulman, G. L., Astafiev, S. V., McAvoy, M. P., d' Avossa, G., & Corbetta, M. (2007). Right TPJ deactivation during visual search: Functional significance and support for a filter hypothesis. *Cerebral Cortex*, *17*(11), 2625–2633, doi:10.1093/cercor/bhl170.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, *15*(4), 3215–3230.

Van Beilen, M., Renken, R., Groenewold, E. S., & Cornelissen, F. W. (2011). Attentional window set by expected relevance of environmental signals. *PLoS One*, *6*(6), e21262, doi:10.1371/journal.pone.0021262.

Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Informatics Association*, *8*(5), 443–459.

van Rootselaar, A. F., Maurits, N. M., Renken, R., Koelman, J. H., Hoogduin, J. M., Leenders, K. L., & Tijssen, M. A. (2008). Simultaneous EMG-functional MRI recordings can directly relate hyperkinetic movements to brain activity. *Human Brain Mapping*, *29*(12), 1430–1441, doi:10.1002/hbm.20477.

Vig, E., Dorr, M., & Barth, E. (2011). Learned saliency transformations for gaze guidance. In *IS&T/SPIE electronic imaging* (pp. 78650W–78650W). San Francisco, CA: SPIE.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Zhang, L., Tong, M. H., & Cottrell, G. W. (2009). SUNDAy: Saliency using natural statistics for dynamic analysis of scenes. In *Proceedings of the 31st annual cognitive science conference*. Amsterdam, The Netherlands.

Appendix A

Simulations

To investigate the properties of our measure of priority, we conducted two simulations. First, we varied the cluster spread of the reference fixations. Second, we varied the cluster spread of both the reference and random fixations. In the simulations, priority was calculated for a grid of observed fixations located across the screen. Fixations were located 8 pixels apart in both the x and y direction (Figure A).

Two sets of artificial fixations were generated, a random set and a reference set. The reference fixations

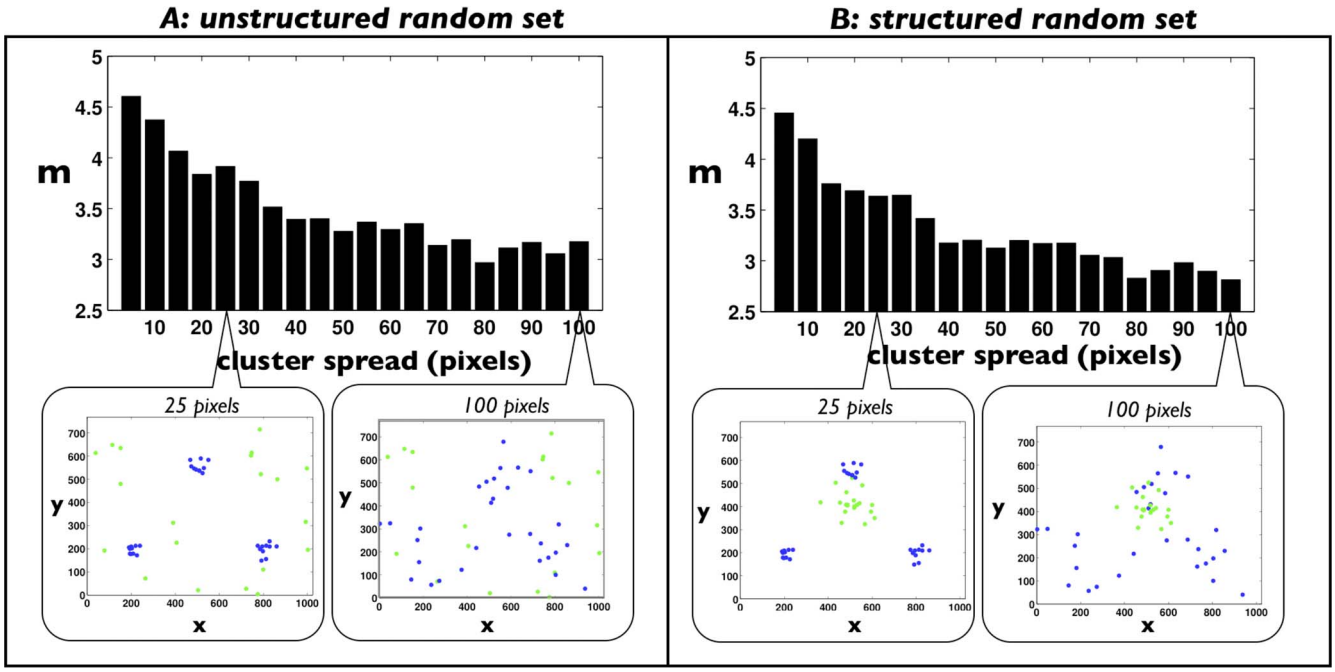

Figure A. Behavior of *m* for different cluster spread. The upper figures show fuzzy parameter *m* related to the cluster spread of the reference set, where the random set is uniformly distributed across the screen (left) and where a strong center bias is present in the random set (right). The lower figures denote two examples with cluster spread of the reference set to either 25 pixels (left) or 100 pixels (right) for both types of random set. Reference fixations are denoted with blue dots and random fixations are indicated with green dots.

were sampled from a random uniform distribution combined with two Gaussian point spread functions at two different locations, of which the full width at half maximum (cluster spread) was varied systematically (5–100 pixels with steps of 5 pixels). For the first simulation, the random set was sampled from a random uniform distribution only. In the second simulation, the random set was created in the same way as the reference set. The number of fixations for both sets was varied systematically (1, 2, 4, 8, 16, 32, 64, 128, 512, and 1,024 fixations).

The results of both simulations show a decreasing trend for $\mu(x)$ as cluster spread of the reference set increases (uppermost bar graphs in Figure 3; only results from the first simulation are depicted). This decreasing trend implies that the priority measure is adaptive for the cluster spread values present in the reference set. More scattering of the reference fixations (larger cluster spread) results in a less skewed weighting of distances to each reference fixation. In other words, the priority measure considers fixations located nearby reference fixations more important when the reference fixations are less scattered. The change of spatial distributions for the random set did not affect our viewing priority.

Based on these simulations, we conclude that the behavior of our measure of priority is adaptive for both the distribution of reference fixations and the distribution of random fixations.

Appendix B

Activation map for viewing priority where subjects were fixating at the center shown with lower threshold

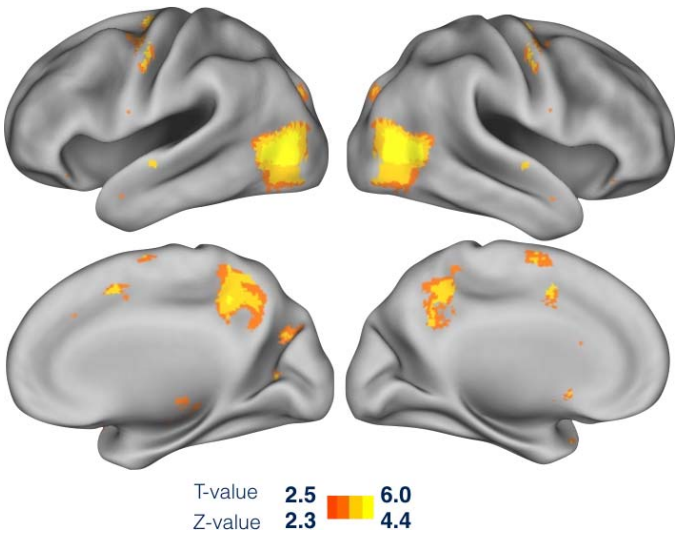

Figure B. Maps illustrating viewing priority assessed on functional data acquired while subjects were fixating centrally during the viewing of movies. Average results for 18 participants; maps thresholded at $p < 0.01$ uncorrected.