



## ORIGINAL ARTICLE

# The Neural Mechanisms of Prediction in Visual Search

Eelke Spaak<sup>1</sup>, Yvonne Fonken<sup>1,2</sup>, Ole Jensen<sup>1</sup> and Floris P. de Lange<sup>2</sup>

<sup>1</sup>Donders Institute for Brain, Cognition, and Behaviour, Centre for Cognitive Neuroimaging, Radboud University, 6525 EN Nijmegen, The Netherlands and <sup>2</sup>Helen Wills Neuroscience Institute, University of California at Berkeley, Berkeley, CA 94720, USA

Address correspondence to Eelke Spaak. Email: eelke.spaa@donders.ru.nl

Eelke Spaak and Yvonne Fonken contributed equally to this work.

## Abstract

The speed of visual search depends on bottom-up stimulus features (e.g., we quickly locate a red item among blue distractors), but it is also facilitated by the presence of top-down perceptual predictions about the item. Here, we identify the nature, source, and neuronal substrate of the predictions that speed up resumed visual search. Human subjects were presented with a visual search array that was repeated up to 4 times, while brain activity was recorded using magnetoencephalography (MEG). Behaviorally, we observed a bimodal reaction time distribution for resumed visual search, indicating that subjects were extraordinarily rapid on a proportion of trials. MEG data demonstrated that these rapid-response trials were associated with a prediction of (1) target location, as reflected by alpha-band (8–12 Hz) lateralization; and (2) target identity, as reflected by beta-band (15–30 Hz) lateralization. Moreover, we show that these predictions are likely generated in a network consisting of medial superior frontal cortex and right temporo-parietal junction. These findings underscore the importance and nature of perceptual hypotheses for efficient visual search.

**Key words:** expectation, MEG, perception, rapid resumption, visual search

## Introduction

Humans have a remarkable ability to quickly find a single item in a cluttered visual world, filled with other, distracting, items. Visual search has been the subject of intense study (Wolfe 1998). Early research in visual search focused mostly on examining how bottom-up stimulus properties determined visual search efficiency (Treisman and Gelade 1980). However, visual search efficiency is also strongly shaped by our prior knowledge of the scene, either derived from scene context (Biederman 1972; Bar 2004) or from past experience with a scene (Chun and Jiang 1998; Stokes et al. 2012). Both of these factors provide a rich source of predictive information about the visual world that can be used to guide and optimize visual selection (Summerfield and de Lange 2014).

Recent evidence for the role of prediction in visual search has come from interrupted serial visual search tasks (Lleras et al. 2005, 2007). In a serial visual search task, subjects are briefly presented with a search array, consisting of a target among

distractors. This search array is repeated after a short delay, until the subject makes a judgment on this target. The crucial behavioral finding is that subjects are sometimes much faster in responding to a search display that they had been presented with before, a phenomenon known as rapid resumption. It has been suggested that trials showing rapid resumption involve a perceptual prediction about the upcoming stimulus that is absent on trials without rapid resumption (Enns and Lleras 2008).

Several things remain unclear from these psychophysical experiments, however. First, what is the nature of the prediction? Do initial glimpses give rise to a spatial prediction (“This location may contain a target”), thus prioritizing a particular part of visual space for subsequent processing (Chun and Jiang 1998; Stokes et al. 2012)? Such a prioritization would likely be reflected in a hemispheric lateralization of posterior alpha-band (8–12 Hz) activity (Worden et al. 2000; Thut et al. 2006). Do subjects form a hypothesis about the target identity, based on initial glimpses (“The target may be an inverted T”)? Given the appropriate target–response mapping, such a hypothesis would be reflected in

lateralized beta-band (15–30 Hz) activity over motor cortex (Pfurtscheller and Lopes da Silva 1999; Donner et al. 2009; de Lange et al. 2013). Second, the neural implementation of the biasing of visual search by prior information is unclear. Which areas may be the sources and beneficiaries of resource allocation by prior information during visual search?

In this study, we tackle these questions by recording brain activity using magnetoencephalography (MEG) while subjects performed a serial visual search task. Foreshadowing our results, we replicate previous behavioral studies (Lleras et al. 2005, 2007) and find that, on a proportion of trials, subjects are much faster in responding to a display that has been presented before, resulting in a bimodal response time distribution. Furthermore, we find that brain activity prior to these fast subsequent responses is markedly different from that prior to slow (“normal”) subsequent responses. Specifically, we find evidence for implicit predictions concerning both spatial location and target identity. Finally, neural activation in medial superior frontal cortex and right temporo-parietal junction (TPJ) appears to be involved in the generation of these sensory predictions that bias subsequent visual search.

## Methods

### Participants

Twenty healthy subjects (14 female; average age 22; range 18–31) participated in the study, which was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the local ethics committee (CMO region Arnhem-Nijmegen). One subject was excluded because of excessive head movement during the MEG recording, leaving 19 subjects in the study. The subjects were paid for participating. All subjects were right-handed individuals with normal or corrected-to-normal vision, and had no history of psychiatric or neurological illnesses (based on self-report).

### Stimuli and Experimental Design

The experimental stimuli used consisted of a search display with 15 distracters (L-shape, rotated at random multiples of 90°) and 1 target (normal or inverted T-shape), which were presented using a PC running the MATLAB-based Psychophysics toolbox software (Brainard 1997). Each trial started with a fixation cross (1000 ms), followed by the search array stimulus (100 ms) and a blank (1300 ms). The stimulus-blank routine was repeated until the subject responded, with a maximum of 4 repeats (Fig. 1A). The task for the subject was to indicate the orientation of the target, giving a left-hand response for an upright T, or a right-hand response for an inverted T. The trial ended with a feedback screen (500 ms, green or red fixation cross for a correct/incorrect response) and a dimmed fixation cross (1000 ms), signifying a moment of rest for the subject. The subject was instructed to respond as fast and as accurate as possible, to blink as little as possible and to keep fixating on the fixation cross in the middle of the screen for the whole duration of the experiment.

The search display consisted of 8 objects in both the left and the right visual field. The stimulus array comprised 10° of visual angle, while each object subtended 1° of visual angle. The exact position of the objects was randomly generated within these spatial constraints, with an equal occurrence of the target in the left and right visual fields. Before the actual experiment, subjects engaged in 3 practice blocks (50 trials each) outside the MEG and 1 practice block (20 trials) within the MEG environment.

Hereafter, the subjects performed 10 blocks (50 trials each) of the task while MEG measurements were acquired.

For all reported MEG analyses, the time axis is such that  $t = 0$  s corresponds to the onset of the display at which the subject identified the target. By consequence,  $t = -1.4$  s corresponds to the onset of the previous display. Trials in which subjects identified the target after the first display were not included in the MEG analyses, nor were trials for which subjects made an incorrect response. When results refer to “fast” or “slow” responses or trials, “fast” always refers to reaction times below the median of that subject, whereas “slow” refers to reaction times above the median.

### Experimental Equipment

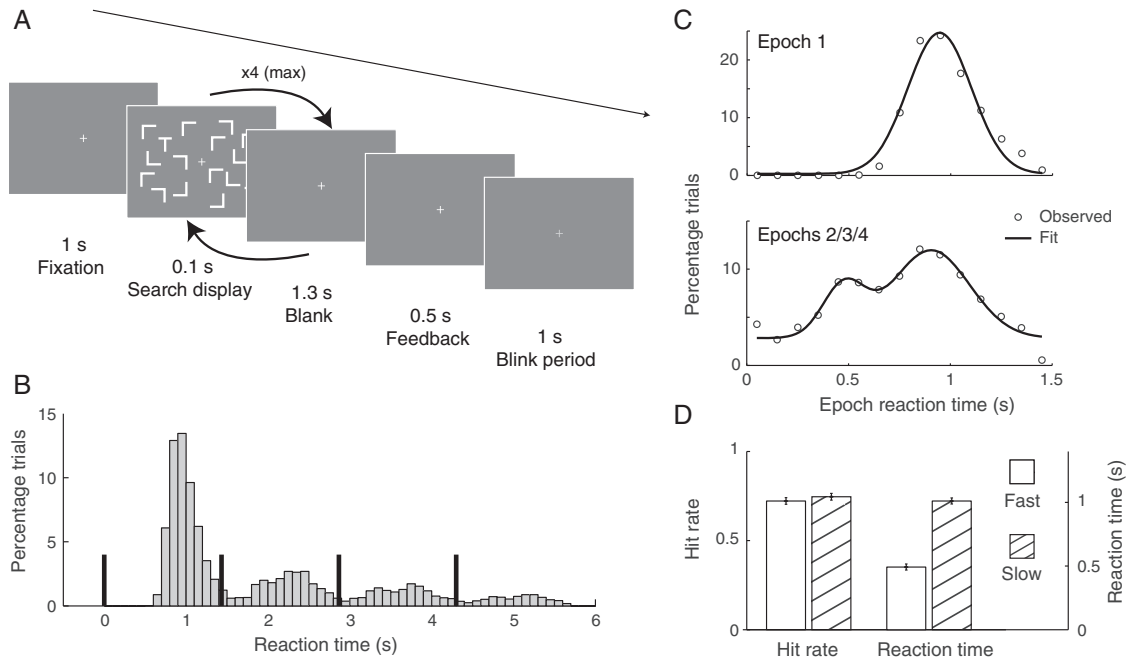
Stimuli were presented by back-projection onto a semi-translucent screen by an EIKI LC-XL100L projector. The projection measured 46 cm in width, and had a resolution of 1024 × 768 pixels. Subjects were seated in a magnetically shielded room, at 80 cm distance from the projection screen. Throughout the experiment, MEG was recorded using a 275-channel axial gradiometer CTF MEG system. An online low-pass filter (cutoff 300 Hz) was applied, and the data were digitized at 1.2 kHz. In addition, we continually recorded subjects' gaze position using an SR Research Eyelink 1000 eye-tracking device, to monitor fixation during the task and to record eye blinks for offline artifact rejection. Also for artifact rejection purposes, the electrocardiogram (ECG) was recorded using three 10 mm Ag–AgCl surface electrodes. After the MEG experimental session, structural magnetic resonance imaging (MRI) images were obtained from all subjects using a 1.5-T Siemens Magnetom Avanto system.

### MEG Data Preprocessing

All MEG data were analyzed using the FieldTrip toolbox (Oostenveld et al. 2011) and custom-written scripts for MATLAB, version R2012a. Artifact rejection was done for each subject individually. First, trials with muscle or MEG jump artifacts were identified and removed from the data using a semi-automatic routine. Then, using independent component analysis (ICA; Bell and Sejnowski 1995; Jung et al. 2000), we removed eye-movement and cardiac-related activity from the MEG signals, by comparing ICA output with the eye tracker and ECG recordings. Finally, the data were inspected visually to remove any remaining artifacts that were not identified by these automated procedures. After artifact rejection, data were resampled offline to 400 Hz (after applying an anti-aliasing filter), to speed up subsequent analyses. After artifact rejection and restricting the analyses to epochs >1 with a correct response,  $160 \pm 17$  (mean ± standard error across subjects) trials were used for the MEG analyses.

### Frequency-Domain Source Analysis

All reported MEG analyses were conducted in source space. To obtain source estimates, we always first conducted a linear frequency domain (Gross et al. 2001). The beamformer algorithm computes a spatial filter from the cross-spectral density (CSD) matrix of the data and a lead field matrix. To obtain the lead fields for each subject, we constructed a realistically shaped single-shell head model based on the individual anatomical MRI (Nolte 2003), after spatially co-registering the MRI to sensor space MEG data by identifying fiducials in the nasion and the 2 ears. Each brain volume was divided into a grid of points spaced 8 mm apart, and warped to the template Montreal Neurological



**Figure 1.** Experimental design and behavioral results. (A) Structure of one trial. After a 1-s fixation period, subjects were presented with a search array for 0.1 s, consisting of 15 L-shaped distractors and 1 T-shaped target. The subject was instructed to respond with a button press as soon as possible when he/she felt confident whether the T was upright or inverted. The search display was presented iteratively, with 1.3 s blanks in between displays, up to 4 times, or until the subject made a response. (B) Grand average reaction time histogram across all epochs. Vertical black bars indicate the onset of each of the 4 displays. A clear difference is visible between the distribution after the first display and those after the subsequent displays. (C) Grand average reaction time distributions, normalized separately for the first and subsequent epochs. Circles indicate the observed data, and solid line indicates the best-fitting single (epoch 1) or double (epochs 2/3/4) Gaussian curve. (D) Reaction time and hit rate, separately for trials classified as having a fast or slow response. Reaction time is different by construction, whereas no significant difference was observed for hit rate.

Institute (MNI) brain. The lead field was calculated for each grid point using a realistic volume conductor model (Nolte 2003).

We selected data from a time window preceding the search display at which subjects identified the target, ranging between  $-1$  and  $-0.1$  s, and subjected this to a Fourier transform. This particular time window was chosen such that it did not include transient activity related to the processing of the previous display (which had its onset at  $-1.4$  s). For the reconstruction of alpha sources (Fig. 2A), we estimated activity at 10 Hz, and applied a set of multiple Slepian tapers to obtain a frequency smoothing of 2 Hz, resulting in an activity estimate in the band of 8–12 Hz. For the reconstruction of beta activity (Fig. 3A), we again used a Fourier transform after applying multiple Slepian tapers, now to obtain an activity estimate in the band of 16–30 Hz ( $23 \pm 7$  Hz smoothing). From the computed Fourier spectra, the CSD was computed. We computed the CSD using trials from all conditions combined, and applied the resulting filters to the conditions separately.

For visualization, the beamformer results and associated statistics were interpolated onto the single-subject MNI template brain using smudge interpolation. The percentage change of  $A > B$  is computed as  $2 \times (A - B) / (A + B) \times 100\%$  (% change in figure legends refers to power % change).

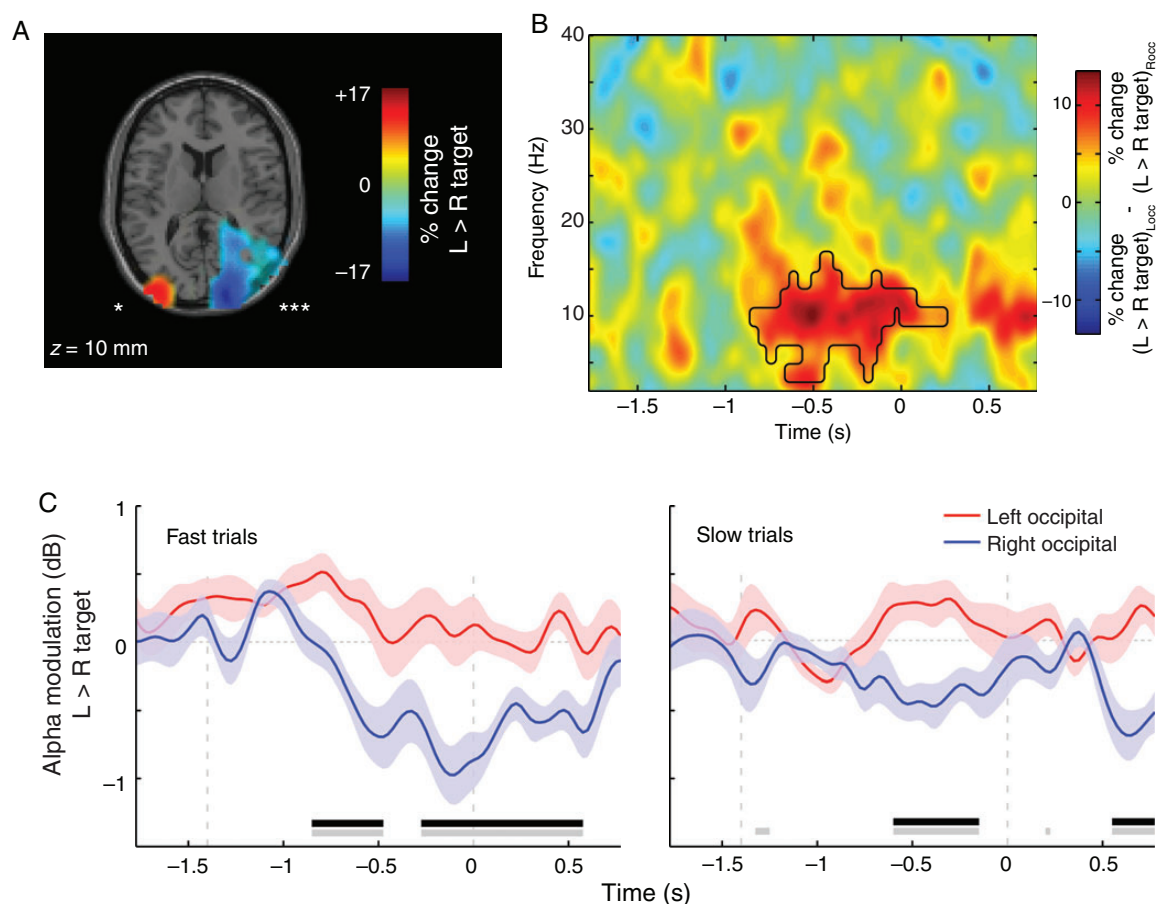
### Time-Domain Source Analysis and Time-Frequency Analysis

For the grid points resulting in a significant source space frequency-domain difference, we computed a time-domain beamformer to obtain temporal source activity estimates. Specifically, we computed a linearly constrained minimum variance (LCMV; Van Veen et al. 1997) beamformer spatial filter, using

the same lead field matrix as computed for the DICS beamformer. The LCMV beamformer uses the time-domain covariance, rather than the CSD, to compute the spatial filter. The covariance was computed from the entire epoch length ( $-2$  to  $+1$  s), after subtracting the mean and fitting and subtracting a linear trend. As for the frequency-domain beamformer, we computed the covariance from all conditions combined. Applying the LCMV spatial filter to our data resulted in single-trial estimates of time-resolved current density at the significant grid points in 3 orthogonal orientations. Effectively, this beamformer is sensitive to all frequencies—as opposed to the above analyses that focused on induced responses in the alpha and beta bands. We subsequently resolved the induced responses in terms of frequency bands using a time-frequency analysis of the reconstructed responses (see below). Subsequently, computed power values were averaged over grid points and orientations to get a single estimate per source cluster.

For the signal thus reconstructed at each grid point, and for each orientation, we computed a time-frequency representation (TFR) of power (Figs 2B and 3B). To obtain the TFR of power, we used a sliding time window Fourier transform, moving over our trials' time axes in steps of 25 ms. The time window was always 400 ms long and multiplied with a Hanning taper of equal length. Estimates were obtained in steps at frequencies between 2 and 40 Hz, in steps of 2 Hz. For visualization purposes, the TFR maps were smoothed by applying a two-dimensional spline interpolation.

For the time courses of alpha and beta modulation (Figs 2C and 3C), we averaged the (unsmoothed) TFR of power in the frequency bands of 8–12 and 16–30 Hz, respectively. Modulation is expressed in decibel (dB), and defined as  $\text{mod} = 10 \times \log_{10}(\text{pow}_{\text{left}} / \text{pow}_{\text{right}})$ , where  $\text{pow}_{\text{left}}$  and  $\text{pow}_{\text{right}}$  refer to the power of trials



**Figure 2.** Anticipatory posterior alpha modulation. (A) Beamformer results for anticipatory alpha power, interpolated onto the single-subject MNI template anatomy (shown is a transverse slice at MNI  $z = 10$  mm). Plotted activity is masked with cluster-corrected  $P < 0.05$ . The image is displayed using neurological convention (left hemisphere displayed on the left). \* $P < 0.05$ ; \*\*\* $P < 0.005$ . (B) TFR of power in reconstructed source time series. Shown is the difference in alpha modulation between the left and right occipital cortex. Outline, cluster-corrected  $P < 0.05$ . Locc, left occipital; Rocc, right occipital. (C) Time course of alpha modulation, shown separately for fast (left) and slow (right) trials. Anticipatory alpha modulation (i.e., a significant difference between the 2 traces in each panel) is visible for fast and slow trials, but is much more pronounced for the fast trials, and persists during stimulus processing. Error shading reflects unbiased within-subjects corrected standard error (Cousineau 2005; Morey 2008). Horizontal bars underneath the curves reflect uncorrected (gray) and cluster-corrected (black) significant differences at  $P < 0.05$ .

with either left or right target (in the case of visuospatial alpha modulation) or left or right response hand (in the case of motor beta modulation).

### Statistical Tests

To statistically quantify the robustness of our results, we performed cluster-based permutation tests (Nichols and Holmes 2002; Maris and Oostenveld 2007) across subjects. Specifically, for each voxel (either time/frequency, time, or  $x/y/z$  voxels), we computed a difference metric between conditions. When raw power formed the input to the statistical test, we used the normalized difference;  $\text{diff} = (A - B)/(A + B)$ . When the input variables were already expressed as some relative metric (e.g., alpha/beta modulation in dB), a  $t$ -score was used instead. The full map of this voxel-level metric was subjected to a standard randomization procedure, correcting for multiple comparisons, and testing the null hypothesis of exchangeability of conditions (Nichols and Holmes 2002; Maris and Oostenveld 2007).

### Exploratory Whole-Brain Analysis

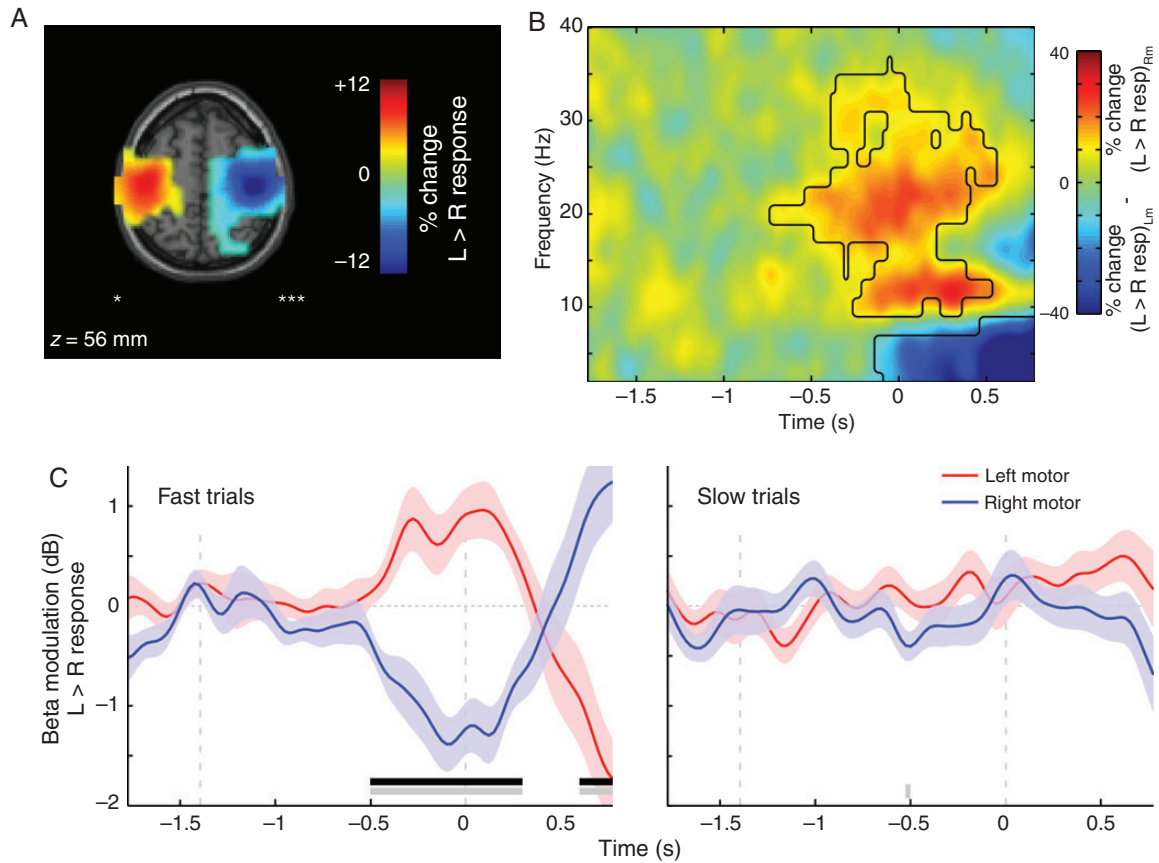
For the analysis of our MEG results, we first examined specific and well-established neurophysiological signals: Lateralization of

visuospatial alpha-band power (in relation to target location) and somatomotor beta-band power (in relation to target identity and associated button press). In addition to these analyses, which were motivated by the question of the nature of predictive signals (spatial and identity), we were also interested in the potential source of the predictive signals. For this analysis, we computed a TFR of power (as described above) for all MEG sensor time courses, after computing an approximation of the vertical and horizontal planar gradient. We then compared these TFR maps between fast and slow responses to identify the sources of anticipatory or predictive processing that set the context for the upcoming search display. To identify the associated neuronal sources, we then reconstructed the induced responses in source space as above. Note that here we established the significance in sensor space and did not need to repeat the randomization testing in source space.

### Results

We presented human subjects ( $n = 19$ ) with a visual search array, consisting of 8 stimuli in both the left and right visual fields (Fig. 1A). One of these was an upright or inverted T-shaped target stimulus, and the others were L-shaped distractor stimuli. The





**Figure 3.** Anticipatory motor beta modulation. All panels analogous to Figure 2, but computed for the beta frequency band and in motor cortex source clusters. A clear anticipatory motor beta modulation is present for the fast trials, but not for the slow trials.

subjects were instructed to identify whether the target was upright or inverted by pressing a button with either the left (upright) or right (inverted) hand, as soon as they had found the target. The search array was presented briefly for 100 ms, followed by a 1300-ms blank screen, and repeated up to 4 times, or until the subject made a response. We refer to a single presentation of the search array as a “display,” and to the time window consisting of a display plus the blank that followed it as an “epoch,” in line with the existing literature. A “trial” thus refers to a set of between 1 and 4 epochs. We will occasionally use the term “glance” to refer to the visual processing of a single display (thus do not mean an overt eye movement by the term). During the entire experiment, we recorded brain activity using MEG. Time courses reported for all MEG analyses follow the convention that  $t = 0$  s corresponds to the onset of the display, in which the subjects identified the target (with  $t = -1.4$  s thus corresponding to the onset of the previous display). Trials in which the subject identified the target during the first display were not used for the MEG analyses, nor were trials with incorrect responses.

### Rapid Resumption During Serial Visual Search

In just over half of all trials, subjects responded after seeing the display only once, so during the first epoch (number of trials =  $266 \pm 22$ , mean  $\pm$  standard error across subjects). In the remaining 216  $\pm 22$  trials, subjects responded after the second ( $109 \pm 8$ ), third ( $69 \pm 9$ ), or fourth ( $37 \pm 10$  trials) display. For the trials with responses during the first epoch, the mean reaction time was  $985 \pm 14$  ms. On average, subjects were significantly

faster when they responded during a later epoch, with a mean reaction time of  $752 \pm 21$  ms (paired-sample  $t_{(18)} = 10.00$ ,  $P < 10^{-8}$ ; reaction times for all epochs, correct responses: 2:  $775 \pm 26$ ; 3:  $752 \pm 23$ ; 4:  $687 \pm 45$  ms). Importantly, the reaction time distribution showed a clear single peak around the mean reaction time for the first epoch, whereas this distribution was broader for the later epochs (Fig. 1B). When inspecting the reaction time distribution normalized separately for the first and later epochs, a clear bimodal distribution was visible for the later epochs, but not for the first (Fig. 1C). We formally tested this bimodality by fitting both a single Gaussian curve and a sum of 2 Gaussians to the reaction time distributions. We found that for the later epochs, the sum of 2 Gaussians fitted the data significantly better than the single Gaussian (Wald  $F_{7,15} = 4.24$ ,  $P = 0.045$ ), while this was not the case for the first epoch ( $F_{7,15} = 2.07$ ,  $P = 0.18$ ). [The Wald test controls for the trivial increase in explained variance because of the increase in degrees of freedom of the model used (Fox 1997).] We thus confirm a clear distinction between trials that show rapid resumption and trials that do not, confirming previous findings (Lleras et al. 2005).

For every subject, we defined trials showing rapid resumption, or “fast” trials, as those having a reaction time below the median for that subject, whereas “slow” trials are those trials having a reaction time above the median. Fast and slow trials are only defined for epochs 2, 3, and 4. This resulted in a mean reaction time for fast trials of  $492 \pm 27$  ms, and a mean reaction time for slow trials of  $1010 \pm 27$  ms. The mean reaction time for slow trials was not significantly different from the mean reaction time in epoch 1 ( $t_{(18)} = 1.11$ ,  $P = 0.28$ ), indicating that the time course of

visual search for a previously-seen display, without rapid resumption, is similar to that of visual search to a display that has not been seen before. Although hit rates were significantly higher for epoch 1, hit rate =  $0.88 \pm 0.019$ , than for later epochs, hit rate =  $0.73 \pm 0.027$ ,  $t_{(18)} = 6.10$ ,  $P < 10^{-5}$ , there was no significant difference between hit rate for fast and slow responses ( $t_{(18)} = 0.89$ ,  $P = 0.38$ , Fig. 1D; hit rates for all epochs: 2:  $0.78 \pm 0.030$ ; 3:  $0.74 \pm 0.030$ ; 4:  $0.65 \pm 0.028$ ), demonstrating that fast responses were not simply fast guesses due to a more liberal decision criterion.

### Rapid Visual Search Is Associated with a Preselection of Relevant Visual Space

The visual search targets could be present in either the left or right visual field. Therefore, we reasoned that if rapid visual search is caused by a preselection of the relevant location where the target is expected to be found (based on earlier glances), we should see stronger modulation of visuospatial preparatory processes for fast than slow responses. To investigate this, we looked at alpha-band modulation over occipital and occipito-parietal cortex. It is a well-established finding that alpha (8–12 Hz) activity in posterior cortex is modulated in a retinotopically specific manner (i.e., alpha power decreases over the visual cortex contralateral to the attended/stimulated hemifield, relative to the ipsilateral cortex), when subjects covertly direct their visuospatial attention (Worden et al. 2000; Thut et al. 2006). Furthermore, it is known that the extent of this alpha modulation is predictive of subsequent target detection performance (Thut et al. 2006; Händel et al. 2011). Specifically, alpha-band modulation should be evident before the display at which subjects identified the target, if part of the “perception cycle” is completed before that display (as is hypothesized to be the case for the fast trials).

In line with this, we found significant alpha modulation (i.e., a stronger alpha suppression in cortex contralateral to the target) in occipital cortex during the 900-ms period before display re-appearance (Fig. 2A; whole-brain cluster-corrected permutation test left cluster  $P = 0.035$ , right cluster  $P = 0.0020$ ). A time-frequency analysis on source space signals extracted from the bilateral occipital clusters is shown in Figure 2B.

Crucially, the anticipatory modulation of alpha differed between fast and slow trials. The alpha modulation (left vs. right target) is plotted separately for the left and right hemisphere, and for fast and slow trials, in Figure 2C. We found that anticipatory modulation occurred for both fast and slow responses. The anticipatory alpha-band modulation was stronger for the fast than for the slow trials (modulation in the left vs. right hemisphere, cluster-corrected permutation test  $P = 0.003$ ). We found a significant interaction effect on alpha modulation (averaged over  $-1$  to  $0$  s prestimulus) between hemisphere and response speed ( $F_{1,18} = 12.7$ ,  $P = 0.002$ ). We repeated the analysis depicted in Figure 2C using only those trials for which the subject responded in the second epoch (as opposed to those trials for which the response occurred during all later epochs 2, 3, or 4). The results of this control analysis did not differ qualitatively from those reported here, thus ruling out the duration of exposure to the stimulus array (over and above the effect of epoch 1 vs. the later epochs) as a potential confound.

### Rapid Visual Search Is Associated with a Preselection of Target Identity

Having established that spatial location is a part of the perceptual predictions being made during rapid visual search, we next asked

if also information concerning target identity could be observed before the response display. Because our subjects were instructed to respond with a left-hand button press when they perceived the target as upright, and with a right-hand response for an inverted target, motor preparatory processes provided a window into the preselection of target identity and its corresponding motor plan (Donner et al. 2009; de Lange et al. 2013). It has previously been shown that lateralization of beta-band (16–30 Hz) oscillatory activity observed over motor areas is a sensitive marker of preparation of either the left or the right response hand, with activity building up gradually during the accumulation of sensory evidence that is linked to a particular response (Donner et al. 2009).

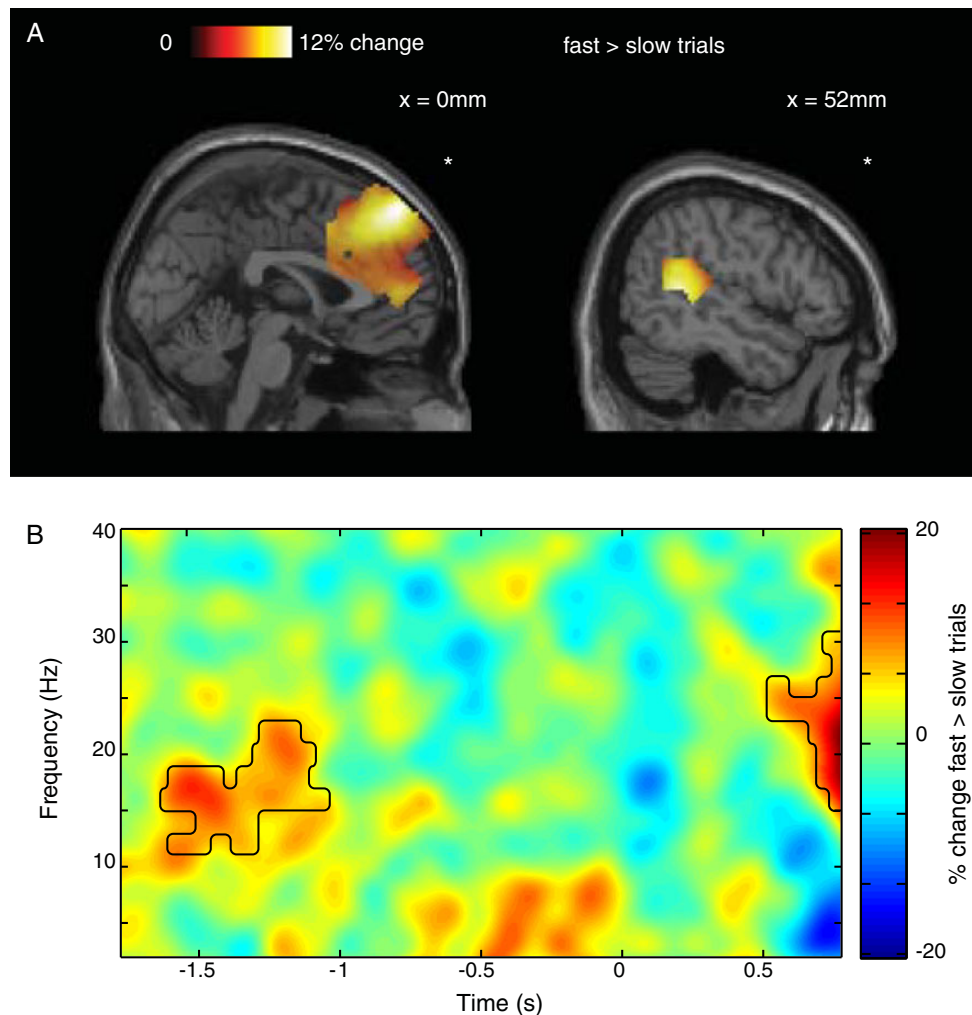
Indeed, we observed significant preparatory modulation of beta power (i.e., a stronger beta suppression contralateral to the response hand) over the motor cortices (Fig. 3A; whole-brain, cluster-corrected permutation test left cluster  $P = 0.01$ , right cluster  $P = 0.001$ ) during the 900-ms period before display re-appearance. A time-frequency analysis on source space signals extracted from the bilateral central clusters is shown in Figure 3B.

Crucially, this prestimulus beta-band modulation was markedly different for fast trials, compared with slow trials. The beta-band modulation (upright vs. inverted target) is plotted separately for fast and slow trials in Figure 3C. Significant preparatory beta-band activity was observed starting at  $-0.5$  s for the fast responses ( $P = 0.001$ ), whereas no preparatory beta-band activity was observed for the slow responses. Again, we repeated the analysis depicted in Figure 3C using only those trials for which the subject responded in the second epoch, and found qualitatively identical results, thus ruling out the duration of exposure to the stimulus array as a potential confound. These findings are consistent with the notion of a perceptual prediction concerning target identity being generated before the onset of the display to which subjects responded, with the confirmation of this prediction following only after the onset of this display. This temporal dissociation between the prediction and confirmation phase then explains the occurrence of rapid resumption.

### Superior Frontal Cortex and Right Temporo-Parietal Junction Are Associated with Prediction Generation

The above analyses focused on the “nature” of the perceptual prediction: Is rapid visual search associated with a preselection of the relevant visual space, and/or the relevant target identity and associated motor program? We find empirical evidence for both of these preselection mechanisms. Next, we asked whether we could localize the “source” of the perceptual prediction: Is there any neural activity that distinguishes rapid from slow visual search and that precedes the biasing signals of visual and motor activity?

To answer this question, we contrasted oscillatory power at the sensor level in fast versus slow trials, in the time window corresponding to the processing of the display prior to the response epoch ( $-1.5$  to  $-1$  s). We found a significant difference in the 10- to 20-Hz frequency band in this time window (sensor-level, cluster-corrected permutation test  $P = 0.037$ ), which originated from 2 spatial clusters, as identified by beamformer source analysis. These clusters were located on the medial surface of the superior frontal cortex and in the right TPJ (Fig. 4A). The time window of this effect,  $t = -1.5$  to  $-1$  s, corresponds to the presentation of the search display preceding the search display at which subjects identified the target. As such, neural activity in these 2 brain regions around the time of stimulus processing dissociated between a perceptual prediction being generated or not.



**Figure 4.** Neural sources active during the generation of a perceptual prediction. (A) Sources that are significantly stronger activated for fast versus slow responses after display  $n$ , during the processing of display  $n-1$ . Shown are coronal slices at MNI  $x = 0$  (left panel) and 52 (right panel) mm, revealing activation in medial SFG (left) and right TPJ. Activation maps are masked with cluster-corrected  $P < 0.05$ . \* $P < 0.05$ . (B) TFR of power in reconstructed source time series for the frontal source cluster. Shown is relative change between fast versus slow trials. Outline, cluster-corrected  $P < 0.05$ .

## Discussion

Visual search is often greatly facilitated by prior knowledge, such as context knowledge or previous acquaintance with a visual scene. In this study, we examined the nature of predictive information that leads to rapid resumption in visual search, as well as the neural sources that may generate this predictive information. We presented subjects with a repeated visual search array, and confirmed previous reports that studied interrupted visual search (Enns and Lleras 2008): When subjects resume a visual search, their reaction time distribution shows a bimodal pattern, compared with a unimodal response pattern for initial visual search. Neural data suggest that the markedly fast visual search times that can be seen at the early peak of the bimodal distribution are due to several predictive processes that are generated prior to the display at which subjects identified the target. Both information concerning target location and target identity were present in these perceptual hypotheses: We found significant preparatory alpha modulation over visual cortex and significant preparatory beta modulation over motor cortex. Furthermore, we found that medial superior frontal cortex and the right TPJ might be responsible for generating these perceptual predictions.

The perceptual prediction contained information about the spatial location of the target. This was evident from the alpha modulation, which was significantly elevated before the onset of the responded-to display. The alpha modulation was elevated throughout a much longer time window for trials with a fast response than those with a slow response. No prestimulus alpha modulation was observed for trials where there was no response (data not shown). The presence of anticipatory alpha modulation indicates an active preparation of the subjects' visual system for the upcoming target. This is in line with the existing literature reporting a crucial role of the preservation of target location for rapid resumption to occur (Zoest et al. 2007). It is known that the allocation of visual spatial attention modulates the topographical organization of alpha activity, when attention is directed in response to an endogenous cue (Worden et al. 2000; Thut et al. 2006). Our study did not employ such an endogenous cue; rather, the attentional effect was driven by the target stimulus itself. However, the time course of the effect (starting for the fast trials ~600 ms, and for the slow trials ~1 s after stimulus onset) suggests that this effect is mediated by a top-down anticipatory drive, and does not reflect a purely exogenous attentional shift.

Not just the location of the target, but also its identity was apparent as a prestimulus prediction. We found clear prestimulus beta modulation over the motor cortex in the direction of the correct target identity and the corresponding response. This modulation was only present for subsequent “fast” trials, that is, in trials where a perceptual prediction was generated. No beta modulation was observed in trials without such a prediction. Build-up of choice-predictive activity in motor cortex during a perceptual decision-making task is a well-established phenomenon (Donner et al. 2009) and can also be elicited by prior sensory expectations (de Lange et al. 2013). Because we observed a clear dissociation between fast and slow responses, we believe that the proper interpretation of the beta-band modulation here is as a reflection of an actively generated top-down prediction or expectation. If the beta-band activity reflected only the passive accumulation of evidence, this would have resulted in a unimodal reaction time distribution (as we observe for the first search display), rather than the bimodality that we observed for resumed visual search. Our finding of a prediction concerning target identity being generated before stimulus onset is in line with psychophysical results, demonstrating that rapid resumption is abolished when target identity is changed between subsequent displays (Jungé et al. 2009).

For displays with subsequently fast responses, we found greater activity in the right TPJ and the medial superior frontal gyrus (SFG) during the processing of the previous display, compared with slow responses. In other words, activity in these regions during processing of display  $n-1$  was predictive of whether subjects responded fast or slow after stimulus display  $n$ . Interestingly, these 2 nodes are part of 2 different frontoparietal networks typically associated with attentional processes. The right TPJ is considered to be part of the ventral attentional network, whereas (part of) the SFG is a node in the dorsal attentional network (Corbetta and Shulman 2002). The dorsal network is thought to be involved in the direction and maintenance of top-down driven attention (e.g., in response to an endogenous cue), whereas the ventral network is more involved in the redirection of attention in response to salient external stimuli. Our observed right TPJ and SFG activation can be interpreted in light of these frontoparietal attentional networks: During processing of the search display, the right TPJ may on some trials signal the presence of the task-relevant target stimulus. Thus, the ventral attentional network of which it is part may execute its “circuit breaker” role to shift attention in the direction of this stimulus. Because the stimulus array is followed by a relatively long delay, visuospatial attention needs to be endogenously maintained to prepare the sensory system for the upcoming next display. This task is performed by the dorsal attentional network, of which the SFG is part. Consistent with this idea, it has been shown that the dorsal network controls spatial attention through a modulation of posterior alpha activity (Capotosto et al. 2009), which we also observed. Furthermore, it was previously shown that synchronization in a frontal-posterior network is associated with unconscious learning during a visual search task (Chaumon et al. 2008); it is plausible that this network consists of the same areas that we reported to be active. The frontal eye fields (FEFs) form the frontal node in the dorsal attentional network. Although the extent of our activated cluster is consistent with it involving the FEF, the inherent uncertainty of the spatial resolution of MEG does warrant some caution in interpreting the functional neuroanatomy.

The fact that activity in 2 areas typically associated with attentional processes differentiates between the presence and absence of an expectation about upcoming stimuli might prompt

the question to what extent our findings truly reflect “expectational” versus “attentional” processes. It is likely that, in the present paradigm, expectation serves to guide attention. The aim of this study was not to separate influences of expectation and attention, but rather to show how implicit expectations about stimuli, based on previous glances, serve to guide visual search, and by which neural mechanisms this facilitation occurs. We found that this facilitation occurs partly through a prioritization of certain parts of visual space where the target is expected to appear, in other words, through spatial attention. In this sense, some of the neural findings we report likely correspond to attentional processes. However, in view of the experimental set-up, which did not include any attentional cues, we can conclude that it must have been an “expectation” about the upcoming stimulus which served to guide attention. To our knowledge, it has not been previously shown that such a rapidly formed and implicit expectation can guide attention. It is interesting to note that, although expected stimuli have previously been reported to result in an attenuated neural response (e.g., Summerfield et al. 2008; Kok, Jehee, et al. 2012; Summerfield and de Lange 2014), the effect of expectation is sometimes reversed when the expected stimulus is actually relevant/attended, such that the expected stimulus elicits “higher” neural activity (Kok, Rahnev, et al. 2012). This finding is in line with our interpretation that expectation likely serves to guide attention in the current experiment.

Another crucial question is whether our results pertain to a “perceptual” process. A possible alternative explanation of the rapid resumption phenomenon could be that subjects have actually already identified the target on display  $n-1$ , yet consciously decide to wait until the next display to be absolutely sure of their decision. However, if this were the case, subjects would be expected to perform above chance when they are forced to respond during a trial for which no response had been made yet. In contrast to this notion, Lleras et al. (2005) found that when a display is presented, but, unexpectedly to the subjects, not repeated, performance is at chance for trials during which subjects chose to withhold their response until a subsequent display. This suggests that the predictive mechanisms that lead to rapid resumption in visual search are implicit, rather than strategic. However, it should be noted that we did not include a forced-response condition in the present experiment. Therefore, we cannot conclude with certainty that the neural mechanisms we observed are reflecting the unconscious process identified by Lleras et al. (2005). Future work is necessary to further disentangle conscious and unconscious predictive processing in visual perception.

Do subjects actually make an informed prediction, or could our results be explained by them randomly directing their attention to the right location at the right time? If the attentional spotlight (by random wandering) happens to be at the location of the target at the onset time of the display, subjects might be expected to respond faster than when attention happens to be directed away from the target. This phenomenon could also result in a bimodal distribution of reaction times. However, this account dictates that also responses after the first display should show such a bimodality. In contrast, we find a bimodal reaction time distribution only after second and later displays. Therefore, we can conclude that the biasing of spatial attention and target identity-related response activity only occurs in response to visual information (i.e., the first display). This visual information is therefore likely the source of the spatial and target identity predictions being generated. Our results thus speak against the hypothesis that “visual search has no memory”, as has been previously suggested (Horowitz and Wolfe 1998).



Summarizing, based on our results, how is visual search helped by prior glances? We find that on some trials, subjects find the target immediately after the first display, and make their decision. This corresponds to all 3 phases (activate, predict, and confirm) of the “perception cycle” being completed after only a single display (Enns and Lleras 2008). On other trials, a sub-threshold signal reflecting the target’s location and identity is perceived after the first (or a subsequent) display (the “activate” phase). This is associated with activation of the stimulus-driven (ventral) and top-down controlled (dorsal) attentional networks. In turn, the dorsal attentional network effects a prioritization of the visual field where the target is expected to occur, and a prioritization of the motor decision associated with the identity of the target (the “predict” phase). These biasing processes before the target display onset then lead to a rapid identification of the target, as only the third (“confirm”) phase of the perception cycle needs to be completed. In conclusion, this study elucidates which features are involved in perceptual prediction during visual search, by showing that subjects can generate a spatially and feature-specific perceptual prediction prior to the visual search display at which the target is identified.

## Funding

This work was supported by the Netherlands Organization for Scientific Research (NWO) VICI Grant #453-09-002 (O.J. and E.S.), a James S. McDonnell-Foundation (JSMF) Scholar Award (F.d.L.), and Fulbright and Howard Hughes Medical Institute fellowships (Y.F.).

## Notes

*Conflict of Interest:* None declared.

## References

- Bar M. 2004. Visual objects in context. *Nat Rev Neurosci.* 5: 617–629.
- Bell AJ, Sejnowski TJ. 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7:1129–1159.
- Biederman I. 1972. Perceiving real-world scenes. *Science.* 177:77–80.
- Brainard DH. 1997. The psychophysics toolbox. *Spat Vis.* 10:433–436.
- Capotosto P, Babiloni C, Romani GL, Corbetta M. 2009. Frontoparietal cortex controls spatial attention through modulation of anticipatory alpha rhythms. *J Neurosci.* 29:5863–5872.
- Chaumon M, Schwartz D, Tallon-Baudry C. 2008. Unconscious learning versus visual perception: dissociable roles for gamma oscillations revealed in MEG. *J Cogn Neurosci.* 21:2287–2299.
- Chun MM, Jiang Y. 1998. Contextual cueing: implicit learning and memory of visual context guides spatial attention. *Cognit Psychol.* 36:28–71.
- Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci.* 3:201–215.
- Cousineau D. 2005. Confidence intervals in within-subjects designs: a simpler solution to Loftus and Masson’s method. *Tutor Quant Methods Psychol.* 1:42–45.
- De Lange FP, Rahnev DA, Donner TH, Lau H. 2013. Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci Off J Soc Neurosci.* 33:1400–1410.
- Donner TH, Siegel M, Fries P, Engel AK. 2009. Buildup of choice-predictive activity in human motor cortex during perceptual decision making. *Curr Biol.* 19:1581–1585.
- Enns JT, Lleras A. 2008. What’s next? New evidence for prediction in human vision. *Trends Cogn Sci.* 12:327–333.
- Fox J. 1997. *Applied Regression Analysis, Linear Models, and Related Methods.* Newbury Park, CA, USA: Sage Publications.
- Gross J, Kujala J, Hämäläinen M, Timmermann L, Schnitzler A, Salmelin R. 2001. Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc Natl Acad Sci USA.* 98:694–699.
- Händel BF, Haarmeier T, Jensen O. 2011. Alpha oscillations correlate with the successful inhibition of unattended stimuli. *J Cogn Neurosci.* 23:2494–2502.
- Horowitz TS, Wolfe JM. 1998. Visual search has no memory. *Nature.* 394:575–577.
- Jungé JA, Brady TF, Chun MM. 2009. The contents of perceptual hypotheses: evidence from rapid resumption of interrupted visual search. *Atten Percept Psychophys.* 71:681–689.
- Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, Sejnowski TJ. 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology.* 37:163–178.
- Kok P, Jehee JFM, de Lange FP. 2012. Less is more: expectation sharpens representations in the primary visual cortex. *Neuron.* 75:265–270.
- Kok P, Rahnev D, Jehee JFM, Lau HC, de Lange FP. 2012. Attention reverses the effect of prediction in silencing sensory signals. *Cereb Cortex.* 22:2197–2206.
- Lleras A, Rensink RA, Enns JT. 2007. Consequences of display changes during interrupted visual search: rapid resumption is target specific. *Percept Psychophys.* 69:980–993.
- Lleras A, Rensink RA, Enns JT. 2005. Rapid resumption of interrupted visual search. New insights on the interaction between vision and memory. *Psychol Sci.* 16:684–688.
- Maris E, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods.* 164:177–190.
- Morey RD. 2008. Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutor Quant Methods Psychol.* 4:61–64.
- Nichols TE, Holmes AP. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp.* 15:1–25.
- Nolte G. 2003. The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys Med Biol.* 48:3637–3652.
- Oostenveld R, Fries P, Maris E, Schoffelen J-M. 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci.* 2011:1–9.
- Pfurtscheller G, Lopes da Silva FH. 1999. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol.* 110:1842–1857.
- Stokes MG, Atherton K, Patai EZ, Nobre AC. 2012. Long-term memory prepares neural activity for perception. *Proc Natl Acad Sci USA.* 109:E360–E367.
- Summerfield C, de Lange FP. 2014. Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci.* 15:745–756.
- Summerfield C, Trittschuh EH, Monti JM, Mesulam M-M, Egner T. 2008. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci.* 11:1004–1006.
- Thut G, Nietzel A, Brandt SA, Pascual-Leone A. 2006. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci.* 26:9494–9502.

- Treisman AM, Gelade G. 1980. A feature-integration theory of attention. *Cognit Psychol.* 12:97–136.
- Van Veen BD, van Drongelen W, Yuchtman M, Suzuki A. 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Trans Biomed Eng.* 44:867–880.
- Wolfe JM. 1998. Visual search. In: Pashler H, editor. *Attention*. London, UK: University College London Press.
- Worden MS, Foxe JJ, Wang N, Simpson GV. 2000. Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. *J Neurosci.* 20:RC63.
- Zoest WV, Lleras A, Kingstone A, Enns JT. 2007. In sight, out of mind: the role of eye movements in the rapid resumption of visual search. *Percept Psychophys.* 69:1204–1217.