Parsing a Mental Program: Fixation-related Brain Signatures of Unitary Operations and Routines in Natural Visual Search

Juan E Kamienkowski[¶,1,2], Alexander Varatharajah[¶,3], Mariano Sigman[4], and Matias J Ison*[,5]

[1] Laboratory of Applied Artificial Intelligence, Computer Science Institute, School of Exact and Natural Science, University of Buenos Aires – National Council of Science and Technology, Buenos Aires, Argentina

[2] Physics Department, School of Exact and Natural Science, University of Buenos Aires, Buenos Aires, Argentina

[3] Department of Engineering, University of Leicester, Leicester, United Kingdom

[4] Torcuato Di Tella University, Buenos Aires, Argentina

[5] School of Psychology, University of Nottingham, Nottingham, United Kingdom

¶ These authors contributed equally to this work.

* Corresponding author

E-mail: Matias.Ison@nottingham.ac.uk

Short title: Parsing Mental Programs in Visual Search: An EEG - Eye Movements Study

Parsing a Mental Program: Fixation-related Brain Signatures of Unitary Operations and Routines in Natural Visual Search

Abstract

Visual search involves a sequence or routine of unitary operations (i.e. fixations) embedded in a larger mental global program. The process can indeed be seen as a program based on a while loop (while the target is not found), a conditional construct (whether the target is matched or not based on specific recognition algorithms) and a decision making step to determine the position of the next searched location based on existent evidence. Recent developments in our ability to co-register brain scalp potentials (EEG) during free eye movements has allowed investigating brain responses related to fixations (fixation-Related Potentials; fERPs), including the identification of sensory and cognitive local EEG components linked to individual fixations. However, the way in which the mental program guiding the search unfolds has not yet been investigated. We performed an EEG and eye tracking co-registration experiment in which participants searched for a target face in natural images of crowds. Here we show how unitary steps of the program are encoded by specific local target detection signatures and how the positioning of each unitary operation within the global search program can be pinpointed by changes in the EEG signal amplitude as well as the signal power in different frequency bands. By simultaneously studying brain signatures of unitary operations and those occurring during the sequence of fixations, our study sheds light into how local and global properties are combined in implementing visual routines in natural tasks.

Keywords

Visual Search; EEG; Eye Movements; Attention; Natural Scenes; Oscillations

1. Introduction

Most daily tasks –for example reading this sentence– involve a sequence of operations. During the last few decades, there has been an accumulation of knowledge about how the brain computes unitary operations. In contrast, the mechanisms by which unitary operations are assembled in programs and routines in the brain have seen comparatively very little investigation [1–7]. One prominent example of a mental program comes from natural viewing, where two levels of processing must be integrated in real-time: visual processing of each stimulus, and the integration of information along the sequence to achieve the goal of the whole task. However, little is known about brain activity in natural viewing as obtaining EEG potentials is intrinsically difficult because eye movements heavily contaminate brain signals.

In fixed-gaze scenarios, sequential-decision making has been one specific case in which the iteration of unitary operations of information updating has been studied in human neurophysiology [5]. Lange and colleagues showed a gradual reduction of cortical activity with each sample: activity was inversely related to the accumulated evidence, which was interpreted as a reflection of top-down influence on sensory processing [8]. These results built on single neuron studies in macaques, reporting that neuronal activity was modulated by the evidence accumulated throughout a sequence of simple operations in neurons in the parietal cortex [9] and in more complex visual tasks by recordings in the primary visual area [10]. In humans, early ERP studies have shown that in fixed-gaze oddball experiments the amplitude of the P3 component, emerging when the target is detected, is modulated by the inter-stimulus interval (ISI) [11–13]. In many studies, this has been related to the concept of expectancy and surprise and led to the prediction that P3-like components could act as a cortical index to a discrete sequence of accumulated processes [7]. Beyond discrete event-related potentials, some experimental evidence suggests the involvement of low-frequency oscillations (theta, alpha and beta) in integrative brain processes, while high-frequency oscillations (gamma-band oscillations) are sought to reflect the circuit-level mechanisms mediating local encoding processes (either sensory or motor), although both could involve neuronal activity within the same cortical region [14]. For instance, top-

down attention has a different spectral profile from the local encoding of sensory features in primary visual cortex [15], and decision-making has a different spectral profile from that of motor planning [16–19]. Large modulations in the alpha-band have been associated to both event-related desynchronization (ERD) and synchronization (ERS). The most common observation is that brain areas that are activated during a task exhibit ERD, whereas areas associated with task irrelevant and potentially interfering processes exhibit ERS. Thus, the latter process is usually associated with an inhibition, and the former one with the release of that inhibition. This has been observed, for instance, in tasks that vary stimulus modality (e.g., Visual vs. Auditory) [20], stimulus processing domain (e.g., related to the ventral vs. dorsal processing stream) [21] or stimulation side (e.g., stimulation of the right vs. left visual hemifield) [22,23]. The changes in alpha-band frequency power are also progressive, as shown by the temporal spectral evolution (TSE) waveforms in many tasks. For instance, when a warning signal preceded a target or non-target in an oddball task, the alpha-band oscillations not only responded to the identity of the stimulus but also to the expectancy of the upcoming stimulus [24]. Desynchronization of alpha/beta activity has also been reported when the anticipation of an impending target increases [25].

Eye movement and modeling approaches to visual search tend to focus on general planning of the scanpath, and how image properties modulate both fixation duration and position distributions [26–32]. Other experimental approaches include fixed-gaze neuroimaging studies with a single spatially spread stimuli presentation, or sequentially centered stimuli presentations. Single presentation EEG studies tend to focus on the neural basis of spatial covert attention and its mechanisms [33,34], while serial visual presentations (fixed-gaze) are mostly used to study the neural correlates of target detection/identification with overt visual attention [13,35]. Regardless of the methodological approach, much of what is known in the field comes presenting ‘synthetic’ stimuli, ranging from simple bars and sinusoidal gratings to other typical laboratory stimuli including pre-segmented objects and faces. Consequently, there has been an increasing interest in visual search of

natural scenes as the complexity of natural scenes may provide a greater depth of the behavior for investigation [36], particularly in free-viewing visual search [37].

Recent studies have started to combine EEG and eye movements [38–42]. They have shown that it is possible to reliably measure brain activity related to the visual information retrieved from each fixation, which has begun to close the gap between these approaches. This opened the possibility to combine these lines of research and study human neurophysiology of visual overt exploration of natural scenes. In particular, previous studies have investigated brain activity elicited on each fixation, showing that a clear P3 arises from both synthetic [40,43–45] and natural [46] stimuli. Moreover, the P3 component is observed both when a target stimulus is embedded in the middle [40] or at the end of the stream of fixated items [43,45–47], and it is robust enough to be detected in a trial-by-trial basis [43,45–47]. This component has been used as part of a brain computer interface (BCI) system for assistive technologies [47], as well as to investigate other "real world" applications such as understanding information systems [48], or the hazard perception driving test [49]. Most studies have focused on the comparison between target and distractor responses (but see [44] for sequential differences among distractors) [33,34]. A few co-registration studies have also focused on eye movements over natural scenes but in scene perception and memorization, and described the early processing of the fixated patches [41,50–52]. In a recent study [46], we showed that there is a high similarity between fERPs recorded in an eye movement visual search task and ERPs recorded in a fixed-gaze experiment with similar stimuli. This was reflected both in early visual potentials, but also in a late potential that discriminated targets from distractors. One pitfall of that study is that participants were explicitly instructed to perform long fixations. This aimed at eliciting comparable number of long epochs with fixations to targets and distractors at the expense of ecological validity, since in natural viewing the processing of each single stimulus has to be combined with the information integrated throughout the task.

Here, we aimed to investigate the existence of brain signatures for unitary operations and mental programs in a completely free viewing visual search task. To date, this has only been

addressed in very few studies and they all used fixed-gaze paradigms. We hypothesized that local processing in unitary operations under fixed-gaze would be conserved in natural viewing, and could be captured by fixation-related components. Furthermore, we also hypothesized that signatures of global integration of information along the task would be reflected in modulations in the P3 amplitude and low-frequency oscillations. We propose a schematic framework to link the identified electrophysiological signatures with underlying integrative processes in natural viewing, which could form the basis of future theoretical frameworks of gaze control integrating electrophysiology and eye movements in natural viewing.

2. Materials and Methods

2.1. Participants

Seventeen subjects participated in the experiment (13 male/4 female; aged between 21-31 years). All subjects were *naïve* to the objectives of the experiment, had normal or corrected to normal vision, and provided written informed consent according to the recommendations of the declaration of Helsinki to participate in the study.

2.2. Apparatus

Stimuli were presented in a 21-inch cathode ray tube (CRT) monitor with a screen resolution of 1024 x 768 pixels and at a refresh rate of 75Hz. Subjects sat 60cm from the monitor in a chair, with their heads stabilized using a specially designed 'cheek rest' (to avoid EEG artifacts from muscular activity from the jaw) and responses were made on a standard 'qwerty' keyboard. All stimuli were were presented in MATLAB (MathWorks 2000) using the Psychophysics Toolbox extensions [53,54].

2.3. Stimuli

In the database there were 60 gray-scale images. The images were of football crowds in stadiums, each image was 800x768 pixels (23.4 by 22.4 degrees of visual angle) and contained between 23-35 distractors (30.68 mean average) (Fig S1B,D). The luminance profile of each image was equalized and the pictures were converted to gray-scale to constrain the variability of visual salience across global displays (i.e. to avoid characteristics of the image that could be particularly salient). From each image, 3 faces of size 80x80 pixels were chosen as targets and the rest were marked as distractor faces (Fig S1C,D). The selected target locations did not follow any specific pattern and included different angles, expressions and genders. The spatial distribution of all faces and particularly the targets covered the whole display.

2.4. Experimental Procedure

We followed a similar procedure as in Kaunitz et al. (2014) [46]. Briefly, at the beginning of each trial subjects pressed the space bar and were presented with a target face for 3 s (Fig 1A). The target faces were first cropped to a square of size 2.04° x 2.04°. To prevent subjects from facilitating their visual search based on the size of the target face, in each trial we rescaled the original target face to a random size between 1.75° x 1.75° and 2.63° x 2.63°. After this time a fixation dot was presented on the screen at a random location (Fig S1A). Subjects had to fixate at the dot location for 1 s for the image of a crowd to appear on the screen (Fig 1A). The task was to search for the target face within the crowd and to fixate on it for 1 s once they have found it (Fig 1A-C). Trials ended when subjects found the target or after 20 s of visual search. The 60 images were presented in pseudo-random order as a block. Between blocks subjects took 5 min resting breaks. In each block the target face for each image was different from previous blocks. In total subjects performed 180 trials (3 different targets per crowd image for the whole experiment).

The main difference from Kaunitz et al. (2014) was that in the present study participants were neither trained to perform slower fixations in a previous session, nor were the instructions biased

towards accuracy instead of speed [46]. Subjects were simply left to explore and search for the target as they naturally would. The rationale of the procedure in the previous experiment was to get a large enough sample of distractors to obtain measurable fERPs. In the present experiment we aimed to explore naturally occurring fixations, and we show that even in this situation it is possible to observe robust fERPs.

2.5. Eye movements and EEG data acquisition

Eye movements were registered with an EYELINK 1000 system (SR Research, Ontario, Canada). The ET was used in binocular mode with stabilised-head and sampling rate of 500Hz in each eye. Saccades were detected using an adapted version of velocity-based Engbert and Kliegl's algorithm [55,56]; using the parameters described in Kamienkowski et al. (2012) [40]. Only saccades larger than 1 degree were kept for the analyses of the data, as saccades below this threshold were considered microsaccades [37]. For all the experiments a drift correction was made every 10 trials, and a recalibration of the ET every 60 trials (before the beginning of a new block).

The EEG data was recorded with a 64-channel 10-20 montage using the Biosemi Active-Two System (Biosemi, Amsterdam, Holland) at 1024 Hz. The Data were imported into Matlab using the EEGLAB toolbox [57] with linked mastoids as the reference. For fERPs, the datasets that were created were down-sampled at 256 Hz and band-pass filtered between 0.1 – 40 Hz (six order elliptic filter). The start of the fixation on the distractor or target face, identified from the eye tracker data, was taken as the onset of the trial. The brain responses for the target from each crowd image were analyzed, as were the fixations to distractors. For the comparisons between targets and distractors, only fixations that lasted longer than 400ms were kept and the EEG data was aligned to fixation onset and epoched between [-0.2 0.8] seconds from the start of the fixation. For constructing the fERPs related to distractors early processing, we kept fixations longer than 200ms. In both cases, we excluded fixations to distractors larger than 1000ms, which represented (0.6 ± 0.1)% of the fixations to distractors. For the fERP local analyses, we applied a baseline correction to each epoch in the time

window [-200 -100]ms from fixation onset. For the global fERP analyses (Fig. 4), no baseline correction was used.

EEG and ET data were synchronized as in Kaunitz et al. (2014) [46]. Briefly, an analog card was used to convert the digital eye position into analog voltage channels. The temporal offset between the signals was corrected by realigning the eye tracking data to the frontal electrodes that exhibited a sharp saccadic spike potential shortly after saccade onset.

Data are fully available on reasonable request.

2.6. Matching Procedure

Fixations to targets and distractors were match-selected, based on the eye-movement properties of each fixation [40,58]. This was to avoid any baseline differences created from the eye movements, so that the fERPs for targets and distractors could be compared without artifactual components affecting the results. In short, we used as matching parameters the preceding saccade horizontal (*dx*) and vertical (*dy*) amplitudes, as well as its duration (*dt*). We implemented K-nearest neighbors (KNN) algorithm as a robust matching method. This algorithm finds the 'nearest neighbor' for an mx-by-n matrix X in each point of an my-by-n matrix Y. The method is exhaustive and uses replacement; first calculating the distance of each point and then finding the smallest distance. Once matched, the element is placed back in the pool to be matched for distance again. A standardized Euclidean distance metric (see Equation 4.1) was used, where x_s is a column vector from X that corresponds to y_t a column vector from Y:

$$d_{st} = (x_s - y_t)V^{-1}(x_s - y_t)' \qquad (4.1)$$

Here, V is the n-by-n diagonal matrix whose jth diagonal element is $s(j)^2$, where s is the vector with the sample standard deviation; this scales the difference between rows x_s and y_t by dividing the corresponding elements by the standard deviation. Following the matching procedure, the

different parameters matched were shown to have similar distributions across both target and distractor condition (see Fig S2). The matching procedure has shown that the parameters have no significant differences in each of the three variables used: Saccade Duration (D: 32.0ms ([24.0 ms 38.0 ms]), T: 32.0 ms ([24.0 ms 38.0 ms]), Wilcoxon rank-sum test: p=0,96), Horizontal Saccade Amplitude (D: 4.7 deg ([2.3 deg 8.1 deg]), T: 4.6 deg ([2.3 deg 8.2 deg]), Wilcoxon rank-sum test: p=0,86), and Vertical Saccade Amplitude (D: 2.4 deg ([1.1 deg 4.7 deg]), T: 2.5 deg ([1.2 deg 4.8 deg]), Wilcoxon rank-sum test: p=0,81).

In earlier studies, explicit instructions were given to participants in order to increase the length of fixations, producing more predictable and "matchable" eye movements [40]. Here we found that it is possible to match properties with unrestricted free-viewing eye movements in a robust way.

2.7. Single-Trial Analysis

Single-trial ERPs were extracted using a denoising algorithm based on wavelet decomposition [59–61]. The first step was to project the single-trial traces into the wavelet space. Then, the wavelet coefficients related to the evoked responses were selected automatically for each channel and subject using the algorithm implemented in the *ep_den* package [59]. Finally, the ERPs were denoised by reconstructing the signal using only those wavelet coefficients.

2.8. Time-Frequency Analysis

For the Time-Frequency analysis, the EEG data sampled at 1024 Hz was band-pass filtered at 0.1 – 95 Hz (butterworth), and band-stop filtered between 49 – 51 Hz (FIR). The data were epoched into trials (including 10 s before the presentation of the stimuli and 10 s after the end of the trial). Then, Infomax ICA was calculated only on successful trials. The artifactual ICs related to muscular, erroneous electrodes and eye movements were selected and removed based on 'ADJUST' [62] and 'EyeCatch' [63] methods and supervised by an expert. Across all participants we removed (8 ± 5)

independent componenets (mean ± stdev). Due to the potential confound of removing aspects of (fixation-related) cognitive processes of interest, we were particularly conservative in discarding any ICs, for instance the sole appearance of an alpha peak was an argument for keeping the IC.

Time-Frequency transform was performed using a method based on the convolution of the signal with a complex wavelet ('mtmconvol' method) from the FieldTrip toolbox [64], using Hanning tapers. The time was sampled on windows of 0.5 secs, covering the whole trial with a step of 0.1 secs. The frequency was sampled from 0.5 to 30 Hz with a step of 0.5 Hz.

Spectral profiles, defined as the change in the cortical power spectrum due to a functional process, have been proposed to be a functional link between the MEG/EEG signals and cognition [14,65]. Spectral profiles account for changes of the frequency spectrum with a given experimental manipulation, stimulus variation or task progression. In our case, we aimed to explore the cognitive correlates of progression of the task by using their fingerprints on the spectral profile, i.e. we calculated the correlation between the frequency power and the fixation rank to distractors during the visual search.

In particular, for the present analysis we focused on two frequency bands: theta ([4 8] Hz) and alpha ([8 13] Hz). Thus, the power of frequencies falling into those bands were averaged for each electrode and time window. Then, the trials were epoched into fixations, averaging for each fixation the power estimated for all the windows that fall into the fixation. Thus, we ended up with two values of power for each fixation (sorted by fixation rank) and electrode, and we averaged those values first within participants and then across participants for the electrodes of interest.

3. Results

3.1. Experimental paradigm and behavior

The time course of an exemplary trial is shown in Fig 1A. Across all subjects, we obtained a total of 19,722 fixations larger than 50 ms. The overall distribution of the eye movement variables

during the search presented a typical pattern. The mean and standard deviation of the fixation duration were (240 ± 103) ms (Fig 1E). The 5% lower limit of the saccade amplitude was 1.28 degrees (Fig 1C), and the percentage of saccades that were actually smaller than the target face (2.04 degrees) was 18.66%. Participants found the target most of the times -*P(target found) = (0.86 ± 0.02)*- and performed very few fixations to the target before the one that finished the trial -percentage of trials with re-fixations to the target = *(4.3 ± 0.6)%*-. Altogether, this suggests that subjects were successful at performing this task and did not typically perform two consecutive fixations on the same face (Fig 1B,C). Another important global variable of the search is the number of fixations performed. The median number of fixations needed to find the target (before the last fixation to the target) was 6 (interquartile range: [3 13]) (Fig 1D). To assess the influence of individual differences in behavioral performance, we compared the variance in two dependent variables (Response Time over correct trials –RT_c- and probability of finding the target, P_T) across participants (N=17) and compared it with the ones obtained across trials (N=180). We found that both the variances in the mean RT_c and P_T across participants: (var(RT_c|part)=0.36, and var(P_T|part)=0.004) were smaller than the ones across trials: (var(RT_c|trials)=1.99, and var(P_T|trials)=0.012). To test the hypothesis of lower variability across participants (H1), we statistically compared the observed variance across participants with the ones obtained from a distribution of 100,000 surrogates, created by randomly taking subsamples of 17 images (same as the number of participants). We determined the p-value as the fraction of surrogates which showed a smaller variance than the original test statistic. The null hypothesis of equal variances were rejected with p-values: P(RT_c|part >(RT_c|trials)=0.00024 and P(P_T|part > P_T|trials)=0.01374.

3.2. Fixation event-related potentials in unrestricted visual search

The dynamics of fERP responses to targets and distractors, summarized in Fig 2A, were broadly consistent with what has been observed in previous studies using synthetic stimuli or prolonged fixations: A strong P1 at occipital electrodes, a central positivity at ~170ms which

resembles the Vertex Positive Potential (VPP) in response to faces [66], and a large target-specific P3 peaking after 300ms spread from central electrodes [35,40,46].

The VPP was confirmed by analyzing all fixations to distractors larger than 200 ms (N=10,466), changing the reference from linked mastoids to average, and observing that the central positive peak turns into a parieto-occipital negativity, i.e. the N170, with a slight lateralization to the right (Wilcoxon rank-sum test: PO7 vs. PO8 amplitudes in the [150 200] ms time window: $p<0.05$, zval = -2.00, Cohen's d (N = 34) = -0.34) [62,63] (Fig S3).

In order to control for differences in eye movements for the comparison between targets and distractors, we considered: First, only the fixations to distractors longer than 400 ms -where we analyzed only the first 400 ms-; and second, we matched the properties of the incoming saccade -the duration, and the horizontal and vertical amplitudes- (Fig S2). As expected, we found that the strongest difference appeared after 300ms and was more prominent in the centro-parietal electrodes but widely spread across most of the scalp (Fig 2B,C, cluster-based permutation test with $\alpha_{cluster}<0.01$ [67]). Interestingly, consistent significant differences emerge across several electrodes much earlier (around 170 ms), which suggestively matches the latency (Fig 2A, black bars; the limits of the early significant interval in the Cz electrode are [117, 207] ms) and scalp distribution of the VPP (Fig 2C). Further investigation of the waveforms for fixations that were longer than 400 ms showed no significant differences with shorter fixations (between 300 ms and 400 ms (Fig S4)).

3.3. Emergence of properties of the P3 component in free-viewing

In fixed-gaze oddball experiments, two major subcomponents of the P3, the P3a and the P3b are typically reported [68]. The P3a is elicited at the more frontal electrode sites and is larger for novel stimuli, whereas the P3b component is elicited in central and parietal regions, and is larger for the detection of a target. Importantly, a modulating property that positively affects the amplitude of the P3 is the ISI [11–13]. This has been related to the concepts of expectancy and surprise. As the participant is anticipating a target to be presented, a build-up of expectancy is sought to occur; this

effect is usually related to the P3b subcomponent of the P3 [35]. Conversely, the surprise effect is usually related to the P3a subcomponent of the P3 [13,35]. Therefore, a question raised as to whether fERPs in free-viewing visual search tasks show properties that relate to classic concepts of expectancy and surprise.

In the current design, there was no control of the ISI -i.e. the interval between fixations- as the participants can freely decide when and where to fixate next. However, the number of distractors presented before the target has also been used to study the properties of the P3 component [13]. In the current setup, the number of fixations made prior to finding the target could give an indication of the complexity of a scene and indicate how difficult it was to find the target. Hence, we can make a parallel with an oddball paradigm -a fixed-gaze sequential visual search- [13].

As a first step, we collapsed the trials into two categories based on the number of fixations needed to find the target: 'Short' (less than 7 fixations; N = 865) and 'Long' trials (more or equal than 7 fixations; N = 1030). Significant differences between these categories were found in frontal, central and parietal midline electrodes (Fz: $p<0.00001$, zval = 6.61, Cohen's d (N = 1895) = 0.15; Cz: $p<0.00001$, zval = 5.77, Cohen's d (N = 1895) = 0.13; Pz: $p<0.00001$, zval = 4.80, Cohen's d (N = 1895) = 0.11) but not in the occipital electrodes (Oz: $p=0.42$, zval = 1.25, Cohen's d (N = 1895) = 0.03) (Fig 3A).

The amplitude of the P3 component on the central electrode showed a continuous decrease as a function of the number of fixations to the target (Fig 3B). Moreover, there was a significant negative correlation of the single-trial P3 amplitudes with the number of fixations to the target (Fig 3C; Pearson's correlation coefficient R = -0.11, $p<0.00001$). This decrease is also in agreement with a gradual reduction of the surprise effect.

3.4. Global properties of fERPs

In the previous section we started the investigation with local properties of the fERPs, such as their relation with face processing and target-face detection, and ended with an analysis of more

global properties such as the dependence of the P3 with the length of the trial (where the length is the number of fixations performed) -and thus, the number of faces seen-. The dependence of the target detection on the number of previous fixations led us to investigate how the development of the whole trial is encoded in each distractor's fERP, before the target was detected, and how the target response was built-up.

In order to examine slow changes across the whole trial, we studied the relationship between fixation rank (the number of fixations up to the one being considered) and baseline activity -i.e. the activity before the saccade- when no baseline correction was applied. To investigate the encoding of each stimulus, we also examined the relationship between fixation rank and the amplitude of the P1 component with baseline correction. Fig 4A,B shows a strong dependence of the baseline activity on the fixation rank of the distractor. This dependence appears more prominent in the centro-frontal electrodes, although significant correlations were observed along midline channels (Pearson's correlation coefficient $R(Fz) = -0.95$, $p<0.000001$; $R(Cz) = -0.92$, $p<0.00001$; $R(Pz) = -0.90$, $p<0.00001$; $R(Oz) = -0.85$, $p<0.0001$). In order to rule out that these correlations reflected the time elapsed since the start of the trial (which correlates with the fixation rank), we compared the correlations of the baseline activity with time using different window sizes to group fixations along the search -between 100ms and 500ms-. The correlations with time were always smaller than the correlations with the fixation rank, failing in many cases to reach significance (Fig S5). This suggests that global changes in baseline activity are related to the sequence of fixations as processing units, and not with the mere time elapsed. Conversely, the encoding of each distractor stimulus did not change with the progression of the trial; as it can be inferred from the constant amplitude of the P100 as function of fixation rank (Fig 4C,D; $R(Oz) = -0.27$, $p=0.33$). Thus, there were global changes in baseline and target detection, but not in the encoding of an individual stimulus.

3.5. Global changes in frequency spectrum

Changes of the frequency power with the fixation rank of the distractors, measured as the correlation between these two variables, showed intervals of positive and negative correlations (Fig 5). These intervals, although not significant for all frequencies, fit nicely into the standard classification of EEG low frequency bands (delta, theta, alpha, and beta) for the different electrodes (Fig 5A). Significant correlations appeared in the four frequency bands at different electrodes (Fig 5A, top panel; $p < 0.01$, Bonferroni corrected for multiple comparisons for 240 tests, i.e. 60 frequency steps x 4 electrodes).

A large positive correlation between theta power and fixation rank, and a decrease in the (upper) alpha band appeared in the occipital electrode (Fig 5B, lower panel). This was also evident in the time-frequency plots (where time is represented in terms of fixation rank; Fig 5C, lower panel). These changes are usually observed together [65], and are sought to reflect increasing demands of visual attention. In contrast, the parietal, central and frontal electrodes exhibited a pronounced negative correlation in the alpha frequency band –i.e. ramping towards the presentation of the target–, but no significant correlation on the theta band (Fig 5B,C). This effect is slightly spread towards the beta band, and vanished beyond 23Hz; which is consistent with the broad frequency range that Donner and Siegel (2011) attribute to the more integrative cognitive processes. In particular, the negative correlation of the alpha frequency band is consistent with the findings of Klimesch and collaborators [24,65], who suggest that the alpha band must be negatively modulated by the expectancy of an upcoming stimulus.

4. Discussion

We developed a free-viewing paradigm that allowed us to investigate the brain correlates of unrestricted visual search with natural images. We showed that the local dynamics following each stimulus/foveation exhibits the fixed-gaze ERP components (P1, N170/VPP, P3). These brain events

are then signatures of the specific unitary operation performed at each stage of a mental program. A global analysis of the brain activity throughout the task allowed us to investigate signatures related to the accumulation of evidence during the task. In particular, we showed an increased demand of visual attention, changes in the expectancy of the stimulus and robust sequential effects from the analysis of baseline and the spectral profile. In what follows, we discuss our results in the context of local dynamics of unitary operations, global dynamics and mental programs, their interaction and the implications that our findings might have for future concurrent EEG and eye movement studies.

4.1. Local Dynamics: fERPs during Natural Tasks

Brain activity elicited after each fixation during visual search has been studied with both synthetic [40,43–45,47] and natural [46] stimuli. These studies have described a robust P3-like potential evoked by the object that the participant is looking for. In our experiment we observed a significant difference between the signals in response to targets and distractors, extending our previous finding using the same stimuli [46]. Interestingly, in the present study the difference between targets and distractors became significant about ~100ms earlier (~170ms vs. ~250ms). Two possible explanations could account for this difference: Firstly, both protocols differ in the speed-accuracy tradeoff that governed the search. In our previous experiment, participants were instructed -and briefly trained- to perform the search slowly; prioritizing the identification of the fixated face before moving on. In the current experiment, participants were only instructed to search for the targets, which led to shorter fixations. Nevertheless, we observed very few misses on the target (situations in which the participant fixated in the target but continued searching, as they failed to identify the target), as this occurred in (4.3 ± 0.6) % of the trials per subject -(7.7 ± 1.1) trials-. Thus, we can speculate that, in order to perform the search at this pace, participants must take an early decision on whether they have to continue moving their eyes or stop. The early detection that we observed in the fERPs could be the brain correlate of this process, although the full identification could occur later. Secondly, in the present study, as in Kamienkowski et al. (2012) but not in Kaunitz et al. (2014), we matched the

fixations to targets and distractors by the eye movement properties before performing statistical comparisons. This procedure could allow the detection of more subtle differences between these signals, which may have otherwise been indistinguishable in Kaunitz et al. (2014).

As in previous studies with faces [69,70], we observed the VPP / N170 potential in response to faces, which was confirmed by changing the reference from linked mastoids to average, which led to a shift from a central positive peak (VPP) to a parieto-occipital negativity (N170), with a slight lateralization to the right [69]. These results further support the validity of EEG and eye movement co-registration in the study of neural correlates of natural visual processing, and conversely extend the previous knowledge to more natural environments. However, an in-depth understanding of face processing in natural environments will require specific future experiments using faces and other objects embedded in the same natural images, to allow a comparison between faces and other stimuli.

Studies focusing on the early detection of different objects within rapid presentations of natural complex scenes -in which a cue is presented before the stimulus and rapid responses are required- have shown that identity stimulus information (contingent upon recognition) is able to be conveyed quite early, ranging from 150 to 300 ms [71–73]. In line with this, one recent study reported an early target related potential in free-viewing, but not fixed-gaze conditions [40], while another study found that it is possible to identify the target location even when fixating on the target for less than 10 ms in a free-viewing search, but not in a fixed-gaze condition [74]. The short timings for target identification can be related to extrafoveal detection. Target identification would trigger a saccade towards the target location, or less strictly, it would suggest that there is an ongoing computation of the probability to find the target that primes a given location [31]. This is consistent with an extrafoveal detection of the target in an easy visual search, and with an early foveal detection in a difficult visual search [75]. Since high attentional engagement is naturally present in free-viewing, from fixation onset or even earlier [76,77], another possible theoretical explanation for this early identification in free-viewing is the temporal modulation of attention [40,78,79]. Indeed, previous behavioral experiments have shown homogeneous performance of target detection within

almost the entire fixation but, strikingly, confidence judgments varied along the fixations. These results suggest that, while it is possible to detect a target accurately in the first tens of milliseconds of a fixation, the integration process continues along the whole fixation [79]. Moreover, previous studies have shown that each fixation is consistent with an attentional episode, where saccades accentuate the episodic borders [78], in agreement with a discrete structuring of attention; sampling information from temporal episodes during which several items can access the encoding process [80–82].

Previous EEG experiments have shown that temporal attention has an effect on the amplitude and latency of the N2 [83], the P3 component [84], and on the amplitude of the P1 component. Targets appearing at attended moments, close to a cue, evoke a larger P1 [83]. Furthermore, early components seemed to be affected only in highly demanding perceptual processing tasks [83]. These studies suggest that high attentional engagement could fasten the processing of the stimulus, and that change is measurable in the EEG signal as a modulation in early components. Following this idea, we can speculate that early target detection is more prominent in highly demanding perceptual processing tasks. We previously reported early target detection in tasks involving prolonged fixations when participants were asked to detect a subtle change in a synthetic stimulus [40], but not when they had to produce long fixations to find a hidden target face in a crowded scene [46]. The results presented in this study, with significant early target detection in free-viewing, might be associated to a higher attentional engagement occurring in a free-viewing visual search task with natural images. Additionally, they are in accordance with perceptual load theories, which state that earlier responses are expected on higher perceptual load tasks, subtle differences in crowded synthetic stimulus or faster searches in natural scenes; since they all involve higher attentional resources [85].

Furthermore, the early differences found in this natural visual search paradigm are in line with fast saccade response time (SRT) experiments, in which participants are asked to decide between two stimuli, and respond making a saccade towards the target as fast as possible. In these experiments, participants are able to respond in less than 150 ms and starting as early as around 100 ms to detect human faces and animals [86,87].

4.2. Global Dynamics: Monitoring the progress of the task

We found a significant decrease in the baseline amplitude in central-frontal regions, which followed the progression of the search. Similar gradual changes in the baseline were observed in other tasks in which participants required to accumulate some evidence in order to achieve a decision [5]. In our case, the amount of evidence is difficult to estimate in a direct way. Current visual search models based on eye movements tend to predict the sequence of fixations based on the construction of saliency maps on static images (i.e. disregarding the amount of evidence gained in each fixation) [88,89]. One interesting exception is the model of Najemnik and Geisler (2005), which estimates how the probability of finding the target at the end of the search changes in every fixation [31]. Although the fixation rank is naturally correlated with time, we found that the correlations between the baseline activity and time were smaller than those between the baseline activity and fixation rank. This suggests that global changes in baseline activity are indexing the sequence of processing steps within the task, rather than the time elapsed since the start of the trial.

To better understand the neural and cognitive processes involved in the slow dynamics throughout the trial, we explored the contribution of different frequency bands to the EEG signal. In particular, we calculated the correlation between the power in narrow bands and the fixation rank to build a spectral profile as proposed by Donner and Siegel (2011) [14]. Consistent with fixed-gaze studies, the broad decrease in alpha and increase in occipital theta activity could suggest a surge in the resources involved in visual attention [65]. For instance, an increase in theta activity in the occipital cortex was reported when viewing images that were later remembered [90]. A negative modulation in the alpha band has been previously related to the expectancy of an upcoming stimulus [24,65,91], which comes from an increasing probability of finding the target over time. This is consistent with our results, since we observed a negative correlation in the alpha frequency band -i.e. ramping towards the presentation of the target- mainly over the parietal, central and frontal electrodes. Interestingly, the negative correlation we observed spread beyond the alpha range, up to 23 Hz in the centro-parietal

sites. This broad frequency range is sometimes referred to as “low frequencies” (e.g., see [14]) and has been particularly linked to cognitive integrative processes.

4.3. Interactions between Local and Global Dynamics

Despite the fact that estimates of “surprise” and “expectancy” cannot be directly derived from our experiment, we can consider that shorter trials will likely be more associated with surprise. On the other hand, the more fixations a subject made within a trial, the more the target will be expected. Therefore, we hypothesized that trials that required very few fixations to find the target will have a larger surprise associated, while long trials will have a larger expectancy. Our results suggest a stronger effect of surprise over expectancy in the P3, since the difference in the amplitude favored shorter trials. Moreover, this result was supported by the more frontal distribution of shorter versus longer trials, consistent with the observations of Polich (2007) [35]. These late changes were observed despite the fact that the early feedforward visual processing of each stimulus, tagged by the P1 component, was preserved throughout the search. Together with the fact that low frequencies, but not high frequencies, correlated with fixation rank, this suggests that only the integrative cognitive processes are related to the progression of the search.

An alternative explanation for the influence of ISI on P3 in fixed-gaze experiments is the Recycling Cycle Hypothesis [11,13], which suggests that the P3 component is reduced with short ISIs because the system requires time to recover from recent ERP production. In our experiment, this would mean that the P3 is reduced in later fixations only because the underlying generation processes lack the resources to fully recover within a multiple-fixation-trial. However, this alternative explanation is unlikely for three reasons: First, participants only saw one target, which implies that the resources related to target processing should be fully available for the whole trial. Second, we have previously shown that the information regarding the salient target in an attentional blink task recovers rapidly in successive fixations, since saccades accentuate episodic boundaries of the temporal episodes [78]. It follows that the alternative explanation would be unlikely even if more than one

target was present. Third, resources related to distractor processing should already vanish before 300 ms -according to mainstream theories such as the Global Workspace Theory, which proposes early processing of a new stimulus could start in parallel [92–94]-. This is also in agreement with the episodic theory of attentional deployment, which proposes that information is sampled from temporal episodes of about 400 ms [81,82].

4.4. Towards an integrative framework in natural viewing

From an algorithmic perspective –in line with Marr's levels at which the brain as an information-processing device must be understood [95]– the schematic diagram shown in Fig 6 can be used to put our findings into a more general framework. At the bottom level (visual encoding in Fig 6), every fixation produces the typical sequences of visual ERPs towards a face, including the P1 and N170/VPP, irrespective of the fixation rank. These fERPs share the topographies and latencies of their fixed-gaze counterparts (see also [40,46]), thus the initial feedforward processing of the stimulus in each fixation seems to be conserved and is independent of the task. The eye movement control level (motor prep./exec in Fig 6) can be separated into the parallel processing of the "when" and "where" streams plus an inhibition, as proposed by the well-established model of saccade generation of Findlay and Walker (1999) [96]. Some existent models already include these ideas, for instance, in the CRISP model [97], the "when" stream is constantly generating saccades that could be inhibited, while the "where" is implemented as a two-stage saccade programming composed by an initial, labile stage that is subject to cancellation (inhibition of the "when" stream), followed by a non-labile stage [97,98]. Thus, there is a continuous crosstalk between saccade preparation and visual-cognitive processing, which allows difficulties in online visual and cognitive processing to immediately inhibit (i.e. delay) saccade initiation, leading to longer fixation durations. The early target detection potential that we observed could be interpreted as a cancellation signal during the first labile stage of the saccade preparation, where the next saccade is canceled at early latencies leading to longer fixations

[97]. Successful early target detection is followed by a global P3 component, reflected in higher and later activity, which also resembles the P3 recorded in fixed gaze conditions [35,40,46].

Interestingly, the preparation of the upcoming saccade could be accompanied with a modulation of cover attention towards the potential landing site. These processes have previously been studied with crowded scenes in fixed-gaze reaction time experiments [32], and in saccade reaction time experiments, which allowed studying the interaction between covert attention and saccade execution [33,34]. However, it is important to distinguish these processes from the overt allocation of attention that follows the eye movement. The interplay between covert and overt allocation of attention is one of the fundamental open issues for which the co-registration of brain activity and eye movements could largely contribute in future studies.

At the top level (global integration in Fig 6), the integration of the information about the whole image acquired so far could involve, for instance, computing the target position probabilities based on existent evidence, which necessarily includes information about the surface already explored as well as past events and the current position [31,58,99]. These integrative processes also serve to build a general model of the visual scene, and to guide future eye movements, as another crosstalk between the saccade preparation and visual-cognitive processing. Although it could run on top of visual encoding and saccade preparation processes, it should interchange information with them. Hence, it should be in pace with the fixations. This accumulated evidence is indexed by changes in baseline, which scales with the fixation rank better than with time itself, and is further reflected by changes in alpha/theta band oscillatory activity and modulations in the amplitude of the P3 component.

It is well established that human behavior in complex tasks is strongly shaped by individual differences in certain capacities and strategies. For instance, measures of visual working memory (VWM) capacity are strongly related to performance in both visual search without eye movements [100,101] and constrained eye movement paradigms [102]. Also, VWM capacity correlates with slow potentials such as the contralateral delay activity (CDA) [103] that have a similar behavior as the

baseline changes during search. These estimates of VWM capacity are potentially appealing because they have been proven to be very robust within a single participant (for instance, they could yield a test-retest reliability of 0,77 after 1.5 years later) and consistent predictors of many complex visual tasks [101]. By comparing the variance in two behavioral dependent variables, we showed that the major source of variability in this experiment were the trials and not the participants. However, future work on both complex visual search models should include VWM or other performance measures, which could influence markers of cognition and decision-making, as co-variables.

4.5. Alternative theoretical frameworks

The main focus of the present paper is related to the deployment of overt visual attention rather than covert visual attention, since eye movements are an essential part of behaviour. Although these processes could share some mechanisms, these are not necessarily the same [104]. Some of these shared mechanisms could be part of the global processes of the task. For instance, the Competitive Guided Search model [105] -which is an updated version of the Guided Search model [32,106]-, has included a module that evaluates whether to quit the search or continue. This "quit unit" establishes an accumulation process that runs in parallel with the actual search, and competes to terminate the task. However, since in our task participants were not able to finish the trial before finding the target or reaching the 20 secs maximum trial period and there was not a forced-choice for target present/absent, response times are not as relevant as in other experiments so it is unlikely that a quit unit would play a major role. Interestingly, the integrative framework we present here could potentially combine more than one integrative processes in the global workspace, leaving space for contribution from other theoretical models.

4.6. Concluding remarks

Our findings contribute to the study of human physiology and cognition in natural environments in three fundamental ways: First, it presents robust fERPs in unconstrained visual search with real-world images. This extends previous fERP work but also highlights potential

differences, as early target detection was only previously found with synthetic stimuli, giving a more complete understanding of how the brain processes unitary operations in natural viewing. Second, by focusing on brain oscillations at different frequencies and modulation of evoked components throughout the trial, we were able to confirm that specific brain signatures, previously identified for fixed-gaze paradigms as indexes of accumulation of evidence, can be robustly obtained in natural viewing. Finally, it allows us to introduce a data-driven framework to link oscillations with the underlying mental program.

The schematic diagram that we have considered here is, of course, a simplification of all the complexities involved in natural viewing everyday scenes. However, we believe it can be a starting point for considering general neurophysiology data-driven frameworks to elucidate how complex mental operations are performed in natural viewing.

5. Acknowledgments

We are grateful to Dr Lisandro Kaunitz for invaluable assistance in the creation of the stimuli and data acquisition and Dr Markus Bauer for helpful comments on the manuscript. Part of this project was funded by the EPSRC (EP/1016899/1).

6. Author contributions

J.E.K., and M.J.I. designed the experiments with contributions from M.S. A.V. conducted the experiments. J.E.K. and M.J.I. developed the analysis approach. J.E.K. and A.V. analyzed the data with contributions from M.J.I. J.E.K., and M.J.I. wrote the paper with input from all authors.

7. Additional information

Supplementary information is available for this paper.

8. Competing interests

The authors declare no competing interests.

9. References

1. Ullman S. Visual routines. Cognition. 1984;18: 97–159. doi:10.1016/0010-0277(84)90023-4

2. Fujii N, Graybiel AM. Time-varying covariance of neural activities recorded in striatum and frontal cortex as monkeys perform sequential-saccade tasks. Proc Natl Acad Sci U S A. 2005;102: 9032–9037. doi:10.1073/pnas.0503541102

3. Roelfsema PR. Elemental operations in vision. Trends Cogn Sci. 2005;9: 226–233. doi:10.1016/j.tics.2005.03.012

4. Jubault T, Ody C, Koechlin E. Serial Organization of Human Behavior in the Inferior Parietal Cortex. J Neurosci. 2007;27: 11028–11036. doi:10.1523/JNEUROSCI.1986-07.2007

5. Lange FP de, Jensen O, Dehaene S. Accumulation of Evidence during Sequential Decision Making: The Importance of Top–Down Factors. J Neurosci. 2010;30: 731–738. doi:10.1523/JNEUROSCI.4080-09.2010

6. Zylberberg A, Dehaene S, Roelfsema PR, Sigman M. The human Turing machine: a neural framework for mental programs. Trends Cogn Sci. 2011;15: 293–300. doi:10.1016/j.tics.2011.05.007

7. Zylberberg A, Paz L, Roelfsema PR, Dehaene S, Sigman M. A neuronal device for the control of multi-step computations. Pap Phys. 2013;5: 1–15. doi:10.4279/PIP.050006

8. Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. Nature. 1999;400: 233–238. doi:10.1038/22268

9. Yang T, Shadlen MN. Probabilistic reasoning by neurons. Nature. 2007;447: 1075–1080. doi:10.1038/nature05852

10. Moro SI, Tolboom M, Khayat PS, Roelfsema PR. Neuronal Activity in the Visual Cortex Reveals the Temporal Order of Cognitive Operations. J Neurosci. 2010;30: 16293–16303. doi:10.1523/JNEUROSCI.1256-10.2010

11. Fitzgerald PG, Picton TW. Temporal and sequential probability in evoked potential studies. Can J Psychol. 1981;35: 188–200.

12. Woods DL, Courchesne E, Hillyard SA, Galambos R. Recovery cycles of event-related potentials in multiple detection tasks. Clin Neurophysiol. 1980;50: 335–347. doi:10.1016/0013-4694(80)90001-2

13. Gonsalvez CJ, Polich J. P300 amplitude is determined by target-to-target interval. Psychophysiology. 2002;39: 388–396. doi:10.1017/S0048577201393137

14. Donner TH, Siegel M. A framework for local cortical oscillation patterns. Trends Cogn Sci. 2011;15: 191–199. doi:10.1016/j.tics.2011.03.007

15. Siegel M, Donner TH, Oostenveld R, Fries P, Engel AK. Neuronal Synchronization along the Dorsal Visual Pathway Reflects the Focus of Spatial Attention. Neuron. 2008;60: 709–719. doi:10.1016/j.neuron.2008.09.010

16. Donner TH, Siegel M, Oostenveld R, Fries P, Bauer M, Engel AK. Population Activity in the Human Dorsal Pathway Predicts the Accuracy of Visual Motion Detection. J Neurophysiol. 2007;98: 345–359. doi:10.1152/jn.01141.2006

17. Gross J, Schmitz F, Schnitzler I, Kessler K, Shapiro K, Hommel B, et al. Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. Proc Natl Acad Sci U S A. 2004;101: 13050–13055. doi:10.1073/pnas.0404944101

18. Hipp JF, Engel AK, Siegel M. Oscillatory Synchronization in Large-Scale Cortical Networks Predicts Perception. Neuron. 2011;69: 387–396. doi:10.1016/j.neuron.2010.12.027

19. Pesaran B, Nelson MJ, Andersen RA. Free choice activates a decision circuit between frontal and parietal cortex. Nature. 2008;453: 406–409. doi:10.1038/nature06849

20. Foxe JJ, Simpson GV, Ahlfors SP. Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. Neuroreport. 1998;9: 3929–3933.

21. Jokisch D, Jensen O. Modulation of Gamma and Alpha Activity during a Working Memory Task Engaging the Dorsal or Ventral Stream. J Neurosci. 2007;27: 3244–3251. doi:10.1523/JNEUROSCI.5399-06.2007

22. Thut G, Nietzel A, Brandt SA, Pascual-Leone A. α-Band Electroencephalographic Activity over Occipital Cortex Indexes Visuospatial Attention Bias and Predicts Visual Target Detection. J Neurosci. 2006;26: 9494–9502. doi:10.1523/JNEUROSCI.0875-06.2006

23. Worden MS, Foxe JJ, Wang N, Simpson GV. Anticipatory Biasing of Visuospatial Attention Indexed by Retinotopically Specific α-Bank Electroencephalography Increases over Occipital Cortex. J Neurosci. 2000;20: RC63–RC63.

24. Klimesch W, Doppelmayr M, Russegger H, Pachinger T, Schwaiger J. Induced alpha band power changes in the human EEG and attention. Neurosci Lett. 1998;244: 73–76. doi:10.1016/S0304-3940(98)00122-0

25. Bauer M, Stenner M-P, Friston KJ, Dolan RJ. Attentional Modulation of Alpha/Beta and Gamma Oscillations Reflect Functionally Distinct Processes. J Neurosci. 2014;34: 16117–16125. doi:10.1523/JNEUROSCI.3474-13.2014

26. Deubel H, Schneider WX. Saccade target selection and object recognition: Evidence for a common attentional mechanism. Vision Res. 1996;36: 1827–1837. doi:10.1016/0042-6989(95)00294-4

27. Findlay JM, Brown V, Gilchrist ID. Saccade target selection in visual search: the effect of information from the previous fixation. Vision Res. 2001;41: 87–95.

28. Findlay JM. Saccade Target Selection During Visual Search. Vision Res. 1997;37: 617–631. doi:10.1016/S0042-6989(96)00218-0

29. Henderson JM. Human gaze control during real-world scene perception. Trends Cogn Sci. 2003;7: 498–504.

30. Ludwig CJH, Gilchrist ID. Stimulus-driven and goal-driven control over visual selection. J Exp Psychol Hum Percept Perform. 2002;28: 902–912. doi:10.1037/0096-1523.28.4.902

31. Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. Nature. 2005;434: 387–391. doi:10.1038/nature03390

32. Wolfe JM. Guided search 4.0. Integrated Models of Cognitive Systems. Oxford University Press; 2007.

33. Ehinger BV, König P, Ossandón JP. Predictions of Visual Content across Eye Movements and Their Modulation by Inferred Information. J Neurosci. 2015;35: 7403–7413. doi:10.1523/JNEUROSCI.5114-14.2015

34. Weaver Matthew D., Hickey Clayton, van Zoest Wieske. The impact of salience and visual working memory on the monitoring and control of saccadic behavior: An eye- tracking and EEG study. Psychophysiology. 2017;54: 544–554. doi:10.1111/psyp.12817

35. Polich J. Updating P300: An integrative theory of P3a and P3b. Clin Neurophysiol. 2007;118: 2128–2148. doi:10.1016/j.clinph.2007.04.019

36. Mack SC, Eckstein MP. Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. J Vis. 2011;11: 9–9. doi:10.1167/11.9.9

37. Otero-Millan J, Troncoso XG, Macknik SL, Serrano-Pedraza I, Martinez-Conde S. Saccades and microsaccades during visual fixation, exploration, and search: Foundations for a common saccadic generator. J Vis. 2008;8: 21–21. doi:10.1167/8.14.21

38. Dimigen O, Valsecchi M, Sommer W, Kliegl R. Human Microsaccade-Related Visual Brain Responses. J Neurosci. 2009;29: 12321–12331. doi:10.1523/JNEUROSCI.0911-09.2009

39. Dimigen O, Sommer W, Hohlfeld A, Jacobs AM, Kliegl R. Coregistration of eye movements and EEG in natural reading: Analyses and review. Journal of Experimental Psychology: General. American Psychological Association; 2011. p. 552. doi:10.1037/a0023885

40. Kamienkowski JE, Ison MJ, Quiroga RQ, Sigman M. Fixation-related potentials in visual search: A combined EEG and eye tracking studyFixation-related potentials in visual search. J Vis. 2012;12: 4–4. doi:10.1167/12.7.4

41. Ossandón JP, Helo AV, Montefusco-Siegmund R, Maldonado PE. Superposition Model Predicts EEG Occipital Activity during Free Viewing of Natural Scenes. J Neurosci. 2010;30: 4787–4795. doi:10.1523/JNEUROSCI.5769-09.2010

42. Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. Transient Induced Gamma-Band Response in EEG as a Manifestation of Miniature Saccades. Neuron. 2008;58: 429–441. doi:10.1016/j.neuron.2008.03.027

43. Brouwer A-M, Reuderink B, Vincent J, Gerven MAJ van, Erp JBF van. Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. J Vis. 2013;13: 17–17. doi:10.1167/13.3.17

44. Körner C, Braunstein V, Stangl M, Schlögl A, Neuper C, Ischebeck A. Sequential effects in continued visual search: Using fixation-related potentials to compare distractor processing before and after target detection. Psychophysiology. 2014;51: 385–395. doi:10.1111/psyp.12062

45. Ušćumlić M, Blankertz B. Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. J Neural Eng. 2016;13: 016015. doi:10.1088/1741-2560/13/1/016015

46. Kaunitz LN, Kamienkowski JE, Varatharajah A, Sigman M, Quiroga RQ, Ison MJ. Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task. NeuroImage. 2014;89: 297–305. doi:10.1016/j.neuroimage.2013.12.006

47. Wenzel MA, Golenia J-E, Blankertz B. Classification of Eye Fixation Related Potentials for Variable Stimulus Saliency. Front Neurosci. 2016;10. doi:10.3389/fnins.2016.00023

48. Léger P-M, Sénecal S, Courtemanche F, Guinea AO de, Titah R, Fredette M, et al. Precision is in the Eye of the Beholder: Application of Eye Fixation-Related Potentials to Information Systems Research. J Assoc Inf Syst. 2014;15. Available: http://aisel.aisnet.org/jais/vol15/iss10/3

49. Savage SW, Potter DD, Tatler BW. Does preoccupation impair hazard perception? A simultaneous EEG and Eye Tracking study. Transp Res Part F Traffic Psychol Behav. 2013;17: 52–62. doi:10.1016/j.trf.2012.10.002

50. Simola J, Torniainen J, Moisala M, Kivikangas M, Krause CM. Eye movement related brain responses to emotional scenes during free viewing. Front Syst Neurosci. 2013;7. doi:10.3389/fnsys.2013.00041

51. Simola J, Le Fevre K, Torniainen J, Baccino T. Affective processing in natural scene viewing: Valence and arousal interactions in eye-fixation-related potentials. NeuroImage. 2015;106: 21–33. doi:10.1016/j.neuroimage.2014.11.030

52. Staudigl T, Hartl E, Noachtar S, Doeller CF, Jensen O. Saccades are phase-locked to alpha oscillations in the occipital and medial temporal lobe during successful memory encoding. PLOS Biol. 2017;15: e2003404. doi:10.1371/journal.pbio.2003404

53. Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997;10: 433–436. doi:10.1163/156856897X00357

54. Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis. 1997;10: 437–442. doi:10.1163/156856897X00366

55. Engbert R, Kliegl R. Microsaccades uncover the orientation of covert attention. Vision Res. 2003;43: 1035–1045. doi:10.1016/S0042-6989(03)00084-1

56. Mergenthaler K, Engbert R. Modeling the Control of Fixational Eye Movements with Neurophysiological Delays. Phys Rev Lett. 2007;98: 138104. doi:10.1103/PhysRevLett.98.138104

57. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 2004;134: 9–21. doi:10.1016/j.jneumeth.2003.10.009

58. Kliegl R, Dambacher M, Dimigen O, Sommer W. Oculomotor Control, Brain Potentials, and Timelines of Word Recognition During Natural Reading. In: Horsley M, Eliot M, Knight BA, Reilly R, editors. Current Trends in Eye Tracking Research. Springer International Publishing; 2014. pp. 141–155. Available: http://link.springer.com/chapter/10.1007/978-3-319-02868-2_10

59. Ahmadi M, Quian Quiroga R. Automatic denoising of single-trial evoked potentials. NeuroImage. 2013;66: 672–680. doi:10.1016/j.neuroimage.2012.10.062

60. Navajas J, Ahmadi M, Quiroga RQ. Uncovering the Mechanisms of Conscious Face Perception: A Single-Trial Study of the N170 Responses. J Neurosci. 2013;33: 1337–1343. doi:10.1523/JNEUROSCI.1226-12.2013

61. Quian Quiroga R. Obtaining single stimulus evoked potentials with wavelet denoising. Phys Nonlinear Phenom. 2000;145: 278–292. doi:10.1016/S0167-2789(00)00116-0

62. Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology. 2011;48: 229–240.

63. Bigdely-Shamlo N, Kreutz-Delgado K, Kothe C, Makeig S. EyeCatch: Data-mining over half a million EEG independent components to construct a fully-automated eye-component detector. Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. IEEE; 2013. pp. 5845–5848. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6610881

64. Oostenveld R, Fries P, Maris E, Schoffelen J-M. FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. Intell Neurosci. 2011;2011: 1:1–1:9. doi:10.1155/2011/156869

65. Klimesch W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. Brain Res Rev. 1999;29: 169–195. doi:10.1016/S0165-0173(98)00056-3

66. Jeffreys DA. A face-responsive potential recorded from the human scalp. Exp Brain Res. 1989;78: 193–202. doi:10.1007/BF00230699

67. Maris E, Oostenveld R. Nonparametric statistical testing of EEG- and MEG-data. J Neurosci Methods. 2007;164: 177–190. doi:10.1016/j.jneumeth.2007.03.024

68. Squires KC, Donchin E, Herning RI, McCarthy G. On the influence of task relevance and stimulus probability on event-related-potential components. Electroencephalogr Clin Neurophysiol. 1977;42: 1–14. doi:10.1016/0013-4694(77)90146-8

69. Joyce C, Rossion B. The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. Clin Neurophysiol. 2005;116: 2613–2631. doi:10.1016/j.clinph.2005.07.005

70. Rossion B, Jacques C. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. NeuroImage. 2008;39: 1959–1979. doi:10.1016/j.neuroimage.2007.10.011

71. Carmel D, Bentin S. Domain specificity versus expertise: factors influencing distinct processing of faces. Cognition. 2002;83: 1–29. doi:10.1016/S0010-0277(01)00162-7

72. Johnson JS, Olshausen BA. Timecourse of neural signatures of object recognition. J Vis. 2003;3: 4–4. doi:10.1167/3.7.4

73. VanRullen R, Thorpe SJ. The Time Course of Visual Processing: From Early Perception to Decision-Making. J Cogn Neurosci. 2001;13: 454–461. doi:10.1162/08989290152001880

74. Kotowicz A, Rutishauser U, Koch C. Time course of target recognition in visual search. Front Hum Neurosci. 2010;4. doi:10.3389/fnhum.2010.00031

75. Dias JC, Sajda P, Dmochowski JP, Parra LC. EEG precursors of detected and missed targets during free-viewing search. J Vis. 2013;13: 13–13. doi:10.1167/13.13.13

76. Corbetta M, Akbudak E, Conturo TE, Snyder AZ, Ollinger JM, Drury HA, et al. A Common Network of Functional Areas for Attention and Eye Movements. Neuron. 1998;21: 761–773. doi:10.1016/S0896-6273(00)80593-0

77. Melcher D, Colby CL. Trans-saccadic perception. Trends Cogn Sci. 2008;12: 466–473. doi:10.1016/j.tics.2008.09.003

78. Kamienkowski JE, Navajas J, Sigman M. Eye movements blink the attentional blink. J Exp Psychol Hum Percept Perform. 2012;38: 555–560. doi:10.1037/a0027729

79. Navajas J, Sigman M, Kamienkowski JE. Dynamics of visibility, confidence, and choice during eye movements. J Exp Psychol Hum Percept Perform. 2014;40: 1213–1227. doi:10.1037/a0036321

80. VanRullen R. Perceptual Cycles. Trends Cogn Sci. 2016;20: 723–735. doi:10.1016/j.tics.2016.07.006

81. Wyble B, Potter MC, Bowman H, Nieuwenstein M. Attentional episodes in visual perception. J Exp Psychol Gen. 2011;140: 488–505. doi:10.1037/a0023612

82. Wyble B, Bowman H, Nieuwenstein M. The attentional blink provides episodic distinctiveness: Sparing at a cost. J Exp Psychol Hum Percept Perform. 2009;35: 787–807. doi:10.1037/a0013902

83. Correa Á, Lupiáñez J, Madrid E, Tudela P. Temporal attention enhances early visual processing: A review and new evidence from event-related potentials. Brain Res. 2006;1076: 116–128. doi:10.1016/j.brainres.2005.11.074

84. Miniussi C, Wilding, L E, Coull, T J, Nobre, et al. Orienting attention in timeModulation of brain potentials. Brain. 1999;122: 1507–1518. doi:10.1093/brain/122.8.1507

85. Lavie N, Beck DM, Konstantinou N. Blinded by the load: attention, awareness and the role of perceptual load. Phil Trans R Soc B. 2014;369: 20130205. doi:10.1098/rstb.2013.0205

86. Crouzet SM, Kirchner H, Thorpe SJ. Fast saccades toward faces: Face detection in just 100 ms. J Vis. 2010;10: 16–16. doi:10.1167/10.4.16

87. Kirchner H, Thorpe SJ. Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. Vision Res. 2006;46: 1762–1776. doi:10.1016/j.visres.2005.10.002

88. Borji A, Itti L. State-of-the-Art in Visual Attention Modeling. IEEE Trans Pattern Anal Mach Intell. 2013;35: 185–207. doi:10.1109/TPAMI.2012.89

89. Torralba A, Oliva A, Castelhano MS, Henderson JM. Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. Psychol Rev. 2006;113: 766–786. doi:10.1037/0033-295X.113.4.766

90. Osipova D, Takashima A, Oostenveld R, Fernández G, Maris E, Jensen O. Theta and Gamma Oscillations Predict Encoding and Retrieval of Declarative Memory. J Neurosci. 2006;26: 7523–7531. doi:10.1523/JNEUROSCI.1948-06.2006

91. Freunberger R, Klimesch W, Griesmayr B, Sauseng P, Gruber W. Alpha phase coupling reflects object recognition. NeuroImage. 2008;42: 928–935. doi:10.1016/j.neuroimage.2008.05.020

92. Dehaene S, Naccache L. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. Cognition. 2001;79: 1–37. doi:10.1016/S0010-0277(00)00123-2

93. Dehaene S, Charles L, King J-R, Marti S. Toward a computational theory of conscious processing. Curr Opin Neurobiol. 2014;25: 76–84. doi:10.1016/j.conb.2013.12.005

94. Sigman M, Dehaene S. Brain Mechanisms of Serial and Parallel Processing during Dual-Task Performance. J Neurosci. 2008;28: 7585–7598. doi:10.1523/JNEUROSCI.0948-08.2008

95. Marr D. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Second. MIT Press; 2010.

96. Findlay JM, Walker R. A model of saccade generation based on parallel processing and competitive inhibition. Behav Brain Sci. 1999;22: 661–674; discussion 674–721.

97. Nuthmann A, Smith TJ, Engbert R, Henderson JM. CRISP: A Computational Model of Fixation Durations in Scene Viewing. Psychol Rev. 2010;117: 382–405. doi:10.1037/a0018924

98. Reichle ED, Pollatsek A, Fisher DL, Rayner K. Toward a model of eye movement control in reading. Psychol Rev. 1998;105: 125–157.

99. Koechlin E, Summerfield C. An information theoretical approach to prefrontal executive function. Trends Cogn Sci. 2007;11: 229–235. doi:10.1016/j.tics.2007.04.005

100. Luria, R. and Vogel, E.K. (2011) Visual search demands dictate reliance on working memory storage. J. Neurosci. 31, 6199–6207

101. Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. Trends in cognitive sciences, 17(8), 391-400.

102. Theeuwes, J., Belopolsky, A., & Olivers, C. N. (2009). Interactions between working memory, attention and eye movements. Acta psychologica, 132(2), 106-114.

103. Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. Nature, 438(7067), 500.

104. Hulleman J, Olivers CNL. On the brink: The demise of the item in visual search moves closer. Behav Brain Sci. 2017;40. doi:10.1017/S0140525X16000364

105. Moran R, Zehetleitner M, Müller HJ, Usher M. Competitive guided search: Meeting the challenge of benchmark RT distributions. J Vis. 2013;13: 24–24. doi:10.1167/13.8.24

106. Wolfe JM, Palmer EM, Horowitz TS. Reaction time distributions constrain models of visual search. Vision Res. 2010;50: 1304–1311. doi:10.1016/j.visres.2009.11.002

Figure Captions

Fig 1. Experimental Design. **(A)** The target face (inset) is presented for 3 seconds. Then, a fixation dot appears in a random position for 1 second. The search starts when the fixation dot disappears and the crowds image is presented. Once the subject finds the target face; they fixate on it for 1 second to end the trial. The crowd image shown here is the same as the one shown in panel B. **(B)** Search example: The scan path (black line) and fixations (color dots) are superimposed to the crowd image and the target (highlighted with a brown square). The color of the dots represents the fixation rank of the distractors, -i.e. the fixation number- and the diameter size represents the duration of the fixation. **(C)** Eye traces of the same trial for both horizontal (black) and vertical (gray) positions of the right eye. The vertical red line shows the onset of the fixation to the target. **(D)** Distribution of the number of fixations preceding the target; these are all the fixations preceding the target regardless of their position. **(E)** Distribution of fixation durations made to distractors. **(F)** Distribution of distractor saccade amplitudes.

Fig 2. fERPs for matched eye movement properties across midline channels. (A) Fixations onset at 0ms and the fERPs were baseline corrected [-200 -100]. Channels Fz, Cz, Pz and Oz are shown for

the target (red) and distractor (blue) conditions. The horizontal and vertical saccade amplitude, as well as the saccade duration preceding the fixation of interest; were matched between targets (N=1895) and distractors (N=1895) for fixations longer than 400 ms. The p-values for the significant differences ($p<0.01$) are represented with black bars at the top of each channel plot, and were estimated using a cluster-based permutation test from the FieldTrip toolbox [67]. The boundaries of the significant intervals observed in the panel A insets are: Fz: {[121, 164] ms, [254 400] ms}, Cz: {[117, 207] ms, [238, 400] ms}, Pz: {[148, 215] ms, [230, 400] ms}, Oz: {[184, 203] ms, [281, 297] ms, [309, 324] ms, [332, 348] ms, [363, 400] ms}. **(B)** Significant channels within 30ms time windows centered in {0 ms, 50 ms, 100 ms, 170 ms, 240 ms, 310 ms, 380 ms}, after the cluster-based permutation test. **(C)** Scalp topographies of the difference wave (Targets – Distractors). Each topography corresponds to the average activity of the same intervals as panel (B).

Fig 3. Global effects on target detection. (A) Midline Target P3 Amplitude as a function of Short and Long trials. The amplitude was calculated as the mean average in the [250 400] ms windows extracted from the individual target fERPs, averaged across all trials and subjects. The gray bars represent the Short trials (<7 fixations in a trial, N=865) and the black bars represent the Long trials (≥7 fixations, N=1030). Bonferroni-corrected rank-sum tests showed significant differences for Fz, Cz, and Pz ($p<0.00001$) but not for Oz ($p = 0.4210$). n.s.: p>0,05, and ***: p<0,001. **(B)** fERPs to targets as function of time, for different fixation ranks. **(C)** Denoised single trial P3 amplitudes vs. fixation rank at Cz electrode. Amplitude of the P3 distribution (top left) and fixation rank (botton right) histograms. Two-dimensional histogram of those variables (top right); each bin of the histogram contains single trial amplitudes of the P3. Amplitudes of the P3 and fixation ranks were negatively correlated (Pearson's $R = -0.11$, $p<0.00001$).

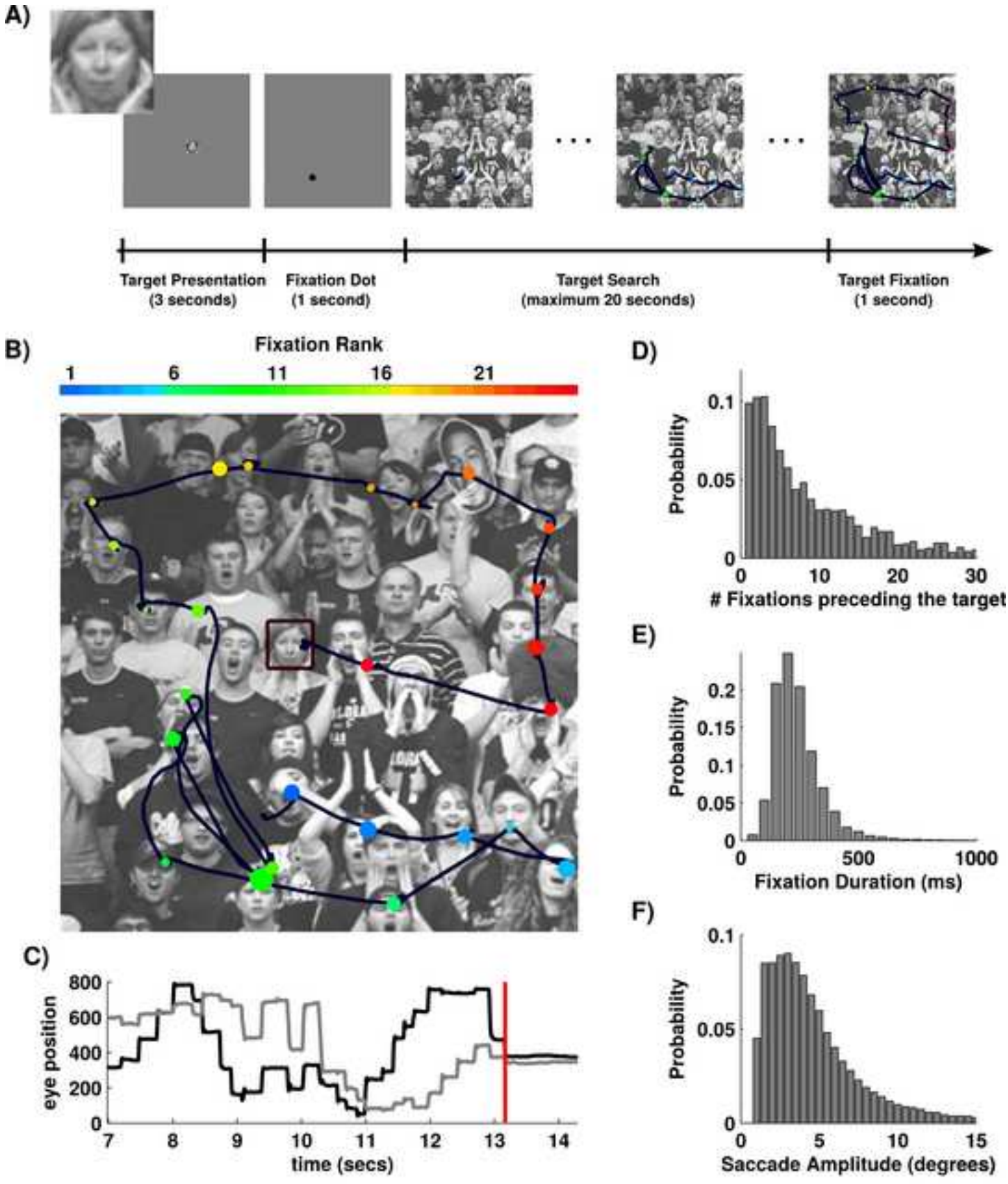
Fig 4. Global effects on baseline and visual processing. (A) Midline fERPs to distractors as function of the fixation rank. fERPs were epoched between [-500 500]ms and no baseline correction were applied in order to observe its slow progression along the trial. From top to bottom: Fz, Cz, Pz, and Oz. The gray shadow in the top panel shows the [-500 -100] ms period analyzed in panel B). Vertical solid lines represent the onset of each fixation, and vertical dashed lines represent the boundaries of each epoch. **(B)** Average amplitude in the [-500 -100] ms period for each electrode. Error bars represent the ±s.e.m. interval. **(C)** Early fERPs to distractors in the electrode Oz as a function of fixation rank. Baseline correction between [-200 -100] ms were applied. The gray shadow shows the [75 125] ms period analyzed in panel D). The vertical solid lines represent the onset of each fixation. **(D)** Average amplitude in the [75 125] ms period for each electrode. Error bars represent the ±s.e.m. interval.

Fig 5. Spectral profile. (A) Correlation between the frequency power during each fixation and the fixation rank as a function of frequency, for each electrode (Fz, Cz, Pz, and Oz). The upper panel shows the Bonferroni-corrected significance of the correlations (considering 240 comparisons, i.e. 60 frequency points and 4 electrodes) and thresholded by 0.01. Red and blue shadows highlight the theta ([3.5 8.0] Hz) and alpha bands ([8.0 13.5] Hz). **(B)** Power in theta and alpha frequency bands as function of fixation rank. Error bars represent the ±s.e.m. interval, and the straight line the general trend. **(C)** Time-Frequency image plots for each electrode, where the timecourse of the exploration is indexed by the fixation rank.

Fig 6. Integrative framework of natural visual search. Schematic framework of the identified electrophysiological signatures and their interactions with key processes in a free-viewing visual search task.

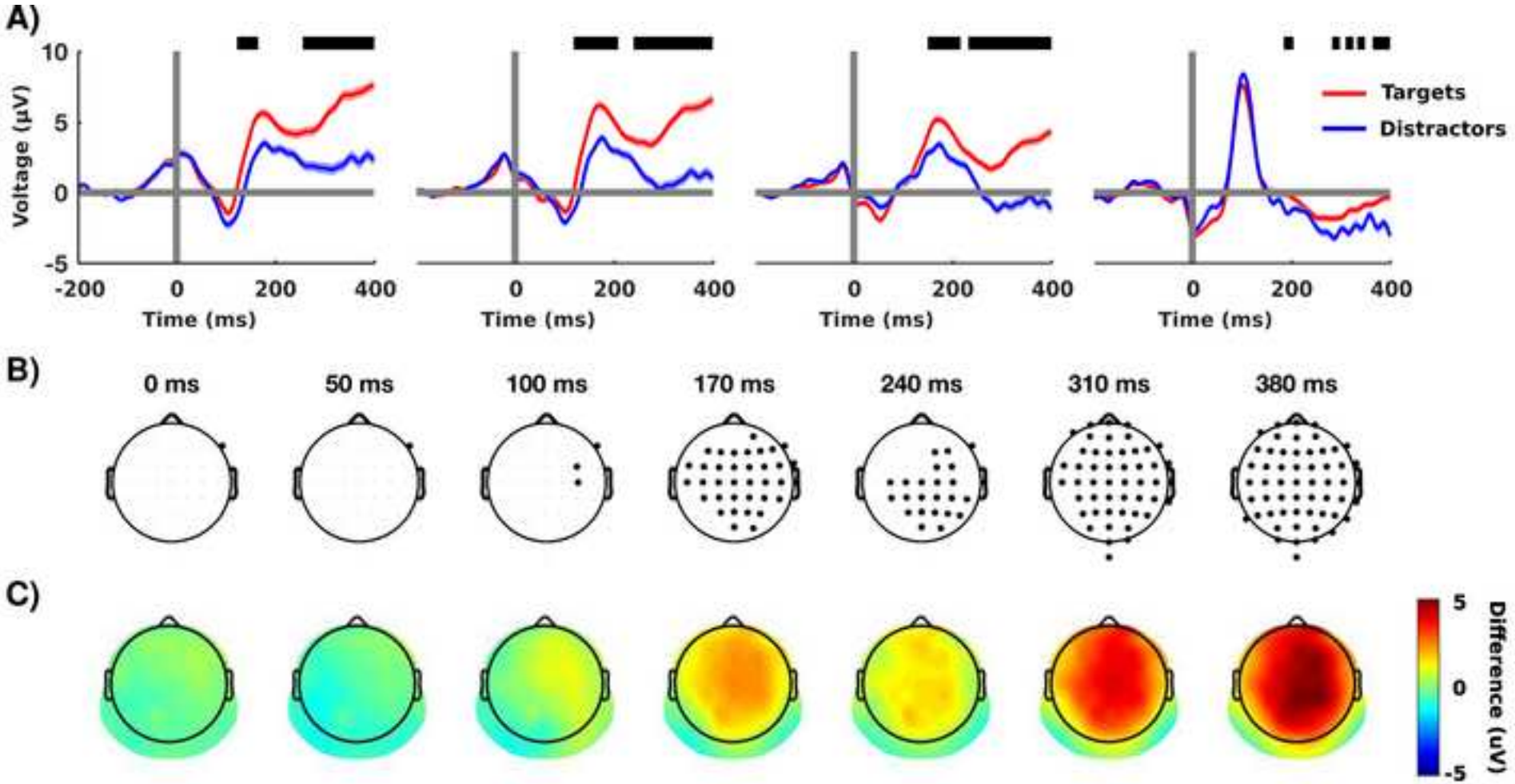
9. Figure

Click here to download high resolution image



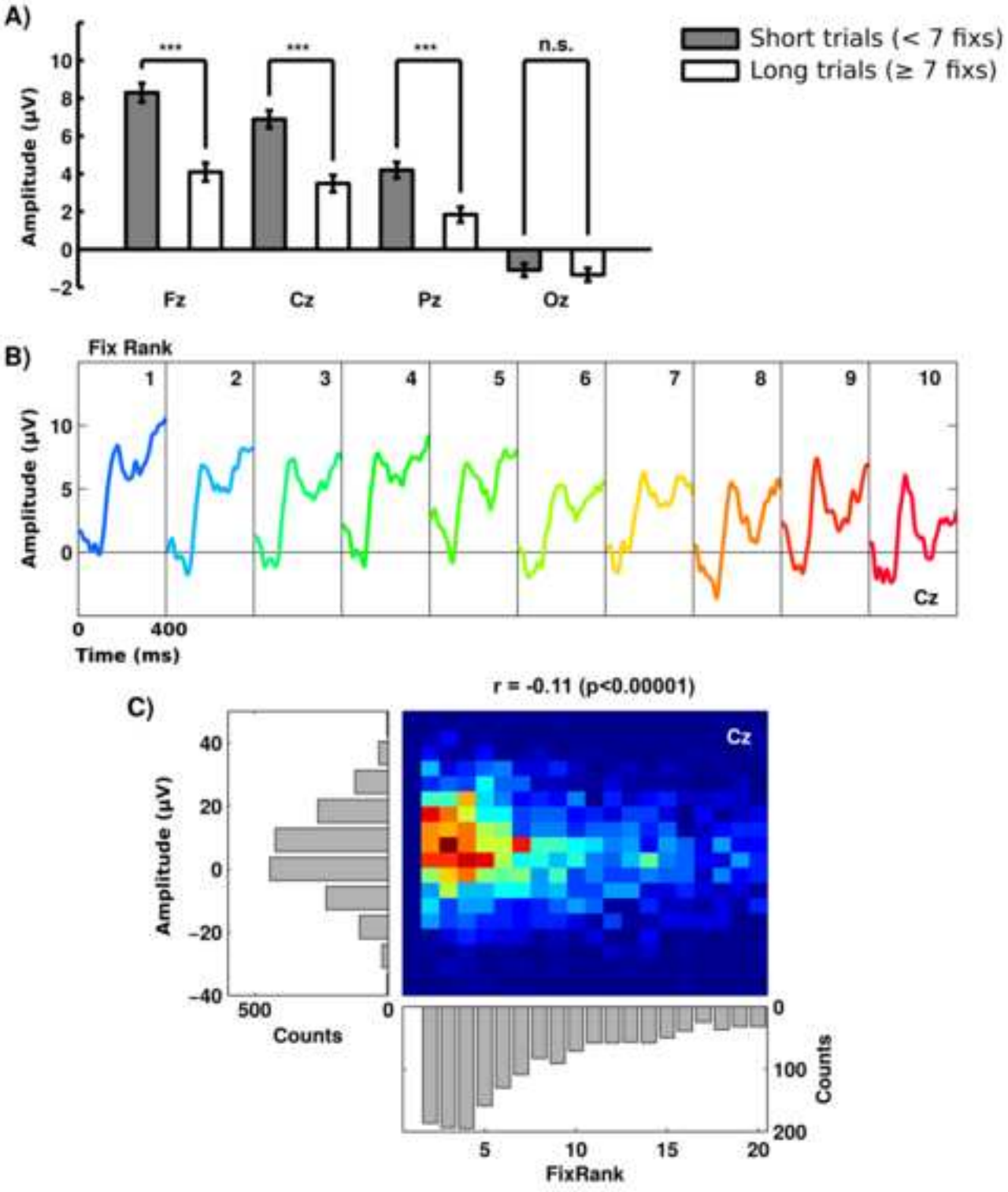
9. Figure 2

Click here to download high resolution image



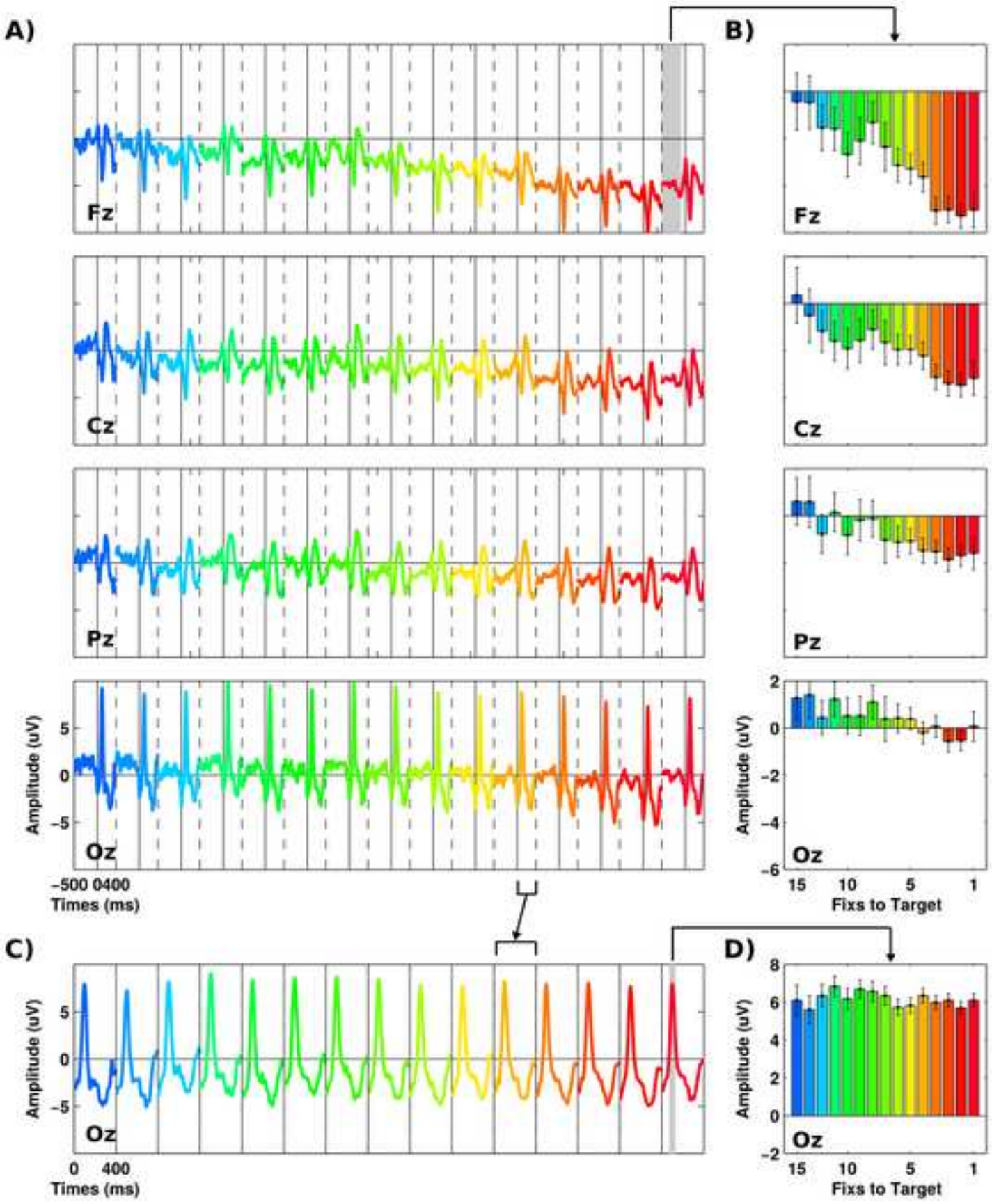
9. Figure 3

Click here to download high resolution image



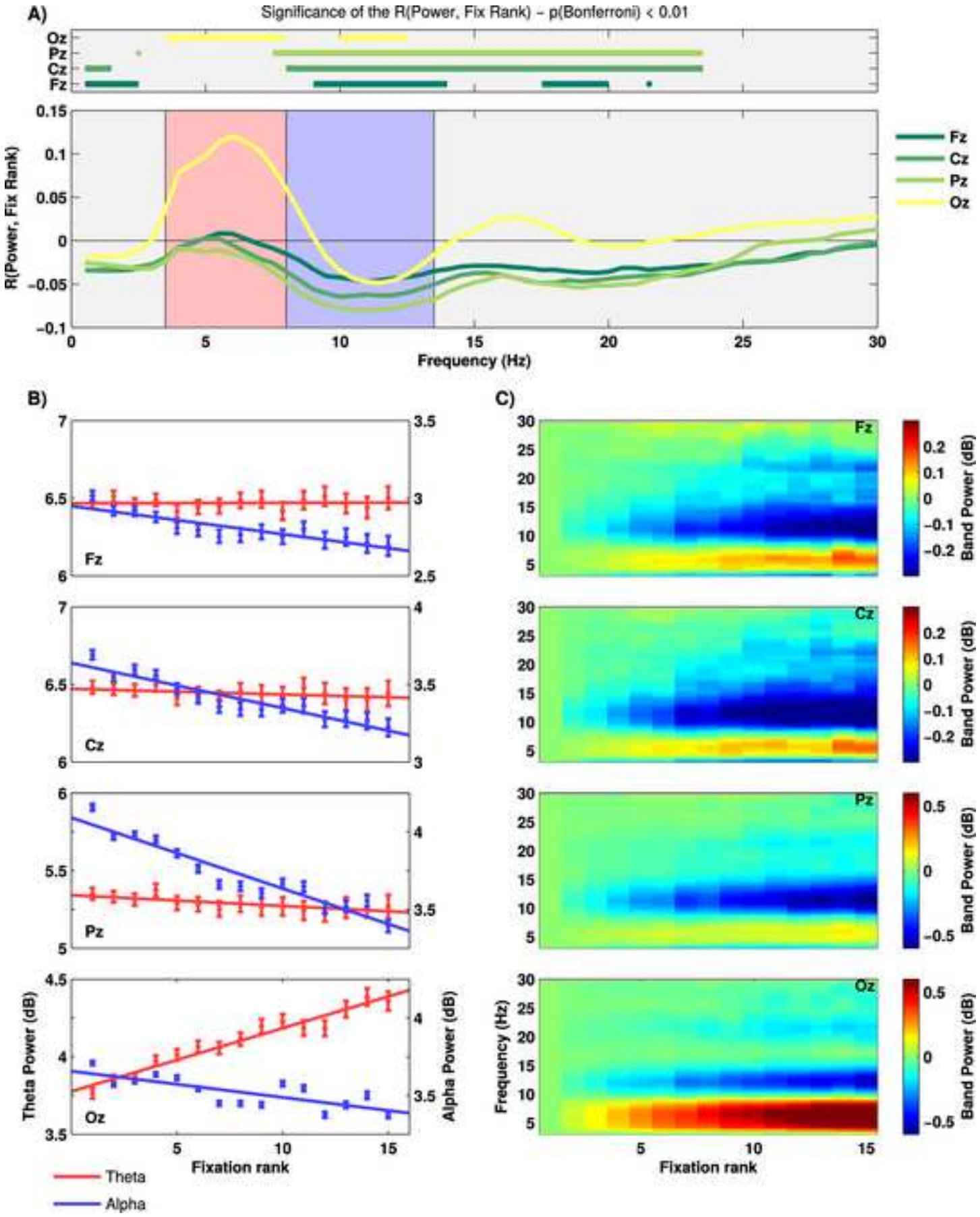
9. Figure 4

Click here to download high resolution image



9. Figure 5

Click here to download high resolution image



9. Figure 6

Click here to download high resolution image