Active search signatures in a free-viewing task exploiting concurrent EEG and eye movements recordings

Damián Care (1,*), María da Fonseca (1,2,*), Matias J. Ison (3,+), and Juan E. Kamienkowski (1,4,5,+)

(1) Laboratorio de Inteligencia Artificial Aplicada, Instituto de Ciencias de la Computación (Consejo Nacional de Investigaciones Científicas y Técnicas - Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires), Argentina
(2) Universidad Nacional de Río Negro, Argentina
(3) School of Psychology, University of Nottingham, United Kingdom
(4) Maestría de Explotación de Datos y Descubrimiento del Conocimiento, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina
(5) Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina

(*) equal contributions
(+) equal contributions

Keywords

visual search, eye tracking, fixation-related potentials, face processing, deconvolutional analysis

Abstract

Early access to categorical visual information plays an important role in real-world behaviour. Under fixed-gaze conditions, there are well-established electrophysiological signatures of categorical processing like the N170 EEG component. Here we aimed to study how the sensitivity to category information extends to a free viewing paradigm and consider how ecologically-relevant variables such as the particular task performed influence the brain activity. We co-registered EEG and eye-tracking to investigate fixation-related potentials to pictures of different categories during visual search and exploration. Participants were asked to “observe” / “search for one target from” an array of faces and objects embedded in random noise. We applied a deconvolution analysis approach to estimate the contribution of the different elements embedded in the task. Consistent with classical fixed-gaze experiments and a handful of free viewing studies, we found a robust categorial effect in occipital and occipitotemporal electrodes. We also report a task effect, which was reflected by a negative occipital activation and a positive frontal activation of the estimates. To further investigate the dynamics along the trial, we defined a trial progression score and found significant effects of trial progression, and the interaction between the trial progression and task effects. Overall, these results generalise the characterisation of early visual face processing to a wider range of experiments and show how a suitable analysis approach allows to discern among multiple neural contributions to the signal, preserving key attributes of real-world tasks.

Introduction

Real-world free viewing tasks require orchestrated sequences of events. When reading a sentence or exploring a visual scene, we move our eyes gazing on different elements of the scene creating a sequence of fixations alternated by saccade movements. The dynamic process that guide human saccadic scanpaths is known to be influenced by bottom-up factors such as the stimulus salience (Itti *et al.*, 1998), but also by top-down factors such as the task (Castelhano *et al.*, 2009), context (Torralba *et al.*, 2006) and the semantic relations within the image (Underwood & Foulsham, 2006). In free-viewing tasks, these factors contribute to the ensemble of eye movements through the visual scene (Devillez, *et al.*, 2015; Devillez *et al.*, 2017).

One of the most informative stimuli we can find in real life is a human face. Indeed, a brief exposure to a human face is enough to gather crucial information such as its identity and emotion, and the EEG signal for faces show a consistent N170, an occipitotemporal component arising about 170 ms after stimulus presentation that is stronger for faces in comparison to objects (Joyce & Rossion, 2005; Rossion & Jacques, 2011).

On the other hand, studies investigating the brain processes underlying free viewing are more scarce, since there are technical challenges associated with the eye movement artifacts in the EEG signal and the self-organised structure of the trial. However, through simultaneous EEG and eye-tracking (ET) recordings, it is possible to examine the neural processing underlying the visual information by extracting the fixation-related potentials (FRPs). Several studies involving free-viewing visual search tasks have shown components that distinguish between targets and distractors when looking for a special character (Kamienkowski *et al.*, 2012; Hiebel *et al.*, 2018) a face in a crowd (Kaunitz *et al.*, 2014), or an object within a natural scene (Devillez *et al.*, 2015). In this context of free-viewing, recent evidence indicates target detection is affected by refixation behaviour (Meghanathan *et al.*, 2020), as well as the processing of semantic integration of the scene information (Coco *et al.*, 2020). More recently, (Auerbach-Asch *et al.*, 2020) have included the dynamic of eye-movements in the context of free-viewing tasks, showing a similar topography of the N170 in comparison with a fixed gaze paradigm.

Most of these studies have manipulated an individual dimension in free viewing. However, real-world free viewing tasks require several concurrent processes to be operating in an orchestrated manner, rendering its analysis a challenge. Deconvolution models have extensively been applied in other fields such as fMRI (Dale & Buckner, 1997) and extended to EEG (Smith & Kutas, 2015). This framework has recently been applied to concurrent EEG and eye movements recordings using the Unfold toolbox (Ehinger & Dimigen, 2019; Dimigen & Ehinger, 2021) to study the impact of pre-saccadic preview over the post-saccadic face processing (Buonocore *et al.*, 2020), as well as face selectiveness of neural activity in free viewing (Auerbach-Asch *et al.*, 2020).

In this study, we apply a deconvolution framework to decipher the contribution of different aspects of the task to free viewing. We designed an EEG and eye movements co-registration paradigm, where faces and objects are embedded in a noisy background, and ask participants to perform two types of trials: *visual search* (VS) trials, in which they are instructed to search for a target, and *exploration* (EX) trials, in which participants observe the display. Our aim is to understand how the category (either a face or an object), the task (visual search or exploration), and the integration of information throughout the trial affect the neural responses.

Materials and Methods

Participants

Twenty-one adult participants performed the experiment. All participants were naïve to the objectives of the experiment, had normal or corrected to normal vision, and provided written informed consent, which was approved by the University of Nottingham School of Psychology Ethics Committee. Data from five participants were excluded: two due to bad performance, two due to poor eye tracking quality, and one due to poor EEG quality even after filtering and applying the ICA procedure (see below). In total, we analysed data from sixteen participants with a median age of 23 y.o. (range = [19, 28] y.o.).

Apparatus and data acquisition

Stimuli were presented in a BenQ XL2420Z monitor with a screen resolution of 1024 x 768 pixels and at a refresh rate of 75Hz. Participants were placed at a distance of approximately 60 cm from the monitor, and responses were introduced through a standard 'qwerty' keyboard. All stimuli were presented in MATLAB (MathWorks, 2000) using the psychophysics toolbox (Brainard, 1997).

EEG activity was recorded using a BioSemi Active-Two system at 1024 Hz with 64 electrode positions on a standard 10-20 montage. Four reference electrodes were used: the right and left mastoids, and the right and left earlobes. After data were recorded, the sampling rate was digitally downsampled to 512 Hz and the signal was re-referenced to the average of the 64 electrodes. Eye-movement tracking (ET) data were recorded using an EyeLink 1000 Plus remote system in monocular mode and a sampling rate of 500 Hz. The EyeLink target sticker was used to ensure movement stabilisation. To integrate the ET and EEG data, the stimulus-generating computer sent shared messages to both the EyeLink and BioSemi systems at different temporal marks according to the experimental procedure.

Experimental procedure

The experiment consisted of 9 blocks containing 24 trials each, making a total of 216 trials; the complete protocol lasted approximately 50 minutes. 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). In each trial, one of two tasks, the *visual search* (VS) and the *exploration* (EX) task were displayed (see Fig. 1). The main phase of each task (VS or EX), lasted 4.5 sec, regardless of having found the target or not (VS). The VS and the EX tasks were run 9 and 3 times, respectively, in each block in random order. The participants completed three eye-map blocks during the session that were used in pre-analysis to overweight artifactual effects. In the eye-map task, two points were displayed at steady positions, and participants were instructed to move their gaze slowly from one point to the other; depending on the trial, the two points were displayed horizontally or vertically and with a large or a small distance between them. Occasionally, a singular point was presented in the centre of the screen and participants were instructed to blink steadily at it. Pulse marks for posterior ET-EEG synchronisation were recorded at the beginning and the end of each trial, as well as at the beginning of the eye-map block. Written instructions were displayed at the beginning of the experiment.

One of the tasks is the VS task consisting of a search for a target image. A picture of either a face or an object was displayed in the centre of the screen for 1.5 seconds. Immediately after, a fixation dot placed at a random position was exhibited on the screen for a lapse that lasted between 1.25 and 1.75 seconds. After the fixation phase, a crowd of 15 stimuli pictures (5 objects and 10 faces, see next section) immersed in a noisy background were exhibited for 4 seconds, and participants were instructed to search for the target image in the crowd. Participants were instructed to retain the information for a 1.5-second period in which a middle grey screen was displayed. Following the retain phase, a question appeared for 1.5 seconds asking the participants to press the left (right) arrow of the keyboard to answer if the target image was present (absent) in the crowd. The target was present in the crowd of items half of the time. A correct answer to this question marked the trial as correct. Target presentation, fixation, retain and question phases were immersed in a middle grey background.

The EX task involved a procedure similar to the VS task. First, an empty box was displayed in the centre of the screen for 1.5 seconds. Immediately after, the same fixation, exploration and retain phases of VS task were presented, with the only instruction for free exploration of items in the crowd. At the end, in the question phase, either a face or an object was displayed in the centre of the screen with a question underneath asking, for example, whether the object is big (left arrow key) or small (right arrow key). This item was always absent in the exploration phase. Before the experiment, the participants classified a bunch of pictures of objects as big or small compared with a football. Trials were tagged as correct in accordance with the reply to this question.

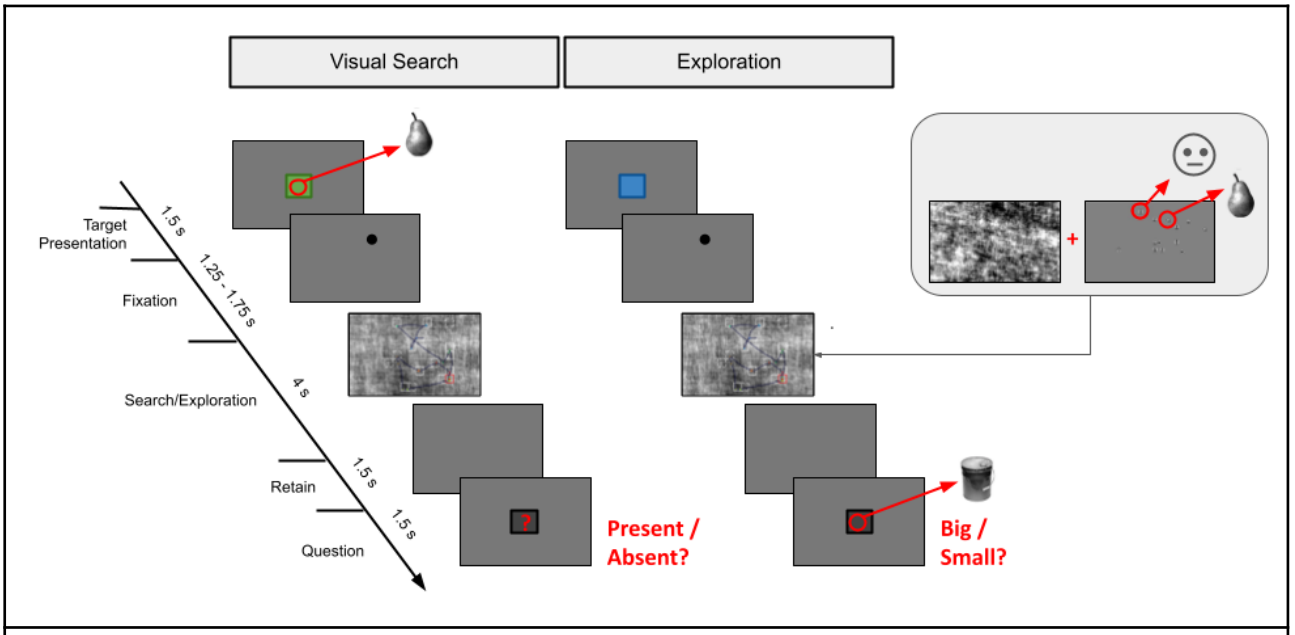

Fig. 1 Experimental paradigm. Time progression of the trial for the Visual Search condition with target presentation and the Exploration condition with a final question unrelated to stimuli. In the sketch we replaced the picture with a face drawing due to the automated verification process conducted by bioRxiv.

Stimuli

A set of 369 face grey-scale images were obtained from the Aberdeen face database[1], Karolinska directed emotional faces (Lundqvist *et al.*, 1998)[2], Yale face database[3], the AT&T face database[4], the Face-Place database[5], and the Aberdeen, the Iranian, the Pain and the UTrecht databases from the Psychological Image Collection at Stirling site (PICS)[6]. Face stimuli were carefully selected to have neutral expressions. Another set of 821 object grey-scale images were selected from the Unique Objects database (Brady *et al.*, 2008)[7]. All stimuli were framed in 250 x 250 pixels boxes and made isoluminant with the background to enhance reliability.

EEG and eye movements pre-processing

EEG data were imported, pre-processed and analysed with MATLAB using the EEGLAB toolbox (Delorme & Makeig, 2004) and custom-made scripts. Firstly, EEG raw data was filtered in the range 0.1-100 Hz and a notch filter was applied at power grid line frequency. Noisy channels with abnormal spectra were interpolated anytime they were less than five contiguous channels. Then, ET and EEG data were synchronised with the EYE-EEG toolbox (Dimigen *et al.*, 2011) using the synchronisation marks. Gaze positions landing outside screen range and blinks were marked as bad data. Eye movements were recalculated using an offline algorithm (Engbert & Mergenthaler, 2006). Then, we applied the OPTICAT procedure (Dimigen, 2020) to the data where an overweighting of presaccadic samples was carried out before the independent component analysis (ICA) training for better correction of myogenic saccadic spike potentials. Afterwards, ICA was calculated with Infomax algorithm (Bell & Sejnowski, 1995) and principal component analysis (PCA) was used prior to training in those cases in which data included interpolated channels. Next, we applied a variance criteria (Plöchl *et al.*, 2012) to recognise components related to ocular artifacts, a manual selection of ICA projections was also performed to exclude components in which OPTICAT failed to automatically recognise ocular movement sources. Lastly, we back-projected the surviving ICA components onto the EEG data high-pass filtered at 0.2 Hz and low-pass filtered at 40 Hz.

For each analysis, in the case of micro-fixations, that is, successive fixations to the same item, we selected the first micro-fixation in the array of successive micro-fixations that exceeded 100 ms duration as the event-related fixation. In the case of non-successive fixations to the same item, we only kept with the first and the second one. We discarded the first fixation of each trial, and the trials with less than three fixations. In order to compare the neural data associated with the fixation behaviour during active search and free exploration, we kept fixations within the correct trials and fixation duration within the range [100, 1000] ms. We defined *active search* (AS) as the set of fixations to non-target items belonging to the VS task with the target present, previous to finding the target (first fixation to target longer than 200 ms).

---

[1] http://pics.stir.ac.uk/2D_face_sets.htm
[2] https://kdef.se/home/about%20akdef.html
[3] http://vision.ucsd.edu/content/yale-face-database
[4] https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[5] http://wiki.cnbc.cmu.edu/Face_Place
[6] http://pics.stir.ac.uk/2D_face_sets.htmlI
[7] http://olivalab.mit.edu/MM/uniqueObjects.html

Single-participant (first level) statistical analysis

For fERP construction, we ran mass univariate regressions on the continuous data with the Unfold toolbox (Ehinger & Dimigen, 2019). We defined the time interval of interest around each fixation onset as the time from 200 ms before to 400 ms after the fixation event; this interval was used to build the design matrix and determine the time limits of the analysis. Fixations marked as bad data were discarded. Deconvolution analysis was achieved through models with the corresponding formula that incorporates the variable of interest as well as those considered relevant to explain overlaps. Then, we applied a baseline correction to the fitted parameters with a baseline period taken from 200 ms before each fixation and ending at fixation onset. The resulting parameters (rFRPs) were treated similarly to FRPs (Smith & Kutas, 2015).

We used contrast defined as treatment coding for categorical effects (e.g. 'task', 'category'), setting one level as a reference level with zero value. Since the intercept is the estimation when all predictors are zero, the interpretation of the parameters depends on the base condition. For instance, deviance at 200 ms in the time course of an estimated parameter would mean an effect at that very latency resulting from the difference between the FRP associated with the considered parameter and the FRP associated with the base condition. For continuous predictors, we interpret the values of the estimated parameters as the change in the predicted FRP per unit of the continuous predictor. Furthermore, interaction parameters give information about non-additive interactions between factors.

To clarify how the interpretation of the regressors depends on the reference level of the categorical effect, we built a synthetic response simulating categorical effects *A*, *B* and *C* and the interaction with a continuous effect *x* (Fig 2). Effect B had a larger positive peak than effect A, while effect C had a larger negative peak than effect A, followed by a square-form continuous interaction at later times. When effect C is coded as 1, the estimation of the base condition response (no increase in the positive and negative peaks and no interaction) is smaller and the parameter C has the expected sign associated with the effect C. On the contrary, switching the reference level for C gives a stronger negative peak in the estimation of the base condition (intercept), since now effect C is added to the base condition, the corresponding parameter has the opposite sign. The parameter β_C for the regressor X_C now estimates the difference between FRPs of the new base condition and the condition when C effect is absent, which is equivalent to effect A.

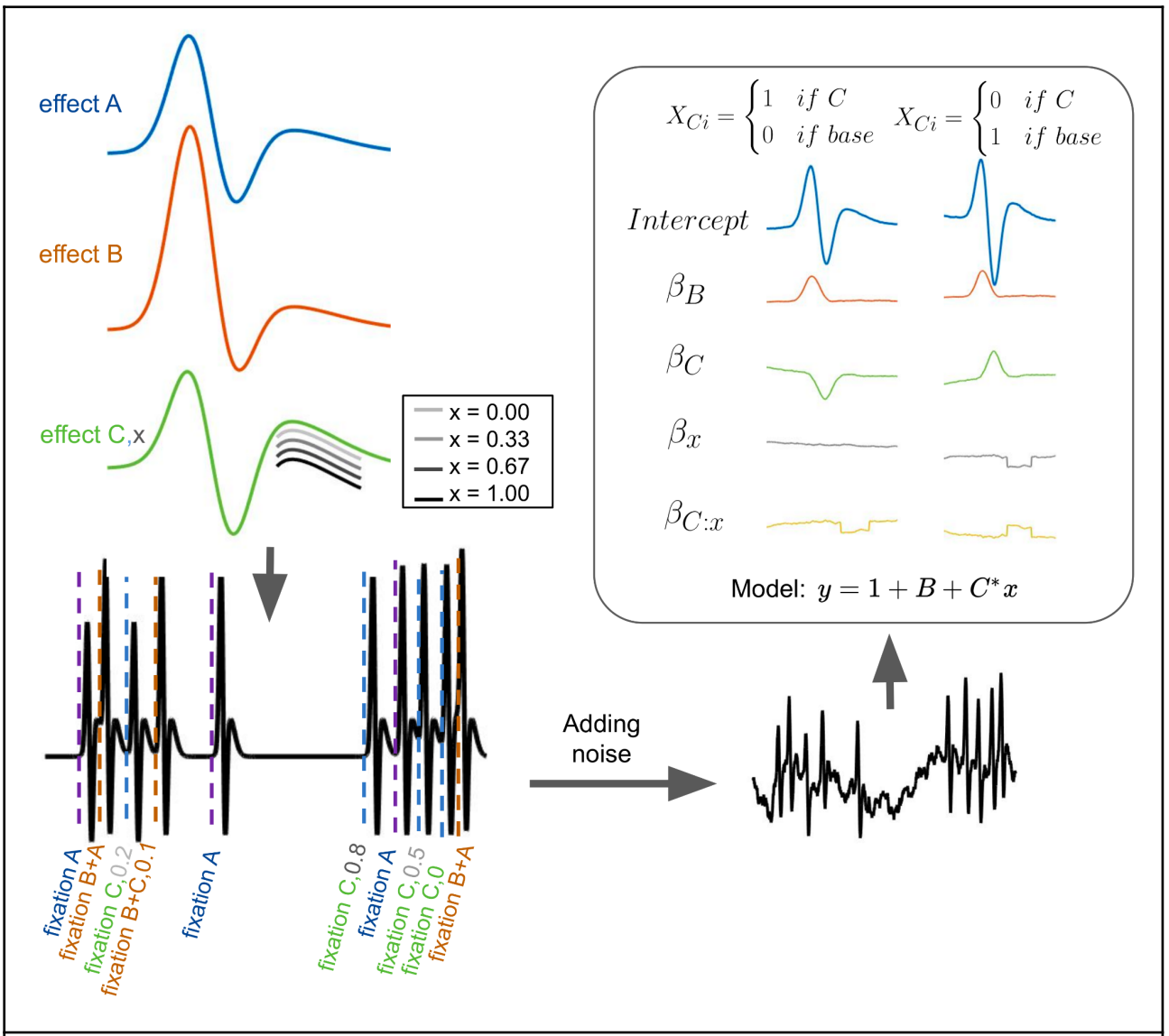

Figure 2. Interpretation of parameters in treatment coding. Synthetic responses A, B and C, x were set along time simulating the distribution of fixation events for an experimental subject. Pink noise was then added to the signal. The synthetic signal can be deconvolved with a linear model, leading to the beta values shown in the top right box. Notice the differences in values and signs, depending on the chosen reference.

The effect associated with the *category* (non-target objects: NTO as a reference level, non-target faces: NTF) on the FRPs was analysed following the model with the formula:

$$EEG \sim 1 + category \quad (1)$$

To analyse the effect related to the *task,* we introduced a dummy variable indicating whether the condition is AS or EX (as reference level). The corresponding formula was:

$$EEG \sim 1 + category + task \quad (2)$$

Then, we incorporated the fixation rank (centred to the mean, and standardised per trial) as a continuous predictor named *trial progression score* (TPS):

$$EEG \sim 1 + category + task + TPS \quad (3)$$

Lastly, we introduced the interaction between the *task* and the *TPS* in order to capture the trial progression score trend in each condition

$$EEG \sim 1 + category + task + TPS + task{:}TPS \quad (4)$$

Group (second-level) statistical analysis.

Second-level statistical analysis was carried out using the threshold-free cluster enhancement method (TFCE) (Smith & Nichols, 2009; Mensen & Khatami, 2013). Cluster correction methods have been used with EEG and MEG signals (Maris & Oostenveld, 2007; Mensen & Khatami, 2013; Bianchi *et al.*, 2019). The TFCE is a permutation test that takes into account the dependence between neighbouring EEG samples, both in time and space, and results in a much less conservative approach than t-testing with Bonferroni correction. It was first introduced as an improvement to cluster-based permutation tests (where an arbitrary threshold value affects the sensitivity of clusters of different sizes and intensities), first applied to fMRI analysis (Holmes *et al.*, 1996; Smith & Nichols, 2009) and later extended to EEG (Maris & Oostenveld, 2007). This method consisted on a transformation of the signal based on a weighted norm of cluster size, that is

$$TFCE(p) = \int_{h=h_0}^{h_p} e(h)^E h^H dh. \quad (5)$$

Afterwards, a thresholded version is compared with an empirically generated distribution using permutations. We used Menser's MATLAB implementation of TFCE (Mensen & Khatami, 2013)[8], which has been shown to yield better rejection for non-significant neighbouring measurements than previous methods (ANOVA and data-driven methods like PCA or ICA).
After TFCE transformation, cluster formation depends on the area supporting each cluster, that is, it takes into consideration the information coming from neighbouring samples and the intensity of the signal, enhancing the initial statistical analysis. This step is driven free of threshold, and therefore, it is less biassed by user decisions.

For each analysis, 2000 random permutations were used. Only *p*-values under 0.05 were shown for visualisation of the TFCE results.

Results

In this study, we developed an experiment involving co-registration of EEG and eye-tracking data. Stimuli were composed by a set of images of objects and faces embedded in a noisy background. During the experiment, participants answered two types of tasks split into different trials: *visual search* (VS) trials, in which they searched for a target, and *exploration* (EX) trials,

[8] http://github.com/Mensen/ept_TFCE-matlab

where they just explored stimuli presented on the screen. By filtering the correct and target-present trials of the VS trials and fixations previous to the target, we studied the *active search* (AS) behaviour and compared it with the behaviour during EX trials. The overall task performance was 78% in VS and 82% in EX trials, and the minimal performance across participants was 62% in VS and 68% in EX trials.

Eye-movements behaviour

The eye movements showed the typical characteristics of free viewing experiments (see Fig. 3). Saccade amplitude and main sequence followed a typical distribution (Fig. 3B,E) associated with the non-linearity in the saccadic system (Lebedev *et al.*, 1996) considered to be present in both types of tasks (Otero-Millan *et al.*, 2008). Saccade direction in VS condition showed a horizontal preference (Fig. 3C) which was slightly expanded in the EX condition (Fig. 3F). Fixation duration distribution was also typical, with a peak at around 170 ms, and a mean duration of (195 ± 110) ms (Fig. 3G). These results were consistent with previous findings in visual search and exploration (Otero-Millan *et al.*, 2008). Fixation duration to targets was longer than to non-target stimuli during VS condition (Fig. 3H). Finally, the mean number of fixations per trial was 8 with a standard deviation of 5.

Sometimes, when a participant fixates a stimulus, microsaccades occur within the fixation (i.e. a multiple fixation). We established a criterion to determine which of those fixations would be used to align the EEG responses. Here, we aim to get the best representation of the response to the stimulus that captured the participant's attention. To that end, we only considered EEG responses to the first fixation within those multiple fixations whose duration was greater than 100 ms. For instance, if the first fixation lasted only 40 ms and the second lasted 150 ms, then we considered the second fixation as the one that would identify the response of interest.

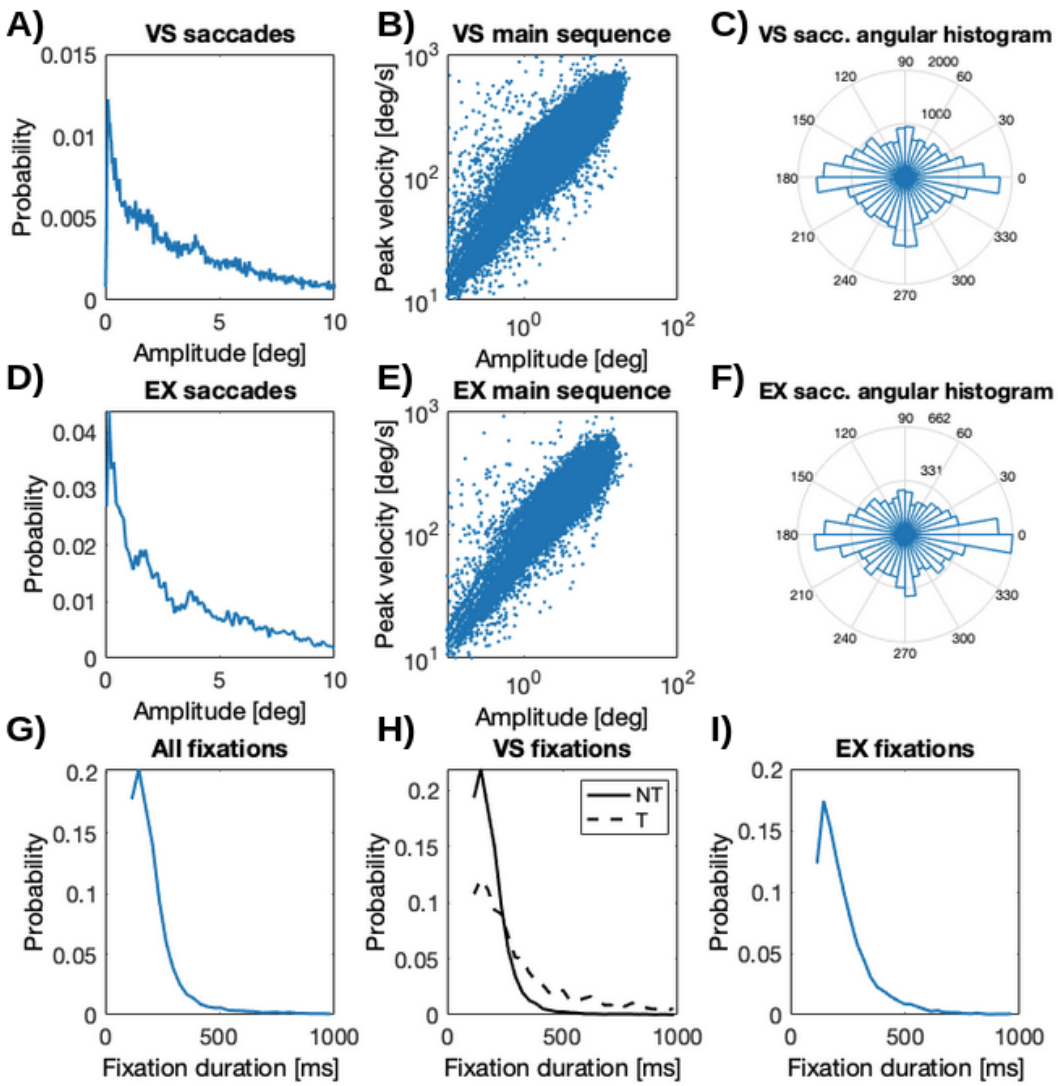

Fig. 3. Eye movement behaviour for all participants. **A. D.** Saccade amplitude distribution in degrees of visual angle, **B. E.** Main sequence of saccades, **C. F.** Saccade angular distributions, **G.** fixation duration distribution for all saccades from all the trials included in the analysis (N=8654), **H.** fixation duration distribution for all saccades from the visual search (VS) trials, separated by identity: target (T, bold line) and non-target (NT, dashed line), and **I.** fixation duration distribution for all saccades from the exploration (EX) trials.

Non-Target Objects and Non-Target Faces

Firstly, we explored differences in the brain response to the category of the fixation content, in particular, we considered fixations to non-target objects' stimuli and compared them with fixations to non-target faces. Because there could be an overlap between the evoked response of consecutive fixation, we applied a regression-based analysis to disentangle the overlapping activity (Ehinger & Dimigen, 2019). We computed the regressor associated with the category predictor (Eq. 1) for each time-channel sample and obtained the time course of the category effect, i.e. the difference between FRPs to faces and objects. Fixations from both tasks, AS and EX, were included in these analyses.

The intercept waveform presented a well-defined P1 component at 100 ms, there was a local minimum between 120 ms and 200 ms, and a second smaller positive component at around 230 ms (Fig. 4C). These activations were present in all occipital and occipitotemporal electrodes (Fig. 4B). The frontal components exposed an inverted polarity (Fig. 4B). On the other hand, the category regressor exhibited an early effect from 130 ms to 300 ms. This effect is compatible with an N170 component since these values are associated with the predictor for face stimuli (Fig. 4C). Similar results were obtained including post-target fixations (see Supp. Fig. 1).

We used a second-level group analysis (TFCE) to assess the statistical significance of FRPs associated with the face predictor (Fig. 4A). We can see significant samples of this effect in central and occipital regions of the scalp that extend consistently from 120 to 350 ms (Fig. 4A,C).

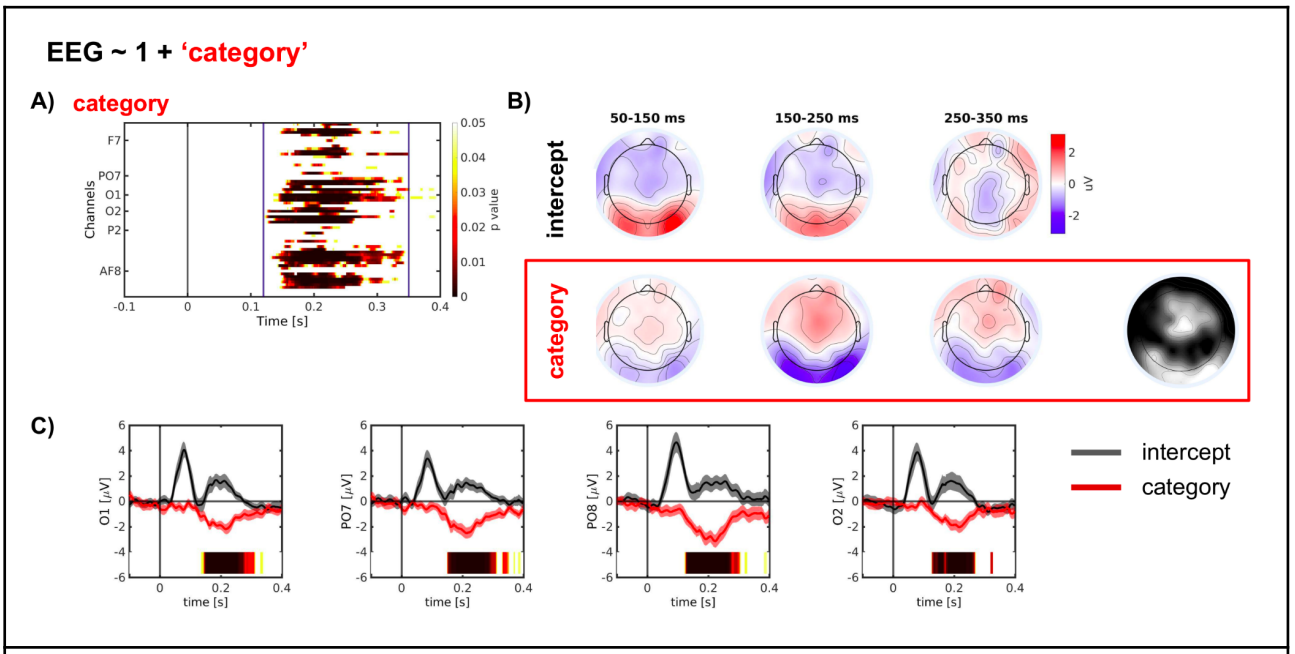

Fig. 4 Time-course of category effect on FRPs for active search and exploration tasks (model Eq. 1). A. Significant samples obtained with the TFCE. **B.** Scalp topography for different stages of the intercept and the category effect (betas). The right scalp topography shows the proportion of significant samples per channel in [120, 350] ms time-window in between 0 (black) and 1 (white). **C.** Mean and s.e.m. across participants of the intercept (black) and the category (red) effect for selected channels [O1, PO8, PO7, O2]. The colour bar below each panel represents the significant intervals after TFCE second-order statistics applied to the face-object (category) effect. The intensity of the bar represents the significance level (p-value), only for the p-values < 0.05, the scale is the same as in panel A. Number of samples (N) = 5127. Number of participants = 16.

Early potentials evoked by task (between trials effect)

To explore how task changes would influence FRPs, we included a binary variable representing the task (AS or EX) in the model (Eq. 2), and computed the time course of the regressors. The task effect on the FRPs showed slightly earlier activations (green curve in panel D of Fig. 5) separated from the content (face/object) effect. Scalp topographies showed that the task effect presented a negative occipital and a positive frontal component. Interestingly, a frontal activation was also present in the task effect, for which the TFCE yielded significant intervals from

80 ms to 250 ms (Fig. 5B,D). Regarding the category effect distribution, it remained essentially unchanged (Fig. 5C) in comparison with the simplest model (Fig. 4C).

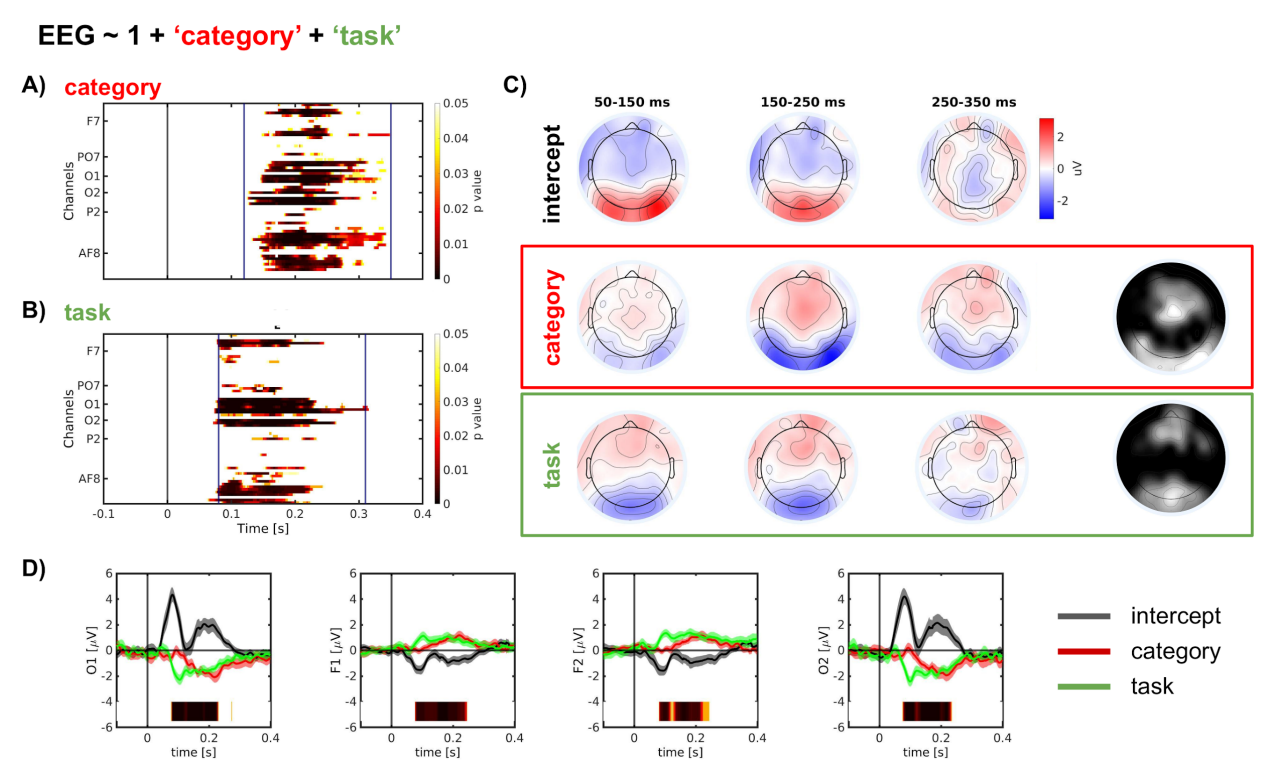

Fig. 5. Time-course of task effect on FRPs for active search and exploration tasks (model Eq. 2). A.B. Significant samples obtained with the TFCE test for the category (**A**) and the task (**B**) effects. **C.** Scalp topography for different stages of the intercept, the category and the task effects (betas). The right scalp topography shows the proportion of significant samples per channel in the time-window [120, 350] ms for the category (red) effect and [80, 310] ms for the task (green) effect, in between 0 (black) and 1 (white). **D.** Mean and s.e.m. across participants of the intercept (black), the category (red) and the task (green) effect for selected channels [O1, F1, F2, O2]. The colour bar below each panel represents the significant intervals after TFCE second-order statistics applied to the AS-EX (task) effect. The intensity of the bar represents the significance level (p-value), only for the p-values < 0.05, the scale is the same as in panel A. Number of fixations = 5127. Number of participants = 16.

Global effects of the task and trial progression on the FRPs

Lastly, we included in the analysis the trial progression score (TPS) as a continuous predictor through the standardised order of each fixation in the trials. To explore how these effects interacted with each other, we fitted a model that also included the interaction between the task and the trial progression score regressors (Eq. 4). The present regression-based analysis showed the time course of the effects that varied between trials and the TPS effect that varied within each trial.

The FRPs of the content and task effects remained virtually unchanged. In the case of the TPS effect, FRP rose up to an almost constant value at about 170 ms for occipital channels, while frontal channels exhibited an opposite behaviour with a smoothly varying pattern around 200 ms (Fig. 6F). Second level statistics results showed a considerable occipital and frontal area of significance for these effects (Fig. 6C) starting at 100ms. Scalp topographies for the TPS effect

showed the distribution of the activations remained until the end of fixations (Fig. 6E). Similar results were obtained when fitting data with a model without the interaction term (Eq. 3; see Supp. Fig. 2).

The interaction effect between the TPS and the task exposed a later activation that followed the tendency of the TPS effect, with positive occipital and negative frontal activation. The second-statistical level results for the interaction betas supported the idea that TPS was more influenced by the active search of a target (AS) in comparison with the free exploration of items (EX) (Fig. 6D).

Interestingly, the intercept presented a similar form through the different models, meaning that the estimation of the neural response to the base condition (objects for model 1, objects + EX for model 2, objects + EX + middle TPS for models 3 and 4) remained largely unchanged.

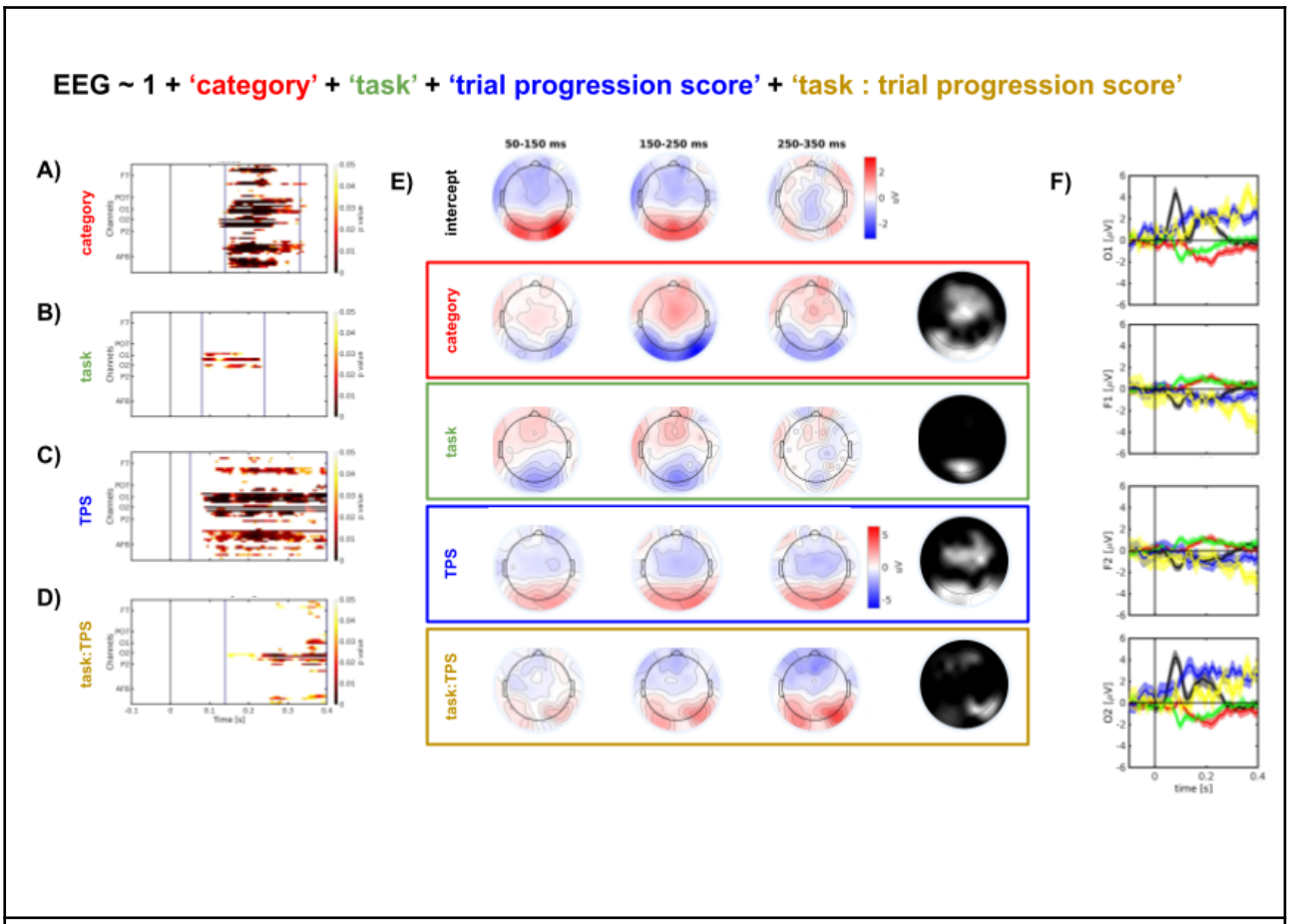

Fig. 6. Time-course of trial progression score (TPS) and the interaction with the task effect on FRPs for active search and exploration tasks (model Eq. 4). A.B.C.D. Significant samples obtained with the TFCE test for the category (**A**), the task (**B**), the TPS effect (**C**), the interaction between TPS and the task (**D**). **E.** Scalp topography for different stages of the intercept and the effects exposed in **A**, **B**, **C**, and **D** panels (intercept, category and task, on the one hand, and TPS and the interaction effect, on the other hand, share the same scale). The right scalp topography shows the proportion of significant samples per channel in the time-window [120, 350] ms for the category (red) effect, [80, 400] ms for the task (green) effect,[50, 400] ms for the TPS effect, and [140, 400] ms for the interaction of the TPS with the task, in between 0 (black) and 1 (white). **F.** Mean and s.e.m. across participants of the task (green), the TPS (blue) and the interaction between the TPS and the task (yellow) effects for selected channels [O1, F1, F2, O2]. Number of fixations = 5127. Number of participants = 16.

Discussion

In the present study, we aimed to characterise how, under free-viewing conditions, visual information processing depends on the content of the stimuli being fixated, the task being performed and its progression. We used a mass univariate deconvolution modelling approach to reduce the effects of overlapping brain activity in successive fixations, such as preparation and execution of eye movements and the visual stimulation itself. We modelled the resulting fixation-locked potentials with linear regressions by sequentially adding different regressors (category of the fixated stimulus, task, and a trial progression score). Cluster-based permutation tests were used to control for multiple comparisons.

We observed a significant category (faces vs. objects) effect that resembles the well-known EEG N170 component seen in fixed-gazed experiments (Joyce & Rossion, 2005; Rossion & Jacques, 2011), fast periodic face stimulations (Liu-Shuang *et al.*, 2014), and, more recently, in an experiment allowing eye movements in controlled displays (Auerbach-Asch *et al.*, 2020). We started with a simple model that only included the category effect. Interestingly, this effect remained almost unchanged (both in latency and topography) when other variables were included in the model. This suggests an effect that is independent of the task and the progression of the trial and reflects the processing of the foveated item. Moreover, given the short latency of the category effect, this resembles the saccadic response times obtained with a brief presentation of stimuli (VanRullen & Thorpe, 2001), which were previously related to the feedforward processing of the stimulus (Thorpe & Fabre-Thorpe, 2001; Thorpe, 2009).

Then, we explored the effects of the task and the trial progression in the model, like variables that could affect FRPs throughout saccades -i.e. global variables-. Firstly, we focused on the differences between active search (AS) and exploration (EX). The effects of the task in the FRPs have been previously explored in the context of modulating task demands on a guided-search task (Ries *et al.*, 2016) and comparing free-viewing visual-search/memorization task versus the recall display (Seidkhani *et al.*, 2017). Ries and collaborators (2016) asked participants to perform a guided-search task, moving their eyes towards the item marked with a red circle until they found the target. At the same time, they were asked to either follow or ignore the auditory information from a 0, 1, or 2-back task. Thus, the concurrent task modulated the resources available for the guided-search task. They observed that higher auditory task demands increased the latency and decreased the amplitude of the P3 in response to the target during the guided-search, while the lambda response showed a decrease in amplitude only in the high task demands condition (Ries *et al.*, 2016). In our case, we manipulated the task itself instead of introducing a concurrent task and exploited the capabilities of the linear models to separate the effect of the early visual responses from the task. We observed that the task showed a significant effect that started early in the fixation (before 100ms) and spread to almost 250ms, while the specific responses to the category remained unaffected. This interval overlaps with the one observed by Seidkhani and collaborators (2017) while comparing the networks that emerge after the fixation onset between a search/encode phase and a recall phase in a change detection task that included free eye movements (Seidkhani *et al.*, 2017). Although this study did not focus on the P3, in line with Ries and collaborators (2016), we observed an important interaction between the task effect and the progression of the task during the active search -i.e. when considering distractors closer to the fixation on the target -.

The task manipulation could also affect the FRPs through an increase in temporal attention or a higher attentional engagement in active search (Corbetta *et al.*, 1998; Correa *et al.*, 2006; Melcher & Colby, 2008; Kamienkowski *et al.*, 2018). For instance, targets appearing at attended

moments in highly demanding perceptual processing tasks have an effect both on the P1 and the N2 components (Correa *et al.*, 2006). However, these previous studies could not disentangle whether this was a modulation of the same component or a new effect, as we showed here by simultaneously analysing both effects. The results presented in this study, with significant early task effects, might be associated with higher attentional engagement, where earlier responses are expected on higher perceptual load tasks in accordance with perceptual load theories (Lavie *et al.*, 2014). Following this idea, we can speculate that early task effects are more prominent in highly demanding tasks such as active search. We previously reported early target detection in tasks involving prolonged fixations when participants were asked to detect a subtle change in a synthetic stimulus (Kamienkowski *et al.*, 2012) or when naturally looking for a face in a crowded scene -i.e. producing fixations in the order of 200-250ms (Kamienkowski *et al.*, 2018), but not when they were instructed to prioritise accuracy in each fixation over time, producing longer fixations to avoid missing the target (Kaunitz *et al.*, 2014). This is also consistent with studies reporting that target identification can be detected from the EEG signal prior to fixation onset (Dias *et al.*, 2013).

We have previously explored the effect of task progression in FRPs, which exhibited signatures related to the accumulation of evidence during the task (Kamienkowski *et al.*, 2018). Here, the task progression showed a massive effect that started around the first 50ms and continued after the 400ms. Before 150ms this effect reflected a local increase only in the occipital region, while from 150ms it also exhibited a decrease in fronto-central electrodes. As mentioned before, this effect did not interact with the categorical response. Still, fixations in active search presented an even stronger pattern as shown by the interaction, providing evidence for the task demand when participants actively search for a target. Indeed, the steady decrease in central-frontal regions following the progression of the search was similar to the significant decrease in the baseline amplitude found in previous visual search free-viewing experiments (Kamienkowski *et al.*, 2018), and to gradual changes in the baseline observed in other tasks in which participants were required to accumulate some evidence in order to achieve a decision, both in humans (de Lange *et al.*, 2010) and monkeys (Yang & Shadlen, 2007). Yang and Shadlen showed that single-neurons in the lateral intraparietal cortex (LIP) - the monkey analogue of the human posterior parietal cortex (PPC) - accumulated probabilistic evidence presented sequentially until a saccadic response was made (Yang & Shadlen, 2007). Moreover, the population activity in the dorsolateral prefrontal cortex (DLPFC) and the caudate nucleus (CN) in monkeys suggest that the brain may have an inherent tendency to represent time as part of ongoing task-specific information processing (Jin *et al.*, 2009). Thus, neural circuits might build time representations to support different functions when needed. These circuits can make available all of the elements needed to form on demand new associations between events and precisely timed actions, which would have major advantages for brain mechanisms related to action control. Interestingly, besides evidence coming from single neuron animal studies for temporal coding (reviewed in (Eichenbaum, 2014)), the existence of time cells in the human temporal lobe has also recently been reported (Eichenbaum, 2014; Umbach *et al.*, 2020). The great variety of experimentally observed patterns, with dominant phasic responses riding on ramping activity or peak responses after the first event or towards the end could be interpreted as a precondition for fine-scale time encoding with population activity (Jin *et al.*, 2009).

Furthermore, Cueva and collaborators specifically analysed neural activity in monkey LIP recorded during delays in four experiments to understand the mechanisms underlying working memory and time encoding. They found that time can be decoded with higher precision in tasks where timing information is relevant and that neural trajectories are always observed to be low-dimensional. These results showed that the linear ramping component of each neuron's firing

rate strongly contributes to the slow timescale variations that make decoding time possible (Cueva *et al.*, 2020). Interestingly, other studies on sequential tasks were focused on parsing individual operations in complex tasks, for instance, unfolding processing stages underlying arithmetic calculations, but their figures also presented a steady change of baseline activity throughout the task which is not part of the individualised components (Pinheiro-Chagas *et al.*, 2019). Moreover, in our experiment, we not only found that the task itself has an effect on the FRPs, but also in relation to the progression of the task. The stronger changes along the task in active search relative to exploration could also be due to increased working memory or general cognitive demands. Previous works showed how visual working memory capacity is strongly related to performance in both visual search without eye movements (Luria & Vogel, 2011; Luck & Vogel, 2013) and constrained eye movement paradigms (Theeuwes *et al.*, 2009).

The present study describes the effects of the task in two timescales: first, the local processing of information from each fixation using fRPs and then, the global processing of information throughout the trial, using mass univariate deconvolution modelling to reduce the effects of overlapping brain activity in successive fixations. These results stress the importance of the goal in free-viewing experiments, and the interaction between different timescales in complex sequential tasks as it occurs when sampling almost any scene -static or dynamic- in the real world. Future work should explore more deeply this interaction. One possibility would be to parametrically manipulate a relevant cognitive variable, such as asking participants to memorise different numbers of items (Wolfe, 2012). Another alternative would be to manipulate the task's goal by finishing trials in different ways (e.g. button press). Altogether, EEG and eye movement co-registration is an ideal combination to explore the interaction between events and processes in different timescales and hierarchies of a complex task. This is particularly relevant when studying behaviour in natural environments, and is arguably a necessary next step in neurocognitive research.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

D.C. was supported by the CONICET, M.J.I was supported by the University of Nottingham, J.E.K. was supported by the CONICET and the University of Buenos Aires (UBA). J.E.K received research grants from CONICET (PIP 11220150100787CO) and ANPCyT (PICT 2018-2699). J.E.K. and M.J.I. received an award from ARL (Award W911NF1920240). We thank Susannah Bilsborough and Joseph Conyers for their collaboration with the data acquisition and Anthony Ries for insightful discussions.

Code and Data availability statement

The code and sample data are available at *[a LINK that will be available upon publication]*. Raw data will be available upon request.

Author contributions

M.J.I. and J.E.K. designed and programmed the experiment, and collected the data. D.C., M.J.I. and J.E.K. preprocessed the EEG and eye-tracking data. D.C. and M.d.F. performed the EEG and eye-tracking analysis. D.C., M.d.F., M.J.I and J.E.K. discussed the results and wrote the manuscript. All authors reviewed the manuscript.

References

Auerbach-Asch, C.R., Bein, O., & Deouell, L.Y. (2020) Face Selective Neural Activity: Comparisons Between Fixed and Free Viewing. *Brain Topogr.*, **33**, 336–354.

Bell, A.J. & Sejnowski, T.J. (1995) An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.*, **7**, 1129–1159.

Bianchi, B., Shalom, D.E., & Kamienkowski, J.E. (2019) Predicting Known Sentences: Neural Basis of Proverb Reading Using Non-parametric Statistical Testing and Mixed-Effects Models. *Front. Hum. Neurosci.*, **13**, 82.

Brady, T.F., Konkle, T., Alvarez, G.A., & Oliva, A. (2008) Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*,.

Brainard, D.H. (1997) The Psychophysics Toolbox. *Spat. Vis.*, **10**, 433–436.

Buonocore, A., Dimigen, O., & Melcher, D. (2020) Post-Saccadic Face Processing Is Modulated by Pre-Saccadic Preview: Evidence from Fixation-Related Potentials. *J. Neurosci.*, **40**, 2305–2313.

Castelhano, M.S., Mack, M.L., & Henderson, J.M. (2009) Viewing task influences eye movement control during active scene perception. *J. Vis.*, **9**, 6.1–15.

Coco, M.I., Nuthmann, A., & Dimigen, O. (2020) Fixation-related Brain Potentials during Semantic Integration of Object–Scene Information. *Journal of Cognitive Neuroscience*,.

Corbetta, M., Akbudak, E., Conturo, T.E., Snyder, A.Z., Ollinger, J.M., Drury, H.A., Linenweber, M.R., Petersen, S.E., Raichle, M.E., Van Essen, D.C., & Shulman, G.L. (1998) A common network of functional areas for attention and eye movements. *Neuron*, **21**, 761–773.

Correa, A., Lupiáñez, J., Madrid, E., & Tudela, P. (2006) Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Res.*, **1076**, 116–128.

Cueva, C.J., Saez, A., Marcos, E., Genovesio, A., Jazayeri, M., Romo, R., Salzman, C.D., Shadlen, M.N., & Fusi, S. (2020) Low-dimensional dynamics for working memory and time encoding. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 23021–23032.

Dale, A.M. & Buckner, R.L. (1997) Selective averaging of rapidly presented individual trials using fMRI. *Hum. Brain Mapp.*, **5**, 329–340.

de Lange, F.P., Jensen, O., & Dehaene, S. (2010) Accumulation of evidence during sequential decision making: the importance of top-down factors. *J. Neurosci.*, **30**, 731–738.

Delorme, A. & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods*, **134**, 9–21.

Devillez, H., Guerin-Dugue, A., & Guyader, N. (2015) How task difficulty influences eye movements when exploring natural scene images. *2015 23rd European Signal Processing Conference (EUSIPCO)*,.

Devillez, H., Guérin-Dugué, A., & Guyader, N. (2017) How a distractor influences fixations during the exploration of natural scenes. *J. Eye Mov. Res.*, **10**.

Devillez, H., Guyader, N., & Guérin-Dugué, A. (2015) An eye fixation-related potentials analysis of the P300 potential for fixations onto a target object when exploring natural scenes. *J. Vis.*, **15**, 20.

Dias, J.C., Sajda, P., Dmochowski, J.P., & Parra, L.C. (2013) EEG precursors of detected and missed targets during free-viewing search. *Journal of Vision*,.

Dimigen, O. (2020) Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. *Neuroimage*, **207**, 116117.
Dimigen, O. & Ehinger, B.V. (2021) Regression-based analysis of combined EEG and eye-tracking data: Theory and applications. *J. Vis.*, **21**, 3.
Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A.M., & Kliegl, R. (2011) Coregistration of eye movements and EEG in natural reading: analyses and review. *J. Exp. Psychol. Gen.*, **140**, 552–572.
Ehinger, B.V. & Dimigen, O. (2019) Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ*, **7**, e7838.
Eichenbaum, H. (2014) Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.*, **15**, 732–744.
Engbert, R. & Mergenthaler, K. (2006) Microsaccades are triggered by low retinal image slip. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 7192–7197.
Hiebel, H., Ischebeck, A., Brunner, C., Nikolaev, A. R., Höfler, M., & Körner, C. (2018). Target probability modulates fixation-related potentials in visual search. *Biological psychology*, **138**, 199-210.
Holmes, A.P., Blair, R.C., Watson, J.D., & Ford, I. (1996) Nonparametric analysis of statistic images from functional mapping experiments. *J. Cereb. Blood Flow Metab.*, **16**, 7–22.
Itti, L., Koch, C., & Niebur, E. (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, **20**, 1254–1259.
Jin, D.Z., Fujii, N., & Graybiel, A.M. (2009) Neural representation of time in cortico-basal ganglia circuits. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 19156–19161.
Joyce, C. & Rossion, B. (2005) The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin. Neurophysiol.*, **116**, 2613–2631.
Kamienkowski, J.E., Ison, M.J., Quiroga, R.Q., & Sigman, M. (2012) Fixation-related potentials in visual search: a combined EEG and eye tracking study. *J. Vis.*, **12**, 4.
Kamienkowski, J.E., Varatharajah, A., Sigman, M., & Ison, M.J. (2018) Parsing a mental program: Fixation-related brain signatures of unitary operations and routines in natural visual search. *Neuroimage*, **183**, 73–86.
Kaunitz, L.N., Kamienkowski, J.E., Varatharajah, A., Sigman, M., Quiroga, R.Q., & Ison, M.J. (2014) Looking for a face in the crowd: fixation-related potentials in an eye-movement visual search task. *Neuroimage*, **89**, 297–305.
Lavie, N., Beck, D.M., & Konstantinou, N. (2014) Blinded by the load: attention, awareness and the role of perceptual load. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, **369**, 20130205.
Lebedev, S., Van Gelder, P., & Tsui, W.H. (1996) Square-root relations between main saccadic parameters. *Invest. Ophthalmol. Vis. Sci.*, **37**, 2750–2758.
Liu-Shuang, J., Norcia, A.M., & Rossion, B. (2014) An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, **52**, 57–72.
Luck, S.J. & Vogel, E.K. (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences*,.
Luria, R. & Vogel, E.K. (2011) Visual search demands dictate reliance on working memory storage. *J. Neurosci.*, **31**, 6199–6207.
Lundqvist, D., Flykt, A., & Öhman, A. (1998). Karolinska directed emotional faces. *Cognition and Emotion*.
Maris, E. & Oostenveld, R. (2007) Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods*, **164**, 177–190.
Meghanathan, R.N., van Leeuwen, C., Giannini, M., & Nikolaev, A.R. (2020) Neural correlates of task-related refixation behavior. *Vision Res.*, **175**, 90–101.
Melcher, D. & Colby, C.L. (2008) Trans-saccadic perception. *Trends Cogn. Sci.*, **12**, 466–473.
Mensen, A. & Khatami, R. (2013) Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics. *NeuroImage*,.
Otero-Millan, J., Troncoso, X.G., Macknik, S.L., Serrano-Pedraza, I., & Martinez-Conde, S. (2008) Saccades and microsaccades during visual fixation, exploration, and search: foundations for a

common saccadic generator. *J. Vis.*, **8**, 21.1–18.
Pinheiro-Chagas, P., Piazza, M., & Dehaene, S. (2019) Decoding the processing stages of mental arithmetic with magnetoencephalography. *Cortex*, **114**, 124–139.
Plöchl, M., Ossandón, J.P., & König, P. (2012) Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. *Front. Hum. Neurosci.*, **6**, 278.
Ries, A.J., Touryan, J., Ahrens, B., & Connolly, P. (2016) The Impact of Task Demands on Fixation-Related Brain Potentials during Guided Search. *PLoS One*, **11**, e0157260.
Rossion, B. & Jacques, C. (2011) *The N170: Understanding the Time Course of Face Perception in the Human Brain*. Oxford University Press.
Seidkhani, H., Nikolaev, A.R., Meghanathan, R.N., Pezeshk, H., Masoudi-Nejad, A., & van Leeuwen, C. (2017) Task modulates functional connectivity networks in free viewing behavior. *Neuroimage*, **159**, 289–301.
Smith, N.J. & Kutas, M. (2015) Regression-based estimation of ERP waveforms: II. Nonlinear effects, overlap correction, and practical considerations. *Psychophysiology*, **52**, 169–181.
Smith, S.M. & Nichols, T.E. (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage*, **44**, 83–98.
Theeuwes, J., Belopolsky, A., & Olivers, C.N.L. (2009) Interactions between working memory, attention and eye movements. *Acta Psychologica*,.
Thorpe, S.J. (2009) The Speed of Categorization in the Human Visual System. *Neuron*,.
Thorpe, S.J. & Fabre-Thorpe, M. (2001) Seeking Categories in the Brain. *Science*,.
Torralba, A., Oliva, A., Castelhano, M.S., & Henderson, J.M. (2006) Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*,.
Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B.E., Sperling, M., & Lega, B. (2020) Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 28463–28474.
Underwood, G. & Foulsham, T. (2006) Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology*,.
VanRullen, R. & Thorpe, S.J. (2001) The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.*, **13**, 454–461.
Wolfe, J.M. (2012) Saved by a log: how do humans perform hybrid visual and memory search? *Psychol. Sci.*, **23**, 698–703.
Yang, T. & Shadlen, M.N. (2007) Probabilistic reasoning by neurons. *Nature*, **447**, 1075–1080.