NeuroImage 89 (2014) 297–305

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task

Lisandro N. Kaunitz [a], Juan E. Kamienkowski [b,c], Alexander Varatharajah [a,d], Mariano Sigman [b], Rodrigo Quian Quiroga [a,d], Matias J. Ison [a,d,*]

[a] Department of Engineering, University of Leicester, Leicester LE1 7RH, UK
[b] Laboratorio de Neurociencia Integrativa, Departamento de Física, FCEyN UBA and IFIBA, Conicet, Pabellón 1, Ciudad Universitaria, 1428 Ciudad Autónoma de Buenos Aires, Argentina
[c] Universidad Diego Portales, Santiago, Chile
[d] Centre for Systems Neuroscience, University of Leicester, Leicester LE1 7QR, UK

ARTICLE INFO

Article history:
Accepted 6 December 2013
Available online 14 December 2013

Keywords:
Visual search
EEG
Faces
Natural scenes
Oddball

ABSTRACT

Despite the compelling contribution of the study of event related potentials (ERPs) and eye movements to cognitive neuroscience, these two approaches have largely evolved independently. We designed an eye-movement visual search paradigm that allowed us to concurrently record EEG and eye movements while subjects were asked to find a hidden target face in a crowded scene with distractor faces. Fixation event-related potentials (fERPs) to target and distractor stimuli showed the emergence of robust sensory components associated with the perception of stimuli and cognitive components associated with the detection of target faces. We compared those components with the ones obtained in a control task at fixation: qualitative similarities as well as differences in terms of scalp topography and latency emerged between the two. By using single trial analyses, fixations to target and distractors could be decoded from the EEG signals above chance level in 11 out of 12 subjects. Our results show that EEG signatures related to cognitive behavior develop across spatially unconstrained exploration of natural scenes and provide a first step towards understanding the mechanisms of target detection during natural search.

Introduction

A central goal in cognitive psychology and visual neuroscience is to understand how we perceive real-world scenes (for a review see Eckstein, 2011). Real-world scenarios typically include several salient features and thus natural vision involves sophisticated mechanisms to efficiently allocate foveal resources (Itti and Koch, 2000). Several processes, such as behavioral goals, motivational state, and the spatial properties of the visual scene, govern saccadic scan paths during free-viewing tasks. But can classical event related potentials (ERPs) be reliably measured during visual search tasks that involve complex and unconstrained spatial distributions of ocular trajectories? EEG recordings typically involve flashing stimuli at fixation to avoid the large artifacts that eye movements introduce in the ERPs. For this reason, the registration of EEG during eye-movement exploration tasks of natural scenes has been largely avoided in the past, posing a potential difficulty to the study of human vision in more ecological environments.

Recent reports have shown that it is feasible to concurrently record EEG and eye movements (Kamienkowski et al., 2012a, 2012b; Ossandón et al., 2010; Plöchl et al., 2012). However, in order to restrict eye-movement contaminations, these studies were run in much simplified scenarios compared to the exploration of natural scenes; namely they involved controlled saccade tasks (Brouwer et al., 2013; Dandekar et al., 2012; Kazai and Yagi, 1999; Thickbroom and Mastaglia, 1985; Thickbroom et al., 1991; Yagi, 1981), reading paradigms (Dimigen et al., 2011; Marton and Szirtes, 1988a, 1988b) or visual search tasks with artificial stimuli (Kamienkowski et al., 2012a, 2012b). To our knowledge, only two studies have focused on fixation event-related potentials (fERPs) during the free-viewing of natural images (Graupner et al., 2007; Ossandón et al., 2010). Due to the difficulty in obtaining long fixations (i.e. long EEG traces without contamination of eye movements), these studies did not deal with long latency components typically associated with cognitive processing.

In the present study we sought to understand the full range of events that unfold during the visual exploration of natural scenes. Subjects had to find a hidden target face in a crowded scene while we simultaneously recorded EEG and eye movements. Before the experiment we trained subjects to avoid making fixations of short duration while searching. In this way, we were able to obtain relatively long fixations during the experiment, which allowed the analysis of late cognitive components without contamination of eye movements. Moreover, we designed a

* Corresponding author at: Department of Engineering, University of Leicester, Leicester LE1 7RH, UK. Fax: +44 116 252 2619.
E-mail address: mji3@le.ac.uk (M.J. Ison).

1053-8119
http://dx.doi.org/10.1016/j.neuroimage.2013.12.006

fixed-gaze experiment akin to a classical oddball paradigm, in order to systematically compare the ERPs obtained in our visual search task with the classic ERPs of paradigms at fixation. This allowed a simple and direct way to compare fERPs and ERPs without the need of ocular correction methods (such as those based on Independent Component Analysis). Developing objective validation criteria (Ossandón et al., 2010; see also Dimigen et al., 2012) would be particularly problematic in this experiment since subjects performed saccades of any size and in any direction.

Our results show that known EEG signatures related to cognitive functions in fixed-gaze paradigms are also present in more ecological settings. Interestingly, we also show that a direct comparison between ERPs and fERPs yields differences in relation to their latency and topography. Furthermore, we show that the information contained in the cognitive fERPs can be used to discriminate target detection in single trials. Altogether, our work provides new insights into the dynamics of brain processes during visual exploration of natural scenes.

Materials and methods

Participants

Twelve subjects (10 males and 2 females, ages 21–31 years old) participated in the experiments. 10 subjects completed both tasks while 2 subjects completed only the eye-movement visual search task. All subjects were naive to the objectives of the experiment, had normal or corrected to normal vision and gave written informed consent according to the recommendations of the declaration of Helsinki to participate in the study.

Stimuli

The image database contained 60 gray scale images of crowds at stadiums downloaded from Internet or obtained at football stadiums. Images were 800 × 768 pixels in size and each one contained between 23 and 35 distractor faces (30.68 faces on average). From all the faces in each image 3 were chosen as targets. Images were made isoluminant in order to avoid areas of increased saliency.

Experimental procedures

Stimuli were presented on a 21″ Iiyama CRT monitor, with a screen resolution of 1024 × 768 pixels and a refresh rate of 75 Hz. Participants sat in a comfortable chair inside a darkened room at 60 cm from the screen, their heads stabilized via an in-house chin rest. All experiments were implemented in MATLAB (Mathworks, Natick, MA) using the Psychophysics toolbox (Brainard, 1997). Manual responses were collected with a standard keyboard. Ocular responses were obtained from the eye position of subjects via the on-line information provided by an eye tracker. During task execution in both experiments, we simultaneously recorded EEG and eye movements.

Visual search experiment

At the beginning of each trial subjects pressed the space bar and were presented with a target face for 3 s. On each trial we resized the original target face to a random value between 2 × 2° and 3 × 3°. This prevented subjects from using target size to guide their visual search strategy. After this time a fixation point was presented on the screen at a random location. Subjects needed to fixate at the new dot location for 1 s for the image of a crowd to appear on the screen. The subjects' 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. 1). 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. The target faces varied in size from 2° to 4° across trials to prevent subjects from making inferences about the face position on the following image presentation. In total subjects performed 180 trials (3 different targets per crowd image for the whole experiment). Before the experiment started, subjects were trained to search the target without rushing and gave them an indicative pace with a metronome clicking at 1 Hz. The metronome was *only used during their training session*, not during the actual experiment; *it* served to train subjects not to rush during the visual search. During the experiment, we provided subjects with visual feedback at the end of a trial only if they had produced less than 2 fixations of at least 0.5 s throughout the trial. The feedback consisted of the sentence "too fast" shown on a gray background screen. The rationale of this was to encourage larger fixation times in order to study late latency fERPs related to cognitive processes, as described below. The images used during the training sessions were not used during the experiment. On 59% of trials (980/1663 trials, over all subjects and experiments) subjects made at least 2 fixations of 0.5 s to distractors. In total, subjects made 1561 fixations to targets and 4655 fixations to distractors. Using these simple instructions we obtained: 1) longer fixations than in other visual search experiments, and 2) less redundancy between fixations (less number of repeated fixations on the same faces), very common when subjects are allowed to freely explore without any instruction. Both properties of the eye movements were important for the analysis: longer fixations opened the possibility to observe clean late evoked potentials and low redundancy prevented fixations in which the subject fixated at the target but might have not identified it.

Visual Oddball experiment

Subjects had to fixate at the center of the screen, where target and distractor images were flashed in pseudo-random order (Fig. 1, Supplementary Materials). From each image of crowds in the dataset we extracted 11 faces. We extracted these faces by cutting a rectangular area of 2 × 2°, keeping the target face original size. Before each trial we selected 1 of these images of faces as the target and the other 10 as distractors. In each trial the target face was presented before the beginning of the trial for 3 s as in the visual search experiment. The sequence of 11 faces was then presented, each one for 0.5 s with a random inter-stimulus interval in the range of 0.2–0.3 s. Subjects were asked to fixate constantly on the sequence of images. In total we presented subjects with 220 trials consisting in 3 blocks of 60 trials each and one final block of 40 trials. Between blocks subjects took 5 min resting breaks. The target was present in the sequence on 80% of the trials (180 trials). Targets appeared with the same frequency at any position from 2 to 11 in the sequence and were never presented as the first image in the sequence. The subjects' task was to report with a keyboard press at the end of the trial whether they had detected the target in the sequence of images.

Eye movements and EEG recordings

Eye movements were registered with an EYELINK 1000 system (SR Research, Ontario, Canada). The eye tracker was used in binocular mode with stabilized-head and sampling rate of 500 Hz in each eye. Saccades and fixations were detected using an adapted version of the velocity-based Engbert and Kliegl's algorithm (Engbert and Kliegl, 2003) using the parameters described in Kamienkowski et al. (2012a, 2012b). We only kept saccades larger than 1° for the analyses of the data. We considered as fixations to targets all those fixations that landed on an area of 2 × 2° of visual angle from the center of the target face. For all the experiments we ran drift corrections every 10 trials and a recalibration of the eye tracker every 60 trials (before the beginning of a new block). The nine-point calibration was kept with an average error below 1° (typically below 0.5°). EEG data were recorded on a standard 64-channel 10–20 montage using a Biosemi Active-Two System (Biosemi, Amsterdam, Holland) at 1024 Hz. Data was imported into MATLAB with EEGLAB toolbox (Delorme and Makeig, 2004) using