The P300 potential for fixations onto target object when exploring natural scenes during a visual task after denoising overlapped EFRP

Hélène Devillez*, Emmanuelle Kristensen*, Nathalie Guyader*, Bertrand Rivet* and Anne Guérin-Dugué*

Abstract

Electroencephalography (EEG) studies have largely reported the P300 Event Related Potential (ERP) elicited by target stimulus processing compared to non-target stimulus. These studies used constrain experimental paradigms during which participants did not move their eyes. However, during more ecological paradigms where participants can move their eyes, the P300 potential might be more difficult to identify due to the overlapped potentials elicited by consecutive ocular fixations. In this study, we use the xDawn algorithm for denoising the overlapped Eye-Fixation Related Potentials (EFRP), and we observe the P300 potential elicited on the first, but also on the consecutive subsequent fixations landed on the target stimuli.

I. Introduction

One of the most studied tasks is the visual search task for a specific target amongst distractors. The formulation of this task is easy to control, as well as the visual complexity of targets and distractors. This visual search task allows studying a specific Event Related Potential (ERP): the P300 potential. This potential is a positive component with a maximal amplitude in the centro-parietal region and a latency between 250 and 500 ms after the stimulus onset (for a review, see [1]). It is elicited when observers detect a rare stimulus among distractors [1, 2, 3]. Such potential is associated with discrimination, categorization, selection, matching processes and decision-making [4, 5]. This P300 potential is elicited only when participants are actively involved in the task (for example, when counting the number of occurrence of the rare stimulus) and are asked to take a final decision on the stimulus.

Experimental paradigms used in ERP studies limit the eye movements of participants asking them to not move their eyes because saccades and blinks create artefacts in the EEG signal. However in real-life situations, eyes are always moving to explore visual scenes. For few years, researchers have been interested in the co-registration of eye movements and EEG signals. This technic allows the analysis of neural activities, while having access to the time at which eye positions are inside the target object. Hence, potentials related to specific fixation onsets, i.e. Eye Fixation Related Potential (EFRP) are studied in function of the fixation positions in the scene and/or according to their rank during the exploration.

However with this methodology, two types of noise need to be taken into account, and even further corrected for EFRPs analysis. First, the EEG signal has to be corrected from the strong artefacts created by the eye movements: blinks and saccades. Second, due to the short Inter-Stimulus Intervals (ISI) between two successive fixations, a possible overlap in time can exist between the early potentials elicited by a fixation with the late potentials elicited by the previous fixations, and conversely between the late potentials with the early potentials elicited by the subsequent fixations.

For the ocular artifacts, authors classically apply an Independent Component Analysis (ICA) on the EEG signals. This correction is frequently used as a pre-processing before the EEG analysis and more specifically in EFRP analyses [6, 7, 8]. However such pre-processing does not account for the overlap problem of the potentials elicited by consecutive fixations. To our knowledge, only four studies have used the co-registration of eye movements and EEG signals to access the P300 potential when participants gazed at the target during visual search task. In these studies, authors did not use ICA to correct eye movements but they only analyzed the EEG signals for fixations with long duration [9, 10, 11, 12]. The authors reported a differentiating cortical activity between the detection of target and non-target objects between 480 and 500 ms after fixation onset in central, parietal and parieto-occipital regions, interpreted as a P300 potential. More recently, a study using real-world scenes also reported a late potential associated to P300 [12]. Participants were asked to find a target face in images of crowds. This target was presented before the scene presentation. Observers were trained before to make very long fixations (duration longer than 500 ms).

The overlap in time problem has been studied in the context of ERP experiments when the latencies of the neural potentials are in the same range than the ISI. An algorithm, called "Adjar" for "adjacent response algorithm", was proposed to estimate overlapped signals, and then, to correct the neural potential [13]. Up to now, this algorithm is not used for EFRP studies. A very recent study [14] has shown the difficulty of using this algorithm to correctly estimate EFRP noised by overlapping. The main reason comes from the temporal distributions of ISI called here Inter-Fixation Intervals (IFI) which are too "concentrated" in the observation window of the potential of interest. To cope with this difficulty, we have used another approach, with the xDawn algorithm [15].

The P300 potential was studied from EEG and eye tracking data co-registration during a visual search task using natural scenes. The EEG signal was preprocessed by ICA for the ocular artifacts correction. Unlike previous studies, we did not have any constraint concerning eye movements; participants moved their eyes making fixations with various durations; all the fixations of interest and their previous and subsequent fixations were considered by applying the xDawn algorithm.

Participants were asked to explore a scene and to answer if an object was or not present. When the object was present in the scene and participants correctly answered, we expected to observe the P300 potential from the first fixation onto the target object. We also compared the P300 potentials elicited

This work was supported by a grant from the ANR project "Gaze-EEG" 2009-2013, a grant from the French "Ministère de la Recherche et de l'Enseignement Supérieur" funding the PhD of H. Devillez and a grant from the CNRS funding the PhD of E. Kristensen.

*Affiliation: Gipsa-lab (Grenoble, Images Parole Signal Automatique), CNRS, Université Grenoble Alpes, 11 rue des Mathématiques, Grenoble Campus, BP46, F-38402 Saint Martin d'Hères Cedex

Preprint submitted to 7th International IEEE EMBS Conference on Neural Engineering. Received November 24, 2014.

by the first fixation onto the target object with the two previous outside and the two subsequent fixations inside the target. Two ways of neural potentials estimation are compared, without or with overlapping correction (classical EFRP vs xDawn).

II. Denoising overlappped EFRP by xDawn algorithm

The objective of the xDawn algorithm was the design of adapted spatial filters to maximize the signal-to-noise-ratio before classification. Brain Computer Interfaces are the basic application context of this algorithm. For the overlap problem, only the first part of the xDawn algorithm is used.

Let us consider an observation window time at the onset of a Fixation Of Interest (FOI). In this window, the observed neural activity $x_i(t)$ for the trial i can be modelled by:

$$x_i(t) = \sum_{k=1}^{K} s_i^k\left(t - \tau_i^k\right) + n_i(t), \qquad (1)$$

where $s_i^k(t)$ denotes the response time-locked on a particular FOI whose latency is τ_i^k, and $n_i(t)$ denotes the outgoing brain activity which is not related to fixations. K is the *a priori* number of different neural elicited responses supposed to occur during the observation window. After averaging across trials, equation (1) can be rewritten as:

$$x(t) = \sum_{k=1}^{K} s^k(t) * D^k(t) + n(t), \qquad (2)$$

with $*$ the convolution product, and $D^k(t)$ the distribution of latencies for each class of responses extracted from the fixation onsets. By least square minimization, one can estimate in a close form, the K average responses time-locked on the respective FOIs. See [14] for the implementation details.

III. Material and methods

A. Participants

Thirty-nine healthy adults participated in the experiment (22 female and 17 male; age range: 20-36; M=24.69; SD=3.49). Data of five other subjects were discarded from the analysis due to technical problems in data recordings. All participants had a normal or corrected-to-normal vision. The study was approved by the local French ethics committee of the "Pôle Grenoble Cognition"[1]. All participants gave their written and informed consents prior to the experiment.

B. Apparatus

Stimuli were displayed onto a 20-inch ViewSonic CRT monitor located 57 cm from participants, with a resolution of 768 × 1024 pixels and a refresh rate of 75 Hz. Scenes subtended 30 × 40 degrees of visual angle.

Eye movements were recorded using the Eyelink 1000 (SR Research) and sampled at 1000 Hz, for both eyes. Head was stabilized using a chin rest. A 9-point calibration routine was carried out at the beginning of each session and was repeated every 20 trials or when the drift correction, performed every 10 trials, reported a mean error above 0.5°.

The EEG activity was recorded using 32 Ag/ AgCl unipolar active electrodes positioned according to the extended 10-20 system [16]. The right earlobe and FCz electrodes were used respectively as reference and ground. Data were amplified using a g.GAMMAsys gtec system (g.tec, Inc.) and sampled at 1200 Hz. An analog band-pass filter (0.01-100 Hz) and a 50 Hz notch filter were applied online.

C. Stimuli

The stimuli consisted of 240 color pictures (various indoor and outdoor scenes). Scenes did not contain any person.

D. Experimental procedure

Participants performed four 20-minute sessions, but only the results for one session are discussed here[2]. Within each session, 60 scenes were randomly displayed. The experiment was designed with SoftEye software [17] to respect a precise temporal sequence for the stimuli presentation, to generate a trigger signal for synchronizing the two data flows (EEG and eye movements). Trials were composed of five successive displays. Each trial started with a question asking whether an object was or not present in the scene. Then a white central fixation cross was displayed for 800 to 1200 ms. When participant stabilized his/her gaze on the central fixation (stabilization for 100 ms in a square of 50 pixels around the central fixation), a scene was displayed for 4 s. The question was recalled after the scene display with the two possible answers, yes or no. Participant gave his/her answer by pressing on a mouse button. The response screen was displayed until the participant's answer. Each trial ended with a grey screen for 1 s. If the gaze was not stabilized, or if the participant did not correctly answer, the trial was not taken into account in the analysis.

IV. Data analysis

A. Data preprocessing

Saccades (and consequently fixations) were automatically detected by the Eyelink software using three thresholds: velocity (30 degrees/s), acceleration (8000 degrees/s²) and saccadic motion (0.15 °). We analyzed the data for the dominant eye of each participant and only fixations with a duration between 50 and 1000 ms.

Using EEGlab software™ [18], EEG data were re-sampled at 1000 Hz. With the same sampling rate, ocular and EEG data were synchronized offline using trigger signal. Noisy channels (T7, T8, TP9 and TP10) were detected by visual inspection and removed from analysis. EEG data were segmented into epochs from 500 ms before the scene onset to 4000 ms after. Segments were visually inspected offline and those containing muscular activity or non-physiological artifacts were rejected. Ocular artifacts were then corrected with a principal com-ponent analysis (number of channels minus one component) followed by an ICA (infomax ICA) [19]. Again, a visual ins-pection was performed. If ocular artifacts were not corrected, epochs were removed from the analysis. Finally, on average 53 ± 7.1 trials (maximum of 60 trials) were kept for analysis.

Fixations were off-line tagged according to their location. For each scene, a Region Of Interest (ROI) was defined by a bounding box around the target object. We tagged the *first*

[1] http://www.grenoblecognition.fr/index.php/ethique/ethique-et-cognition

[2] The whole experiment consisted of four different tasks, only the visual search task is reported here.

Preprint submitted to 7th International IEEE EMBS Conference on Neural Engineering. Received November 24, 2014.

fixation landed inside the ROI ("target"). We also tagged the two previous ones ("target-2", "target-1") which were *outside* the ROI and the two subsequent ones ("target+1", "target+2") which were *inside* the ROI. We finally considered five FOIs.

B. Temporal signal segmentation

To correct EFRP for overlap problem we considered a model with six classes of potentials ($K=6$): the different neural responses elicited by the five FOIs, and a sixth class gathering all the responses elicited by non-tagged fixations possibly inside the observation window, but spatially outside the ROI. These temporal windows were chosen inside the initial segments [- 500; 4000 ms], time-locked to the onsets of the "target" fixations (FOI onto the target). These windows have to be large enough to contain (i) the four other FOIs (two before and two after the FOI onto the target) and (ii) the possible P300 potential (that classically appears 250 to 500 ms after stimulus onset). Summing up the four means of each IFI (see Figure 1B) and considering a range of 600 ms to ensure the complete P300 potential observation, the size of the observation window was chosen: [-700; 1200 ms].

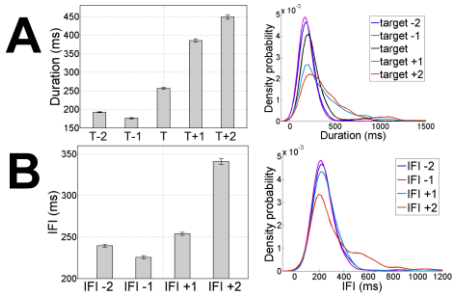

Figure 1. **A**. Mean fixation duration and distribution of fixation durations for the five FOIs in ms. **B**. Mean IFI and distribution of the four IFI in ms.

B. Data analysis

For EFRPs analysis, epochs of [-200; 600 ms], time-locked to the onsets of the five FOIs were extracted from the each observation window, for each participant. EEG activity during the period [-200; -100 ms] before the fixation onset was used for the baseline correction. For each participant, EEG epochs were averaged for each FOI and each electrode. The five grand averages were obtained by averaging across participants.

For xDawn analysis, for each participant, all the EEG signals inside the observation windows were modeled by equation (2) and neural potentials of each FOI were obtained at the algorithm output. These five potentials were described for a period of [-200; 600 ms], and the same baseline correction as for EFRPs, was applied. Then, the five grand averages were computed across participants.

We focused on the Cz electrode and on the time period of [250; 600 ms], which corresponded to the electrode and the time period where the P300 potential was generally observed in ERP and EFRP studies [20, 1, 11, 12].

For both EFRP and xDawn analyses, we calculated the mean "relative" amplitude of the P300 potential over the period [250; 600 ms]. We computed the "relative" amplitude rather than the raw amplitude because differences were observed between the conditions at the beginning of the analyzed period, around 250 ms (Figure 3). A reference period was used to estimate the mean amplitude of the signal for each participant, each FOI and each analysis (classical EFRP and xDawn). Finally, this interval was averaged across participants, for each condition and each algorithm. The times for the xDawn version are written in parentheses for the five FOIs: "target-2": [255(285); 315(320) ms]; "target-1": [250(250); 270(285) ms]; "target": [260(270); 280(285) ms]; "target+1": [285(285); 305(310) ms]; "target+2": [255(230); 300(300) ms]. The "relative" amplitude of the P300 wave was the difference between the raw amplitude and the mean amplitude computed over the reference time interval.

V. Results

All comparisons were done using a repeated measures ANOVA with the *Condition* ("target -2", "target-1", "target", "target+1" and "target+2") and the *Data Type* (EFRP data or xDawn data) (for the EFRP analysis only) as within-subject factors. Multiple comparisons were assessed with Bonferroni post-hoc tests. For all statistical analyses the significant level of tests was set at 0.05.

A. Correct answers

Participants successfully performed the task (84.50 ± 0.20 % of correct answers).

B. Eye movement data

We calculated the mean fixation durations for the five FOIs and the mean of the four IFIs ("IFI-2" between "target-2" and "target-1"; "IFI-1" between "target-1" and "target"; "IFI+1" between "target" and "target+1" and "IFI+2" between "target+1" and "target+2") (Figure 1).

Fixation duration: The analysis revealed that the *Condition* significantly influenced the fixation duration ($F(4,152) = 927.88$, $p<.001$) (Figure 1A). Specifically, we observed an increase of the fixation duration between previous fixations and the others: "target-2" and "target-1" fixations were significantly shorter that "target", "target+1" and "target+2" fixations ($p<.005$).

Inter-stimuli intervals: The analysis revealed that the *Condition* significantly influenced the IFI between fixations ($F(3,114)=2045,2$, $p<.001$) (Figure 1B). Like for fixation durations, we observed a significant increase of IFI along the exploration from "IFI-1" to "IFI+2" ($p<.001$).

C. EFRPs

On Figure 2, the topographic maps represent the amplitude of the EFRP signal on the period [0; 600 ms] from the onset of the fixation for the five FOIs, for xDawn estimations. The grand averages of the each potential are represented on Figure 3A and Figure 3B, for the EFRP waves (dotted lines) and for the xDawn ones (solid lines).

The analysis revealed that the *Condition* significantly influenced the relative integrated amplitude of the P3 wave ($F(4,152)=5.55$; $p<.001$). The interaction *Condition* × *Data Type* was also significant ($F(4,152)=6.20$; $p<.001$). No significant effect was observed for the Data Type ($F(1,38)=0.32$; ns) (Figure 4). Specifically for EFRP data, this amplitude was higher for "target" compared to "target-2"

(p<.005), to “target +1” (p<.05) and to “target +2” (p<.001) and also higher for “target-1” compared to “target+2” (p<.001). For xDawn data, this amplitude was higher for “target” compared to “target-2” (p<.005) and “target-1” (p<.05). We observed that the xDawn overlapping correction specifically impacted the amplitude for “target-1” and “target+2” conditions. This amplitude was lower for xDawn data compared to EFRP for “target-1” (p=.05) but on the contrary, it was higher for xDawn data compared to EFRP for “target+2” (p<.001).

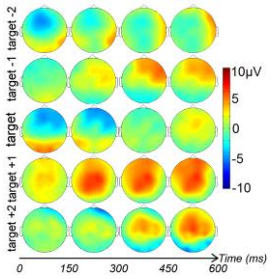

Figure 2. Topographic maps between 0 and 600 ms from the fixation onset for xDawn data.

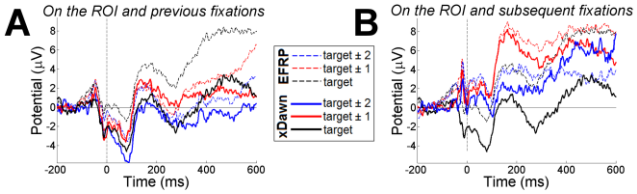

Figure 3. A. EFRPs for the target fixation and the two previous fixations and B. EFRPs for the target fixation and the two subsequent fixations: EFRP data (dotted lines) and xDawn data (solid lines).

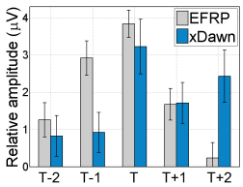

Figure 4. Mean relative amplitude of the P300 wave in [250; 500 ms] for the target fixation, the two previous fixations and the two subsequent fixations: EFRP data (grey) and xDawn data (blue).

VI. Discussion and Conclusion

Due to overlaps, the amplitude of late potentials estimated by the classical EFRP analysis is retained in time, while they are clearly identified in time thanks to the xDawn estimation. The topographic maps show this potential that starts around 250 ms with a maximal amplitude around 550 ms in the parieto-central region. This potential identified as the P300 component is elicited by the visual input on the *consecutive fixations landed on the ROI* during the visual search task. This result confirms ERP studies on the P300 potential, but it has been obtained on ecological experimental contexts thanks to eye tracking and EEG data co-registration, and a careful denoising of overlapped observed ERP signals.

Acknowledgment

A part of the soft development was realized by R. Phlypo, N. Tarrin, A. Andreev and G. Ionescu.

References

[1] J. Polich, "Neuropsychology of P300," *Oxford handbook of event-related potential components,* pp. 159-188, 2012.

[2] S. Sutton, M. Braren, J. Zubin and E. R. John, " Evoked-potential correlates of stimulus uncertainty," *Science,* vol. 150, no. 3700, pp. 1187-1188, 1965.

[3] E. Donchin, "Surprise!… surprise?," *Psychophysiology,* vol. 18, no. 5, pp. 493-513, 1981.

[4] T. W. Picton, "The P300 wave of the human event-related potential," *Journal of clinical neurophysiology,* vol. 9, no. 4, pp. 456-479, 1992.

[5] T. Hruby and P. Marsalek, " Event-related potentials-the P3 wave," *Acta Neurobiologiae Experimentalis,* vol. 63, no. 1, pp. 55-63, 2002.

[6] S.-T. Graupner, B. M. Velichkovsky, S. Pannasch and J. Marx, "Surprise, surprise: Two distinct components in the visually evoked distractor effect," *Psychophysiology,* vol. 44, no. 2, pp. 251-261, 2007.

[7] J. Ossandón, A. Helo, R. Montefusco-Siegmund and P. Maldonado, "Superposition model predicts EEG occipital activity during free viewing of natural scenes," *The journal of Neuroscience,* vol. 30, no. 13, pp. 4787-4795, 2010.

[8] A. R. Nikolaev, C. Nakatani, G. Plomp, P. Jurica and C. van Leeuwen, "Eye fixation-related potentials in free viewing identify encoding failures in change detection," *Neuroimage,* vol. 30, no. 13, pp. 1598-1607, 2011.

[9] G. Healy and A. F. Smeaton, "Eye fixation related potentials in a target search task," in *Engineering in Medicine and Biology Society, EMBC, Annual International Conference of the IEEE*, IEEE, 2011, pp. 4203-4206.

[10] J. Kamienkowski, M. Ison, R. Quian Quiroga and M. Sigman, "Fixation-related potentials in visual search: a combined EEG and eye tracking study," *Journal of Vision,* vol. 12, no. 7, pp. 1 - 20, 2012.

[11] A. M. Brouwer, B. Reuderink, J. Vincent, M. A. van Gerven and J. B. van Erp, "Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials," *Journal of vision,* vol. 13, no. 3, p. 17, 2013.

[12] L. N. Kaunitz, J. E. Kamienkowski, A. Varatharajah, M. Sigman, R. Q. Quiroga and M. Ison, "Looking for a face in the crowd: Fixation-related potentials in an eye-movement visual search task," *NeuroImage,* vol. 89, pp. 297-305, 2014.

[13] M. G. Woldorff, "Distorsion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction," *Psychophysiology,* vol. 30, pp. 98-119, 1993.

[14] E. Kristensen, A. Guerin-Dugué and B. Rivet, "Comparison between Adjar and xDawn algorithms to estimate Eye-Fixation Related Potentials distorded by overlapping," submitted to the 7th Int. IEEE/EMBS Conf. on Neural Engineering.

[15] B. Rivet, A. Souloumiac, V. Attina and G. Gibert, "xDAWN algorithm to enhance evoked potentials: application to brain–computer interface," *Biomedical Engineering, IEEE Transactions on,* vol. 56, no. 8, pp. 2035-2043, 2009.

[16] H. Jasper, "The ten-twenty electrode system of the International Federation," *Electroencephalogr Clin Neurophysiol ,* vol. 10, p. 371–375, 1958.

[17] G. Ionescu, N. Guyader and A. Guérin-Dugué, "SoftEye software.," *IDDN. FR,* vol. 1, no. 000, 2009.

[18] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods,* vol. 134, no. 1, pp. 9 - 21, 2004.

[19] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation,* vol. 7, no. 6, pp. 1129-1159, 1995.

[20] A. P. Fornaryova Key, G. O. Dove and M. J. Maguire, "Linking brainwaves to the brain: an ERP primer," *Developmental Neuropsychology,* vol. 27, no. 2, pp. 183-215, 2005.