Estimation of overlapped Eye Fixation Related Potentials: The General Linear Model, a more flexible framework than the ADJAR algorithm

Emmanuelle Kristensen
Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble France
CNRS, GIPSA-Lab, F-38000 Grenoble France
Univ. Grenoble Alpes, GIPSA-Lab, 11 rue des Mathématiques
Grenoble Campus, BP 46, 38000 Grenoble France

Bertrand Rivet
GIPSA-Lab

Anne Guérin-Dugué
GIPSA-Lab

The Eye Fixation Related Potential (EFRP) estimation is the average of EEG signals across epochs at ocular fixation onset. Its main limitation is the overlapping issue. Inter Fixation Intervals (IFI) - typically around 300 ms in the case of unrestricted eye movement- depend on participants' oculomotor patterns, and can be shorter than the latency of the components of the evoked potential. If the duration of an epoch is longer than the IFI value, more than one fixation can occur, and some overlapping between adjacent neural responses ensues. The classical average does not take into account either the presence of several fixations during an epoch or overlapping. The Adjacent Response algorithm (ADJAR), which is popular for event-related potential estimation, was compared to the General Linear Model (GLM) on a real dataset from a conjoint EEG and eye-tracking experiment to address the overlapping issue. The results showed that the ADJAR algorithm was based on assumptions that were too restrictive for EFRP estimation. The General Linear Model appeared to be more robust and efficient. Different configurations of this model were compared to estimate the potential elicited at image onset, as well as EFRP at the beginning of exploration. These configurations took into account the overlap between the event-related potential at stimulus presentation and the following EFRP, and the distinction between the potential elicited by the first fixation onset and subsequent ones. The choice of the General Linear Model configuration was a tradeoff between assumptions about expected behavior and the quality of the EFRP estimation: the number of different potentials estimated by a given model must be controlled to avoid erroneous estimations with large variances.

Introduction

Seminal works (Gastaut, 1951); (Barlow & Ciganek, 1969) combined electrooculography (EOG) and electroencephalography (EEG) to gain understanding of the consequences of eye movements on EEG activity. From their work, the Eye-Fixation Related Potential (EFRP) technique was developed to provide greater insight into mechanisms related to eye movements, and the time course of the continuous cognitive processing involved in experi-

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

mental tasks. This technique requires joint electroencephalography (EEG) and eye-tracking acquisition. Its fundamental difference from the popular Event-Related Potential (ERP) technique is that the neural response extracted is synchronized with ocular fixation rather than with the onset of a stimulus. In the context of a visual exploration experiment, for example, the ERP is the neural response that is time-locked with image onset, whereas the EFRP is the neural response elicited at fixation onset. When the EFRP technique is employed, the cognitive processes involved and their timelines are explored by encoding visual information through tasks of greater ecological validity (everyday tasks). During a classical ERP experiment, participants are required to gaze at a given location on the screen to avoid eye movement artifacts, whereas in an EFRP experiment they can explore the visual scene freely. During a reading experiment using the ERP approach, for instance, the text is displayed word by word on the screen in the same location to avoid ocular artefacts, whereas in the EFRP approach, the whole text is displayed at the same time and participants can move their eyes freely (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011), (Frey, et al., 2013).

EFRP estimation is based on the average across epochs at fixation onset. Because the Signal to Noise Ratio is low, the visual task is repeated several times to provide a sufficient number of epochs. However, the average obtained leads to an unbiased EFRP estimation if, and only if, a single potential is evoked inside an epoch. This condition cannot often be fulfilled when estimating EFRP. Inter-Stimuli Intervals can be controlled with ERP technique. By contrast, Inter-Fixation Intervals (IFI), which are the sum of the duration of a current fixation and of the subsequent saccade, depend on the oculomotor pattern of each participant. IFI duration (typically around 300 ms) can be shorter than the latency of the components of the evoked potentials. Overlap between several evoked potentials inside a single epoch is, therefore, a major limitation of the EFRP technique (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011); (Baccino, 2011); (Nikolaev, Meghanathan, & van Leeuwen, 2016). Different strategies have been adopted to address this issue:

- In (Kaunitz, et al., 2014), participants were trained to fix the target for one second. In this way, only one fixation occurred per epoch. However, this strategy reduces the ecological validity of the protocol and can only be used in specific experiments.

- In (Brouwer, et al., 2013) in a visual search task, and (Ries, Touryan, Ahrens, & Connolly, 2016) in a guided search task, fixations shorter than 500 ms were excluded from the analysis on the P300 component (an epoch lasted 500 ms to be free of eye movements). In (Kamienkowski, Ison, Quiroga, & Sigman, 2012) a similar selection was implemented for the same reason, with a threshold of 550 ms, in a visual search task. Only a part of all epochs recorded is used to estimate the EFRP. Data loss is a drawback in this procedure, which is an ad-hoc procedure designed to avoid eye movement within the latency window of the component of interest (component P300 in the three studies above).

- The matching of eye movement characteristics is the most common strategy (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011); (Devillez, Guyader, & Guérin-Dugué, 2015); (Dias, Sajda, Dmochowski, & Parra, 2013); (Fischer, Graupner, Velichkovsky, & Pannasch, 2013); (Kamienkowski, Ison, Quiroga, & Sigman, 2012); (Nikolaev, Meghanathan, & van Leeuwen, 2016). This technique compares ocular data that are similar in size, direction, duration, etc. It allows the distortions due to overlap between different experimental conditions to be counterbalanced, but not corrected.

- In addition to the matching technique, Dias and colleagues used a subtraction technique to correct distortions due to saccade response overlap (Dias, Sajda, Dmochowski, & Parra, 2013). Dandekar and colleagues (Dandekar, Privitera, Carney, & Klein, 2012) also developed a well-known linear regression method, the General Linear Model (GLM) (Kiebel & Holmes, 2003); (Dale, 1999) to extract the potentials elicited by ocular saccades of different sizes and orientations which were affected by the overlapping issue.

- In (Kristensen, Guerin-Dugué, & Rivet, 2017), a regularized GLM was studied and compared to the classical estimation by averaging, and was evaluated to estimate EFRP during the free exploration of visual scenes, irrespective of fixation rank.

EFRP estimation is the central question in this study. It requires the correction of distortions due to response overlaps elicited by adjacent fixations. We addressed this question from two perspectives.

In the first of these, we compared three classical algorithms in a common framework where identical evoked

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

potentials overlapped. In order to do so, we compared these three approaches using real data from joint EEG and eye-tracking recordings during a free visual exploration experiment: (1) the classical estimation by averaging time-locked EEG signals, (2) the popular Adjacent Response (ADJAR) algorithm (Woldorff, 1993), developed for the ERP technique and (3) the GLM configured to deconvolve evoked potentials with temporal overlaps. These algorithms were compared to estimate the potentials elicited by fixations in the middle of visual exploration, with the assumption that potential was the same irrespective of fixation rank. Moreover, in accordance with this aim, these algorithms were compared on their ability to deconvolve identical overlapped potentials, and not on their ability to estimate different kinds of potential per epoch.

In the second perspective, based on these results, two case studies focused on the GLM in order to choose the best match between GLM configuration and targeted objectives. Unlike in the first perspective, the GLM was used to estimate not only a unique potential, but two or three different potentials. The two case studies deal with the estimation (i) of the EFRP at the first fixation onset and (ii) of the ERP at image onset. In methodological studies (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011), (Nikolaev, Meghanathan, & van Leeuwen, 2016), specifying the first EFRP was recommended for at least two reasons: (1) differences in ocular features, (2) influences of neural activity at stimulus onset. In line with these recommendations, in the first case-study, a three-class GLM was implemented to estimate the first EFRP compared to the potential elicited by the second and following fixations. In this case-study, the potential elicited by image presentation was also included in the model since this potential overlaps the EFRP at the very beginning of exploration. The potential evoked at image presentation was estimated in the second case study. We were interested in the impact of the differentiation of the first fixation on the estimation of the ERP at image onset. To study this, the three-class GLM was contrasted with the two-class GLM, in which only the ERP at image onset and the EFRP irrespective of fixation rank were estimated. The aim of these two case-studies was to show that an efficient estimation by GLM is produced by a tradeoff between assumptions of expected behavior and the model's parsimony: the number of different potentials (i.e. number of classes for the GLM configuration) estimated by a GLM must be carefully set to avoid a large estimation variance. It is the expression of the classical tradeoff between bias and variance.

Materials and Methods

In this Section, the methods used to estimate the EFRP are detailed, then a description of our experimental data follows.

Eye Fixation Related Potential Estimation

We present three methods of estimating EFRP: the classical average of epochs, known as the Average method, the ADJAR algorithm in the context of EFRP and the GLM. To avoid confusion, we define here the main terms of the description of algorithms. Depending on the objective of the EFRP study, a temporal interval of interest, called the *window of interest*, has to be decided. This interval can, for example, be between -200 ms and 600 ms to include activities before fixation onset (saccadic potentials), and also to include activity for early and late components. In order to do so, EFRP is estimated during the *estimation window,* which may be either the same interval or a larger one which includes the window of interest, depending on the estimation method, as described below. EFRP estimation is based on a set of epochs time-locked on the event of interest (stimulus presentation, fixation onsets, or saccade onsets). The time interval of the epoch (*epoch window*) therefore has to be defined in relation to the estimation window. The choice of the other two intervals (estimation window and epoch window) will be explained for each of the three methods in accordance with the window of interest.

Average estimation and overlapping issue

Let us consider the given i^{th} fixation during a trial. The observed neural response $x_i(t)$ time-locked on this i^{th} fixation onset can be written as:

$$x_i(t) = a(t) + n_i(t) \tag{1}$$

where $a(t)$ is the potential evoked at this fixation onset, and $n_i(t)$ a noise corresponding to ongoing brain activity. The estimation of EFRP $a(t)$ by averaging is based on the implicit assumption that the neural potential $a(t)$ is the same for each eye fixation. For each repetition, the signal is time-locked at fixation onset, segmented into epochsand then averaged. The temporal epoch window must be long enough to include the latency and the whole temporal evolution of the potential of interest. The neural response $x_i(t)$ is observed throughout the epoch. The EFRP $a(t)$ is estimated by averaging all epochs. Thus, for the estimation by averaging, there is no distinction between the temporal

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

epoch window, the estimation window and the window of interest. The number of epochs must be large enough to significantly cancel out ongoing brain activity and to increase the Signal to Noise Ratio. Let E denote the number of epochs[1]. In this way, $a(t)$ is estimated by averaging across these epochs:

$$\hat{a}^{Av}(t) = \frac{1}{E}\sum_{i=1}^{E} x_i(t). \tag{2}$$

With equation (1), only a single evoked potential per epoch is taken into account. Consequently, this model provides an unbiased estimator if IFI duration is longer than the latency window of the components of interest in the evoked potential. If this condition is not fulfilled, more than one fixation occurs inside the same epoch, there is some overlapping between adjacent responses and the estimation is biased. To take these overlaps into account in a linear way, equation (1) is rewritten as:

$$x_i(t) = a(t) + a_{-1}\left(t - \tau_i^{(-1)}\right) + \cdots \tag{3}$$
$$+ a_{-Fpi}\left(t - \tau_i^{(-Fpi)}\right)$$
$$+ a_1\left(t - \tau_i^{(1)}\right)$$
$$+ \dots + a_{Fsi}\left(t - \tau_i^{(Fsi)}\right) + n_i(t)$$

where F_{pi} (respectively F_{si}) is the number of previous (resp. subsequent) fixations which occur during the ith epoch. $a_f(t)$ is the potential elicited at the fth fixation's onset ($f > 0$ means subsequent and $f < 0$ means previous). $\tau_i^{(f)}$ is the timestamp of the fth fixation: $\tau_i^{(f)} < 0$, and $f = -1, \dots, -F_{pi}$ for previous fixations and $\tau_i^{(f)} > 0$, and $f = 1, \dots, F_{si}$ for subsequent fixations inside the ith epoch. By averaging across the epoch, equation (3) highlights a summation of convolution products (*) between the evoked potentials and the normalized distribution of the timestamps $\left\{\tau_i^{(f)}, i = 1..E\right\}$:

$$\hat{a}^{Av}(t) = a(t) + Q_{-1}(t) * a_{-1}(t) + \cdots \tag{4}$$
$$+ Q_{-F_p}(t) * a_{-F_p}(t)$$
$$+ Q_{+1}(t) * a_{+1}(t) + \cdots$$
$$+ Q_{+F_s}(t) * a_{+F_s}(t) + n(t)$$

where $Q_f(t) = \frac{1}{E}\sum_{i=1}^{E} \delta(t - \tau_i^{(f)})$ is the normalized[2] distribution of the timestamps at the onset of the fth rank fixation and $n(t)$ is the average of ongoing brain activity ($n(t) = \frac{1}{E}\sum_i n_i(t)$). $-F_p = \min_i(-F_{pi})$ (respectively $F_s = \max_i F_{si}$) is the minimum (respectively maximum) rank of the previous (resp. subsequent) fixations across all epochs. This last equation can be recast in the following compact form:

$$\hat{a}^{Av}(t) = a(t) + \sum_{\substack{f=-Fp \\ f \neq 0}}^{F_s} Q_f(t) * a_f(t) + n(t) \tag{5}$$

$$\hat{a}^{Av}(t) = a(t) + \sum_{f=-1}^{-F_p} ov_{p,f}(t) + \sum_{f=+1}^{+F_s} ov_{s,f}(t) + n(t)$$

$$ov_{p,f}(t) = Q_f(t) * a_f(t), f \leq -1$$

$$ov_{s,f}(t) = Q_f(t) * a_f(t), f \geq 1$$

where $ov_{p,f}(t)$ (resp. $ov_{s,f}(t)$) is called the previous (resp. subsequent) response overlaps. This equation highlights the overlapping issue: i.e. the term $\sum_{\substack{f=-Fp \\ f \neq 0}}^{F_s} Q_f(t) * a_f(t)$ does not necessarily decrease to 0 even if E increases to infinity.

The next two subsections present the ADJAR and GLM methods designed to deal with the overlapping issue.

Estimation by the ADJAR algorithm

In the context of ERP estimation time-locked on external events, Woldorff proposed the ADJAR algorithm (Woldorff, 1993) to correct distortion of the ERP due to the overlap from temporally adjacent responses. In this article, the algorithm is described in the context of EFRP estimation.

The underlying assumptions are:

1. Each evoked potential is identical, irrespective of fixation rank (i.e. $a_{-1}(t) = a_{+1}(t) = a(t)$, with $a(t)$, the underlying response);

---

[1] The number of epochs is the number of fixation times the number of trials, or simply the number of trials, if only one fixation of interest occurs per trial, such as for example a fixation on a target during a visual search task.

[2] The normalized distribution is defined by the contribution of each timestamp τ_i^f divided by E. $\delta(t)$ is the Dirac impulse.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

2. During each epoch, only overlaps from immediately adjacent fixations are considered (i.e. $|f| = 1$). In other words, during each epoch, overlaps from earlier and later fixations are not considered (i.e. when $|f| > 1$);

3. At the second-order, the amplitude of higher order adjacent responses is already considered to be negligible ($Q_{+1}(t) * Q_{+1}(t) * a(t)$, $Q_{-1}(t) * Q_{-1}(t) * a(t)$). See below for details;

4. Noise due to ongoing activity is not considered.

Based on these assumptions, equation (4) becomes:

$$\hat{a}^{Av}(t) = a(t) + ov_p(t) + ov_s(t) \qquad (6)$$

where $ov_p(t) = ov_{p,-1}(t)$ and $ov_s(t) = ov_{s,+1}(t)$ are called the first-order, previous, and subsequent response overlaps, respectively. It is important to note that Q_{-1}(t) and Q_{+1}(t) are normalized (by the number of epochs): for each of them the integral of their contribution is equal to one. The distribution Q_{-1}(t) (resp. Q_{+1}(t)) is computed from the timestamps of the previous (resp. subsequent) fixation extracted from each epoch. The epoch must be sufficiently large to include both the previous and the subsequent fixation onsets. The estimation window for the ADJAR algorithm must be slightly larger than the window of interest, to take into account the border effects of the iterative procedure using convolutions. From a practical point of view, a tapering window is applied at each iteration step to manage these border effects. Schematic illustrations of these windows, describing the ADJAR algorithm implementation in this study, are shown in the appendix.

From the distributions Q_{-1}(t) and Q_{+1}(t), the first-order response overlaps $ov_p(t)$ and $ov_s(t)$ are estimated iteratively. At each iteration, these estimates are then subtracted from the average $\hat{a}^{Av}(t)$, to progressively improve the estimation of $a(t)$. After the kth iteration, the estimations of response overlaps are updated as:

$$\widehat{ov}_s^k(t) = Q_{+1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_p^{k-1}(t) \right) \qquad (7)$$

$$\widehat{ov}_p^k(t) = Q_{-1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_s^k(t) \right) \qquad (8)$$

with $\widehat{ov}_p^0(t) = 0$, as the initialization for the first iteration. The estimation of $a(t)$ after the kth iteration is given by:

$$\hat{a}^k(t) = \hat{a}^{Av}(t) - \widehat{ov}_p^k(t) - \widehat{ov}_s^k(t) \qquad (9)$$

The iterative procedure stops when the estimate $\hat{a}^k(t)$ no longer changes between two consecutive iterations. Let $\hat{a}^{Adjar}(t)$ denote the estimate given at the last iteration. The development of the earlier iterations is shown in the appendix. As a result, the final estimation is equal to:

$$\hat{a}^{Adjar}(t) = a(t) - ov_{ss}(t) - ov_{pp}(t) \qquad (10)$$
$$+/- \sum high\ order\ overlaps$$

where $ov_{pp}(t) = Q_{-1}(t) * Q_{-1}(t) * a(t) = Q_{-1}(t)^{(*2)} * a(t)$ and $ov_{ss}(t) = Q_{+1}(t)^{(*2)} * a(t)$ denote second-order response overlaps. Other terms appear in the iterative development of $\hat{a}^{Adjar}(t)$, for example: $Q_{-1}(t)^{(*k)} * Q_{+1}(t)^{(*k)} * a(t)$. They are called higher order response overlaps, and contribute positively or negatively to the final estimation. The ADJAR algorithm estimates and iteratively removes first-order response overlaps. Thus only second and higher order response overlaps remain. However, upon successive convolutions, higher order response overlaps are naturally progressively shifted outside the estimation window as the iteration number increases. Consequently, convergence is confirmed if convolution shifts the distributions of the second-order terms ($Q_{-1}(t)^{(*2)}$ and $Q_{+1}(t)^{(*2)}$) far enough outside the estimation window to render them negligible for all following iterations. In other words, convergence towards the underlying potential $a(t)$ is only possible if distortions in the measurements are due solely to first-order response overlaps. This assumes that the second-order response overlaps are negligible. These conditions are in line with assumptions 2 and 3 presented above. However these conditions may or may not be fulfilled, as they depend on the temporal ranges of the distributions $Q_{-1}(t)$ and $Q_{+1}(t)$ in relation to the epoch duration. This condition will be discussed in the context of EFRP in the section entitled "Estimation of EFRP during exploration". Assumptions 2 and 3 are, moreover, directly linked to convergence towards an unbiased estimation of the evoked potential. In the appendix, we show how these assumptions constitute strong limitations of the ADJAR algorithm. A modified ADJAR algorithm was designed to take into account previous and subsequent events within an epoch, but the iterative procedure did not converge.

The ADJAR algorithm is a trade-off between the accuracy of the model and the quality of the estimation. On the one hand, using only first-order response overlaps provides a convergent algorithm leading to a biased estimator of the response overlaps. On the other hand, taking into

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

account all contributions to ensure better modeling of the response overlaps leads to a divergent algorithm.

Estimation by the General Linear Model

The GLM (Kiebel & Holmes, 2003) (Dale, 1999) was proposed to manage the overlapping issue, taking into account all fixations that occur during an epoch.

Equation (3) can be rewritten in a matrix form:

$$\forall i \in \{1, \dots, E\}, \qquad x_i = \sum_{f=1}^{F_i} D_{f,i} \cdot a_f + n_i \qquad (11)$$

where x_i represents EEG samples observed during the ith epoch, $x_i = [x_i(1), \dots, x_i(N_e)]^\dagger \in \mathbb{R}^{N_e}$ with N_e the number of samples and $n_i = [n_i(1), \dots, n_i(N_e)]^\dagger \in \mathbb{R}^{N_e}$ with $.^\dagger$ the transpose operator. $a_f \in \mathbb{R}^{N_{a_f}}$ is the vector of the response time-locked on the onset of the fth fixation and N_{a_f} is the length of the response $a_f(t)$. Each potential $a_f(t)$ is estimated within an estimation window equal to the chosen window of interest. F_i is the number of fixations inside the epoch. Finally, $D_{f,i} \in \mathbb{R}^{N_e \times N_{a_f}}$ is the Toeplitz matrix for the fth fixation and the ith epoch, defined by its first column with entries that are all equal to zero except $\tau_i^{(f)}$ which is equal to one, with $\tau_i^{(f)}$ the onset of fth fixation in the ith epoch. Consequently, all epochs can be concatenated to obtain:

$$x = \sum_{f=1}^{F_{all}} D_f a_f + n \qquad (12)$$

with x equal to $[x_1^\dagger, \dots, x_E^\dagger]^\dagger \in \mathbb{R}^N$ and D_f equal to $[D_{f,1}^\dagger, \dots, D_{f,E}^\dagger]^\dagger \in \mathbb{R}^{N \times N_{a_f}}, f = 1 \dots F_{all}$, where N is the total number of samples and $N = N_e \times E$ and F_{all} the total number of fixations.

Equation (12) can be recast as:

$$x = D a + n \qquad (13)$$

with D equal to $[D_1, \dots, D_{F_{all}}]$ and a is the concatenation of the evoked potentials such as $a = [a_1^\dagger, \dots, a_{F_{all}}^\dagger]^\dagger$. The solution given by the least square minimization is:

$$\hat{a}^{Glm} = (D^\dagger D)^{-1} D x \qquad (14)$$

where $\hat{a}^{Glm}$ is the concatenation of all estimates, such as $\hat{a}^{Glm} = [\hat{a}_1^{Glm\,\dagger}, \dots, \hat{a}_{F_{all}}^{Glm\,\dagger}]^\dagger$.

The GLM configuration is presented with F_{all} fixation classes according to their rank inside the epoch. However, the fixations can be grouped in different ways depending on the assumptions made about the differentiation of the potentials evoked by consecutive fixations during the experimental task.

It is well known that the quality of estimation by the GLM is linked to the condition number of the $D^\dagger D$ matrix that has to be inverted to obtain the final estimation (equation (14)). The higher the condition number is, the more singular the $D^\dagger D$ matrix is, and consequently the less accurate the estimation. Conversely, the condition number is small if $D^\dagger D$ is a diagonal dominant matrix. Using GLM, Bardy and colleagues (Bardy, Dillon, & Dun, 2014) showed that the condition number of the $D^\dagger D$ matrix was linked to the amount of jitter on the Inter Stimuli Intervals in their experiments using sequences of auditory stimuli. In the context of our study, this result applies to the amount of jitter between the timestamps $\{\tau_i^f, i = 1 \dots E\}$ of events between each class.

Two cases will be considered in the following sections: the estimation of EFRP in the middle of visual exploration, and that of evoked potentials at the beginning.

Experimental data

Real data was used from an experiment on the free exploration of natural scenes in order to evaluate EFRP estimation using the three methods described in the previous subsection. Three conditions were included in the experiment: free exploration, image categorization, and visual search (Devillez, Guyader, & Guérin-Dugué, 2015). However, for the purpose of this study, only the condition involving free exploration was used.

Participants

Thirty-nine healthy adults participated in the experiment (22 women and 17 men; age range: 20-36; mean = 24.69; std = 3.49). The data from five other participants were removed because of technical problems during the recording process. All participants had normal or corrected-to-normal vision. The study was reviewed and approved by the local French ethics committee of the "Pôle Grenoble Cognition". All participants gave their written and informed consent prior to the experiment.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Apparatus

Visual scenes were displayed on a 20-inch ViewSonic CRT monitor located 57 cm from the 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 with a video-based infrared eye-tracking system (EyeLink® 1000, SR Research Ltd., Ontario, Canada) and sampled at 1000 Hz, for both eyes. The 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 (Jasper, 1958). The right earlobe and FCz electrode were used respectively as reference and ground. Data were amplified using a g.GAMMAsys gtec system (g.tec, Inc.) and sampled at 1200 Hz using the g.USBAmp (g.tec, Inc.). An analog band-pass filter (0.01-100 Hz) and a 50 Hz notch filter were applied online.

Stimuli

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

Experimental procedure

Participants performed four 20-minute sessions, but only the results for one session, (free exploration condition) are discussed here. Sixty scenes were randomly displayed within each session.

The experiment was designed using the SoftEye software (Ionescu, Guyader, & Guérin-Dugué, 2009) to control (i) the timescale for displays, (ii) the eye-tracker and (iii) the sending of synchronization triggers to both devices, i.e. the EEG and eye-tracker. In the free exploration condition, trials were composed of three successive displays. Each trial started with a white central fixation cross, which was displayed for 800 to 1200 ms. Once participants had stabilized[3] their gaze for 100 ms on a square of 50 pixels around the central fixation cross, a scene was displayed for 4 s. Each trial ended with a grey screen for 1 s.

Data preprocessing

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

Eye movement and EEG signals were synchronized offline, on the basis of triggers sent simultaneously to both the EEG system and the eye-tracker. EEG data were then re-sampled at 1000 Hz (eye-tracker sampling rate). Visual inspection revealed that channels T7, T8, TP9 and TP10 were too noisy for the majority of participants. We therefore decided to cancel these four channels for all participants. Using EEGlab software (Delorme & Makeig, 2004), EEG data were segmented into trials ranging from 500 ms before scene onset to 4000 ms afterwards. The segments obtained were visually inspected offline, and those containing muscular activity or non-physiological artifacts were rejected. Ocular artifacts were then corrected using a principal component analysis (number of channels minus one component) followed by ICA (infomax ICA) (Bell & Sejnowski, 1995). A visual inspection was performed once again. If ocular artifacts had not been corrected, epochs were removed from the analysis. One participant was excluded because of a high number of deleted trials (37). Four other participants were removed because of high variance across trials. The individual inter-trial variance (averaged on times and channels) was computed for each subject. Four subjects had the highest individual average variance with a relative gap of more than 60% with the average variance. This was confirmed by the empirical variance of these individual variances across subjects[4]. In the end, trials from thirty four participants were retained for analysis. On average, 47.7 ± 6.3 trials were kept for analyses. The maximum number of trials per participant was 57, and the minimum number was 30.

Fixations were tagged off-line according to their rank, from the first to the last fixation. The beginning of the first

---

[3] If the gaze failed to stabilize, the trial was not taken into account in the analysis.

[4] The respective contribution of these four subjects to the empirical variance was more than double their expected contribution, which should have been equal to one over the number of subjects, if all subjects contributed equally. The four values of the contribution to the empirical variance for the removed datasets were 2.23, 3.42, 5.64, and 6.63 divided by the number of subjects.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

fixation started after the scene onset, and the last fixation was such that this fixation ended before the scene offset.

Results

This section is divided into three parts. The first part presents the behavioral results which were used as a basis to configure estimation methods for the evoked potentials. The second part deals with the comparison between the linear regression method (GLM) and the ADJAR algorithm to estimate EFRP in the middle of the free visual exploration task. Finally, and using the same datasets, the results in which the GLM is configured to estimate potentials evoked at the beginning of exploration, i.e. the potential elicited at image onset, as well as the EFRP at the beginning of exploration are presented in the third part. Based on the results presented in part two, the ADJAR algorithm was not implemented for comparison in part three.

Behavioral results

Firstly, we present the behavioral results on ocular activity across the whole trial, and then at the start of visual exploration, i.e. for earlier fixations.

Eye movement during the free exploration task.

Table 1 summarizes the averages of the main ocular features evaluated throughout the entire duration of the trials (number of fixations, first saccade latency, fixation duration, saccade amplitude, saccade duration, inter-fixation interval duration 4 s).

Eye movement on the first fixations

Table 2 summarizes the averages of the main ocular features, but only for the first five fixations. These early fixations were analyzed specifically because, in the section entitled "Estimation by GLM of evoked potentials at the beginning of exploration", we were interested in the potential evoked at image onset, i.e. at the very beginning of the task.

A repeated measures analysis of variance was used, with fixation rank (five ranks) as the within-participant factor, and multiple comparisons were assessed using the Bonferroni correction. The statistical results on fixation duration revealed that this was significantly lower for the first rank than for subsequent ranks (F(4, 165) = 8.04, p <

0.001). The same result was observed for the incoming saccade amplitude determined from ranks one to five (F(4, 165) = 18.35, p < 0.001); incoming saccade amplitude for the first fixation was smaller than for the following ones. But the differences in saccade duration across the rank were not significant (p = 0.64). It has been established that a saccade's duration is linearly related to its amplitude (van Beers, 2007). We therefore expected a significant difference for the first saccade duration. The result obtained could be explained by the fluctuations in individual correlations between duration and amplitude observed. This provided a relative standard deviation (ratio of the standard deviation to the mean) that was larger for the duration than for the amplitude of the first saccade.

To sum up, we found that the first fixation differed from the following ones in terms of fixation duration and incoming saccade amplitude. Free exploration of a scene began with a fixation cross at the image's center. At the very beginning of the exploration, the first fixation remained close to the image's center[5], contributing to the so-called central fixation bias usually observed in scene viewing (Tatler, 2007).

Luminance and contrast at the first fixations

The early potential (lambda component) is modulated by the physical properties of the regions gazed at (Gaarder, Krauskopf, Graf, Kropfl, & Armington, 1964), (Hopfinger & Ries, 2005), (Ossandón, Helo, Montefusco-Siegmund, & Maldonado, 2010). Table 3 summarizes the average image features evaluated for the regions gazed at during the first fixations. Three luminance features were evaluated based on the region gazed at during each fixation. The features were evaluated for regions restricted to the foveal region: the average luminance, the contrast computed by the standard deviation of the local luminance of the fixated region and the absolute value of the difference of the mean luminance between two regions on consecutive fixations (for a given fixation the absolute value of the luminance difference lies across the incoming saccade).

A repeated measures analysis of variance was used with fixation rank (5 ranks), as within-participant factors, and multiple comparisons were assessed using the Bonferroni correction. Statistical results of the local luminance standard deviation revealed that the differences across the fixation ranks were significant (F(4, 165) = 5.42, p <

---

[5] Before image presentation, the participant's gaze was stabilized on a central fixation cross.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

0.001): the standard deviation of the luminance of the first region gazed at during the first fixation was significantly higher than that of subsequent fixations. On average, for all trials and subjects, the local statistics on the luminance (mean and contrast, evaluated here by standard deviation) were similar for all fixation ranks except the first one. The local contrast at the first fixation position was higher than at later ones. This first fixation was located near the screen center and the higher contrast might be explained by bias due to the center driven composition of usual image databases.

Table 1 Statistical summary: number of fixations, first saccade latency, fixation, saccade and Inter-Fixation Interval durations and saccade amplitude, based on individual means (std), during free exploration

Number of fixations	First saccade latency [ms]	Fixation duration [ms]	Incoming saccade amplitude [°deg]	Incoming saccade duration [ms]	IFI duration [ms]
12.60 (1.27)	266.85 (44.57)	243.40 (26.29)	7.18 (0.97)	45.06 (6.48)	288.69 (26.97)

Table 2 Statistical summary for the first five fixations: latency, fixation duration, amplitude and duration of the incoming saccade, based on individual means (std), at the start of free exploration

Rank / Feature	Fixation latency [ms]	Fixation duration [ms]	Incoming saccade amplitude [°deg]	Incoming saccade duration [ms]
Fixation 1	307.65 (46.14)	209.57 (38.93)	5.03 (0.79)	40.80 (9.52)
Fixation 2	561.50 (80.85)	241.98 (36.40)	6.46 (1.13)	44.30 (13.69)
Fixation 3	847.58 (108.18)	243.65 (26.03)	6.28 (1.26)	43.06 (12.11)
Fixation 4	1135.42 (126.89)	243.01 (20.56)	6.94 (1.39)	43.15 (7.58)
Fixation 5	1421.69 (145.30)	241.76 (26.21)	7.40 (1.38)	44.25 (7.78)

Table 3 Statistical summary for the first five fixations, of the image features of the fixated region: mean local luminance, local contrast computed by the standard deviation of the local luminance and absolute value of the difference of the mean luminance of fixated regions before and after the incoming saccade, based on individual means (std), at the start of free exploration

Rank / Feature	Luminance	Contrast	Luminance difference on saccade
Fixation 1	115.09 (7.54)	44.34 (3.02)	42.27 (4.13)
Fixation 2	112.78 (8.44)	41.99 (3.07)	42.90 (5.48)
Fixation 3	114.58 (7.72)	41.42 (2.38)	41.25(7.07)
Fixation 4	115.63 (6.74)	42.29 (2.43)	41.65 (6.97)
Fixation 5	113.98 (7.13)	41.99 (3.10)	41.79 (5.84)

Estimation of EFRP during exploration: comparison between Average, ADJAR and GLM estimations

In this section, we present the results of the comparison of the three methods of EFRP estimation during a free visual exploration task. Firstly, the ADJAR algorithm is analyzed alone to provide details of its estimation procedure. We go on to compare three estimates of the EFRP: (i) the

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

ADJAR algorithm, (ii) the GLM and (iii) the classical average on time-locked signals, as a baseline method, even though it is known that this last estimation is biased by overlap.

Explanation of the configuration of the three algorithms

The three methods were compared only in relation to the overlapping issue, and not in relation to their ability to tackle different kinds of potentials per epoch. To set a fair benchmark for the three methods (Average, ADJAR and GLM), the epochs were selected in the middle of visual exploration. This ensured that the hypothesis of a single EFRP elicited irrespective of fixation rank was acceptable.

The following methodology was designed:

- For the estimation method by average and in line with the basic requirement for the ADJAR algorithm, it was assumed that each fixation elicited the same potential regardless of its rank during exploration. During the free exploration of these scenes, there were no specific spatial loci (e.g., for instance people, faces, incongruent objects, etc.) which could elicit specific potentials. Only fixations in the middle of visual exploration were selected (see below).

- The observed x_i signal for the ith trial was time-locked at the onset of the nth fixation. For each participant and each scene presentation, a rank n was randomly selected from 3 to 9 three[6] times, on the basis of a uniform draw. The delay between the onset of the visual stimulus and the onset of the first fixation should be large enough to allow the temporal overlap between the potentials evoked by these two events to be ignored. Based on methodological studies (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011); (Nikolaev, Meghanathan, & van Leeuwen, 2016), a minimum delay of 700 ms was chosen and the choice of the third fixation led to an average (std) latency of 847.58 (108.18) ms, justifying the lower bound value of 3. Moreover, a ninth fixation with an outgoing saccade on the visual scene before the end of the trial occurred in all trials, justifying the upper bound value of 9. This choice resulted in an average

(std) number of epochs per participant of 165.26 (7.11).

- The window of interest for the EFRP ranged from -250 ms to 600 ms in relation to fixation onset. This window is of a typical configuration, designed to evaluate the waveform both before the saccade and after the saccade for the early and late components of the EFRP. For the estimations by average and by GLM, the EFRP ($a^{Av}(t)$ and $a^{Glm}(t)$ respectively) were then estimated on an estimation window equal to the window of interest. However, the estimation window for the ADJAR algorithm was slightly larger to include the border effect of the convolution product with a tapering window. For the EFRP $a^{Adjar}(t)$, the estimation window was therefore defined as being from -450 ms to 800 ms in relation to fixation onset. The parametrization of the algorithm is shown in the [appendix](#) describing the ADJAR implementation for this study.

- Accordingly, for the estimation by average, the epoch window was set to the chosen estimation window (i.e. from -250 ms to 600 ms in relation to fixation onset), for all participants. For the two other algorithms (ADJAR, GLM), the epoch window was defined with a common criterion: it had to be long enough to include at least the temporal description of three evoked responses on adjacent fixations (a current fixation at rank n, a previous fixation at rank $n-1$ and a subsequent fixation at rank $n+1$)[7]. Since each participant had his/her own oculomotor pattern, the IFI distributions differed from one participant to another Thus, the epoch window was defined for each participant, (1) to include the contributions of the potentials elicited by the fixation at rank $n-1$, as well as by that at rank $n+1$, (2) to obtain an almost constant number of fixations inside the epochs and consequently (3) to standardize as far as possible the contributions of the responses that overlap across epochs. For this purpose and for each participant, the epoch window was defined as: $[-250-\tau; 600+\tau]$ ms, with τ being the sum of the individual mean and standard deviation of the IFI values distribution. The average $\overline{\tau}$ value for all participants was equal to

---

[6] The number of random draws was fixed at three as a balance between a large number of epochs and the memory demanded for computation.

[7] For the ADJAR algorithm, the epoch must be sufficiently large to include both the previous and the subsequent fixation

onset. With a larger interval, this criterion remains satisfied. The distributions $Q_-(t)$ and $Q_+(t)$ are not impacted by a larger epoch.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

406.37 ms, at around 410 ms[8]. The epoch window was, therefore, on average [-660; 1010] ms, and the number of fixations which occurred within an epoch was on average (std), 5.8 (1.07). The average (std) onset of immediate previous fixations was equal to -279.51 (22.26) ms, and the average (std) onset of immediate subsequent fixation was equal to 288.43 (24.85) ms.

In the next section, we look at the specificity of the ADJAR algorithm, and of the GLM configuration.

The ADJAR algorithm was applied to the estimate $\hat{a}^{Av}(t)$ on the estimation window from -450 ms to 800 ms, and included the interval of interest from -250 to 600 ms. For the iterative procedure, the stopping criterion C was defined as the relative power between two consecutive iterations became lower than a given threshold: $C = \dfrac{\left\| \hat{a}^k(t) - \hat{a}^{k-1}(t) \right\|^2}{\left\| \hat{a}^k(t) \right\|^2} < 10^{-4}$. The estimate by ADJAR, $\hat{a}^{Adjar}(t)$, was obtained after 21 (6) iterations on average (std).

For GLM implementation, we worked on the assumption that the same potential was elicited at each fixation onset. We therefore looked at only one class: the $a(t)$ potential. The onsets of all fixations (not just adjacent ones) during the epochs took into account computation of the Toeplitz D matrix, and then estimated the evoked potential. The estimate $\hat{a}^{Glm}(t)$ was given by equation (14). This was applied to each subject and for all electrodes. The dimension of the x matrix of the observed EEG signals was time-locked at fixation onset (rank n) and was $E \times N_e$ rows and N_{el} columns. E was the number of epochs for a given subject, N_e the number of samples of an epoch (here $1010 + 660 + 1 = 1671$) and N_{el} the number of electrodes (28). The dimension of the $\hat{a}^{Glm}$ matrix was N_a rows and N_{el} columns, with N_a the number of samples of the estimation window ($600 + 250 + 1 = 851$).

Results with the ADJAR algorithm

The ADJAR algorithm was applied to estimate the potential elicited at fixation onset during free exploration,

since this potential was only overlapped by those elicited by adjacent fixations.

In Figure 1a, normalized distributions of the timestamps of adjacent fixations are plotted for all participants, namely $Q_{-1}(t)$ and $Q_{+1}(t)$ and the envelopes[9] of these distributions are represented after convolution, to give $Q_{-1}(t) * Q_{-1}(t)$ and $Q_{+1}(t) * Q_{+1}(t)$. The support of these second-order distributions was around [-1000; -150] ms and [270; 1120] ms, respectively, and intersected with the estimation interval. The supports of the resulting higher-order distributions were progressively enlarged, and their contributions decreased progressively as iterations increased.

Figure 1b shows the grand averages after convergence of the algorithm. This figure illustrates the three final estimations given by the algorithm: $\hat{a}^{Adjar}(t)$, $\widehat{ov}_p(t)$ and $\widehat{ov}_s(t)$. Previous ($ov_{pp}(t)$) and subsequent ($ov_{ss}(t)$) second-order response overlaps were also estimated[10] by convolution with the final estimate $\hat{a}^{Adjar}(t)$ and are presented in Figure 1b.

The period [-200; -100] ms before 0 (fixation onset) was used for the baseline correction, In common with other EFRP studies (Kamienkowski, Ison, Quiroga, & Sigman, 2012), (Fischer, Graupner, Velichkovsky, & Pannasch, 2013), we used the period [-200; -100] ms before 0 (fixation onset) for baseline correction. The saccadic spike potential takes place during the period [-100; 0] ms, due to saccade generation (Keren, Yuval-Greenberg, & Deouell, 2010). This period therefore has to be excluded from the baseline computation. Finally, the period [-200; -100] ms corresponds to the end of the previous fixation and is free of any activity related to saccade generation.

In Figure 1b, all waveforms are illustrated in the estimation window [-450; 800] ms. Estimations during the first and the last 200 ms were biased by the return to zero of the tapering window. These parts were, as a result, excluded from the [-250; 600] ms window of interest. The contribution of the subsequent second-order response overlap ($\widehat{ov}_{ss}(t)$) was not corrected in this estimation. The

---

[8] On average, $\bar{\tau}$ is equal to the sum of the average of individual IFI means and that of individual IFI standard deviations. Computed between the 3rd and the 9th fixation, these averages were respectively 291.72 ms and 114.65 ms. $\bar{\tau} = 291.72 + 114.65 = 406.37 \approx 410$ ms.

[9] The distribution envelope was chosen for reasons of clarity when graphs were superimposed on the same figure.

[10] $\widehat{ov}_{ss}(t) = Q_{+1}(t) * Q_{+1}(t) * \hat{a}^{Adjar}(t)$; $\widehat{ov}_{pp}(t) = Q_{-1}(t) * Q_{-1}(t) * \hat{a}^{Adjar}(t)$

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

final estimate was under-estimated from 0 ms (positive values for $\widehat{ov}_{ss}(t)$). The contribution of the previous second-order response overlap ($\widehat{ov}_{pp}(t)$) was also not corrected. The final estimate was thus over-estimated by up to 400 ms (negative values for $\widehat{ov}_{pp}(t)$), see equation (10). Some of the previous and subsequent second-order response overlaps were not negligible in the estimation window, and contradicted the third assumption of the ADJAR algorithm. The response overlaps were increasingly smoothed by these successive convolution operations, as after a low-pass filtering. However, we must bear in mind that previous and subsequent second-order response overlaps remain in all iterations up to the final convergence.

The baseline (before 0ms) of the waveform $\hat{a}^{Adjar}(t)$ presented in Figure 1b, is therefore impacted by the second-order previous response overlap. The return to zero (after 400 ms) was distorted by subsequent second-order response overlaps. Despite these distortions, Figure 1b shows that the potential was mainly composed of the lambda response (latency around 100 ms). The estimated first-order subsequent response overlap (red line; $\widehat{ov}_s(t)$), increased between -260 ms and 280 ms, and attained its maximum value maximum between 300 and 370 ms. These latencies were in line with the sum of the lambda component latency and the average onset of the subsequent fixation ($\approx$ 80+288=368 ms). In other words, $ov_s(t)$ had a greater impact on the late components of the evoked potential than on the early ones, here the lambda wave. The large negative deviation from 200 ms observed for $\hat{a}^{Adjar}(t)$ was the result of neglecting subsequent second-order response overlaps. The estimated first-order previous overlap response (blue line; $\widehat{ov}_p(t)$), was a negative wave, with a minimum of around 100 ms, i.e. the same latency as the lambda wave. The temporal evolution of this wave was mainly due to the highly negative $\hat{a}^{Adjar}(t)$ deviation observed at the end of the window of interest.

Consequently, $ov_p(t)$ had a high influence on the estimation of the lambda wave, and its estimation was highly biased, since previous second-order response overlaps were not considered. Finally, the lambda wave estimation was erroneous.

Therefore, in the context of EFRP estimation, the third hypothesis of the ADJAR algorithm was not validated. The distortions due to second-order response overlaps were not taken into account and were not corrected when estimating the evoked potential, and this resulted in a major bias.
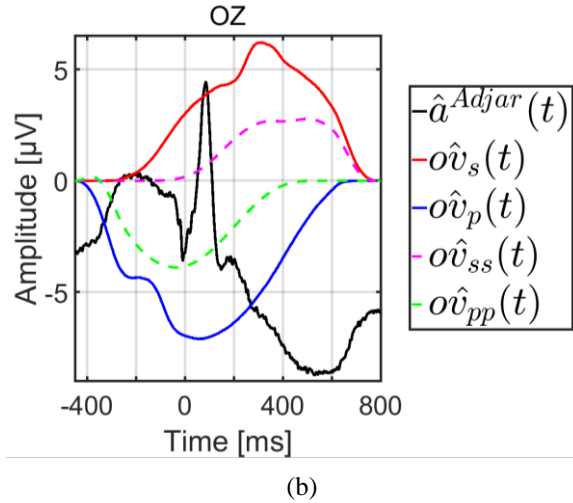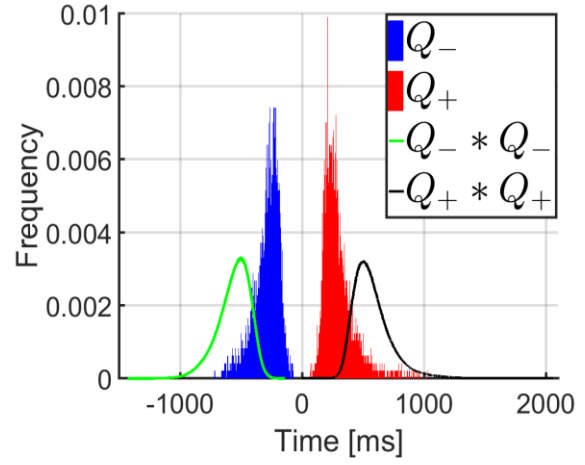

(a)

(b)

Figure 1 (a) Normalized distributions $Q_{-1}(t)$, $Q_{+1}(t)$ and the envelope of these distributions after convolutions: $Q_{-1}(t) * Q_{-1}(t)$, $Q_{+1}(t) * Q_{+1}(t)$; (b) evoked potential estimated by the ADJAR algorithm ($\hat{a}^{Adjar}(t)$), the first-order ($\widehat{ov}_p(t)$, $\widehat{ov}_s(t)$) and second-order ($\widehat{ov}_{pp}(t)$, $\widehat{ov}_{ss}(t)$) response overlaps on Oz electrode, at convergence

Comparison of estimations by average, the ADJAR algorithm and the GLM

In this section, we look at the three algorithms used to estimate an EFRP. The grand averages for the three methods are plotted in Figure 2, with a zoomed plot between -200 and -100 ms. For the Average method, the estimation was performed without taking into account the potentials elicited by previous and subsequent fixations. Only adjacent fixations were considered for the ADJAR estimation, because this algorithm is only able to account for adjacent

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

fixations. The second underlying assumption in the section entitled "Estimation by the ADJAR algorithm" illustrates this. By contrast, the GLM is able to account for all response overlaps. Consequently all fixations inside the epoch were considered for the GLM estimation.

As expected, each method elicited a clearly visible lambda wave (of around 100 ms). The three estimates were evaluated according to two qualitative criteria. The first of these was potential stabilization during the baseline period, and the second was the return to zero of the potential amplitude at the end of the estimation period. In any study of EFRP under experimental conditions, stabilization, i.e. no drift, during the baseline period is needed to obtain a reliable estimation for comparison. We expected a progressive return to zero of the amplitude of the evoked potential at the end of the estimation period. Indeed, neural activities during this late period became less and less synchronized with fixation onset.
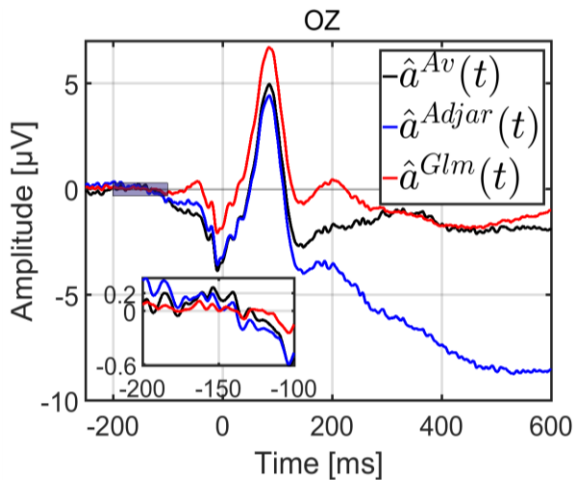

Figure 2 Grand average for the three methods: Average ($\hat{a}^{Av}(t)$), ADJAR ($\hat{a}^{Adjar}(t)$), and GLM ($\hat{a}^{Glm}(t)$) on the OZ electrode; the inset is a zoomed plot between -200 and -100 ms

Statistical comparison of the variance during the baseline interval, between -200 and -100 ms, revealed that the variance of the estimate $\hat{a}^{Glm}(t)$ was lower than the variance of estimates $\hat{a}^{Av}(t)$ and $\hat{a}^{Adjar}(t)$ ($F(2, 99) = 13.86$, $p < 0.001$, and Bonferroni adjustment for multiple comparisons). The zoomed plot can be seen in Figure 2. The baseline correction for this period was therefore erroneous for both the ADJAR algorithm and the Average method. As far as the return to zero at the end of the estimation window was concerned, this criterion was only met for the GLM estimate (after 450 ms). Results with a longer estimation window [-400; 1550] ms are presented in the appendix to confirm this statement.

Estimation by GLM of evoked potentials at the beginning of exploration

Based on the same trials as mentioned previously, the epochs here were time-locked at stimulus onset. We looked at potentials elicited at the beginning of visual exploration: potentials elicited at the first fixation, and the event-related potential at image onset. In this subsection, we discuss the configuration of the GLM depending on the potential of interest. Two GLM configurations are detailed. We then present the condition number of the $D^{\dagger}D$ matrix as a global indicator of the expected estimation quality. Finally, the results for the different models are presented (according to) potentials of interest.

Configuration of selected models

As detailed previously, in order to take into account the response overlap for the estimation of potentials of interest, several GLM configurations can be chosen. These are based on different assumptions.

With a two-class GLM, the potential elicited at each fixation onset was assumed to be the same irrespective of fixation rank. Thus, for a given trial i, the observed signal $x_i(t)$ at image onset can be written as:

$$x_i(t) = s(t) + \sum_{f=1}^{F_i} a_{1_+}\left(t - \tau_i^f\right) + n_i(t) \tag{15}$$

with $s(t)$, the evoked potential at stimulus (image) onset, $a_{1_+}(t)$ the evoked potential at each fixation onset, τ_i^f, the timestamp of the onset of the fth fixation rank, occurring in the ith epoch, and $n_i(t)$ brain activity unrelated to the task. Consequently, for all trials and for each participant, using the same notation as in the section "Estimation by the General Linear Model", the GLM can be expressed as:

$$x = D_s s + D_{1_+} \cdot a_{1_+} + n \tag{16}$$

The potentials $s(t)$ and $a_{1_+}(t)$ were estimated from equation (14), where $\hat{a}^{Glm}$ is the vertical concatenation of $\hat{s}$, and $\hat{a}_{1_+}$, and D is the horizontal concatenation of D_s and D_{1_+}. See Table 4 (left column).

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

However, analysis of statistical behavior (see section entitled "Behavioral results") revealed significant differences between the first fixation -in terms of duration, incoming saccade amplitude, contrast of the gazed foveal region- and subsequent fixations. Part of the EFRP is explained by the lambda potential which reflects the visual change in the image retina due to the saccade (Barlow & Ciganek, 1969) (Gaarder, Krauskopf, Graf, Kropfl, & Armington, 1964); (Thickbroom, Knezevic, Carroll, & Mastaglia, 1991). At fixation onset, the lambda potential depends not only on low level image features (luminance and contrast of the area gazed at), but also on the incoming saccade amplitude (Yagi A. , 1979), (Ossandón, Helo, Montefusco-Siegmund, & Maldonado, 2010), (Thickbroom, Knezevic, Carroll, & Mastaglia, 1991). As we were aware of these modulations early in the case of EFRP, we were able to consider an alternative GLM as a supplementary class for the first EFRP. Thus, equation (3) can be rewritten as:

$$x_i(t) = s(t) + a_1(t - \tau_i^1) + \sum_{f=2}^{F_i} a_{2_+}(t - \tau_i^f) \qquad (17)$$
$$+ n_i(t)$$

with $a_1(t)$ being the potential elicited specifically at the onset of the first fixation, and $a_{2_+}(t)$, the potential elicited at the onset of the second and subsequent fixations. The equation of the corresponding GLM is then:

$$x = D_s s + D_1 . a_1 + D_{2_+} . a_{2_+} + n \qquad (18)$$

The potentials $s(t)$, $a_1(t)$, and $a_{2_+}(t)$ were estimated by equation (14), where $\hat{a}^{Glm}$ is the vertical concatenation of $\hat{s}$, $\hat{a}_1$, and $\hat{a}_{2_+}$, and D is the horizontal concatenation of D_s, D_1, and D_{2_+}. See Table 4 (right column).

All results were obtained using the following configurations for the epochs:

- We chose an epoch duration which was long enough to include at least the temporal description of the first two evoked responses on adjacent fixations (the first and second ones). This epoch was defined as follows: [-200; 800 + τ] ms, with τ being the sum of the average of the latency of the second fixation (561.50 ms), and its standard deviation (80.85 ms): τ = 561.50 + 80.85 = 642.35 ≈ 700 ms. See Table 2. Consequently for all trials and participants, the temporal interval of epochs was [-200 ; 1500] ms from image onset. On average

(std), the number of fixation onsets in an epoch was 5 (1). The number of epochs per subject was equal to the number of trials.

- The duration of the estimation window for the evoked potential $s(t)$ at image onset was the same as that of the epoch.

- The estimation window for the evoked potential at fixation onset ($a_1(t)$, $a_{1_+}(t)$, $a_{2_+}(t)$) was from -200 ms to 800 ms.

Table 4 Summary of the two GLM configurations

Two-class Model	Three-class model
$x = D_s s + D_{1_+} a_{1_+} + n$	$x = D_s s + D_1 a_1$ $+ D_{2_+} a_{2_+} + n$
s: Class one, Potential evoked at image onset a_{1_+}: Class two, EFRP at all fixation onsets at the beginning of exploration	s: Class one, Potential evoked at image onset a_1: Class two, EFRP at the 1st fixation onset a_{2_+}: Class three, EFRP at the 2nd and following fixations
Resolution by equation (14) : $\hat{a}^{Glm} = (D^\dagger D)^{-1} Dx$	
$x \in \mathbb{R}^{N \times N_{el}}$, with N the total number of samples ($N = N_e \times E$), E the number of epochs for a given subject, N_e the epoch size ($N_e = 1500 + 200 + 1 = 1701$ samples) and N_{el} the number of electrodes ($N_{el} = 28$)	
$\hat{a}^{Glm} \in \mathbb{R}^{(N_s + N_a) \times N_{el}}$, with $N_s = N_e$, $N_a = 800 + 200 + 1 = 1001$ samples	$\hat{a}^{Glm} \in \mathbb{R}^{(N_s + 2.N_a) \times N_{el}}$, with $N_s = N_e$, $N_a = 800 + 200 + 1 = 1001$ samples
$D = [D_s, D_{1_+}]$, with $D \in \mathbb{R}^{N \times (N_s + N_a)}$	$D = [D_s, D_1, D_{2_+}]$, with $D \in \mathbb{R}^{N \times (N_s + 2.N_a)}$

Overall estimation of quality: condition number of the $\mathbf{D}^\dagger\mathbf{D}$ *matrix*

Using the GLM, potentials were estimated by equation (14). This required computation of the inverse of the $D^\dagger D$ matrix. For the two-class GLM, the D matrix was composed as $D = [D_s, D_{1_+}]$ (equation (16)), and for the three-class GLM, the D matrix was composed as $D = [D_s, D_1, D_{2_+}]$ (equation (18)). The $D^\dagger D$ matrix is a block matrix: a 2 x 2 block matrix for the two-class GLM and a 3 x 3 block matrix for the three-class GLM.

In general, the condition number of the $D^\dagger D$ matrix is taken as an indicator (or a warning) of an estimation's reliability. With a greater condition number, there is a risk that the noise $n(t)$ on observed data might be amplified on

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

the solutions $a(t)$ of the mean square error problem (equation 13). More specifically, on average, the variance of the estimate (expressed in the appendix) depends on the inverse of $D^\dagger D$ and thus on its condition number.

A bootstrap estimator with 10 000 replications was implemented for statistical assessment of the condition number of the $D^\dagger D$ matrix for each model, and each participant.

On average (std), the condition number (CN) for the three-class model ($CN = 1446.70\ (834.36)$) was higher than that of the two-class model ($CN = 707.36\ (322.26)$), and much higher than that of the model by average ($CN = 1$)[11]. The first class gathered together events at image onset and the associated timestamps all equaled zero. Therefore, for the two-class model, the jitter on the timestamps between the two classes was accurately represented by the variability of the timestamps of events within the second class. In other words, the condition number for this model was directly linked to the variability of all fixation onsets, irrespective of their rank. In the case of the model with three classes, this set of fixation onsets was split into two parts. The first fixation (second class) was separated from subsequent ones (third class). The timestamp variability for the first fixation onset was therefore smaller than that of the following fixation onsets. It was for this reason that on average, the condition number for the two-class model was better (lower value) than for the three-class model.

Figure 3[12] illustrated the distribution of condition numbers based on individual means obtained by bootstrap estimation. Because of the link, on average, between the condition number and the estimation variance, the result concerning the condition number of the $D^\dagger D$ matrix for each model ought to predict a larger variance for the three-class than for the two-class model and for the model established by averaging (one-class model). In view of the tradeoff between bias and variance, this decrease should be associated with an increase in estimation bias. In practice, since a high-level condition number was only a warning, the bootstrap estimator of the variance of evoked potentials was systematically computed in order to obtain a quantitative criterion for the assessment of estimates.

Figure 3 Averaged condition number of the $D^\dagger D$ matrix for the two GLM configurations, with two or three classes, based on individual means computed with a bootstrap estimator.

In the next two parts, we study the two GLM configurations in order to estimate different potentials of interest. In the subsection "EFRP at first fixation onset", the potential of interest is the EFRP at the first fixation onset. The three-class model is used; the EFRP at the first fixation onset is compared to EFRPs at subsequent onsets, in order to evaluate the specificity of the first fixation. In the following subsection "Evoked potential at image onset", the average and the two GLM configurations are compared to assess the impact of the first EFRP estimation on the estimation of the ERP at image onset.

Table 5 summarizes the content of the two following subsections.

Table 5 Summary of model choices related to the potential of interest for the two following subsections

Subsection name	Potential of interest	Estimation method
EFRP at first fixation onset	$a_1(t)$	Three-class model
Evoked potential at image onset	$s(t)$	i. Average ii. Two-class model iii. Three-class model

---

[11] The estimation by average corresponds to the one-class GLM estimation, considering only one event of interest inside each epoch. In this case, the $D^\dagger D$ matrix is a diagonal matrix with all diagonal values equal to the number E of epochs. Its condition number is the lowest, i.e. equal to one ($CN = 1$).
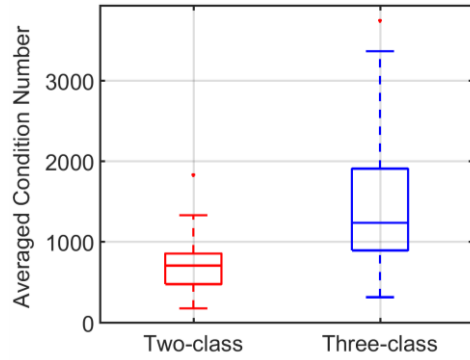
[12] The values of the condition number of the two-class GLM, and the three-class GLM were much higher than 1. This explains why the condition number for the model by averaging was not plotted on the graph of Figure 3.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

EFRP at first fixation onset

We considered the EFRP estimation at the first fixation onset. We showed that ocular features (incoming saccade amplitude, fixation duration) and the local luminance contrast of the foveal region gazed at were different for the first fixation than for the following ones (see above in section "Behavioral results"). This justifies establishing a specific class to estimate the first EFRP, and is in line with methodological studies (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011), (Nikolaev, Meghanathan, & van Leeuwen, 2016).

In the case of the three-class GLM (Table 4, right column), the potential of interest $a_1(t)$ was associated with the second class. The first and third classes provided an unbiased estimation of the EFRP at the first fixation by taking into account the influence of the ERP $s(t)$ at image onset as well as overlaps of the response $a_{2_+}(t)$ at subsequent fixations.
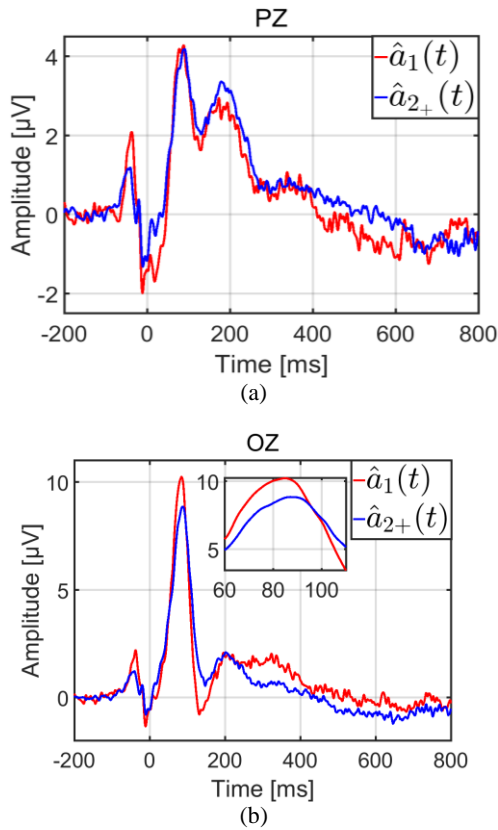

Figure 4 Grand average of the EFRP at the first fixation $\hat{a}_1(t)$ (red line) and at the second and subsequent fixations $\hat{a}_{2_+}(t)$ (blue line), estimated by the three-class GLM (a) on PZ and (b) OZ electrodes

Using the three-class GLM, Figure 4 illustrates the estimation of the first EFRP $\hat{a}_1(t)$, and of subsequent ones $\hat{a}_{2_+}(t)$, on PZ and OZ electrodes. In line with common practice, these estimates were corrected with a baseline of between -200 and -100 ms.

The component of interest was the lambda response between 70 ms and 90 ms. Inside this temporal interval, the difference between was not significant on the PZ electrode according to a t-test: $t(33) = 0.49, p = 0.63$. However, on the OZ electrode, the same difference was significant: $t(33) = 2.69, p = 0.01$. This result confirmed the assumption that the first fixation would have to be differentiated from following ones. As Figure 5 shows, the variances of both estimates were also different, that of $\hat{a}_1(t)$ being larger than $\hat{a}_{2_+}(t)$.
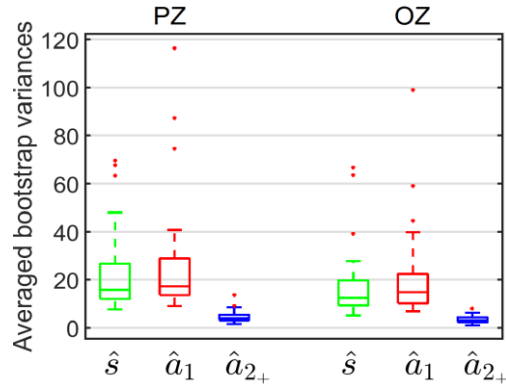

Figure 5 Averaged bootstrap variances of the ERP at image onset $\hat{s}(t)$, the EFRP at the first fixation $\hat{a}_1(t)$ and at the second and subsequent fixations $\hat{a}_{2_+}(t)$, estimated by the three-class GLM, on PZ and OZ electrodes

The bootstrap variances were statistically analyzed using a repeated measures ANOVA with the class ($\hat{s}(t)$, $\hat{a}_1(t)$, $\hat{a}_{2_+}(t)$) and the electrode (PZ, OZ) as within-participant factors. Multiple comparisons were assessed with Bonferroni post-hoc tests. The statistical results revealed significant differences according to class ($F(2,66) = 36.62, p < 0.001$), to electrodes ($F(1,33) = 10.71, p = 0.002$) and to both ($F(2,66) = 7.51, p = 0.001$). In the light of the main effect on electrodes, the variance of the estimate $\hat{a}_{2_+}(t)$ was lower than the variances of the estimates $\hat{s}(t)$ and $\hat{a}_1(t)$ which were similar. These results were expected. Firstly, $\hat{a}_{2_+}(t)$ was estimated from more samples (on average four fixations with a rank greater than one per epoch), than $\hat{a}_1(t)$ (a single first fixation per epoch). Secondly, $\hat{s}(t)$ and $\hat{a}_1(t)$ were estimated from only one event per epoch, for each potential.

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Evoked potential at image onset

The potential of interest was the potential elicited at image onset, i.e. the $s(t)$ waveform, or the s vector, in equations (16) and (18). After image display, the onset of the first fixation occurred just after the first saccade, on average (std) at 307.65 (46.14) ms (cf. Table 2). The potentials elicited by this ocular event and by subsequent ones, provided distortions by overlapping on this potential of interest.

For comparison $\hat{s}^{(0)}(t)$ expresses the ERP estimated by the average, $\hat{s}^{(1)}(t)$ and $\hat{s}^{(2)}(t)$ the ERP estimated by the GLM

with two (equation (16)) and three (equation (18)) classes, respectively.

Figure 6a illustrates the estimations of the potential evoked at image onset by the three methods, after a base line correction on OZ and PZ electrodes from -200 to 0 ms. Firstly, early and late components were observed for up to 600 ms before a return to a stabilized level. Interestingly, this stabilized level at the end of the segment was highest for the classical estimation $\hat{s}^{(0)}(t)$, showing that residual activities from all potentials elicited at fixation onsets provided on average a positive bias.
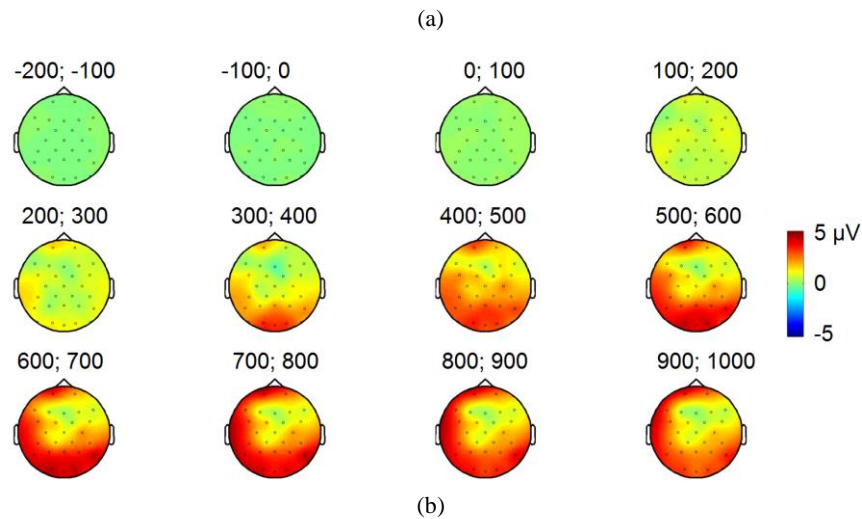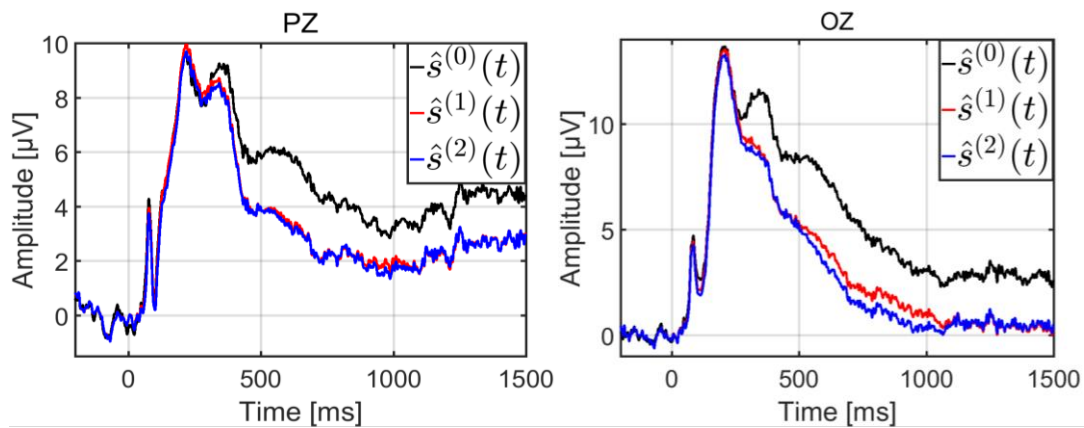

(a)

(b)

Figure 6 (a) Evoked potential at image onset on the midline PZ and OZ electrodes for the three estimations, by average $\hat{s}^{(0)}(t)$, by the two-class GLM $\hat{s}^{(1)}(t)$ and by the three-class GLM $\hat{s}^{(2)}(t)$; (b) Topographic maps of the difference between the two estimates: $\hat{s}^{(0)}(t) - \hat{s}^{(1)}(t)$

Journal of Eye Movement Research
10(1):7, 1-28

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Irrespective of the estimation method or choice of model, the P1 early component was clearly identifiable (Figure 6a), with a similar pattern (shape, maximum, latency). In other words, the estimation of the P1 component was not affected by distortions from overlaps. This was as expected, because the latency of this component (around 90 ms) was shorter than the latency of the first saccade (on average (std) 266.85 (44.57) ms, see Table 2. This also explains why the estimation of the P2 component from the latency around 200 ms differed from one method of estimation to another (Average vs GLM).

The large gap between $\hat{s}^{(0)}(t)$ and $\hat{s}^{(1)}(t)$ (or between $\hat{s}^{(0)}(t)$ and $\hat{s}^{(2)}(t)$) occurred from 300 ms on, and a higher maximum for $\hat{s}^{(0)}(t)$ was observed at around 360 ms. The interpretation of these distortions on $\hat{s}^{(0)}(t)$ was derived from the contribution of the potential elicited by the first fixation onset. On average, across all trials, this contribution corresponded to the convolution of the first evoked potential $a_1(t)$ with the distribution of Dirac impulses at the timestamps of the first fixation onset. The average (std) of this distribution among participants was 307.65 (46.14) ms (see Table 2). The result of this convolution was a low-pass version of the first evoked potential $a_1(t)$, with a lambda wave peaking at around 90 ms, then shifted at the average onset of the first fixation. This explained why aless steep maximum was observed on $\hat{s}^{(0)}(t)$ at a latency around 90 +307.65 ≈390 ms

Figure 6b shows the topographic map of the difference between $\hat{s}^{(0)}(t)$ and $\hat{s}^{(1)}(t)$. The positive gap between $\hat{s}^{(0)}(t)$ and $\hat{s}^{(1)}(t)$ started at 300 ms in the occipito-parietal area, and went on to progressively cover the whole scalp almost uniformly, with a gap value of a few microvolts (near 2μV, on PZ and OZ electrodes). In other words, from the $\hat{s}^{(0)}(t)$ estimate, the analysis of components with latencies above 300 ms yielded erroneous interpretations. In contrast, the period ranging from image onset to the first saccade (on average 266.85 ms, see Table 2), was free of eye movement.

Consequently, the estimations obtained by simple averaging and by the GLM were similar, as observed for instance for the early P1 component.

The selection of each of the three models (Table 5) was analyzed in terms of bias and variance of the estimation. Estimation by average ($\hat{s}^{(0)}(t)$) could be expected to provide the most biased estimator of the three. In contrast, by making a distinction between the first EFRP ($a_1(t)$) and

the following EFRP ($a_{2_+}(t)$), the three-class GLM could be expected to provide the estimate with the lowest bias.

Moreover, no significant difference was observed between $\hat{s}^{(1)}(t)$ and $\hat{s}^{(2)}(t)$ estimates (Figure 6a); a similar bias for these two estimates was assumed. However, the assumption for the variance was that the variances for the three-class model would be greater than the variances of the two-class model, because the former is less parsimonious than the latter. These variances were evaluated using 10 000 bootstrap repetitions for each model, each electrode and each participant. Figure 7 shows the bootstrap variances averaged for participants, on PZ and OZ electrodes, for the three estimates of the evoked potential at image onset. These bootstrap variances were statistically analyzed using a repeated measures ANOVA with model choice (Average for $\hat{s}^{(0)}(t)$, two-class GLM for $\hat{s}^{(1)}(t)$, three-class GLM for $\hat{s}^{(2)}(t)$) and electrode (PZ, OZ) as within-participant factors. Multiple comparisons were assessed with Bonferroni post-hoc tests. The statistical results revealed significant differences according to the model ($F(2,66) = 45.47$, $p < 0.001$), to the electrodes ($F(1,33) = 8.09, p = 0.007$) and to both ($F(2,66) = 5.85$, $p = 0.005$). As expected, on both electrodes, the variance of the estimate by the three-class GLM ($\hat{s}^{(2)}(t)$) was greater than the variance of the estimate by the two-class GLM ($\hat{s}^{(1)}(t)$) which was in return greater than that of the estimate obtained by the average ($\hat{s}^{(0)}(t)$).
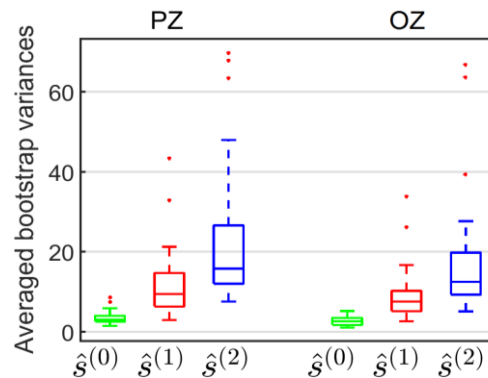

Figure 7 Averaged bootstrap variances for the three estimations of the evoked potential at stimulus onset, on PZ and OZ electrodes: by average $\hat{s}^{(0)}(t)$, by the two-class GLM $\hat{s}^{(1)}(t)$ and by the three-class GLM $\hat{s}^{(2)}(t)$

These results were in agreement with the increased number of parameters for the first, second and third model

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

relative to a fixed number of observed data, and were consistent with the increasing profile of the condition number of the $D^\dagger D$ matrix for the three models (Figure 3).

Discussion

One of the main limitations of EFRP estimation is the overlapping issue between adjacent neural responses. This is due to IFI values which are too low (Dimigen, Sommer, Hohlfeld, Jacobs, & Kliegl, 2011); (Nikolaev, Meghanathan, & van Leeuwen, 2016). The amount of overlapping is linked to the ocular sequences of each participant, and is not controlled by experimental design, except in specific cases (Kaunitz, et al., 2014). To address this difficulty, Kaunitz and colleagues (2014) trained participants to make long fixations. All other things being equal, this increases IFI values and consequently decreases overlap between adjacent potentials.

It has been well-established that the estimation of evoked potentials by averaging time-locked EEG signals is biased in the case of overlapping responses. Woldorff (Woldorff, 1993) proposed an iterative procedure in the context of ERP experiments where the EEG signal is time-locked on external events. This was called the ADJAR algorithm, and was designed to estimate overlap responses from immediately adjacent events, to converge towards the evoked potential of interest. Moreover, regression techniques, especially the GLM (Kiebel, Holmes, 2003), have proved their efficiency in the estimation of evoked overlapping potentials (Dale, 1999); (Dandekar, Ding, Privitera, Carney, & Klein, 2012); (Burns, Bigdely-Shamlo, Smith, Kreutz-Delgado, & Makeig, 2013); (Bardy, Dillon, & Dun, 2014); (Kristensen, Guerin-Dugué, & Rivet, 2017). The ADJAR algorithm and the GLM are applied to experimental data from conjoint EEG and eye-tracking recordings during free visual exploration, and are compared on the basis of their efficiency in relation to overlap responses for the estimation of evoked potentials.

Response overlaps are added linearly to the potential elicited at the time-locked event in both models. Only temporal delays are considered in the estimation of previous or subsequent response overlaps. In the case of such comparisons between the ADJAR algorithm, and the GLM, the same type of evoked responses for all events is considered. This restriction can be removed in some cases for the ADJAR algorithm (Talsma, 2005). However, taking into account various types of evoked potentials during trials is easy when using the GLM. This comparison therefore focuses on overlaps coming from the same type of potential.

In the ADJAR model, only adjacent responses (immediately previous and subsequent fixations) are considered. Thus, for this model, the potential of interest is *assumed* to be overlapped only by potentials elicited by the previous or subsequent event. This strong assumption is at the core of the definition of the ADJAR iterative process. The ADJAR algorithm is based on the definition of the normalized distribution of the timestamps of the fixation onsets of previous and subsequent fixations. This means that the contribution of all events inside each epoch must be the same, and this is a necessary condition of convergence. For this reason *only* two adjacent fixations (the previous and the subsequent one) are taken into account in the epoch. We have shown that this assumption fails in the context of the EFRP estimation. The window of interest is chosen to include early and late components of the evoked potential. This potential is estimated within an estimation window which is slightly larger than the window of interest, and which includes extra time to accommodate the border effects of convolution products. In practice, this extra time is of the same magnitude as the IFI value. Consequently, more than two evoked potentials on fixations overlap the EFRP of interest. Moreover, two main issues can lead to a non-negligible contribution of second-order adjacent response overlaps at convergence: an insufficiently long IFI and a low variability. At the last iteration, the estimated potential remains biased by these second-order responses, for which overlaps are not corrected. To sum up, the ADJAR algorithm appears to be poorly suited to EFRP estimation.

The comparison with the GLM shows a more natural framework in this context. All events can be taken into account inside each epoch without restriction by a closed-form estimation: there is no iterative procedure, and therefore no convergence issue, and no stopping criterion needs to be defined.

In addition, the ability of the GLM to deconvolve different neural responses is illustrated via two situations depending on the potential of interest: the potential elicited (1) at the first fixation onset and (2) at image onset. To this end, we considered two GLM configurations, with two and three classes. For the first model, events were split into two classes: image onset and fixation onsets irrespective of fixation rank. For the second model, the events were split into three classes: image onset, the first fixation onset and

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

the following fixation onsets. The two-class model assumes that the same potential was elicited at fixation onset irrespective of fixation rank. In contrast, the three-class model establishes a distinction between the potential elicited by the first fixation and by subsequent ones, and assumes a specific status or particular features for the first fixation at the beginning of the task. The choice between these two models depends not only on the issue in visual perception, or reading, but also on the quality of estimations. If the potential of interest is the first EFRP, the three-class model is fully justified. If the potential of interest is the one elicited at stimulus onset both models can be used. The choice should be based on additional criteria.

The three-class model is suitable if the potential of interest is the one elicited at the first fixation onset. The estimated potentials corresponding to the first and third class, provide an unbiased estimation of the potential of interest (corresponding to the second class). However, the variances of the estimation of the first EFRP and of the ERP at image onset are of the same magnitude. Moreover, these two variances are greater than the variance of the estimated EFRP at subsequent fixation onsets. This latter potential is estimated using more data (on average 3.9 fixations with a rank greater than one, inside each epoch) than the two former ones (only one event each inside each epoch). The results obtained show that the first fixation differs from the following ones by its significantly shorter duration and by the smaller amplitude of its incoming saccade. The local luminance contrast of the regions gazed at during the first fixation is also significantly higher than for regions gazed at later. With the three-class GLM, the estimated potential elicited at the first fixation onset exhibits a significantly larger amplitude of the lambda wave than the amplitude of this wave elicited by subsequent fixations. Modulation of the lambda wave amplitude by incoming saccade amplitude has been well-established (Yagi A. , 1979), (Thickbroom, Knezevic, Carroll, & Mastaglia, 1991), (Ossandón, Helo, Montefusco-Siegmund, & Maldonado, 2010), (Kaunitz, et al., 2014), (Nikolaev, Meghanathan, & van Leeuwen, 2016). The greater the incoming saccade amplitude is, the larger the lambda amplitude will be. The same holds true for the physical properties of visual stimuli such as illuminance (Scott, Groethuysen, & Bickford, 1967), (Gaarder, Krauskopf, Graf, Kropfl, & Armington, 1964) (Yagi, Imanishi, Konishi, Akashi, & Kanaya, 1998). A smaller first incoming saccade amplitude implies a decrease in the amplitude of the lambda component for the first estimated EFRP. A

greater local contrast of the first region gazed upon is observed, but no significant difference is observed for local luminance. Taken together, the observation of a greater first lambda amplitude cannot be explained by these low-level features. High level factors such as task demand and information processing load also modulate the lambda amplitude (Yagi A. , 1981); (Ries, Touryan, Ahrens, & Connolly, 2016). A higher level of attention may be assumed at the beginning of a task. This could be one interpretation of the greater amplitude of the lambda wave elicited at the first fixation onset, following the same speculations in (Yagi A. , 1981); (Yagi A. , 1982).

The two-class or the three-class model can be designed to estimate the potential elicited at image onset. There is no definitive result on this choice, which depends on the classical tradeoff between bias and variance. A priori, the distinction between the first EFRP and the following ones makes sense based on ocular behavior. This distinction was ignored in the two-class GLM. There is therefore a risk that a biased estimation of the EFRP irrespective of fixation rank, and consequently of the ERP at image onset, might occur because of a poor fit between the model and the observed data. The more distinctive first EFRP is, the greater the level of risk becomes. This risk of a biased estimation is balanced out by a lower variance. In our study, the difference between the estimates for the ERP at image onset given by the two models was not significant, but the variance for the three-class model was higher. This result justifies the choice of the two-class model for the estimation of the potential elicited at image onset.

In the context of EFRP, the GLM is a useful model to estimate overlapped evoked potentials. The GLM also allows the estimation of different neural responses. The main assumption is the linearity of the additive model to take into account different neural responses. The configuration of the number of classes depends on the assumptions concerning the cognitive processes under examination. As far as estimation is concerned, the number of classes results from the classical tradeoff between bias and variance. In order to obtain the best configuration of the GLM, two main questions must be asked: "what is the potential of interest?", and "what are the related potentials which may affect the estimation of the potential of interest?". The answer to the first question is often trivial. However, the answer to the second one is not. This is the outcome of a tradeoff between the parsimony level of the GLM and the quality of estimation. The higher the class number is, the

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

more accurate and less parsimonious the model becomes, and the more the quality of the estimation is affected. Guidelines are presented to help with model selection: the condition number and the estimation variances must be evaluated in relation to the jitter between the timestamps of events. Sufficient jitter on the timestamps within and between classes is necessary to prevent near collinearity in the $D^\dagger D$ matrix. In other words, sufficient jitter is essential to provide reliable estimates (Bardy, Dillon, & Dun, 2014), i.e. to be able to separate the overlapped responses. Estimation of the potentials requires the inversion of the $D^\dagger D$ matrix. The overall quality can be evaluated by the condition number of this matrix. For a given model, greater jitter allows the condition number to be reduced. Consequently the quality of the estimation should be increased by ensuring better separation between the overlapped potentials. From the point of view of model selection, increasing the number of classes is often combined with a decrease in intra-class jitter. Consequently, the condition number should increase, along with the variance estimation. Choosing the model with the highest number of classes means that a lower bias is favored over variance increase. For this reason, we suggest that this variance should be estimated in order to allow informed model selection.

The GLM is a very popular method and has been proposed in numerous studies as a meaningful tool to linearly deconvolve overlapped responses (Lalor, Pearlmutter, Reilly, McDarby, & Foxe, 2006); (Rivet, Souloumiac, Attina, & Gibert, 2009); (Dandekar, Privitera, Carney, & Klein, 2012); (Burns, Bigdely-Shamlo, Smith, Kreutz-Delgado, & Makeig, 2013); (Dias, Sajda, Dmochowski, & Parra, 2013); (Bardy, Dillon, & Dun, 2014), (Smith & Kutas, 2015ab); (Congedo, Korczowski, Delorme, & da Silva, 2016). However, in the context of EFRP, this methodology is not yet widely used (Dandekar, Privitera, Carney, & Klein, 2012); (Dias, Sajda, Dmochowski, & Parra, 2013); (Devillez, Kristensen, Guyader, Rivet, & Guérin-Dugué, 2015), while the overlapping issue remains a major concern. Moreover, in some cases, the assumptions of linear additive mixing of time-invariant responses may appear limited for the estimation of auditory evoked potentials, as discussed in (Bardy, Dillon, & Dun, 2014). Appropriate nonlinear models need to be designed to overcome these limitations. Nevertheless, the Generalized Additive Mixed-Effects Model (GAMM) has recently been proposed to take into account nonlinear relationships between co-variables where necessary (Tremblay &

Newman, 2015). Both the GLM, and the GAMM, are powerful statistical models for the EFRP estimation in complex situations with overlaps and modulations through both low-level (oculomotor behavior, stimulus properties) and high-level features (such as attentional resources, and arousal).

Acknowledgements

References

Baccino, T. (2011). Eye movements and concurrent event-related potentials : eye fixation-related potential investigations in reading. *Eye, 2*, 1. doi:10.1093/oxfordhb/9780199539789.013.0047

Bardy, F., Dillon, H., & Dun, B. V. (2014). Least-squares deconvolution of evoked potentials and sequence optimization for multiple stimuli under low-jitter conditions. *Clinical Neurophysiology, 125*, 727-737. doi:10.1016/j.clinph.2013.09.030

Barlow, J., & Ciganek, L. (1969). Lambda responses in relation to visual evoked responses in man. *Electroencephalography and clinical neurophysiology, 26*(2), 183-193.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation, 7*, 1129-1159. doi:10.1162/neco.1995.7.6.1129

Brouwer, A.-M., Reuderink, B., Vincent, J., van Gerven, A. J., Marcel, van Erp, B. F., & Jan. (2013). Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. *Journal of Vision, 13*, 17-17. doi:10.1167/13.3.17

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Burns, M. D., Bigdely-Shamlo, N., Smith, N. J., Kreutz-Delgado, K., & Makeig, S. (2013). Comparison of averaging and regression techniques for estimating event related potentials., (pp. 1680-1683). doi:10.1109/embc.2013.6609841

Congedo, M., Korczowski, L., Delorme, A., & da Silva, F. L. (2016). Spatio-temporal common pattern: A companion method for erp analysis in the time domain. *Journal of neuroscience methods, 267*, 74-88. doi:10.1016/j.jneumeth.2016.04.008

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Human brain mapping, 8*, 109-114. doi:10.1002/(sici)1097-0193(1999)8:2/3<109::aid-hbm7>3.3.co;2-n

Dandekar, S., Ding, J., Privitera, C., Carney, T., & Klein, S. A. (2012). The fixation and saccade p3. *PloS One, 7*, e48761. doi:10.1371/journal.pone.0048761

Dandekar, S., Privitera, C., Carney, T., & Klein, S. A. (2012). Neural saccadic response estimation during natural viewing. *Journal of neurophysiology, 107*, 1776-1790. doi:10.1152/jn.00237.2011

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*, 9-21. doi:10.1016/j.jneumeth.2003.10.009

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. *Journal of Vision, 15*, 20-20. doi:10.1167/15.13.20

Devillez, H., Kristensen, E., Guyader, N., Rivet, B., & Guérin-Dugué, A. (2015). The P300 potential for fixations onto target object when exploring natural scenes during a visual task after denoising overlapped EFRP., (pp. 1024-1027). doi:10.1109/ner.2015.7146801

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, 13*, 13-13. doi:10.1167/13.13.13

Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. M., & Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: analyses and review. *Journal of Experimental Psychology: General, 140*, 552. doi:10.1037/a0023885

Fischer, T., Graupner, S.-T., Velichkovsky, B. M., & Pannasch, S. (2013). Attentional dynamics during free picture viewing: evidence from oculomotor behavior and electrocortical activity. *Frontiers in Systems Neuroscience, 7*, 7-17. doi:10.3389/fnsys.2013.00017

Frey, A., Ionescu, G., Lemaire, B., López-Orozco, F., Baccino, T., & Guérin-Dugué, A. (2013). Decision-making in information seeking on texts: an eye-fixation-related potentials investigation. *Frontiers in systems neuroscience, 7*. doi:10.3389/fnsys.2013.00039

Gaarder, K., Krauskopf, J., Graf, V., Kropfl, W., & Armington, J. C. (1964). Averaged brain activity following saccadic eye movement. *Science, 146*, 1481-1483. doi:10.1126/science.146.3650.1481

Gastaut, Y. (1951). A little-known electroencephalographic sign: Occipital points occuring during opening of the eyes. *Revue neurologique, 84*(6), 635-640.

Hopfinger, J. B., & Ries, A. J. (2005). Automatic versus contingent mechanisms of sensory-driven neural biasing and reflexive attention. *Journal of cognitive neuroscience, 17*, 1341-1352. doi:10.1162/0898929055002445

Ionescu, G., Guyader, N., & Guérin-Dugué, A. (2009). SoftEye software. *IDDN. FR, 1*.

Jasper, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology, 10*, 371-375.

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. *Journal of Vision, 12*, 4-4. doi:10.1167/12.7.4

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. doi:10.1016/j.neuroimage.2013.12.006

Keren, A. S., Yuval-Greenberg, S., & Deouell, L. Y. (2010). Saccadic spike potentials in gamma-band EEG: characterization, detection and suppression. *Neuroimage, 49*, 2248-2263. doi:10.1016/j.neuroimage.2009.10.057

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Kiebel, S. J., & Holmes, A. P. (2003). *The general linear model.* Academic Press. doi:10.1016/b978-012372560-8/50008-5

Kristensen, E., Guerin-Dugué, A., & Rivet, B. (2017). Regularization and a General Linear Model for Event-Related Potential Estimation. *Behavior Research Methods*, 1-20. doi:10.3758/s13428-017-0856-z

Lalor, E. C., Pearlmutter, B. A., Reilly, R. B., McDarby, G., & Foxe, J. J. (2006). The vespa: a method for the rapid estimation of a visual evoked potential. *Neuroimage, 32*(4), 1549-1561. doi:10.1016/j.neuroimage.2006.05.054

Nikolaev, A. R., Meghanathan, R. N., & van Leeuwen, C. (2016). Combining EEG and eye movement recording in free viewing: Pitfalls and possibilities. *Brain and Cognition, 107*, 55-83. doi:10.1016/j.bandc.2016.06.004

Nikolaev, A., Nakatani, C., Plomp, G., Jurica, P., & van Leeuwen, C. (2011). Eye fixation-related potentials in free viewing identify encoding failures in change detection. *NeuroImage, 56*(3), 1598-1607. doi:10.1016/j.neuroimage.2011.03.021

Ossandón, J. P., Helo, A. V., Montefusco-Siegmund, R., & Maldonado, P. E. (2010). Superposition model predicts EEG occipital activity during free viewing of natural scenes. *The Journal of Neuroscience, 30*, 4787-4795. doi:10.1523/jneurosci.5769-09.2010

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. doi:10.1371/journal.pone.0157260

Rivet, B., Souloumiac, A., Attina, V., & Gibert, G. (2009). xDAWN algorithm to enhance evoked potentials: application to brain--computer interface. *IEEE Transactions on Biomedical Engineering, 56*, 2035-2043. doi:10.1109/tbme.2009.2012869

Scott, D. F., Groethuysen, U. C., & Bickford, R. G. (1967). Lambda responses in the human electroencephalogram. *Neurology, 7*(8), 770-778. doi:10.1212/wnl.17.8.770

Talsma, D. &. (2005). Selective Attention and Multisensory Integration: Multiple Phases of Effects on the Evoked Brain Activity. *Journal of Cognitive Neuroscience, 17*(7), 1098-1114.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7(14)*(4), 1-17.

Thickbroom, G. W., Knezevic, W., Carroll, W. M., & Mastaglia, F. L. (1991). Saccade onset and offset lambda waves: relation to pattern movement visually evoked potentials. *Brain research, 551*, 150-156. doi:10.1016/0006-8993(91)90927-n

Tremblay, A., & Newman, A. J. (2015). Modeling nonlinear relationships in ERP data using mixed-effects regression with R examples. *Psychophysiology, 52*, 124-139. doi:10.1111/psyp.12299

van Beers, R. J. (2007). The Sources of Variability in Saccadic Eye Movements. *The Journal of Neuroscience, 27*(33), 8757– 8770.

Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction. *Psychophysiology, 30*, 98-119. doi:10.1111/j.1469-8986.1993.tb03209.x

Yagi, A. (1979). Saccade size and lambda complex in man. *Physiological Psychology, 7*, 370-376. doi:10.3758/bf03326658

Yagi, A. (1981). Visual signal detection and lambda responses. *Electroencephalography and clinical Neurophysiology, 52*, 604-610. doi:10.1016/0013-4694(81)91434-6

Yagi, A. (1982). Lambda response as an index of visual perception. *Japanese Psychological Research, 24*(2), 106-110.

Yagi, A., Imanishi, S., Konishi, H., Akashi, Y., & Kanaya, S. (1998). Brain potentials associated with eye fixations during visual tasks under different lighting systems. *Ergonomics, 41*, 670-677. doi:10.1080/001401398186838

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

Appendix

This appendix gathers together a number technical comments concerning the ADJAR algorithm, and details of the full expressions of the second and third iterations. Finally, it features the theoretical expression of the variance for GLM estimation.

Description of ADJAR algorithm implementation

The ADJAR computation is based on the distributions $Q_{+1}(t)$ and $Q_{+1}(t)$ that are extracted from the set of epochs. The ADJAR algorithm for each iteration is based on a succession of convolution products. Theoretically, all convolutions used in the ADJAR algorithm involve signals defined from minus infinity to plus infinity. However, in practice, a finite number of samples had to be dealt with. To prevent the border effects of convolution and of signal truncation, the estimation window for the evoked potential is chosen to be larger than its window of interest, and a tapering window $h(t)$ is applied after each convolution throughout the estimation window. In our case, the window of interest is equal to $[T_o; T]$ ($T_o = -250$ ms and $T = 600$ ms) and the larger estimation window is equal to $[T_{start}; T_{end}]$ ($T_{start} = -450$ ms and $T_{end} = 800$ ms). Figure 8a illustrates these windows. For the tapering window $h(t)$, a Tukey window with a coefficient of 0.32 was chosen, instead of the usual Hamming window. The Tukey and Hamming windows are plotted on the estimation window in Figure 9. The Tukey window preserves the signal between -250 and 600 ms (the amplitude is equal to 1), inside the window of interest, unlike the Hamming window.

The convolutions of the waveform by the distributions shift the waveform to the left (convolution by $Q_{-1}(t)$) or to the right (convolution by $Q_{+1}(t)$). Figure 8b illustrates the result of the convolution with the distribution $Q_{+1}(t)$, which is shifted to the right. We consider the signal in the estimation window from T_{start} to T_{end} and the distribution $Q_{+1}(t)$ with a support between $T_{Q_{start}}$ and $T_{Q_{end}}$. The signal after convolution is shifted to the right and its temporal support is larger between $T_{C_1} = T_{Q_{start}} + T_{start}$ and $T_{C_2} = T_{Q_{end}} + T_{end}$. Then, after each convolution, the result is truncated with the tapering window to take into consideration only the signal that remained in the estimation window. The steps at each iteration are defined here:

$$\widehat{ov}_s^k(t) = Q_{+1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_p^{k-1}(t) \right) \quad (19)$$

$$\widehat{ov}_s^k(t) \leftarrow h(t).\widehat{ov}_s^k(t) \quad (20)$$

$$\widehat{ov}_p^k(t) = Q_{-1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_s^k(t) \right) \quad (21)$$

$$\widehat{ov}_p^k(t) \leftarrow h(t).\widehat{ov}_p^k(t) \quad (22)$$

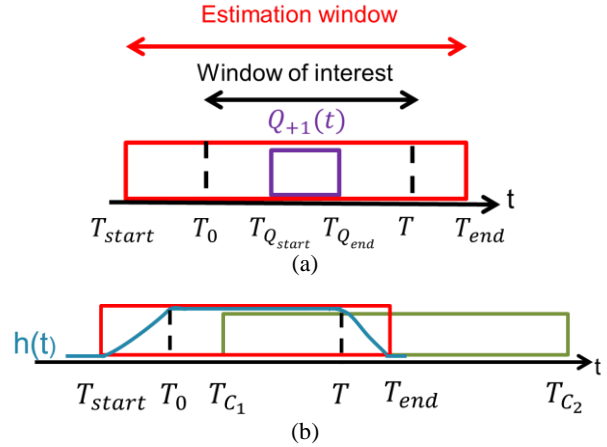where $h(t)$ is a tapering window along the estimation window $[T_{start}; T_{end}]$.

Figure 8 Illustrations of a) the estimation window and the window of interest b) temporal support of the output signal by the convolution with the distribution $Q_{+1}(t)$, between T_{C_1} and T_{C_2}. The tapering window h(t) is in blue
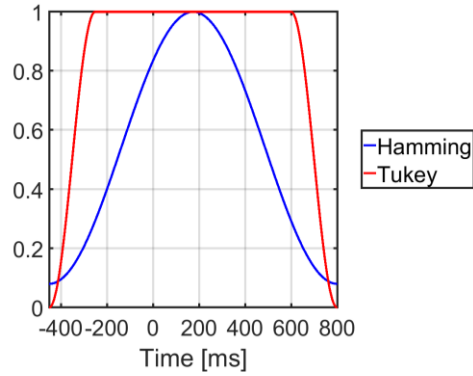

Figure 9 Tukey and Hamming windows

Development of the iterative procedure for the ADJAR algorithm

Adjacent response overlaps ($ov_p(t)$ and $ov_s(t)$) are estimated by an iterative procedure. At convergence, they are subtracted from the evoked potential estimated by the av-

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

erage of the time-locked signal $\hat{a}^{Av}(t)$ to obtain an estimation $\hat{a}^{Adjar}(t)$ of the underlying potential $a(t)$. The model is:

$$\hat{a}^{Av}(t) = a(t) + ov_p(t) + ov_s(t), \tag{23}$$

$$ov_p(t) = Q_{-1}(t) * a(t),$$

$$ov_s(t) = Q_{+1}(t) * a(t),$$

where $Q_{-1}(t)$, respectively $Q_{+1}(t)$, is the normalized distribution of the timestamps at the onset of the previous and subsequent fixations respectively. Starting from $\widehat{ov}_p^0(t) = 0$, equations (7) and (8) in the main part of the article, describe the kth estimation of the response overlaps. We describe here the development of the first interactions in order to show in the final estimation, the particular status of second-order response overlaps highlighted in red in the following equations.

At the first iteration:

$$\widehat{ov}_s^1(t) = Q_{+1}(t) * \hat{a}^{Av}(t) \tag{24}$$

$$\widehat{ov}_s^1(t) = Q_{+1}(t) * a(t) + \textcolor{red}{Q_{+1}(t) * Q_{+1}(t) * a(t)} + Q_{+1}(t) * Q_{-1}(t) * a(t)$$

$$\widehat{ov}_p^1(t) = Q_{-1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_s^1(t) \right) \tag{25}$$

$$\widehat{ov}_p^1(t) = Q_{-1}(t) * a(t) + \textcolor{red}{Q_{-1}(t) * Q_{-1}(t) * a(t)} - Q_{-1}(t) * Q_{+1}(t) * Q_{+1}(t) * a(t) - Q_{-1}(t) * Q_{+1}(t) * Q_{-1}(t) * a(t)$$

$$\hat{a}^1(t) = \hat{a}^{Av}(t) - \widehat{ov}_s^1(t) - \widehat{ov}_p^1(t) \tag{26}$$

$$\hat{a}^1(t) = a(t) - \textcolor{red}{Q_{+1}(t) * Q_{+1}(t) * a(t)} \\ - \textcolor{red}{Q_{-1}(t) * Q_{-1}(t) * a(t)} \\ - Q_{+1}(t) * Q_{-1}(t) * a(t) \\ + Q(t) * Q_{+1}(t) * Q_{+1}(t) * a(t) \\ + Q_{-1}(t) * Q_{+1}(t) * Q_{-1}(t) * a(t)$$

At the second iteration:

$$\widehat{ov}_s^2(t) = Q_{+1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_p^1(t) \right) \tag{27}$$

$$\widehat{ov}_s^2(t) = Q_{+1}(t) * a(t) + \textcolor{red}{Q_{+1}(t)^{(*2)} * a(t)} \\ - Q_{+1}(t) * Q_{-1}(t)^{(*2)} * a(t) \\ + Q_{-1}(t) * Q_{+1}(t)^{(*3)} * a(t) \\ + Q_{+1}(t)^{(*2)} * Q_{-1}(t)^{(*2)} * a(t)$$

$$\widehat{ov}_p^2(t) = Q_{-1}(t) * \left(\hat{a}^{Av}(t) - \widehat{ov}_s^2(t) \right) \tag{28}$$

$$\widehat{ov}_p^2(t) = Q_{-1}(t) * a(t) + \textcolor{red}{Q_{-1}(t)^{(*2)} * a(t)} \\ - Q_{-1}(t) * Q_{+1}(t)^{(*2)} * a(t) \\ + Q_{+1}(t) * Q_{-1}(t)^{(*3)} * a(t) \\ - Q_{-1}(t)^{(*3)} * Q_{+1}(t)^{(*3)} * a(t) \\ - Q_{+1}(t)^{(*2)} * Q_{-1}(t)^{(*3)} * a(t)$$

$$\hat{a}^2(t) = \hat{a}^{Av}(t) - \widehat{ov}_s^2(t) - \widehat{ov}_p^2(t) \tag{29}$$

$$\hat{a}^2(t) = a(t) - \textcolor{red}{Q_{+1}(t)^{(*2)} * a(t)} - \textcolor{red}{Q_{-1}(t)^{(*2)} * a(t)} \\ + Q_{+1}(t) * Q_{-1}(t)^{(*2)} * a(t) \\ + Q_{-1}(t) * Q_{+1}(t)^{(*2)} * a(t) \\ - Q_{-1}(t) * Q_{+1}(t)^{(*3)} * a(t) \\ - Q_{+1}(t) * Q_{-1}(t)^{(*3)} * a(t) \\ - Q_{+1}(t)^{(*2)} * Q_{-1}(t)^{(*2)} * a(t) \\ + Q_{-1}(t)^{(*3)} * Q_{+1}(t)^{(*3)} * a(t) \\ + Q_{+1}(t)^{(*2)} * Q_{-1}(t)^{(*3)} * a(t)$$

And so on. For any k^{th} iteration, the terms $Q_{+1}(t)^{(*2)} * a(t)$ and $Q_{-1}(t)^{(*2)} * a(t)$ remain in the estimate $\hat{a}^k(t)$. The other high-order terms are progressively shifted outside the estimation window, due to the combination of $Q_{-1}(t)$ and $Q_{+1}(t)$. At convergence, the final estimations of the response overlaps are:

$$\widehat{ov}_p^\infty(t) = Q_{-1}(t) * a(t) + \textcolor{red}{Q_{-1}(t)^{(*2)} * a(t)} \\ +/- \sum high\ order\ ov_p \tag{30}$$

$$\widehat{ov}_s^\infty(t) = Q_{+1}(t) * a(t) + \textcolor{red}{Q_{+1}(t)^{(*2)} * a(t)} \\ +/- \sum high\ order\ ov_s \tag{31}$$

At convergence, the final estimate $\hat{a}^{Adjar}(t)$ can be expressed as:

$$\hat{a}^{Adjar}(t) = \hat{a}^{Av}(t) - \widehat{ov}_s^\infty(t) - \widehat{ov}_p^\infty(t)$$

$$\hat{a}^{Adjar}(t) = a(t) - \textcolor{red}{Q_{-1}(t)^{(*2)} * a(t)} - \textcolor{red}{Q_{+1}(t)^{(*2)}} \\ \textcolor{red}{* a(t)} \\ +/- \sum high\ order\ overlaps \tag{32}$$

This last equation corresponds to equation (10) in the main part of the article.

Convergence issue of a modified ADJAR algorithm

It is tempting to adapt the ADJAR algorithm to take into account all previous and subsequent fixations, as they effectively occur during the epoch. Indeed, the basic principle of the ADJAR algorithm is to estimate the evoked

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

potential at the fixation of interest by estimating the overlaps of previous and subsequent responses, taking into consideration only three evoked potentials which occur during an epoch. However in practice, more than three fixations (the fixation of interest, the previous and the subsequent ones) can occur during a single epoch. Consequently, equation (23) is rewritten as:

$$\hat{a}^{Av}(t) = a(t) + Q_-(t) * a(t) + Q_+(t) * a(t) \qquad (33)$$

$$Q_-(t) = \sum_{f=-1}^{-F_p} Q_f(t) = \frac{1}{E} \sum_{i=1}^{E} \sum_{f=-1}^{-F_p} \delta(t - \tau_i^f)$$

$$Q_+(t) = \sum_{f=1}^{F_s} Q_f(t) = \frac{1}{E} \sum_{i=1}^{E} \sum_{f=1}^{F_s} \delta(t - \tau_i^f)$$

where $Q_-(t)$ and $Q_+(t)$ are the distributions of all previous and subsequent fixation onsets across all epochs, respectively. Since the number of fixations per epoch can be greater than three (previous, current and subsequent) in this case, these distributions cannot be normalized to one (i.e. their integral can differ from one). In one epoch, there is one current fixation (the EEG signal is time-locked on this fixation onset), and possibly more than one previous fixation, and also more than one subsequent fixation. The integral of $Q_-(t)$ (resp. $Q_+(t)$) 's elements is equal to the average number of previous (resp. subsequent) fixations occurring in the epochs. This is possibly greater than the number of the epochs, E. By dividing by E, the sum of the contribution of $Q_+(t)$'s elements (or $Q_-(t)$) can be greater than one. In order to take into account all fixations during an epoch, $Q_{-1}(t)$ (resp. $Q_{+1}(t)$) is replaced by $Q_-(t)$ (resp. $Q_+(t)$) in equations (24) and (25) to estimate the current EFRP. This modified ADJAR algorithm was implemented to estimate the EFRP in the middle of visual exploration. During the iterative estimation of the response overlaps, the powers of estimates $\widehat{ov}_p^k(t)$ and $\widehat{ov}_s^k(t)$ increased along iterations because of cumulative convolutions with distributions that were not normalized. Consequently this modified ADJAR algorithm did not converge, as shown in Figure 10 illustrating the ratio expressed in dB, between the power of $\hat{a}^{Adjar}(t)$ and $\hat{a}^{Av}(t)$ against the number of iterations. For the original ADJAR algorithm, this ratio converged to a finite value. However when all fixations inside each epoch were considered, this ratio diverged when the number of iterations increased due to iterative

convolutions with distributions $(Q_-(t), Q_+(t))$, which were not normalized to one.
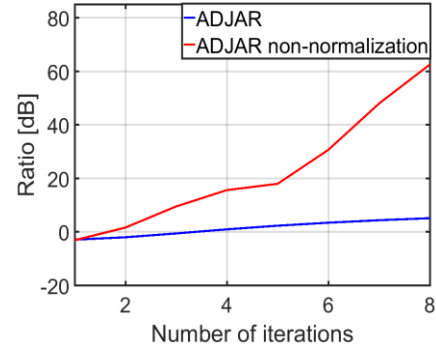

Figure 10 Ratio [dB] between the powers of $\hat{a}^{Adjar}(t)$ and $\hat{a}^{Av}(t)$ against the number of iterations for two cases: blue line for the original ADJAR algorithm and red line for the same iterative algorithm considering all fixations with non-normalized distributions

Theoretical Variance for the GLM estimation

Equation (14) can be developed:

$$\hat{a} = (D^\dagger D)^{-1} D^\dagger (Da + n) \qquad (34)$$

and the variance $Var[\hat{a}]$ is equal to:

$$Var[\hat{a}] = E[(\hat{a} - E[\hat{a}])(\hat{a} - E[\hat{a}])^\dagger]$$

$$Var[\hat{a}] = (D^\dagger D)^{-1} D^\dagger E[nn^\dagger]((D^\dagger D)^{-1} D^\dagger)^\dagger$$

$$Var[\hat{a}] = (D^\dagger D)^{-1} D^\dagger E[nn^\dagger] D (D^\dagger D)^{-1} \qquad (35)$$

with $E[.]$ the mathematical expectation.

The more singular the $D^\dagger D$ matrix is, the higher the variance is.

EFRP Estimation in the middle of visual exploration with a larger estimation window

We present here the results obtained in the same condition as in the subsection in Figure 2, for the estimation of EFRP in the middle of visual exploration, but with a larger estimation window [-400; 1550] ms, for the Average method, and the GLM. With such a large window, the ability of the estimation method to separate the precedent and subsequent evoked potentials from the current one was observed. Concerning the estimate given by the Average method, residual activity from the precedent potential was

Journal of Eye Movement Research
10(1):7, 1-27

Kristensen, E., Rivet, B., & Guérin-Dugué A. (2017)
Estimation of overlapped EFRP: The GLM, a more flexible framework than the ADJAR algorithm

noted between -400 and -300 ms. This activity was not present for the estimate given by the GLM. Moreover, with this window, the expected return to zero at the end of the estimation window was more visible than with a shorter window. This criterion was only met for the estimate given by the GLM.
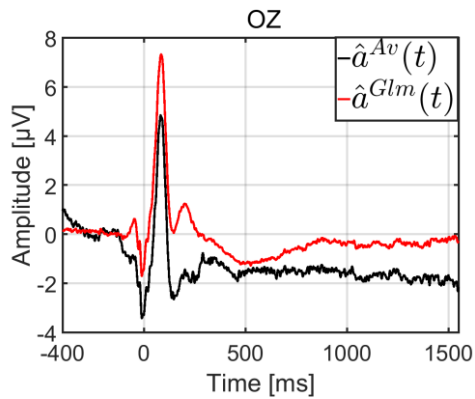

Figure 11 Grand average for the two methods: Average ($\hat{a}^{Av}(t)$), and GLM ($\hat{a}^{Glm}(t)$) on the OZ electrode