Journal of Neuroscience Methods 313 (2019) 77–94

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

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

Improving free-viewing fixation-related EEG potentials with continuous-time regression

Tim Cornelissen[a,*], Jona Sassenhagen[b], Melissa Le-Hoa Võ[a]

[a] Scene Grammar Lab, Department of Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany
[b] FiebachLab Department of Psychology, Goethe University Frankfurt, Frankfurt am Main, Germany

ARTICLE INFO

Keywords:
Eye tracking
Eye movements
EEG
FRP
Regression
rFRP
Free viewing

ABSTRACT

Background: In the analysis of combined ET-EEG data, there are several issues with estimating FRPs by averaging. Neural responses associated with fixations will likely overlap with one another in the EEG recording and neural responses change as a function of eye movement characteristics. Especially in tasks that do not constrain eye movements in any way, these issues can become confounds.

New method: Here, we propose the use of regression based estimates as an alternative to averaging. Multiple regression can disentangle different influences on the EEG and correct for overlap. It thereby accounts for potential confounds in a way that averaging cannot. Specifically, we test the applicability of the rERP framework, as proposed by Smith and Kutas (2015b), (2017), or Sassenhagen (2018) to combined eye tracking and EEG data from a visual search and a scene memorization task.

Results: Results show that the method successfully estimates eye movement related confounds in real experimental data, so that these potential confounds can be accounted for when estimating experimental effects.

Comparison with existing methods: The rERP method successfully corrects for overlapping neural responses in instances where averaging does not. As a consequence, baselining can be applied without risking distortions. By estimating a known experimental effect, we show that rERPs provide an estimate with less variance and more accuracy than averaged FRPs. The method therefore provides a practically feasible and favorable alternative to averaging.

Conclusions: We conclude that regression based ERPs provide novel opportunities for estimating fixation related EEG in free-viewing experiments.

1. Introduction

1.1. Background

Both eye tracking (ET) and electroencephalography (EEG) have a rich history of experimental paradigms and results that have revealed much about the workings of visual perception. Studies are increasingly combining these two techniques into what we will refer to as ET-EEG experiments. Where eye movements (EMs) offer a close insight into the temporal and spatial allocation of attention, EEG is informative about what happens in the brain before, during, and after the eyes land on a certain region of a stimulus. EEG-only paradigms often require observers to not move their eyes and not blink during trials. Such paradigms without eye movements (static EEG experiments) often aim to infer the workings of perception in everyday life. But eye movements are abundant in everyday tasks (e.g. Hayhoe and Ballard, 2005). Allowing eye movements in experiments is a natural step towards studying these processes in ways that generalize better to situations outside the lab.

ET-EEG combinations have been proposed for several purposes. These include Brain-Computer Interfacing (Brouwer et al., 2013; Hale et al., 2008; Shishkin et al., 2016; Slanzi et al., 2017), comparing neural responses under both static and eye movement scenarios (Dandekar et al., 2012a), studying reading under more natural conditions (Dimigen et al., 2011; Hutzler et al., 2013), studying early ERP components related to saccades (Dandekar et al., 2012b; Thickbroom et al., 1991), investigations of visual search (Devillez et al., 2015; Kamienkowski et al., 2012; Kaunitz et al., 2014; Körner et al., 2014), emotional processing (Simola et al., 2013), art perception (Fischer et al., 2013), and many other contexts. As research questions have differed in the past, so have the proposed approaches to ET-EEG paradigms, data processing, and analysis. Here, we focus on recordings

* Corresponding author at: Scene Grammar Lab, Department of Cognitive Psychology, Theodor-W.-Adorno-Platz 6, 60316, Frankfurt am Main, Germany.
E-mail address: workmailfortim@gmail.com (T. Cornelissen).

https://doi.org/10.1016/j.jneumeth.2018.12.010
Received 6 December 2017; Received in revised form 14 December 2018; Accepted 14 December 2018
Available online 24 December 2018

made during *free-viewing*, i.e. conditions under which eye movement behavior was not influenced or restricted by the experimenter in some way other than telling participants to perform a task. Under the current definition, free-viewing could therefore occur in different research interests, from reading to visual search and scene perception.

Because the signal to noise ratio (SNR) of EEG measurements is relatively low, data from a single stimulus presentation usually do not allow inferences about the stimulus related activity of interest. Instead, Event Related Potentials (ERPs) are obtained through averaging (Luck, 2014). To estimate the neural response related to the onset of different eye movements, which is what is addressed in this paper, one can apply the same logic to ET-EEG data. Cutting periods of EEG around the onset of, for instance, fixations and averaging these data periods, yields a fixation related potential (FRP). But changing from ERPs to FRPs, especially in a free-viewing scenario, involves more than merely time locking data epochs to a different kind of event. Trying to obtain fixation-locked EEG responses introduces several analysis challenges. We will present a possible solution to these challenges in the form of regression based ERPs (Smith and Kutas, 2015a, 2015b). We test the applicability of regression based ERP estimation to ET-EEG data, by analyzing a standard visual search task and a scene memorization task.

ERP is a widely accepted and used term, but almost synonymous with averaging. To indicate that the ERP was estimated using regression, Smith and Kutas (2015a, 2015b) propose the abbreviation rERP. A fixation related potential is essentially a sub-category of ERP (the "event" in ERP, after all, does not specify that the event has to be a stimulus onset) and seems almost equally tied to averaging. Therefore, we propose the abbreviation rFRP to indicate regression based estimates of fixation related potentials.

1.2. Possible pitfalls and current solutions

1.2.1. Possible pitfalls in combining eye movements and EEG

Dimigen et al., (2011) point out four main issues in co-registration and analysis of ET-EEG data. First, measuring without interference and the (post-hoc) temporal alignment of both streams of data. Second, the EEG artifact introduced by eye movements. Third, overlapping neural activity caused by eye movements happening in relatively short succession. Finally, neural activity itself might be related to the characteristics of an eye movement. Technical issues can be resolved in different ways depending on make and model of hardware in individual labs, and by using specially designed software (e.g., http://www2.hu-berlin.de/eyetracking-eeg). Dealing with corneoretinal and myogenic eye movement artifact is, to some extent, not new to the field. Although close consideration of the efficacy of different correction methods for "heavily contaminated" EEG data such as those recorded in ET-EEG experiments is warranted, current methods seem adequate for correcting EM artifact in a combined ET-EEG experiment (e.g., Independent Component-based EOG artefact correction; Delorme et al., 2012). The remaining two issues are the ones that are of interest to this paper.

Similar to the onset of a visual stimulus, a fixation will yield a neural response that overlaps with those from the previous and the next fixation. The experimenter does not control the succession rate of the eye movements and therefore does not control the degree of overlap. Therefore, the overlap becomes a possible confound when comparing two conditions. In the example of a visual search experiment, neural responses measured upon target fixations can be compared to those measured upon non-target fixations. The difference will probably be largely due to target recognition (Brouwer et al., 2013; Dandekar et al., 2012a; Devillez et al., 2015; Körner et al., 2014), but could be distorted if targets are fixated systematically longer than non-targets. The shape of this distortion will be hard to predict, due to the fluctuating shape of the overlapping responses.

The fourth issue, the *relationship between eye movement parameters and neural activity*, can be a subject of study in its own right (Kazai and Yagi, 1999; Nikolaev et al., 2016; Thickbroom et al., 1991). When it is not, differing EMs can confound EEG comparisons. When comparing targets and non-targets one might interpret an early difference in FRPs as an indicator of rapid, extrafoveal target recognition. There are, however, early FRP responses that vary with EM characteristics such as saccade amplitude (Dimigen et al., 2011; Nikolaev et al., 2016). What makes accounting for EM confounds more difficult is that despite an increase in research (Nikolaev et al., 2016) we neither know exactly what eye movement characteristics influence EEG, nor how. This holds especially true for late periods of FRPs. Furthermore, having a better description of EM confounds in the EEG would only caution over-interpretation of difference waves, not control for the confounds.

1.2.2. Current solutions to overlapping responses

A popular approach to minimizing overlapping responses is the exclusion of epochs if the fixation offset falls inside the EEG period of interest. This ensures calculation of the component is not affected by the latency of the next saccade and fixation in any condition. Some FRP components are, however, expected to occur relatively late after a fixation onset, for example, the P300 (Polich, 2007) or N400 (Kutas and Federmeier, 2011). Such late components require a minimum fixation duration of > 600 ms. Yet saccades are typically made at a rate of 3–4 per second (Buswell, 1935; Henderson, 2003; 2007; Rayner, 2009; Yarbus, 1967) and in the fields of reading and scene perception such long fixations are rather a-typical. Post-hoc selection of long fixations therefore leads to a biased sample of data that does not reflect typical eye movements and neural responses.

Researchers also restrict analysis to fixations occurring after a minimum amount of time since stimulus onset (e.g., an estimation of the time it takes the EEG to return to baseline after stimulus onset; Dimigen et al., 2011), again discarding otherwise usable data. Alternatively, or additionally, stimuli can be chosen to cause long fixations. Participants can also be instructed (e.g. Kamienkowski et al., 2012) or even trained (Kaunitz et al., 2014) to prolong fixations and prevent refixations of the same regions. Experiments can also be designed so that events of interest occur always at the end of a sequence (Hutzler et al., 2013; Kaunitz et al., 2014). While all these approaches effectively avoid overlapping responses, they are of limited use for the kind of free-viewing ET-EEG experiments we aim to investigate.

1.2.3. Regression based overlap correction

Current solutions focus on analyzing EEG data that do not contain overlapping neural responses at all. Here, we propose to *account* for temporal overlap by explicitly modelling it, in continuous EEG activity (Smith and Kutas, 2015b). Similar methods have previously been applied to ET-EEG data (Dandekar et al., 2012b; Kristensen et al., 2017a, b) and the effectiveness of the rERP framework in correcting for overlap has been shown in simulations (Smith and Kutas, 2015b). Recently a method using the same principals was shown to yield better results than simple averaging in ET-EEG data (Kristensen et al., 2017a) and to be more flexible than the popular ADJAR method (Kristensen et al., 2017b; Woldorff, 1993). The main additional benefit of applying regression as described by Smith and Kutas (2015b) is that it not only accounts for overlapping potentials, but for a wide range of EM-related confounds, through the addition of eye movement parameters as numerical predictors to the same regression models.

1.2.4. Current solutions to eye movement related confounds

In addition to preventing or excluding overlapping neural responses, researchers have aimed to prevent neural responses associated with EM characteristics from confounding FRP analyses. Again, current solutions involve experimental design, e.g. by restricting participants' eye movements, so that data from all conditions contain EMs with similar directions, durations, amplitudes, etc. (Brouwer, Brinkhuis et al., 2014; Dandekar et al., 2012a; Huber-Huber et al., 2016). However, a design that times and limits eye movements is by definition not a solution for investigating free-viewing.

In free-viewing scenarios, researchers have aimed to balance out EM confounds during data processing. Commonly, a test for significant differences in EM parameters across or within participants is employed. A lack of statistically significant differences is then used to argue eye movements are similar enough between conditions (Kamienkowski et al., 2012; Kaunitz et al., 2014; Nikolaev et al., 2016; Devillez et al., 2015). We question this application of null hypothesis significance testing (NHST), because failing to reject the null hypothesis does not confirm it. Instead, confound control should establish the lack of a meaningful influence from confounding independent variables on the dependent one (Sassenhagen and Alday, 2016). Without defining "meaningful" in the context of EM-EEG relationships, the use of NHST for confound control is therefore fundamentally flawed.

Finally, an EM difference between conditions might in itself be reason for an FRP study. To relate findings from static EEG paradigms to eye movement behavior, or to investigate what the neural correlate of an EM effect is, comparing data with minimized EM differences is undesirable.

In sum, current solutions to eye movement confounds involve discarding otherwise informative data, and can lead to the inclusion of a subset of data that are not representative of general eye movement behavior. Moreover, we question whether current solutions using null hypothesis testing actually prevent eye movement confounds.

1.3. The rERP procedure

The alternative proposed here is to account for both temporal overlap and EM related EEG confounds through linear regression. More specifically, we apply the rERP framework proposed by Smith and Kutas (2015a, 2015b; see also Hauk et al., 2009) to ET-EEG data. The framework extends previous regression based approaches, mainly by incorporating numerical predictors (see below) to account for EM confounds. To investigate the use of rERPs in ET-EEG data, we recorded EEG and eye movements simultaneously while participants performed a visual search task. Visual search should yield a distinct difference in EEG responses upon fixation of a target versus fixation of a distracter (the P300 component; Polich, 2007). The recorded data are used to inspect whether the rERP framework provides sensible estimates of confounds to the P300 in the example experiment. Additional data were recorded during a scene memorization task. These data serve as a comparison where applicable and are included for two reasons: Firstly, to show robustness of the rERP method across participants and tasks. Secondly, to show the influence of using stimuli that are more naturalistic and thus vary in physical characteristics more strongly. Although a brief introduction to regression based estimation of ERPs is given below, we refer the interested reader to the original publications (Smith and Kutas, 2015a; 2015b) for a more elaborate and complete description.

1.3.1. A brief description of the rERP framework

The rERP framework offers an alternative to estimating ERPs by averaging. Rather than average epochs across trials at each latency to the time-locking event, a linear model could be fitted at each latency. Within this model, the traditional event (e.g. a stimulus onset or a fixation) becomes a predictor. A simple model would be:

$$y_i = \beta_1 x_{1i} + noise_i$$

Where y is the EEG, i represents a trial number, x_1 is a predictor value and β_1 is a regression coefficient. Here, x_1 only takes the value 1 because a fixation either happened or it did not. β_1 can then also be called the "intercept term" of the fixation event. The y_i values are set to the measured scalp potential at a single electrode, at a single latency, across different time-locked trials. Since the values of y_i and x_{1i} are known, β_1 can be estimated by minimizing the squared noise: We assume Y = XB, where Y represents the EEG data, X a matrix of predictor values, and B a matrix of β values. The β values can then be estimated as $B = (X^TX)^{-1}X^TY$, where X^T is the transpose of X. Estimating β values is also known as fitting the model to the data.

The estimated value for β_1 represents the estimated influence of the predictor x_1 on the EEG. By plotting the different β_1 values (each obtained from fitting the model at a single latency within the epochs) in temporal order, a waveform of β values emerges. This waveform is the regression based ERP, or rERP. When there is no overlap in neural responses, this waveform actually is the mathematical equivalent of an average across time (Smith and Kutas, 2015a). Its unit is the same as the dependent variable, in this case microvolts, and these waveforms can be baselined and statistical tests can be performed on them as if they were an ERP subject average.

1.3.2. The inclusion of eye movements as numerical predictors

So far, FRP averaging was re-defined in regression terms. Yet there are other advantages of linear estimation of ERPs. Multiple predictors can be added to the model and, given sufficient data, it will separate the influence of these predictors. Such as:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_n x_{ni} + noise_i$$

What is more, its predictors need not be 0's or 1's. Linear estimation allows for the inclusion of numerical predictors. The resulting rERP waveform (a so-called slope-rERP) then represents the change in EEG that is related to one unit change in the predictor value. This could, for instance, represent increased negativity in the N400 time window as the cloze probability of a fixated word decreases (cf. Dimigen et al., 2011). Linear estimation of ERPs with intercept terms as well as numerical predictors has been applied before (Ehinger et al., 2015; Rousselet et al., 2010, 2009; Rousselet et al., 2008), and it applies to ET-EEG because it allows entering the rich description of the eye movement behavior measured with the eye tracker, as predictors into one regression model. This is different from using regression only to correct for overlap. The main interest of Dandekar et al. (2012b), for example, was to study the EEG response around a saccade. Although the authors did construct different linear estimates for different saccade sizes, the regression only served the purpose of correcting overlap. Here, we go one step further, mainly to *account* for any EEG changes related to eye movements rather than study them. Only correcting for overlap will therefore not suffice. Analysis also needs to account for changes in neural responses as a function of eye movements. To achieve this, the eye movements are explicitly modelled using numerical predictors. A model that takes into account saccade amplitude, for example, can then be defined as:

$$y_i = \beta_{fixation} x_{fixation_i} + \beta_{saccadeAmplitude} x_{saccadeAmplitude_i} + noise_i$$

This model contains one intercept term for all fixations (rather than bins) under the assumption that different predictors (e.g. incoming saccade amplitude) add to or subtract linearly from this intercept.[1] The influence of saccade amplitude is now estimated in the presence of other predictors, rather than used as a dimension to bin along, or to exclude EEG data points (like when using matching procedures before averaging). Models may include many more eye movement parameters than just saccade amplitude. Similarly, the experimental manipulation can be included as another predictor influencing the same fixation intercept. In a search task, for instance, one would add a predictor that codes whether a fixation fell on a target or not, along with any other predictors thought relevant:

$$y_i = \beta_{fixation} x_{fixation_i} + \beta_{saccadeAmplitude} x_{saccadeAmplitude_i} + \beta_{targetFixation} x_{targetFixation_i} + \ldots + \beta_n x_{ni} + noise_i$$

[1] We will return to the validity of such an assumption and model construction in general in the Discussion section.

From the equation above, it follows that the process of estimating the influence of different predictors minimizes their confounding effects on one another. That is, variance in the data that the model attributes to one predictor cannot be attributed to another predictor - as it would during averaging if there is any correlation between predictors. Variance that the model estimates to be due to eye movement characteristics therefore cannot be attributed to the estimated effect of the fixation being on a target. This is a key difference between averaged FRPs and *r*FRPs. In contrast to linear estimates, FRPs are used to isolate one effect at a time, by binning data under the assumption that the bins only differ along one dimension before averaging and comparing them. When constructing FRPs from free-viewing data, the assumption of bins being different along a single dimension is problematic, as outlined in Section 1.2. Critically, multiple regression does not require data-points to vary along a single dimension, but takes all data into account simultaneously and "disentangles" the influence of multiple predictors in a single model fit. Thus, continuous predictors eliminate the need to categorize a continuous parameter such as saccade size into bins (Dandekar et al., 2012b). We know of only one example in the literature including EM parameters as continuous predictors in the linear estimation of rFRPs (Weiss et al., 2016). The authors, however, gave no indication of the efficacy of the method.

1.3.3. Adding overlap correction

So far, rFRPs were constructed fitting a model per time-point in epoched data. Yet if epochs contain multiple fixations, overlap confounds still exist. The series of models that estimate rFRPs can be transformed into a model that corrects for overlap, by entering the value of the same predictor multiple times, at different latencies (Smith and Kutas, 2015b). One predictor then effectively becomes multiple predictors and multiple models at different time lags become one large model that estimates the influence of the predictors from the continuous data of each electrode. For example, one could include the predictors "fixation at latency -200 ms" to "fixation at latency + 700 ms", with the total number of "fixation ± xms" predictors depending on the sampling rate.

In Fig. 1, this leads to the diagonal lines seen in the design matrix. Each diagonal is the set of predictor values associated with a single eye movement event. The predictor "element fixation onset, latency -200 ms" has value 1 at any sample that is taken 200 ms before the onset of a fixation on a search element and a 0 at any time point that is not. At each time point (best imagined as a horizontal line through the data and the design matrix), the EEG data can now be thought of as the result of multiple fixations, *at different latencies since onset*. Another way to describe the same design matrix is to say that each column is a separate predictor. By adding events at different latencies as separate predictors, the overlap correction becomes another matter of estimating the influences of multiple predictors, which can be achieved by solving the least squares problem just as before. In terms of Fig. 1, this means that we model the EEG data to be the product of the β values and the design matrix (Y = XB in the figure). Because we know the design matrix and the data, we can estimate the β values (b). The rERP for "element fixation onset" is now constructed by plotting the β values of all "element fixation onset ± xms" in temporal order (the bottom waveforms in Fig. 1). By estimating β values per latency, the value of the rFRP can vary arbitrarily per time point, but always and only based on the data.

1.3.4. Testing the efficacy of the rERP framework in ET-EEG data

Given that overlap correction by means of linear regression is effective (Kristensen et al., 2017a, 2017b), and a theoretically sound way to include numerical predictors when estimating rERPs (Smith and Kutas, 2015a) one goal of the current paper is to test whether linear estimates yield sensible results in real, experimental ET-EEG data. To this end we recorded data in a visual search task in which participants searched through an array of abstract elements on a gray background, as well in a scene memorization task with more variable stimuli. To these datasets, we applied the type of linear models described above and detailed below. Fitting such models will yield an rERP for each predictor (e.g. the bottom four waveforms in Fig. 1). The rERP represents the closest linear estimate of a predictor's influence on the EEG associated with the event it is time-locked to. The goal is to *account* for these predictors' influences. Therefore, the key question to answer first is whether the output of fitting the model represents sensible estimates of eye movement influences. To do this, we attempt to isolate the effects of different predictors in the current data, and compare the outputs of a model to such an "isolation" of each predictor.

Ironically, the current gold standard for isolating the effect of a *single* predictor on EEG, is to generate averaged FRPs. FRPs yield less than perfect estimates for the dimension along which bins are constructed, for reasons outlined in previous sections. FRPs are, however, the only standard available, with a fairly good understanding of its imperfections. As an imaginary example, let participants freely make a series of fixations. To infer the influence of vertical position, fixations can be binned, e.g. one bin for fixations above center of screen, and one bin for all fixations below center. Averaging both bins and subtracting their averages yields an isolation of the influence of vertical position:

$$\text{FRP}_{\text{ Fixations above center}} - \text{FRP}_{\text{ Fixations below center}} = \text{FRP Difference Wave}_{\text{ Vertical position}}$$

All the objections previously raised about this kind of comparison still hold (bins might differ along other dimensions than vertical position too, and bins are not corrected for overlap). Still, a linear estimate of the influence of the predictor "vertical position" should resemble the FRP difference wave. Moreover, the rFRP should resemble its paired FRP, but not the FRPs for the other predictors. Below we make such comparisons for each predictor in an rERP model. Results for both search and scene memorization indicate that the rERP framework accounts for eye movement differences in the data in a sensible manner and provides a better estimate of experimental effects than FRPs.

2. Visual Search Task

2.1. Methods

2.1.1. Participants

Sixteen participants recruited at the Goethe University Frankfurt took part in the visual search experiment (9 female, median age 24, min 19, max 37). All were tested for normal or corrected to normal vision and had normal color vision as assessed by the Ishihara test. None reported any neurological disorders. All participants received course credit or financial compensation and had given informed consent according to protocols approved by the ethics committee of the department.

2.1.2. Apparatus

Eye movements were recorded with an EyeLink 1000 desktop mounted eye tracker (SR Research, Canada) at a sampling rate of 500 Hz. Viewing was binocular, but data were recorded from the left eye only. The experiment was run on a computer running Windows 7. Stimulus presentation was controlled by MATLAB (Version 8.1.0.604), making use of the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Subjects were seated in a dimly lit room with their heads fixated in a chinrest, in front of a 24-in. computer screen with a refresh rate of 144 Hz and a 1920 × 1080 resolution. Viewing distance was approximately 65 cm. To prevent EEG artifacts caused by pressure on the frontal electrodes, no forehead rest was used. Instead, frequent drift checks were included and participants were instructed to relax their neck muscles but move as little as possible. The eye tracker was used in remote mode (as opposed to "head fixed" mode), so that small head movements were better compensated and fewer recalibrations were

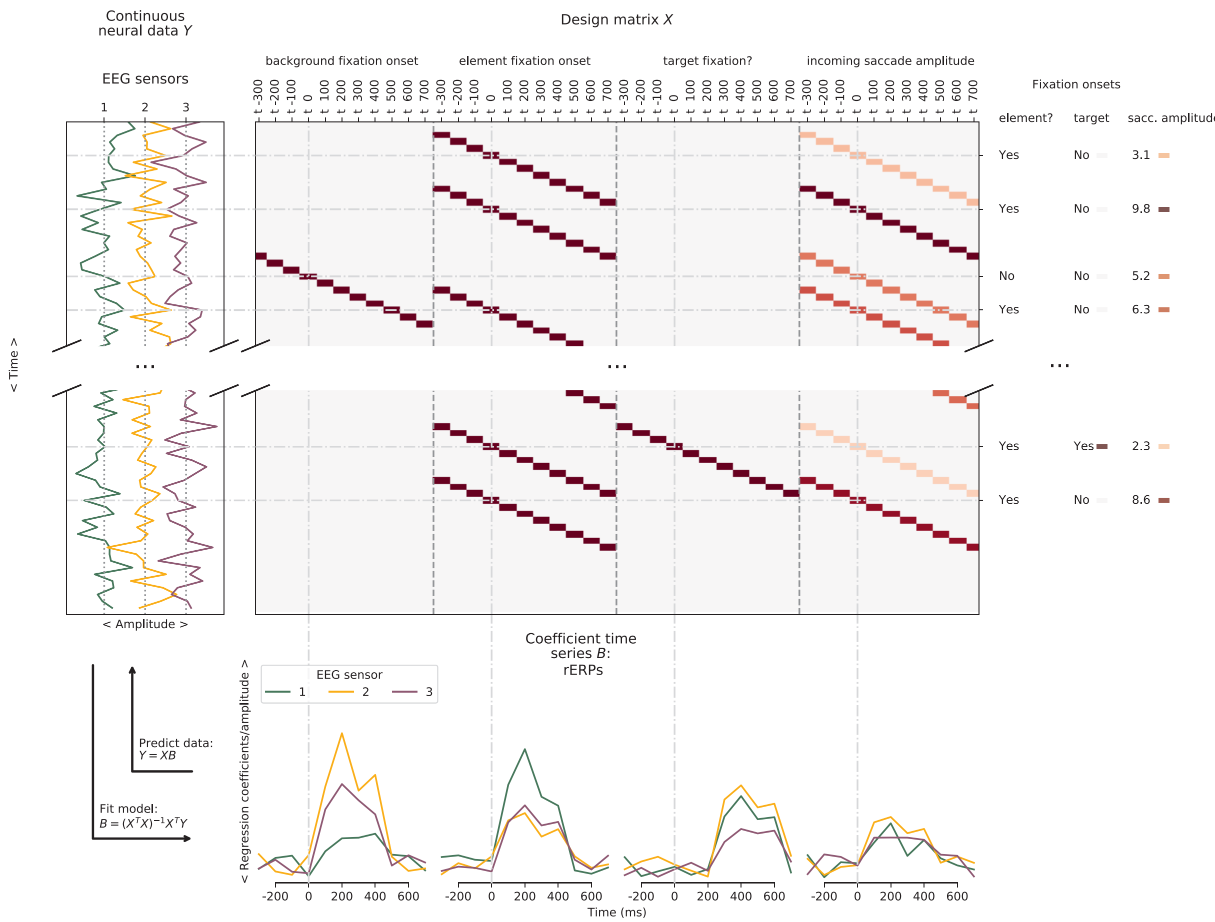

Fig. 1. Visualization of the linear model. The left matrix represents the continuous data, colored positions in the design matrix represent non-zero values, and the bottom four waveforms represent the rERPs that result from fitting the model. EEG responses can also be predicted, by multiplying the design matrix by the regression coefficients (the rERPs). In this case, the obtained time series represent the *predicted* continuous neural data and not the measured data.

required. This came at the cost of slightly higher variable error in the eye tracking signal. This did not visibly influence classification into eye movements event though, and since events rather than raw data served as predictors for rERPs, the added noise was assumed non-problematic. EEG signals were recorded using 64 active electrodes (actiChamp, Brain Products), positioned according to the 10–20 system. All EEG signals were referenced to the mastoids. Electrode positions SO2, LO1, LO2, and FP1 were used to measure EOG activity. All EEG signals were recorded at a sampling rate of 1000 Hz. To allow post-hoc temporal alignment of ET and EEG data, shared triggers were sent to the EEG amplifier and the recording computer of the eye tracker, making use of a parallel port and an Ethernet connection respectively. Synchronization triggers were sent at the onset and offset of every trial.

2.1.3. Stimuli and procedure

For 250 trials plus 5 additional practice trials, participants searched for two identical targets among a grid of 14 search elements. All search elements were crosses, presented on a grey background (RGB values (128,128,128)). Targets consisted of a yellow horizontal bar and a light-grey vertical bar, with the overlap between bars colored in black. Distractors consisted of the same parts, but with the orientation of the bars switched (grey horizontal, yellow vertical). Participants were instructed to search for two crosses with a yellow horizontal bar and press a button on the keyboard as soon as they had located the second target. Additionally, observers were instructed to do so as fast and as accurately as possible. Targets and distractors spanned an angular size of 0.4° by 0.4°, making identification without foveation difficult. All search arrays fit within a 14.5° by 14.5° central portion of the screen. Search element positions were determined by defining a 5 by 5-position hexagonal grid within the stimulus area. In each trial, a random jitter in both horizontal and vertical direction (max 0.8°, ¼ of the minimum distance between grid positions) was added to each position. Then, 14 positions were pseudo-randomly selected from all positions in the hexagonal grid, excluding the position closest to the center of the screen. Trials consisted of a central fixation cross, presented for 1200–1700 ms, followed by the onset of the search array while the fixation cross remained visible. The search array remained on screen until a button was pressed by the observer to indicate that a second target had been found. Upon button-press the search array remained visible for an additional 80–120 ms before disappearing, indicating the start of the next trial (Fig. 2). The experiment was preceded by a 13-point eye tracker calibration and validation procedure. Calibration was deemed successful when the average validation deviation was lower than 0.5° and none of the validation points yielded a larger deviation

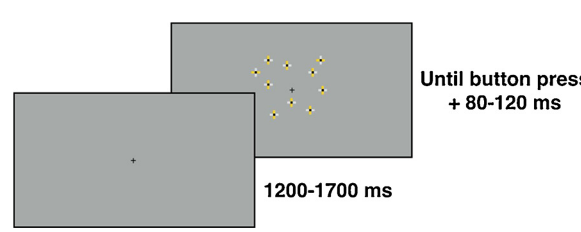

Fig. 2. Schematic depiction of a trial sequence. Elements are enlarged for display purposes and fewer elements are shown than were present during the experiment.

than 0.7°. Calibration and validation were repeated when drift-check was unsuccessful, or additionally at the discretion of the experimenter. An optional short break and a non-optional drift-check for the eye tracker occurred every 30 trials. After a successful drift check (a measured deviation between dot position and gaze of less than 1°), the fixation cross re-occurred, signaling the start of the next trial.

2.1.4. Data pre-processing

Saccades and fixations were extracted from raw gaze data during recording by the EyeLink parser. Velocity and acceleration thresholds were set to the EyeLink default values of 30 °/s and 8,000 °/s^2, respectively. Fixations were attributed to the element nearest to the fixation, with a maximum distance of 2 ° (cf. Hessels et al., 2016). We distinguished first fixations of the first target (whichever of the two targets was fixated first) and the first fixation of the second target. In case only one fixation on a target was made, it was classified as fixation of a second target because the data do not allow us to determine whether a misclassification occurred before or after the single target fixation. Moreover, mislabeling a fixation as *first* target fixation could lead to measuring a P300 FRP contaminated with activity due to button-press (preparation). Only fixations with a duration longer than 100 ms were included when defining target fixations.

EEG data were imported into EEGlab (Delorme and Makeig, 2004). Using the EYE-EEG toolbox (http://www2.hu-berlin.de/eyetracking-eeg), eye movement onsets were temporally aligned with the EEG data. The EEG was then band-pass filtered at 1–40 Hz and saved separately for ICA. Data were also band-pass filtered at 0.1–40 Hz for further analysis. After excluding breaks, overly noisy periods and overall noisy electrodes, ICA was performed on non-epoched, non-down-sampled data using AMICA (Delorme et al., 2012; Palmer et al., 2012). To select independent components (IC's) that reflected eye movement artifact, a variance ratio threshold was applied (Plöchl et al., 2012). For all components, the average variance in activation during saccade periods and during fixation periods was calculated. Any component of which the variance during a saccade was more than 1.1 times as large as during a fixation, was marked as eye movement related (1.1 is the cut-off recommended by Plöchl et al. (2012) and visual inspection of the selected IC's yielded no reason to adapt this). The weights resulting from ICA were then transferred to the less heavily high pass filtered data, followed by exclusion of the IC's marked as artefactual by the variance ratio threshold method. After filtering and artefact correction, further pre-processing and analyses were performed using MNE-Python (Gramfort et al., 2013), version 0.14.1. After importing data into MNE-Python, noisy electrodes were interpolated using the spherical spline interpolation method implemented in MNE python, and signals were down-sampled to 500 Hz. Down-sampling to 250 Hz or 100 Hz yielded almost identical results and made no difference for any of the conclusions drawn. Computations were, however, faster and more memory-efficient at 250 Hz and 100 Hz. At 100 Hz, calculation time was about an order of magnitude shorter than for data down-sampled to 500 Hz.

2.1.5. rERP model definition

The applied model was defined to contain intercept (binary) terms for stimulus onset events, stimulus offset events, button press events and fixation onsets. We distinguished between fixations on search elements and fixations on stimulus background (by including a different intercept term for both). Time-locked to each fixation onset were also the numerical predictors associated with EM parameters. Modeled EM parameters included horizontal fixation position, vertical fixation position, incoming saccade amplitude, the incoming saccade angle sine component, and incoming saccade angle cosine component. Sine and cosine components of the same angles were included as separate predictors to account for circularity. Fixation positions in both horizontal and vertical direction were scaled to range from -1 to 1, with the edges of their range corresponding to the edges of the display and with 0 at the center of the display. Sine and cosine are naturally scaled from -1 to 1. Saccade amplitudes were added in degrees and not centered or scaled, in order to obtain an easily interpretable slope which signifies additional EEG per added degree of incoming saccade amplitude. Statistics and distributions of the eye movement characteristics included as covariates can be found in the supplementary materials. A binary predictor for first fixations of a first search target was added, as were a predictor representing first fixations of a second search target, and an intercept term for button presses. For overlap correction, each predictor was added for latencies ranging from -400 to 1000 ms, pre- to post fixation onset.

2.1.6. rERP artifact rejection

For artifact rejection, continuous data were treated as one second periods. Any period containing voltage fluctuations larger than 100 μV was excluded from analysis. To exclude data, the corresponding time-points were removed from both the data and the predictor matrix before fitting the linear model. Blinks were handled in the same manner as other artifacts, i.e. by removing data and predictors surrounding the blink. Data from 300 ms before blink onset, until 300 ms after blink offset were removed. Artifact rejection, construction of the predictor matrix, and solving the least squares problem were handled by the function *linear_regression_raw()* included in MNE-Python.

2.1.7. Constructing grand average rFRP waveforms

The procedure fits the specified model to the continuous data from each electrode individually, for each participant individually. This results in an rERP for each electrode, per predictor, per participant, similar to what regular averaging would yield (an average per participant per electrode), except a model fit estimates the influence of multiple predictors at once. When averaging across fixations instead, only the influence of the binned "predictor" is estimated (often the difference between two experimental conditions). Just like ERPs or FRPs, rERPs are subsequently averaged across participants to obtain a grand average.

2.1.8. Quantifying model performance

Overall model performance was inspected by means of a five-fold cross validation, within each participant. EEG were split into 5 consecutive equal-length parts. For each of the 5 folds the model was fitted to the remaining 80% of the data. The score for the fold was then obtained by predicting the EEG in the fold from the fitted model coefficients (i.e., the rERPs), and correlating the prediction to the real data. Squaring the correlation coefficient while retaining its sign yields the amount of variance in the real data that can be explained by the model. A score per participant is obtained by taking the median across electrodes within each fold, and then averaging across folds. A score for the whole model is obtained by averaging participant scores.

2.1.9. Creating FRPs

To calculate FRPs, for each participant, 1300 ms epochs were cut around each fixation (−300 ms to 1000 ms). Only epochs from fixations that fell upon a search element were included. Fixations on the target that were shorter than 200 ms were dropped from analysis. Short fixations have been argued to be functionally different from longer ones (e.g. Kingstone and Klein, 1993; Kapoula and Robinson, 1986). Additionally, a higher minimum such as 500 ms would have been preferable to prevent averaging across overlapping responses. However, the cut-off was set to this specific duration because we aimed to measure naturalistic viewing behavior and therefore did nothing to manipulate subjects towards making long fixations. If a 500 ms inclusion criterion was applied, too few fixations were left for a valid comparison to rFRPs. Note that no fixation duration cut-off is needed or was set for rERPs. Any epochs containing blinks (as detected in the ET signal) or voltage fluctuations larger than 100 μV were dropped as artefactual. For EM parameter FRPs, epochs containing target fixations or button presses were excluded.

2.1.10. Binning eye movement parameters

The remaining epochs from each individual participant were then binned, matched and averaged within each participant, several times. The binning each time consisted of dividing the data into two extremes, corresponding to the predictors included in the linear model described above. Epochs associated with fixations falling right vs. left of the middle of the screen (to isolate horizontal viewing position related EEG differences), above vs. below the middle of the screen (vertical position), incoming saccades that contained rightward movement vs incoming saccades moving leftward (cosine), upward incoming saccades vs downward incoming saccades (sine), and finally large vs small incoming saccade amplitudes. Large versus small was determined by taking the 5th to 50th percentile (small) and 50th to 95th percentile (large) of all saccade amplitudes from the participant's included data. For all eye movement-related bins (all the ones mentioned thus far) epochs that overlapped with a target fixation or button press were excluded (see below), as were epochs containing blinks. Finally, fixations on distracters vs fixations on targets were binned.

2.1.11. Matching bins by eye movement parameters

For each pair of bins, an attempt was then made to minimize the difference between the epochs in the two bins, on all eye movement variables (the same ones as included in the linear model) except the variable of interest. To this end a method similar to that proposed by Nikolaev et al. (2016) was implemented. Data from both bins were merged and Mahalanobis Distance (MD) between each epoch and every other epoch was calculated. Then, an iterative procedure was followed. The first step consisted of removing epochs based on MD, with a percentile cut-off and a proximity cut-off (see Nikolaev et al., 2016). The next step was to apply a non-parametric permutation *t*-test on each variable and to repeat both steps in case of any significant differences. For more than half the participants in both tasks, different variables would yield significant differences as the cut-off was lowered, until (almost) no epochs were left for analysis. This brings into question the feasibility of matching epochs along a large number of variables. Ultimately, a one-time percentile cut-off of 95% was applied to attain some degree of matching. For the visual search dataset, more distractor fixation related epochs were obtained than target fixation related ones. For the target and distractor related epochs we therefore followed a procedure in which all target fixation related epochs were kept and used as a reference distribution. The subset of distractor fixations with the smallest MD to the reference distribution was included for analysis. The size of the subset was set to the number of included target fixations. The number of resulting epochs per bin can be found in the supplementary material (Table 1). For each pair of bins the same averaging procedure was then applied: per participant, both bins were averaged and subtracted, yielding a difference wave per participant. Lower predictor values were always subtracted from higher ones. Individual difference waves were averaged to obtain a grand average difference wave per predictor.

2.1.12. Excluding target fixations and button presses from EM related FRPs and rFRPs

For a comparison of FRPs and rFRPs in similar data, the linear model was fit twice: Once with target fixations and button presses excluded and once with these events included (to obtain a linear estimate of the P300). In the rERP approach these events are explicitly modeled and *can* be left in the data. However, in order to compare FRPs and rFRPs based on data as similar as possible, target fixations were also excluded from the data and the predictor matrix when estimating the EM confound rFRPs. Exclusion was accomplished the same way as artifact rejection: by excluding all events within -400 to 1000 ms pre- to post fixation onset. Therefore, in supplementary Table 1 the number of fixations for rFRPs *could* have been higher had target fixations and button presses not been excluded. Conversely, although roughly 200 first target fixations were included in the rFRP estimate of the P300, it was derived from a model that had a much higher number of non-target fixations included, all off which could aid the robust estimation of EM confounds.

2.1.13. Significance testing

Where needed, statistical significance of waveforms was judged with a spatio-temporal cluster-based permutation *t*-test of the target-distractor difference waves against 0. The test corrects for multiple comparisons of different electrodes at different time points by only retaining significant spatio-temporal clusters of the *t*-statistic. In short, the obtained data are used to bootstrap a distribution of clusters of the *t*-statistic under the null hypothesis. The null hypothesis is simulated by assigning a random sign (- or +) to all samples in each bootstrap iteration (2000 iterations). From each permutation iteration, peak cluster values are stored. The p-value for the spatio-temporal clusters of t-statistics obtained from the original data is then determined by the proportion of clusters in the simulated distribution that is larger than the obtained cluster (Maris and Oostenveld, 2007).

2.2. rFRP-FRP Comparison results and preliminary discussion

First, we will compare the predictors for EM confounds to their FRP comparisons. Then, we highlight the issue of pre-fixation activity and baselining. Subsequently, we will compare the rERP for the predictor of experimental interest (fixating a target object) to a traditional, averaged estimate of the same effect.

2.2.1. The goal of comparing rFRPs for eye movement characteristics to FRP difference waves

In the absence of some form of ground truth for comparison, it is impossible to calculate if an rFRP or an FRP is "more correct", even if the nature of regression makes that the rFRPs are less confounded by the shortcomings of FRPs. Therefore, this first comparison does not primarily serve to determine what method is better. Theory tells us the rFRP is a more suitable estimate. The goal is to ascertain that the estimated rFRPs are *sensible* estimates of EM confounds. Establishing that rFRPs estimate confounds well in a typical task like visual search, is a key step in showing the potential of the method for ET-EEG research.

2.2.2. Interpreting rFRPs and FRP difference waves

rFPRs can be interpreted of as follows: There is EEG activity that is common to each fixation, which is expressed in the fixation intercept term. The rFRPs for other fixation related predictors add to this intercept linearly. The estimated intercept term for fixations is shown in Fig. 3.

All obtained rFRP waveforms except for the fixation intercept term and the one for target fixation will represent a slope: the addition to the fixation intercept with each unit increase of the predictor. As an example: Horizontal and vertical position were centered on the middle of the screen with values scaled from -1 (bottom or left of the screen) to +1 (top or right of the screen). To interpret this rERP, one can imagine a fixation *i* landing half-way between the center and the left border of the screen. The predictor for horizontal position here takes value -0.5. Ignoring other predictors, the expected EEG in relation to fixation would comprise the fixation intercept, plus -0.5 times the rFRP for horizontal position, as expressed in the equation below. The same logic applies to vertical position, sine, cosine and saccade amplitude predictors.

$$E\hat{E}G_{fixation_i} = 1 * rFRP_{fixation} - 0.5 * rFRP_{horizontalPosition}$$

FRP difference waves, on the other hand, represent the difference in EEG related to whatever difference in the binned predictor was present between the two bins it stems from. By this logic, the difference waves should be scaled versions of the rFRPs. However, this scaling should only serve as a rule of thumb, as this will only hold if the FRPs are

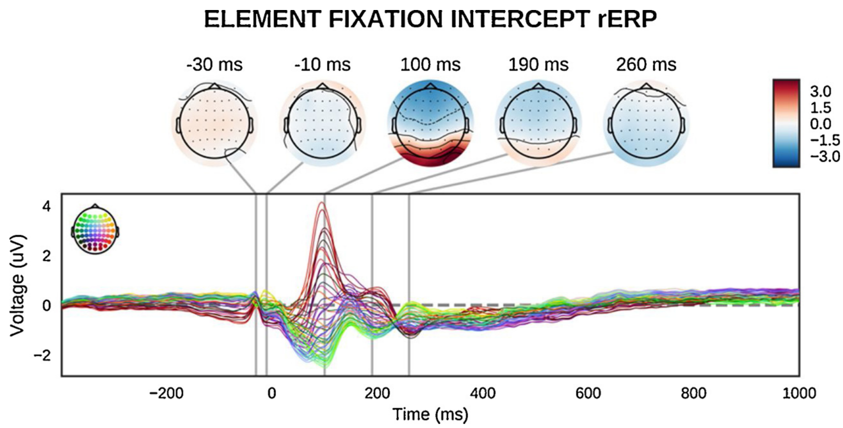

Fig. 3. Grand average intercept rFRP for fixations on search elements.

perfect isolation of one predictor, without overlapping responses (which they are not).

2.2.3. Comparison of FRP difference waves to rFRPs

If the rERP model is accounting for variance due to eye movements in a sensible manner, the rFRP and FRP difference wave for one predictor should resemble one another. More specifically, we expect the rFRP associated with a predictor to resemble the FRP difference wave associated with that same predictor, but to be less similar to the difference waves constructed for other predictors. Absolute voltages should be left out of the comparison, because all presented rFRP waveforms (except the one for target fixation) represent a slope. Fig. 4 depicts the grand average difference wave and grand average rFRP per predictor. Indeed, there is a striking similarity between each rFRP and the associated difference wave. Also, none of the rFRPs bear strong resemblance to FRP difference waves that depict other predictors. To quantify this claim, correlations between both time series were calculated.

For each participant, correlations were obtained by concatenating the data from all electrodes. This was done for both the FRP difference waves and rFRPs. Then, a Pearson correlation coefficient between both series of data-points was calculated, yielding a single correlation coefficient per participant, for each rFRP and its corresponding difference wave. Absolute correlations are less informative here than in the typical case, because we predict that rFRPs and FRPs are similar, but also that rFRPs suffer from fewer confounds and are thus not identical to FRPs. We therefore inspect how correlation coefficients compare to one another. To this end, additional correlations were computed, between the rFRP of each predictor and the difference wave constructed for every other predictor. To meet our criterion, each rERP should have a higher correlation to the difference wave of its own predictor, than to the difference wave for any of the other predictors. Fig. 5 clearly shows this pattern. This, together with Fig. 4, indicates that the linear model yields sensible estimates of the influence of eye movement parameters on the EEG.

2.2.4. Pre-fixation activity

None of the waveforms in Fig. 4 have been baselined. This was done because pre-fixation activity obstructs choosing a neutral baseline period. Baselining through activity could induce spurious voltage changes in later time windows. Lack of a neutral baseline has been recognized as a problem for FRPs (e.g. Nikolaev et al., 2016). Additionally, without overlap correction one cannot tell whether pre-fixation activity is preparatory activity, or an artefact caused by averaging over activity related to previous fixations. In contrast to FRPs, rFRPs are overlap corrected and show notably less pre-fixation activity. This makes it easy to choose a baseline as the researcher sees fit and allows the interpretation of pre-fixation differences between conditions of an experiment.

As expected, both FRP difference wave and rFRP for saccade amplitude (Fig. 4 - 3rd row) resemble an increase in lambda activity as a function of saccade amplitude. The lambda response is characterized by a positive peak in occipital activity, around 90 ms post fixation onset. We briefly highlight this finding because it fits with existing literature (e.g. Thickbroom et al., 1991). Furthermore, as an example of the argument about baselining, the rFRP contains very little activity before fixation onset (t0), whereas the difference wave shows clear signs of contamination by what are probably visual responses related to the previous fixation, as also observed by Dimigen et al. (2011). The likely reason this is visible in a difference wave for saccade amplitude is that saccade amplitude and fixation duration are at least weakly correlated (Nuthmann, 2017). The absence of pre-t0 activation in the rFRP indicates that overlap was successfully corrected, whereas it is clearly present in the FRP difference wave.

2.2.5. P300 effects

In the example search task, the waveform of experimental interest is the one that contrasts target fixations to distracter fixations, expected to yield a P300 effect (Brouwer et al., 2014; Devillez et al., 2015; Kamienkowski et al., 2012; Körner et al., 2014). The P300 to target detection is characterized by a positive signal deflection over occipito-central electrodes (Polich, 2007). A P300 effect of first target fixation is visible in both the rFRP and the target - distracter difference wave (Fig. 6). The positive deflection in the FRP difference wave reaches peak activation earlier. However, with a 200 ms minimum fixation duration cut-off, overlap confounds cannot be excluded. The earlier peak appears in occipital regions, which are also influenced by activity from previous or subsequent fixations. Fixations on the first target lasted 315 ms on average (standard error across participants 18 ms). Additionally, there is a negative deflection in the difference wave shortly after target fixation onset, which is absent from the rFRP. Previous literature provides no reason to expect differences in this early time-window or with this topography.

A cluster-based permutation *t*-test yielded a significant P300 effect in both FRPs and rFRPs, with similar time-courses. The early negativity in the FRP difference wave, however, was not a significant deviation from 0. Baselining the waveforms by subtracting the average activity in the -50 to 150 ms time window did not change this result.

A signal estimate can contain both systematic errors, sometimes called biases, and variable errors, often expressed as variance. Inspection of cross-subject variance reveals that variance is generally higher for the FRPs. As an example, Fig. 7 depicts the P300 effect at a single electrode along with its 95% percentile-bootstrap confidence interval. Before computing the confidence interval each participant's difference-wave was baselined by subtracting the average activity in the -50 to 150 ms time window (cf. Dimigen et al., 2011). A baseline was

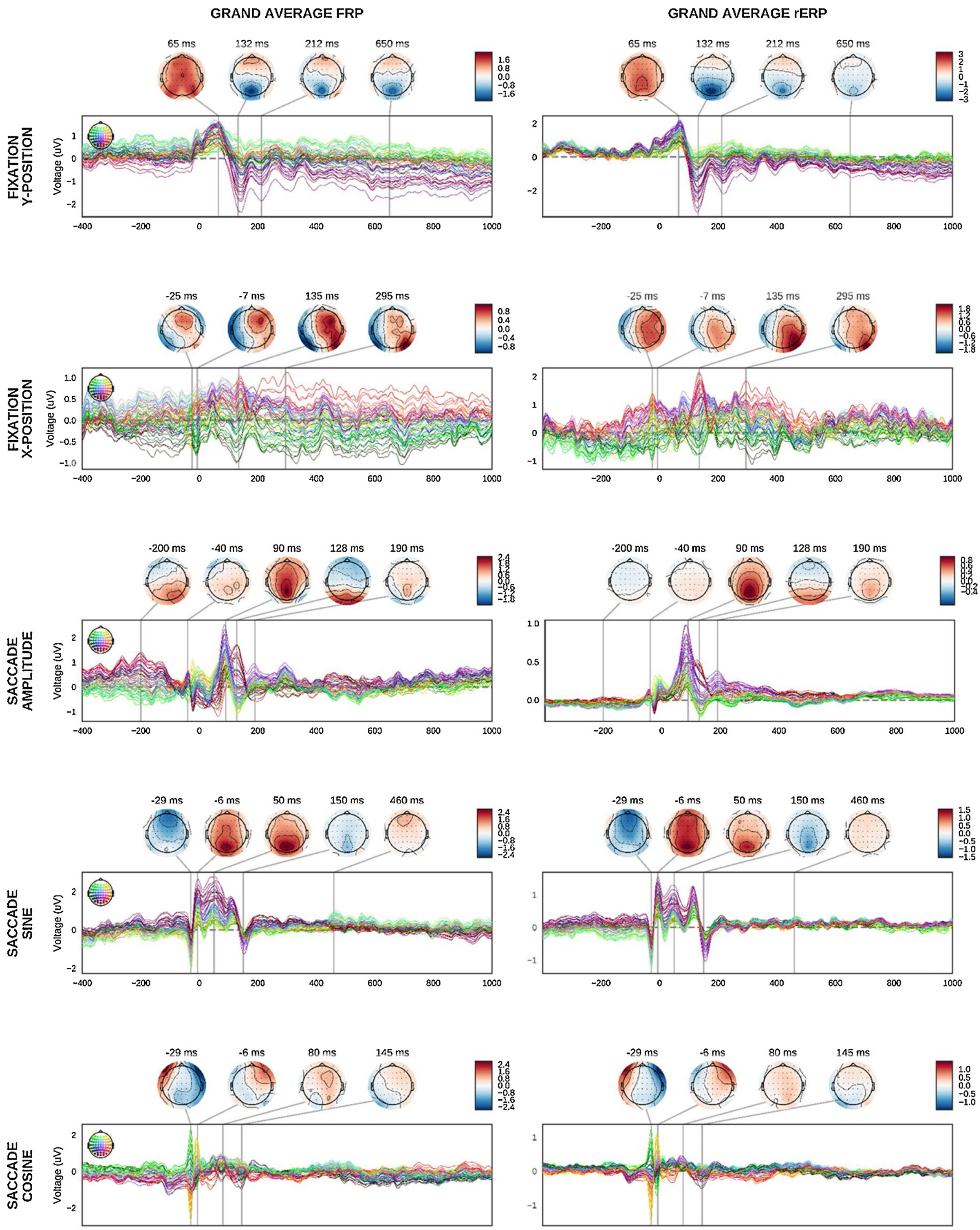

Fig. 4. Grand average rFRP waveforms (right column) and grand average FRP difference waveforms (left column) per predictor.

applied here to prevent the variance in P300 estimation across participants from being obscured by offset differences across participants. Variance is clearly larger in the FRP-based waveforms than in rFRPs. Fig. 7 also shows more clearly the unexpected negativity over occipital electrodes in the difference wave. This shows that the FRP comparison between targets and non-targets has higher variable error than the rFRP, and that the FRP-based waveform also contains a more systematic error that is absent from the rFRP.

2.2.6. A qualitative example of overlap correction: button presses

In the dual-target search task, detection of the second target was followed by a button press. The averaged FRP difference wave associated with the second target fixation (versus a matched set of non-target fixations) is depicted in Fig. 8 (top panel). The observed P300 amplitude is much larger than for the first target fixation (middle panel). Additionally, the difference wave deflects in negative direction for almost all electrodes towards the end of the time-window, rather

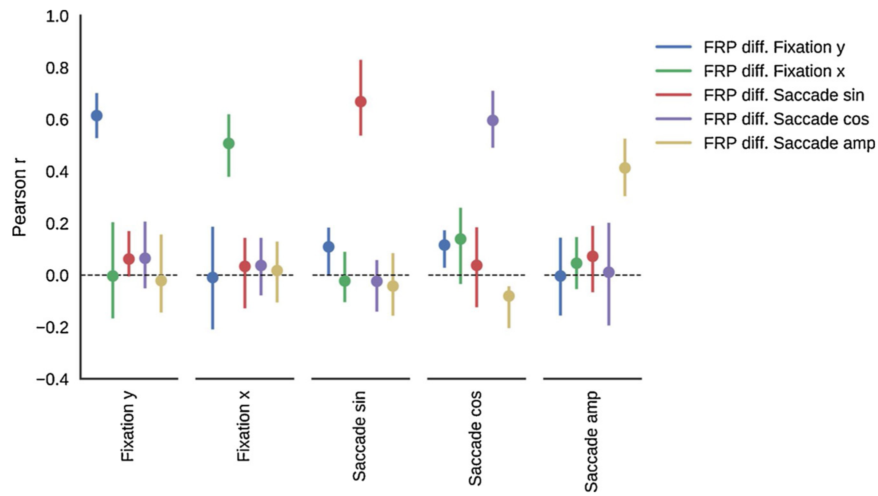

Fig. 5. Mean correlation coefficients across participants, between each rFRP (horizontal axis) and the FRP difference wave for each predictor. Error bars indicate 99% percentile bootstrap confidence intervals (2000 resampling iterations).

than returning to baseline. Both the high amplitude and the continued activity beyond the P300 time window are likely caused by contamination of the FRP with neural activity related to (the preparation of) the button press, which happened an average 751 after target fixation (standard error across participants 52 ms). These findings are comparable to those by Hale et al. (2008). Assuming the neural activity related to a button press is the same each press, and that button presses happen at sufficiently varying latency from fixation onset, regression should correct for the overlap between the button press related response and the response that is evoked by the target fixation. Fig. 8 shows the overlap corrected rFRP of second target fixation (bottom panel). The rFRP is more similar to the difference wave observed in relation to *first* target fixation (middle panel) and returns to zero around 800 ms post-fixation. This highlights an additional benefit of rERPs. In the current task, for example, participants searched for *two* targets. The second target was only included so that manual responses and fixations to at least one target happened sufficiently far apart in time. Although needed here for averaging, overlap correction frees the researcher from these kinds of design considerations.

2.2.7. *Overall model performance*

Overall, the model including target fixations explained 8.7 percent of the variance in the real data. Additionally, Fig. 9 shows the amount of explained variance across the scalp. The fitted model explains a substantial fraction of variance over occipital electrodes, i.e., assuming the rERP model allows the accurate prediction of brain activity, particularly in occipital regions, in the continuous EEG data. This finding fits existing literature, in a sense that occipital regions are also where most (fixation locked) variance in neural activity is expected to occur. Explained variance is indicative of overall model performance. Note, however, that in data with a low SNR, like EEG, successfully estimated but small systematic confounds will hardly influence the amount of explained variance. The same confounds can nevertheless distort fixation locked waveforms. The decision to in- or exclude predictors should therefore not be based on overall explained variance alone (see "Predictor inclusion" under Section 5).

3. Scene memorization task

To test the rERP method on less controlled stimuli than the abstract search arrays, and to inspect its applicability across tasks, we also gathered data from participants viewing photographs of indoor scenes. Participants were instructed to memorize photographs of indoor scenes in preparation for a subsequent memory test. Below, details of the scene memorization experiment are described where they differ from the search experiment. A subset of presented scene photographs included experimentally manipulated objects. Fixations on these objects were excluded before rFRP-FRP comparisons, in the same way target fixations were excluded from the visual search data.

3.1. *Methods*

3.1.1. *Participants*

Twenty-two different participants took part in the scene memorization experiment: 14 female, average age 26 (18–47). As before, participants were screened for exclusion criteria and gave informed consent in accordance with the department's ethics committee.

3.1.2. *Apparatus*

Data were collected and processed with the same equipment, setup geometry, and settings as the visual search task. However, the eye tracker was used in "head-fixed" mode and eye movements were recorded at 1000 Hz. Frequent drift-checks were implemented to ensure head-movement would not cause large systematic inaccuracies in the

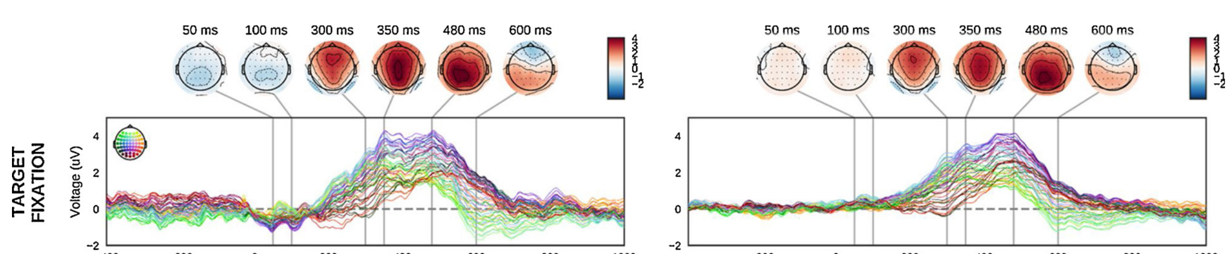

Fig. 6. Grand average waveforms for the effect of fixating a target element. Left: averaged FRP difference wave. Right rERP estimate for a model predictor representing first target fixation.

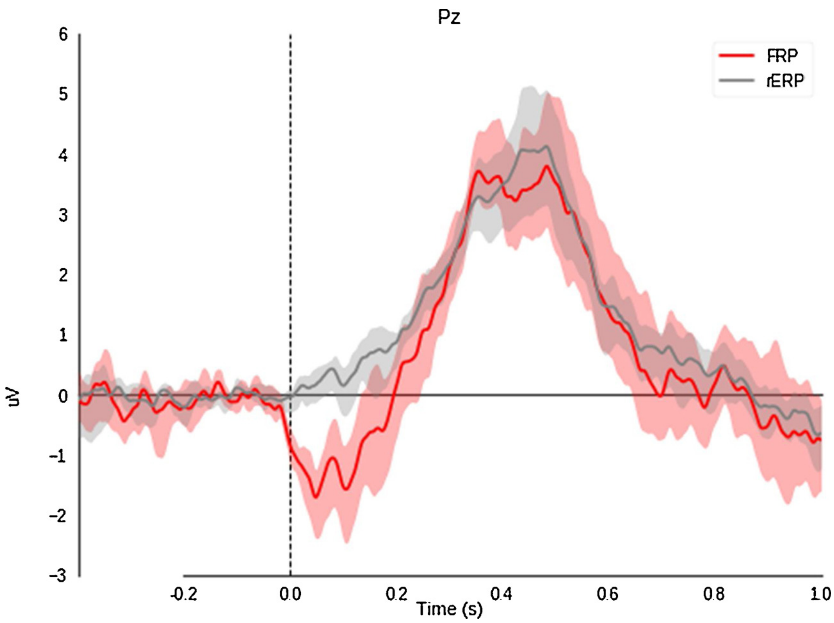

Fig. 7. Grand average FRP difference wave (target-distractor, in red) and rFRP (grey) for target fixation, at electrode Pz. Shaded areas indicate 95% percentile bootstrapped confidence intervals (2000 resampling iterations). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

eye tracker signal.

3.1.3. Stimuli and procedure

Observers viewed real-world scenes for 7 s each in preparation for a memory test. To ensure maintained attention, observers were additionally instructed to monitor for exact repeats of scenes presented before. An additional 20 scenes were presented twice (trials excluded from analysis). The experiment was preceded by a 9-point calibration procedure. Each trial began with a black central fixation dot and a participant-triggered drift-check for the eye tracker (Fig. 10). After a successful drift check (a measured deviation between dot position and gaze of less than 1°) the fixation dot immediately turned green, but remained visible for 800–1100 ms more. Subsequently, a scene was presented and participants moved their eyes around freely to explore it. After 7 s, the scene was replaced by the fixation dot again, or by a response screen asking whether the scene had been an *exact* repeat. Participants were to respond by keypress; then, a red or green square indicated for 1 s whether the given response was correct, before continuing to the next trial.

3.1.4. Artifact rejection

The same procedure of removing events and artefactual data was applied to the scene viewing data. The number of fixations included after exclusion of artefacts is generally higher than for the search task,

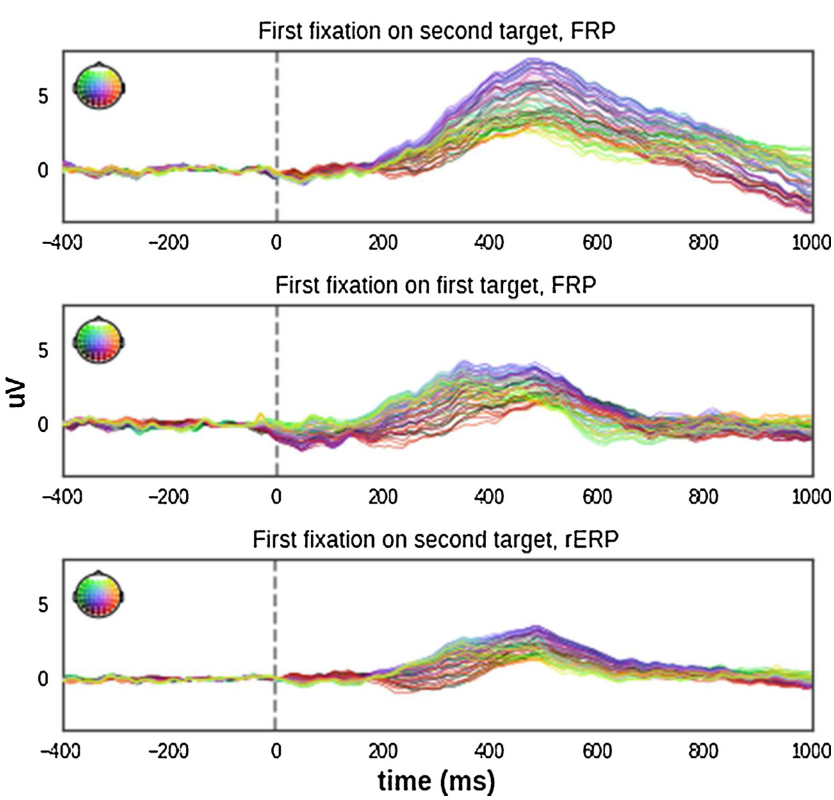

Fig. 8. Top: grand average matched FRP difference wave for second target fixation, likely confounded with overlapping button press-related activity. Middle: grand average matched FRP difference wave for *first* target fixation (Same as in Fig. 4). Bottom: grand average rFRP for second target fixation, corrected for overlapping button-press related activity. For better comparability, waveforms were baselined by subtracting the average activity between -150 and −50 ms before fixation onset.

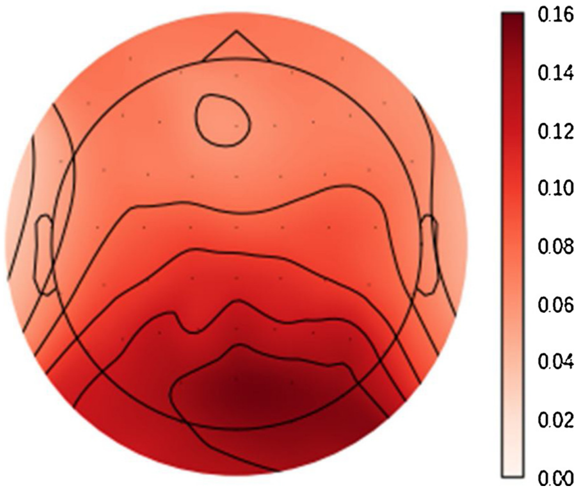

Fig. 9. Average proportion of explained variance as obtained from a cross-validation procedure, per electrode.

and specified in supplementary table 2.

3.1.5. Model definition

All the same predictors were modeled as for the scene viewing experiment, with the exception of target fixations and button presses. Only fixations occurring during scene presentation were included for model fitting.

3.2. rFRP-FRP Comparison results and preliminary discussion

Overall, results are similar to the visual search task: Correlations between matched FRP difference waves and corresponding rFRPs are higher than between the rFRPs and the FRP difference waves for other predictors (Fig. 11). Also similar to the visual search experiment, rFRPs visually resemble matched FRP difference waves, but not one another or any other difference waves (Fig. 12). Contrary to the search dataset, the rFRP and the FRP difference wave for saccade amplitude are strikingly similar and both resemble the rFRP obtained for saccade amplitude in the visual search dataset. Possibly, a larger number of included fixations yielded FRPs with a higher SNR. Fitting a linear model also reveals a more defined influence of horizontal fixation position than in the visual search dataset. The difference could be explained by a higher number of included fixations, leading to a higher SNR, in this case for both the rFRPs and the FRPs. Both FRP difference wave and rFRP for the predictor cosine show lateralized influences not visible in the search task results. We argue that these lateralized occipital influences are due to differences in stimuli and will return to them shortly.

3.2.1. Stimulus properties

There is one noticeable difference between EM confound estimates from the two experiments that seems unlikely to be inherent to differences in subjects and eye movement characteristics. The *scene viewing* rFRP for "saccade cosine" contains a distinct lateralized influence of horizontal saccade direction on occipital EEG activity between 0 and 300 ms after fixation onset, which is not visible in the search experiment "saccade cosine" rFRP (Fig. 13, top and bottom panel respectively). A similar effect can be seen for horizontal fixation position (Fig. 12, 2nd row), but due to the noisy estimate for that predictor in the search task, no comparison across tasks can be made.

To our knowledge, lateralized occipital effects of saccade direction have not been shown before. Moreover, they only occur here in data from more complex stimuli and show an occipital topography. We therefore argue that the effects have characteristics of activity related to visual input, rather than eye rotation only. Indeed, there were large differences in visual input between both tasks. In the scene memorization task, the naturalistic scenes formed one central area of varying visual content on a monochrome background (see also Fig. 10), a situation carefully avoided in the search task. We speculate these lateralized influences are related to where in the visual field the majority of the scene falls as the eyes re-position horizontally. A similar lateralized influence can be seen around 100 ms in the rERP for horizontal fixation *position*, which is directly correlated to the position of the scene in the visual field. This shows that stimulus properties can be estimated from ET-EEG data and that such properties possibly "distort" rFRPs of correlated predictors when not explicitly modelled. Although some EEG variance due to stimulus properties is thereby accounted for, the model could probably control for confounds better if it were extended to include stimulus properties. *How* to model relationships between stimuli and neural responses should be the subject of carefully controlled experiments.

3.2.2. Overall model performance

The cross-validation procedure yielded an average 6.9% of explained variance across participants. Again, the largest proportion of variance was explained for electrodes over occipital regions (Fig. 14). This percentage is somewhat lower than the 8.7% explained variance in the search data. One explanation for this is that stimulus related variance remained partially unexplained.

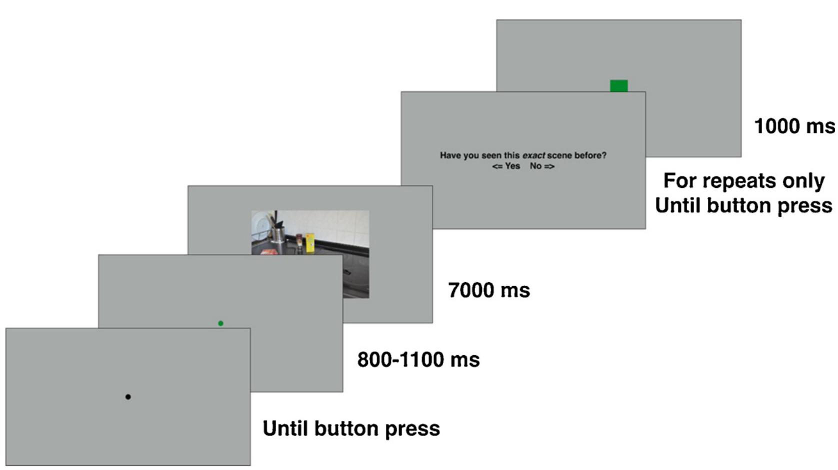

Fig. 10. Schematic depiction of a trial sequence in the scene memorization task.

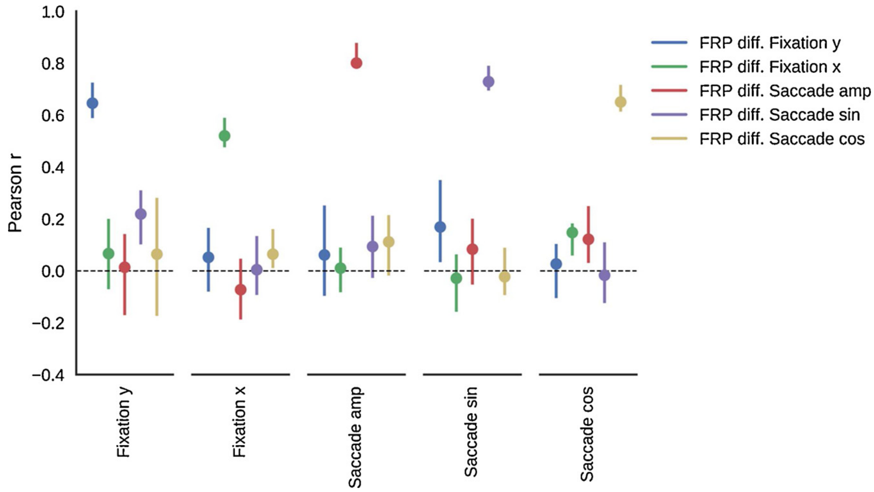

Fig. 11. Mean correlation coefficients across participants, between each rFRP (horizontal axis) and the FRP for each predictor. Error bars indicate 99% percentile bootstrap confidence intervals (2000 resampling iterations).

4. General considerations in rERP model construction

4.1. Predictor inclusion

So far, comparing averaged FRPs to rFRPs has yielded several observations and conclusions.

Firstly, we have established in two different tasks and types of stimuli that a linear regression approach to ERP estimation yields sensible rFRPs when applied to ET-EEG data. rFRP estimation can disentangle influences of different eye movement related confounds from continuous EEG. Also, the effect of the experimental manipulation in the search task could be established with higher accuracy and precision using rERPs. Overlap correction provides several benefits of rFRPs over FRPs. Researchers have more freedom in choosing experimental designs (e.g. ones that require manual responses), and baselining can be done without risking distortions. Given these findings, more intricate models and considerations can be explored.

Above, we specified a model that both seems reasonable and allows comparison of its output to FRPs. To this end, the model has been a simplification. We do not claim it should form a standard. A definitive list of predictors to be included is beyond the scope of this paper, but recommendations and considerations can be discussed.

Firstly, any predictor included should be included based on the reasonable assumption that it has an influence on the EEG. Adding predictors only to exclude *any* kind of confound will increase the risk of defining a model that closely explains the sampled data, but hardly generalizes. Moreover, added predictors will inevitably be attributed *some* variance. Adding a large number of variables without an indication for doing so therefore increases chances that some variation due to an experimental manipulation is attributed to a needlessly included predictor, subsequently increasing the risk of a false negative. Conversely, predictor exclusion requires caution. Any unexplained systematic variance in EEG could be attributed to another predictor that correlates with the one excluded, increasing the risk of a false positive (an example of an excluded predictor is given under “stimulus properties”). Therefore, it might be preferable to include slightly more predictors known or assumed to have an influence rather than leaving some out in fear of overfitting.

4.2. Collinearity

(Multi-)Collinearity means (substantial) correlation between predictors; when predictors are not perfectly correlated it is also known as partial collinearity. For an extensive discussion of collinearity and rERPs, see Smith (2011) and Sassenhagen (2018). Here we focus on ET-EEG rFRPs. Collinearity results in unstable estimates or high β values attributed to one predictor and low values to a collinear predictor. In extreme cases this makes rERP estimates for collinear predictors uninterpretable. In free-viewing, most EM parameters are presumably partially collinear. It is, for instance, reasonable to assume upward saccades will tend to land higher on the screen (see the supplementary materials for predictor correlation matrices for both experiments). Still, fixation position and saccade direction are assumed to have distinct influences on EEG activity and are thus modelled separately. This is unlikely to be problematic for the data presented here for several reasons: Firstly, the effects of collinearity on rERPs can be reduced or overcome by including more data. Eye movements occur at a relatively high frequency and provide many data points (i.e. fixations) in a short recording. Still, in continuous-time regression there is no straightforward way to calculate exactly *how much* extra data is needed to counter the effects of collinearity (Smith and Kutas, 2015b). Secondly, collinearity *only* influences the rERPs of collinear predictors. Collinearity does *not* affect the rERPs of uncorrelated predictors in the same model. Thus, if collinear predictors are included only as covariates, then nothing is lost to collinearity. When not *perfectly* correlated, each collinear predictor should be modeled because each will account for some unique and potentially confounding variance. As for the influence of collinearity on the rFRP for a predictor of interest (e.g. “target fixation”) the predictor of interest should be only weakly collinear with any other modelled predictors. Note that the requirement for the independent variable to not correlate strongly with the experimental manipulation holds for averaging approaches, too. Averaged waveforms, however, will become distorted by such correlations in a possibly more misleading manner, namely by displaying the influence of the confound as an experimental effect. Note that relationship between EM characteristics and neural activity is presented in the current work as a confound to the relationship of interest. When *studying* the influence of different eye movement characteristics on EEG (Dandekar et al., 2012b; Nikolaev et al., 2016), collinearity should be considered with an increased amount of caution (and free-viewing might not be the ideal paradigm).

4.3. Collinearity and overlap correction

Regression based overlap correction has been shown to be effective for fixation-locked neural responses (Kristensen et al., 2017a, 2017b). Theoretically, overlap correction should allow separate inspection of saccade *and* fixation locked components. However, the way overlap is

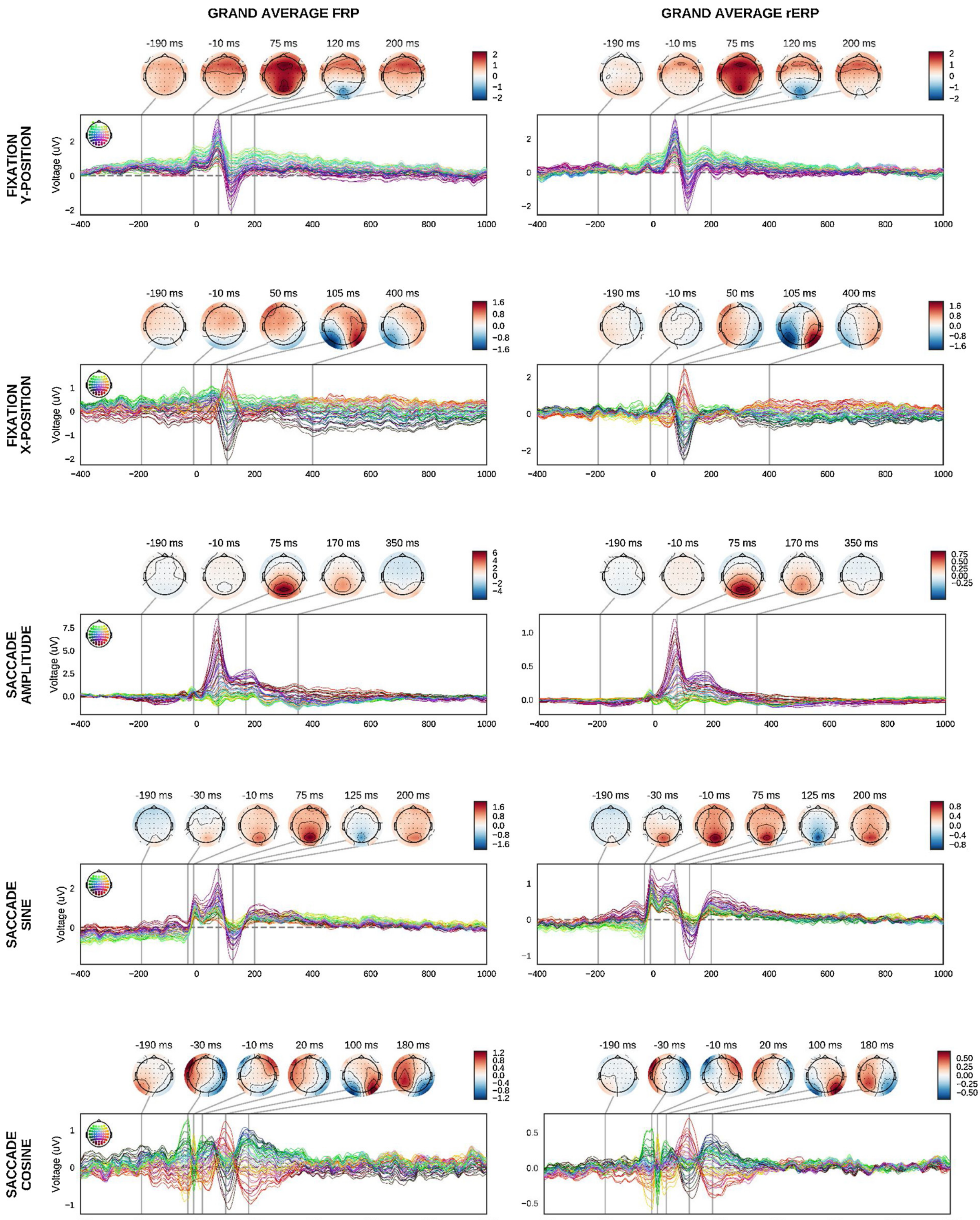

Fig. 12. Grand average rFRP (right column) and FRP (left column) waveforms per predictor.

modeled within the rERP framework incurs a potential limitation. By temporally expanding predictors, overlap correction becomes a matter of disentangling partially collinear predictors (Smith and Kutas, 2015a, 2015b). Such collinearity imposes the same problems and restrictions as described above and is not problematic if it does not concern a predictor of interest. Modelling separate intercepts for saccade- and fixation onsets is an example of modelling two temporally close predictors with zero variation in value (both events get value 1 in the predictor matrix) and relatively little variation in their relative position in time. As expected, the resulting rERPs show clear distortions (Fig. 15). It is impossible to interpret these rERPs in a meaningful way, showing that temporal collinearity *can* hamper overlap correction in ET-EEG data. There are two remedies for this kind of collinearity: As before, including more data reduces distortions. Additionally, collinearity will be reduced by variation in the amount of time between events and by variation in the value of the predictor (Smith and Kutas, 2015b). Maximizing the

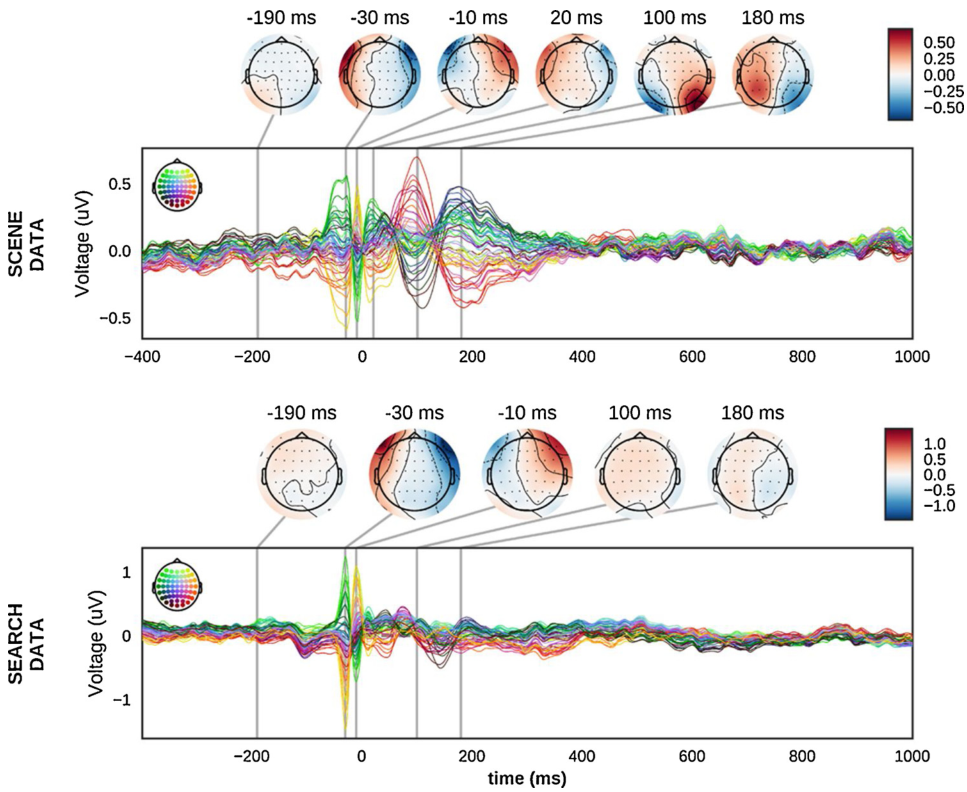

Fig. 13. Grand Average rFRP estimates, visualizing influence of horizontal saccade direction on fixation related EEG. Top panel: scene viewing data. Bottom panel: the same rFRP, derived from the visual search data.

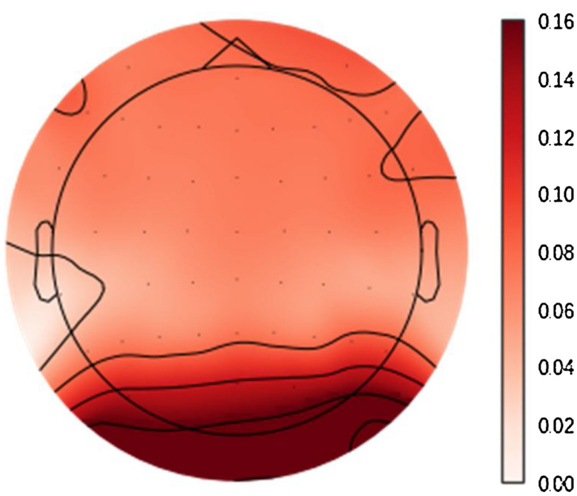

Fig. 14. Average proportion of explained variance per electrode as obtained by a cross-validation procedure.

amount of data one gathers and maximizing predictor variance should therefore be part of any ET-EEG design that one intends to analyze using the method proposed here.

5. Discussion and possible model improvements

We have already highlighted that for experiments involving visually complex stimuli, stimulus properties should probably be modelled along with EM confounds. Below we describe a number of additional considerations.

5.1. Multiple classes versus additional modifying predictors

A key assumption to the application of regression so far is that each fixation evokes the same neural response and that this intercept is affected as the EM related predictors vary in value. Is this assumption always warranted? Short fixations provide an interesting example.

Saccades sometimes fall short of their goal, to be quickly followed by a corrective saccade (Kingstone and Klein, 1993; Kapoula and Robinson, 1986). This invites the question whether short fixations serve the same purpose as longer ones and whether they evoke the same neural responses. Similarly, in the eye tracking literature there is a tendency to exclude fixations with a shorter duration than a certain cut-off from outcome measures (like total gaze duration). Sometimes that cut-off can be as high as 100 ms (Holmqvist et al., 2011). For rERPs, we have implied no minimum fixation duration. Still, we assumed that short fixations do not lead to target recognition (only fixations longer than 100 ms were labeled as fixation on a target element). We do not aim to settle any debate on the functional role of short fixations within this paper. But given sufficient reason to assume "short" fixations functionally differ from longer ones, they should be modeled as an additional class (intercept) rather than modeling *all* fixations with an equal intercept.[2] In a similar manner, we assume stimulus onsets evoke

[2] When neurophysiological data are time-locked to eye movements, classification of different EMs in the ET data becomes even more important than in gaze duration studies, especially for comparability between studies. The built-in eyelink detection algorithm used here only classifies saccades, fixations, and blinks. Other algorithms also classify post-saccadic movements of the pupil (Nyström and Holmqvist, 2010; Nyström et al., 2013), which have been observed to last up to 30 ms and vary with saccade characteristics and with accommodation (Hooge et al., 2015; Nyström et al., 2015). Regardless whether these kinds of eye movements are considered perceptually relevant or eye tracker induced artefact, the employed classification algorithm clearly has a direct influence on the timing of saccade- and fixation onsets. ET-EEG

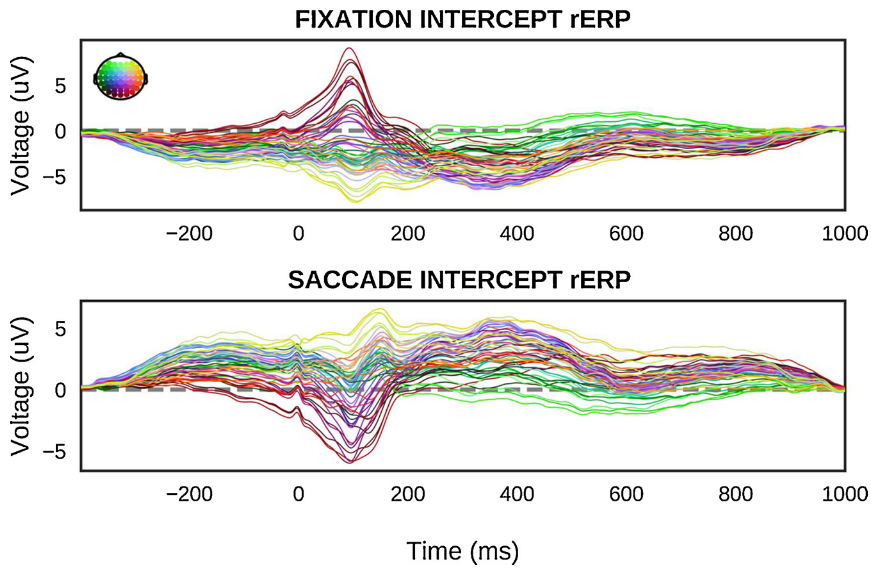

Fig. 15. Top panel: Grand Average rERP for the intercept term of fixation onset. Both fixation locked and saccade locked predictors are included in the model used for estimation. Bottom panel: Grand Average rERP for the intercept term of saccade onset, as estimated by the same model.

the same neural response for every onset. Perhaps for complex scenes this is an over-simplification that ignores large differences in stimulus properties. These are assumptions that deserve close consideration during model construction.

5.2. The assumption of linearity

Finally, one might question the linearity of the relationships between predictors and the EEG. It is impossible that any continuous predictor-EEG relationship is strictly linear. Still, we may assume the relationship approaches linearity closely enough within a certain range of predictor values. Assuming such, and refraining from predictions about predictor values outside the ones measured, non-linearity is not a problem for applying linear regression. For instance, the relationship between spike potential amplitude and saccade size is linear up to roughly 10° (Boylan and Doig, 1989). Within the current experiment stimuli spanned a relatively small area of the screen and saccades > 10° were rare. As such, the assumption of linearity is a very valid one. Some relationships, however, may be non-linear even within the measured predictor range, requiring a model with non-linear relationships to obtain more realistic rERP estimates. Allowing non-linearity does not mean no definition of the shape of the relationship is necessary. Theoretically one could define a model with very lenient assumptions about the shape of any predictor-EEG relationship, and let the shape be determined by the data in the fitting process. However, such flexibility requires large amounts of data to prevent over-fitting; otherwise, overfitting would result in low within-sample error, but high out of sample error (e.g. a very low cross-validated R^2). In general, the increased complexity introduced by fitting any kind of nonlinearity requires an informed choice of parameters. What relations to model as non-linear and what assumptions to make about their shape should therefore, in our opinion, be informed by future research rather than by making nonlinear fits the standard. Much in the same way as collinearity, *mild* overfitting and difficult to interpret rERPs might be less problematic if the nonlinearity is modeled for predictors that serve as controls only. The exploration of ways to model nonlinear relations between eye movements and EEG might be a fruitful and necessary next step in the application of rERPs to ET-EEG data.

A method that could deal with non-linearity and the risk of overfitting in its own way is the use of Generalized Additive Mixed-effects Models/GAMMs (e.g. Wood, 2017). However, one challenge for nonlinear model fitting is computational load. At the time of writing, any model beyond the least-squares fit involves a great drop in computational efficiency. This is exacerbated by the fact that the same model will be estimated for multiple channels. While this can be done very efficiently in the linear least-squares case (because the computationally demanding inversion or factorization of the gram matrix has to be solved only once), it probably makes many more complex choices for coefficient estimation - e.g., GAMMs - computationally infeasible.

Conversely, this peculiarity of the least-squares procedure also implies even complex rERP designs can be estimated quickly with optimized least-squares solver algorithms. The same applies to any transformation of the design matrix or the data which still allows the usage of Ordinary Least Squares; for example, the technique can be employed in the time-frequency domain, obtained coefficient time series/rERPs can be source localized in the same manner as ERPs, and to model nonlinearities polynomial terms can be added just as we have added sine and cosine transforms.

6. Limitations and future directions

As a limitation, single trial data are not available in continuous time regression. This is inherent to the method. Additionally, regression might be less suited for paradigms that yield fewer fixations for analysis. Both simulations (Smith and Kutas, 2015b) and observations within our own lab, indicate that having fewer than 200–300 events (fixations) severely hampers overlap correction. Furthermore, additional studies will have to reveal to what extent a design can limit participants' freedom to move their eyes, and thereby the variance and range of predictor values, before it becomes problematic to regression analysis. Here, we have focused on free-viewing behavior because studying it challenges the use of averaged FRPs.

We see particular opportunities for regression-based methods in richly described stimulus sets such as corpus data in reading (e.g. Kliegl et al., 2004). Such stimulus sets can contain large variance, but also a description of said variance for each fixation. Moreover, there are manifold descriptions of each word along several different dimensions. Such a rich description effectively renders every fixation a potential data-point for evaluation of experimental effects (as opposed to a single target fixation, or a single manipulated object per scene). Regression

(*footnote continued*)
researchers should be aware of this.

can then serve to answer questions about the separate influences of different variables on an outcome measure. As such, it eliminates the need for binning. Although available for reading research, such datasets are lacking for other fields, such as scene perception; a large body of scene photographs with a full physical and semantic description is not available. Still, modelling approaches of a similar nature have already been applied to eye movements, for instance to study the effects of physical stimulus properties on fixation duration or fixation location in complex scenes (Einhäuser and Nuthmann, 2016; Nuthmann, 2017). With the rERP framework and the development of suitable stimulus sets we believe them to be applicable to ET-EEG free-viewing experiments, too.

7. Summary and conclusions

This paper has focused on the analysis of ET-EEG data from free-viewing experiments. Linear estimation is specifically suited for these kinds of data because control over subjects' eye movement behavior is inherently minimal. Such lack of control makes it implausible to construct bins contrasting data from two or more conditions along a single dimension. Adding confounding variables into a regression model as done here is a theoretically sound alternative to binning.

In summary, this paper has provided an introduction to the application of linear least squares regression to fixation related potentials. We have demonstrated that linear estimates of possible confounds in a free-viewing task are sensible. Critically, and as an advantage over averaging, any variance attributed to the confound in linear regression cannot be attributed to an experimental manipulation. Relatedly, we have shown that experimental effects can be estimated with higher accuracy and precision. We have also given multiple examples of how overlap correction is beneficial and critical for analyzing ET-EEG data. As a consequence, baselining can be applied without risking distortions and researchers have more freedom in choosing experimental designs, such as ones that require manual responses shortly after critical fixations. We have obtained similar results from a task with a highly-controlled stimulus (the visual search task) and a task with a highly variable stimulus (scene memorization). In both cases rFRPs for confound control are similar to FRP visualizations created for comparison. Visual comparison across experiments suggests that physical stimulus properties can manifest in rFRPs for other predictors when not explicitly modelled. We propose that extra research is needed to better understand the relations between free-viewing EEG and stimulus properties, especially for complex stimuli.

To indicate limitations, we have shown an example of how collinearity caused by temporal adjacency can hamper overlap correction. Lastly, we have discussed non-linearity. Most predictor/EEG-relationships in free-viewing may well be non-linear. For some we assume linearity over a certain range of predictor values. For predictors and ranges that violate the assumption of linearity, non-linear relationships can be fit within a linear regression model. Even less restrictive predictor-EEG relationships should be defined in a thoughtful manner, and could be a basis for future investigations.

In conclusion: for estimating fixation related EEG in a free-viewing task, rFRPs are favorable over averaged FRPs.

Funding

This work was funded by DFG grant VO 1683/2-1 to MLV, and ERC Consolidator Grant617891 supporting JS.

Acknowledgements

The authors would like to thank Dr. Dejan Draschkow for helpful discussions and comments, as well as Aylin Kallmayer and Daniela Gresch for their help with data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jneumeth.2018.12.010.

References

Boylan, C., Doig, H.R., 1989. Effect of saccade size on presaccadic spike potential amplitude. Invest. Ophthalmol. Vis. Sci. 30 (12), 2521–2527.
Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.
Brouwer, A.M., Reuderink, B., Vincent, J., van Gerven, M.A.J., van Erp, J.B.F., 2013. Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. J. Vis. 13 (3). https://doi.org/10.1167/13.3.17. 17–17.
Brouwer, A.-M., Brinkhuis, M., Reuderink, B., Hogervorst, M., van Erp, J., 2014. Fixation-Related Potentials: Foveal Versus Parafoveal Target Identification.
Buswell, G.T., 1935. How People Look at Pictures.
Dandekar, S., Ding, J., Privitera, C., Carney, T., Klein, S.A., 2012a. The fixation and saccade P3. PLoS One 7 (11), e48761. https://doi.org/10.1371/journal.pone.0048761.
Dandekar, S., Privitera, C., Carney, T., Klein, S.A., 2012b. Neural saccadic response estimation during natural viewing. J. Neurophysiol. 107 (6), 1776–1790. https://doi.org/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. J. Neurosci. Methods 134 (1), 9–21.
Delorme, A., Palmer, J., Onton, J., Oostenveld, R., Makeig, S., 2012. Independent EEG sources are dipolar. PLoS One 7 (2), e30135.
Devillez, H., Guyader, N., Guérin-Dugué, A., 2015. An eye fixation–related potentials analysis of the P300 potential for fixations onto a target object when exploring natural scenes. J. Vis. 15 (13). https://doi.org/10.1167/15.13.20. 20–20.
Di Liberto, G.M., Lalor, E.C., 2017. Indexing cortical entrainment to natural speech at the phonemic level: methodological considerations for applied research. Hear. Res. 348, 70–77. https://doi.org/10.1016/j.heares.2017.02.015.
Dimigen, O., Sommer, W., Hohlfeld, A., 2011. Coregistration of eye movements and EEG in natural reading: analyses and review. J. Exp. Psychol. Gen. 140 (4), 552–572. https://doi.org/10.1037/a0023885).
Ehinger, B.V., König, P., Ossandón, J.P., 2015. Predictions of visual content across eye movements and their modulation by inferred information. J. Neurosci. 35 (19), 7403–7413.
Einhäuser, W., Nuthmann, A., 2016. Salient in space, salient in time: fixation probability predicts fixation duration during natural scene viewing. J. Vis. 16 (11), 13–17. https://doi.org/10.1167/16.11.13.
Fischer, T., Graupner, S.-T., Velichkovsky, B.M., Pannasch, S., 2013. Attentional dynamics during free picture viewing: evidence from oculomotor behavior and electrocortical activity. Front. Syst. Neurosci. 7.
Gramfort, A., Luessi, M., Larson, E., Engemann, D.A., Strohmeier, D., Brodbeck, C., et al., 2013. MEG and EEG data analysis with MNE-Python. Front. Neurosci. 7.
Hale, K.S., Fuchs, S., Axelsson, P., Berka, C., Cowell, A.J., 2008. Using physiological measures to discriminate signal detection outcome during imagery analysis. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 52 (3), 182–186. https://doi.org/10.1177/154193120805200307.
Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W.D., Davis, M.H., 2009. Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. Biol. Psychol. 80 (1), 64–74.
Hayhoe, M., Ballard, D., 2005. Eye movements in natural behavior. Trends Cogn. Sci. 9 (4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009.
Henderson, J.M., 2003. Human gaze control during real-world scene perception. Trends Cogn. Sci. 7 (11), 498–504.
Henderson, J.M., 2007. Regarding scenes. Curr. Dir. Psychol. Sci. 16 (4), 219–222. https://doi.org/10.1111/j.1467-8721.2007.00507.x.
Hessels, R.S., Kemner, C., Boomen, C., Hooge, I.T.C., 2016. The area-of-interest problem in eyetracking research: a noise-robust solution for face and sparse stimuli. Behav. Res. Methods 48 (4), 1–19. https://doi.org/10.3758/s13428-015-0676-y.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J., 2011. Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press.
Hooge, I., Nyström, M., Cornelissen, T., Holmqvist, K., 2015. The art of braking: post saccadic oscillations in the eye tracker signal decrease with increasing saccade size. Vision Res. 112, 55–67.
Huber-Huber, C., Ditye, T., Fernández, M.M., Ansorge, U., 2016. Using temporally aligned event-related potentials for the investigation of attention shifts prior to and during saccades. Neuropsychologia 92, 129–141.
Hutzler, F., Fuchs, I., Gagl, B., Schuster, S., Richlan, F., Braun, M., Hawelka, S., 2013. Parafoveal X-masks interfere with foveal word recognition: evidence from fixation-related brain potentials. Front. Syst. Neurosci. 7.
Kamienkowski, J.E., Ison, M.J., Quiroga, R.Q., Sigman, M., 2012. Fixation-related potentials in visual search: a combined EEG and eye tracking study. J. Vis. 12 (7). https://doi.org/10.1167/12.7.4. 4–4.
Kapoula, Z., Robinson, D.A., 1986. Saccadic undershoot is not inevitable: saccades can be accurate. Vision Res. 26 (5), 735–743.
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. https://doi.org/10.1016/j.neuroimage.2013.12.006.
Kazai, K., Yagi, A., 1999. Integrated effect of stimulation at fixation points on EFRP (eye-fixation related brain potentials). Int. J. Psychophysiol. 32 (3), 193–203.
Kingstone, A., Klein, R.M., 1993. What are human express saccades? Atten. Percept. Psychophys. 54 (2), 260–273.
Kliegl, R., Grabner, E., Rolfs, M., Engbert, R., 2004. Length, frequency, and predictability effects of words on eye movements in reading. Eur. J. Cogn. Psychol. 16 (1-2), 262–284.
Körner, C., Braunstein, V., Stangl, M., Schlögl, A., Neuper, C., Ischebeck, A., 2014. Sequential effects in continued visual search: Using fixation-related potentials to compare distractor processing before and after target detection. Psychophysiology 51 (4), 385–395.
Kristensen, E., Guérin-Dugué, A., Rivet, B., 2017a. Regularization and a general linear model for event-related potential estimation. Behav. Res. Methods 1–20.
Kristensen, E., Rivet, B., Guérin-Dugué, A., 2017b. Estimation of overlapped eye fixation related potentials: the general linear model, a more flexible framework than the ADJAR algorithm. J. Eye Mov. Res. 10 (1), 1–27.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62 (1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123.
Luck, S.J., 2014. An Introduction to the Event-Related Potential Technique. MIT Press.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.
Nikolaev, A.R., Meghanathan, R.N., van Leeuwen, C., 2016. Combining EEG and eye movement recording in free viewing: pitfalls and possibilities. Brain Cogn. 107, 55–83. https://doi.org/10.1016/j.bandc.2016.06.004.
Nuthmann, A., 2017. Fixation durations in scene viewing: modeling the effects of local image features, oculomotor parameters, and task. Psychon. Bull. Rev. 24 (2), 370–392. https://doi.org/10.3758/s13423-016-1124-4.
Nyström, M., Holmqvist, K., 2010. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behav. Res. Methods 42 (1), 188–204.
Nyström, M., Hooge, I., Holmqvist, K., 2013. Post-saccadic oscillations in eye movement data recorded with pupil-based eye trackers reflect motion of the pupil inside the iris. Vision Res. 92, 59–66.
Nyström, M., Andersson, R., Magnusson, M., Pansell, T., Hooge, I., 2015. The influence of crystalline lens accommodation on post-saccadic oscillations in pupil-based eye trackers. Vision Res. 107, 1–14.
Palmer, J.A., Kreutz-Delgado, K., Makeig, S., 2012. AMICA: an adaptive mixture of independent component analyzers with shared components. Swartz Center for Computatonal Neursoscience. University of California San Diego, Tech. Rep.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat. Vis. 10 (4), 437–442.
Plöchl, M., Ossandón, J.P., König, P., 2012. Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data. Front. Hum. Neurosci. 6.
Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148.
Rayner, K., 2009. Eye movements and attention in reading, scene perception, and visual search. Q. J. Exp. Psychol. 62 (8), 1457–1506. https://doi.org/10.1080/17470210902816461.
Rousselet, G.A., Pernet, C.R., Bennett, P.J., Sekuler, A.B., 2008. Parametric study of EEG sensitivity to phase noise during face processing. BMC Neurosci. 9 (1), 98.
Rousselet, G.A., Husk, J.S., Pernet, C.R., Gaspar, C.M., Bennett, P.J., Sekuler, A.B., 2009. Age-related delay in information accrual for faces: evidence from a parametric, single-trial EEG approach. BMC Neurosci. 10 (1), 114.
Rousselet, G.A., Gaspar, C.M., Pernet, C.R., Husk, J.S., Bennett, P.J., Sekuler, A.B., 2010. Healthy aging delays scalp EEG sensitivity to noise in a face discrimination task. Front. Psychol. 1.
Sassenhagen, J., 2018. How to analyse electrophysiological responses to naturalistic language with time-resolved multiple regression. Lang. Cogn. Neurosci. https://doi.org/10.1080/23273798.2018.1502458.
Sassenhagen, J., Alday, P.M., 2016. A common misapplication of statistical inference: nuisance control with null-hypothesis significance tests. Brain Lang. 162, 42–45.
Shishkin, S.L., Nuzhdin, Y.O., Svirin, E.P., Trofimov, A.G., Fedorova, A.A., Kozyrskiy, B.L., Velichkovsky, B.M., 2016. EEG negativity in fixations used for gaze-based control: toward converting intentions into actions with an Eye-Brain-Computer Interface. Front. Neurosci. 10.
Simola, J., Torniainen, J., Moisala, M., 2013. Eye movement related brain responses to emotional scenes during free viewing. Front. Syst. https://doi.org/10.3389/fnsys.2013.00041/abstract.
Slanzi, G., Balazs, J.A., Velásquez, J.D., 2017. Combining eye tracking, pupil dilation and EEG analysis for predicting web users click intention. Inf. Fusion 35 (C), 51–57. https://doi.org/10.1016/j.inffus.2016.09.003.
Smith, N.J., 2011. Scaling up Psycholinguistics. [Ph.D. Thesis]. Retrieved from:. https://vorpus.org/papers/smith-2011-phd-thesis.pdf.
Smith, N.J., Kutas, M., 2015a. Regression-based estimation of ERP waveforms: I. The rERP framework. Psychophysiology 52 (2), 157–168. https://doi.org/10.1111/psyp.12317.
Smith, N.J., Kutas, M., 2015b. Regression-based estimation of ERP waveforms: 3. Nonlinear effects, overlap correction, and practical considerations. Psychophysiology 52 (2), 169–181. https://doi.org/10.1111/psyp.12320.
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 Res. 551 (1), 150–156.
Weiss, B., Knakker, B., Vidnyánszky, Z., 2016. Visual processing during natural reading. Sci. Rep. 6.
Woldorff, M.G., 1993. Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction. Psychophysiology 30 (1), 98–119.
Wood, S.N., 2017. Generalized Additive Models: An Introduction With R.Generalized Additive Models: An Introduction With R.
Yarbus, A.L., 1967. Eye Movements During Perception of Complex Objects. pp. 171–211.