NeuroImage 207 (2020) 116117

Contents lists available at ScienceDirect

NeuroImage

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

Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments☆

Olaf Dimigen

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099, Berlin, Germany

ARTICLE INFO

Keywords:
EEG
Eye tracking
Eye artifact correction
Independent component analysis (ICA)
Fixation-related potentials (FRP)
Beta and gamma oscillations

ABSTRACT

Combining EEG with eye-tracking is a promising approach to study neural correlates of natural vision, but the resulting recordings are also heavily contaminated by activity of the eye balls, eye lids, and extraocular muscles. While Independent Component Analysis (ICA) is commonly used to suppress these ocular artifacts, its performance under free viewing conditions has not been systematically evaluated and many published reports contain residual artifacts. Here I evaluated and optimized ICA-based correction for two tasks with unconstrained eye movements: visual search in images and sentence reading. In a first step, four parameters of the ICA pipeline were varied orthogonally: the (1) high-pass and (2) low-pass filter applied to the training data, (3) the proportion of training data containing myogenic saccadic spike potentials (SP), and (4) the threshold for eye tracker-based component rejection. In a second step, the eye-tracker was used to objectively quantify the correction quality of each ICA solution, both in terms of undercorrection (residual artifacts) and overcorrection (removal of neurogenic activity). As a benchmark, results were compared to those obtained with an alternative spatial filter, Multiple Source Eye Correction (MSEC). With commonly used settings, Infomax ICA not only left artifacts in the data, but also distorted neurogenic activity during eye movement-free intervals. However, correction results could be strongly improved by training the ICA on optimally filtered data in which SPs were massively overweighted. With optimized procedures, ICA removed virtually all artifacts, including the SP and its associated spectral broadband artifact from both viewing paradigms, with little distortion of neural activity. It also outperformed MSEC in terms of SP correction. Matlab code is provided.

Humans actively explore their environment with 2–4 saccadic eye movements per second, or about 10,000 during every waking hour. Although natural vision is fundamentally trans-saccadic, procedures in electroencephalographic (EEG) research have traditionally aimed to minimize oculomotor behavior by requiring sustained visual fixation. In recent years, however, there has been rising interest in measuring brain-electric activity also during unconstrained viewing situations such as reading (e.g. Dimigen et al., 2011; Henderson et al., 2013), scene viewing (e.g. Nikolaev et al., 2011; Ossandon et al., 2010; Simola et al., 2013), visual search (e.g. Brouwer et al., 2013; Kamienkowski et al., 2012; Körner et al., 2014; Ries et al., 2018) or whole-body motion (Soto et al., 2018).

With this approach, eye movements are co-recorded with the EEG and the signal is aligned to the beginning or end of spontaneously occurring eye movements, yielding saccade- or fixation-related potentials (SRPs/FRPs), respectively. Fixations onsets, in particular, provide natural time-locking points to study attentional, cognitive, or affective processes during natural vision, since every fixation triggers a renewed sequence of lambda waves (Evans, 1953; Gaarder et al., 1964; Yagi, 1979), primarily visually-evoked potentials that share many features with those elicited by passive retinal stimulation (Dandekar et al., 2012; Kazai and Yagi, 2003; Kornrumpf et al., 2016; Marton et al., 1985).

Despite their promises, recordings during natural vision are also complicated by serious data-analytical challenges (Baccino, 2011; Dimigen et al., 2011; Nikolaev et al., 2016), the most obvious of which are the voltage distortions produced by rotation of the eye balls, movements of the eye lids, and contraction of the extraocular muscles (Berg and Scherg, 1991; Keren et al., 2010; Picton et al., 2000; Plöchl et al.,

☆ *Author Note.* I wish to acknowledge Maarten De Schuymer who conducted early explorations of the effects of high-pass filtering that helped to initiate this project. I am also grateful to Lisa Spiering for assisting with the MSEC correction, Alexander Mies for collecting the reading data, and Werner Sommer for providing an excellent working environment. Constructive reviews by Leon Deouell, David Groppe and an anonymous reviewer improved the current work. Collection of the one of the datasets was supported by a DFG grant (FG-868-A2). Supporting data and code is found at https://osf.io/ka6qh and www.github.com/olafdimigen/opticat.

E-mail address: olaf.dimigen@hu-berlin.de.

https://doi.org/10.1016/j.neuroimage.2019.116117
Received 18 January 2019; Received in revised form 1 July 2019; Accepted 20 August 2019
Available online 2 November 2019

2012). These ocular artifacts pose inferential hazards because they are not only magnitudes larger than the event-related neural signals, but typically correlated to the experimental condition due to condition differences in saccade size, saccade orientation, or fixation duration. Their complete removal is therefore crucial to avoid misinterpretations.

A multitude of methods has been proposed for ocular correction, including those based on EEG-on-EOG regression, dipole modelling/beamforming, PCA, and other variants of blind source separation (Brunia et al., 1989; Delorme et al., 2007; Gratton, 1998; Ille et al., 2002). Of these, Independent Component Analysis (ICA) is now perhaps most commonly used to remove occasional saccade and blink artifacts in steady-fixation experiments (Delorme et al., 2007; Jung et al., 2000). ICA employs higher order statistics to decompose the EEG into independent components (ICs), linear combinations of the scalp channels weighted to be maximally temporally independent (Stone, 2004). For this purpose, ICA is typically trained on a portion of the recording – the *training data* – containing a sample of the neural and non-neural sources active during the task. Only ICs believed to reflect neural sources are then converted back (i.e., back-projected) to the electrode space, yielding an artifact-corrected version of the EEG.

Although ICA is now also frequently applied in SRP/FRP studies, its performance on heavily contaminated free viewing data has not been systematically evaluated, nor compared to that of alternative methods. Yet there are reasons to suspect that the quality of ICA-based ocular artifact correction may be overestimated in practice, in particular for free viewing applications. The first is that the corrected data is rarely analyzed time-locked to the saccade itself, but often inspected in continuous data (where small residual artifacts are nearly impossible to spot) or in fixation onset-aligned averages (where the saccade onset-locked artifacts are jittered due to variance in saccade duration). Second, many free viewing studies aggregate across saccades of different orientations, meaning that the most salient corneoretinal (CR) artifacts cancel out (partially) in the average, but may still remain present at the single-trial level. Third, the vast majority of SRP/FRP studies so far has focused on posterior scalp sites, where residual artifacts are less obvious. In contrast, facial electrooculogram (EOG) electrodes, which show artifacts most clearly, are often excluded before ICA or not plotted. Finally, few studies have tested whether the correction removes legitimate neurogenic activity.

If the corrected EEG is instead aligned to the onsets of saccades of a single orientation and inspected at frontal sensors, the results are often sobering. Eye muscle-generated saccadic spike potentials (SPs, see below), in particular, are difficult to model with ICA and most existing reports display residual SP artifacts in the plotted waveforms (for arbitrarily picked examples see Dimigen et al., 2012; Henderson et al., 2013; Kamienkowski et al., 2012; Körner et al., 2014; Kornrumpf et al., 2016; Nikolaev et al., 2011; Ossandon et al., 2010; Ries et al., 2018; Simola et al., 2013).

The presence of residual artifacts in most published free viewing studies indicates that correction procedures need to be improved and specifically adapted for this purpose. In case of ICA, this requires addressing three practical problems: (1) How to select and preprocess the training data, (2) how to reliably categorize ICs as ocular/non-ocular, and (3) how to quantify and minimize the distortion of neural activity. The preprocessing of the input data, in particular, has received comparatively little attention, but has been shown to be more important than the choice of the ICA algorithm (e.g. Infomax, FastICA, or AMICA; Zakeri et al., 2014). Rather than proposing yet another correction method, goal of the present study was therefore to validate and adapt the pipeline for the widely-used extended Infomax algorithm (Bell and Sejnowski, 1995; Lee et al., 1999) to natural viewing.

In the following, I will first briefly review the three partially independent mechanisms that generate ocular artifacts. I will then describe the four parameters of the ICA pipeline manipulated in the present study. Finally, I will outline how parallel eye-tracking can be used to quantify correction outcomes more objectively.

1. Three types of ocular artifacts

Corneoretinal (CR) dipole. Due to metabolic activity in the pigmented layer of the retina (Marmor and Zrenner, 1993), each eye ball posesses an electrical gradient of one to several millivolts (depending on ambient illumination level, Marmor and Zrenner, 1993; Young and Sheena, 1988) between its front (cornea) and back (retina). When the eyes rotate, these corneoretinal dipoles also rotate, such that a rightward saccade, for example, generates maximal positive distortions at frontolateral right-hemispheric electrodes (towards which the corneas rotate). For saccades up to about 30°–40° (Shinomiya et al., 2008), CR artifact amplitude increases as a linear function of saccade size (at ~9.5–16 μV per degree; Keren et al., 2010), meaning that propagation factors are similar for small and large saccades of the same orientation. However, artifact topographies are not strictly mirrosymmetric for upward versus downward saccades (Picton et al., 2000; Plöchl et al., 2012) and have also been reported to depend on the participant's absolute screen viewing position (Ai et al., 2016), which of course changes frequently during free viewing.

Eye lids: During blinks, the eye lids slide across the positive corneas, allowing current to flow to the forehead (Matsuo et al., 1975). A smaller blink-like artifact – sometimes called eye "rider artifact" (Lins et al., 1993) – also occurs towards the end of upward or oblique-upward saccades (Barry and Jones, 1965; Plöchl et al., 2012) and is believed to be caused by the eye lids lagging behind the upward-rotating eye balls, temporarily changing their overlap with the positive cornea. The result is a blink-like frontal positivity that begins during upward saccades but outlasts the offset of the saccade (by about 100 ms in Plöchl et al., 2012).

Spike potential: The most difficult-to-correct artifact is the saccadic spike potential (SP), a brief, high-frequency biphasic wave, which ramps up ~5–10 ms before the saccade and reaches its primary peak at saccade onset (Blinn, 1955; Keren et al., 2010; see Carl et al., 2012 for MEG). The main peak of the SP has a focal negative maximum at facial electrodes but is accompanied by a more widespread positive deflection over parietal sites. SP topography also changes with saccade direction, but lateralization of its main peak is opposite to that of CR artifacts such that for a rightward saccade, the frontally-negative main spike is largest near the right eye, whereas the diffuse parietal positive pole shifts towards the left hemisphere. However, the SP is less lateralized than the CR artifact around the eyes, which explains why bipolar EOG montages do not capture this artifact well. Instead, it is largest in a "radial" EOG montage for which facial electrodes are referenced against a centroparietal site. SP amplitude increases with saccade size, although it is unclear whether this relationship is linear (Keren et al., 2010) or not (Boylan and Ross Doig, 1989).

Current scientific consensus holds that the SP most likely reflects myogenic (EMG) activity from the recruitment and only initially synchronous firing of the motor units of the extraocular muscles at saccade onset (Thickbroom and Mastaglia, 1985; Yamazaki, 1968). The SP is not of corneoretinal origin since it precedes rotation of the eye ball and survives removal of the bulbus in patients with an eye ball prothesis but preserved eye muscles (Thickbroom & Mastaglia, 1985). A myogenic rather than cerebral source is also suggested by its frontal generators (Carl et al., 2012; Hipp and Siegel, 2013), its presence in darkness (Riggs et al., 1974) and intramuscular EMG (Yamazaki, 1968) as well as its weakness or absence in intracranial EEG (Jerbi et al., 2009; Kovach et al., 2011; Sakamoto et al., 1991) and patients with extraocular muscle palsy (Thickbroom & Mastaglia, 1985). In contrast, there is little persuasive evidence for cerebral contributions to the SP itself (Balaban and Weinstein, 1985; Berchicci et al., 2012; Parks and Corballis, 2008), although relevant cortical activity may of course take place during the same interval (e.g. Parks and Corballis, 2008).

Spike potentials have received attention because even involuntary microsaccades (<1°) during attempted fixation generate sizeable SPs (Armington, 1978; Dimigen et al., 2009; Yamazaki, 1968; Yuval-Greenberg et al., 2008), which introduce a broadband artifact in the

time-frequency spectrum of the EEG, affecting the low-amplitude beta (~14–30 Hz) and gamma bands (>30 Hz), in particular (Reva and Aftanas, 2004; Yuval-Greenberg et al., 2008). Complete removal of SPs with ICA has proven challenging even for microsaccades (Craddock et al., 2016; Hassler et al., 2011; Hipp and Siegel, 2013; Keren et al., 2010) and ICA often fails to single out the SP in one or more distinct ICs (Hipp and Siegel, 2013; Keren et al., 2010). In addition, even clean SP components can be difficult to spot in scalp maps alone (Plöchl et al., 2012), especially if EOG electrodes were not placed around both eyes, recorded with a bipolar montage, or excluded before ICA (Keren et al., 2010).

Residual SP artifacts pose problems for SRP/FRP analyses in free viewing. First, they distort the typical baseline period in the time and

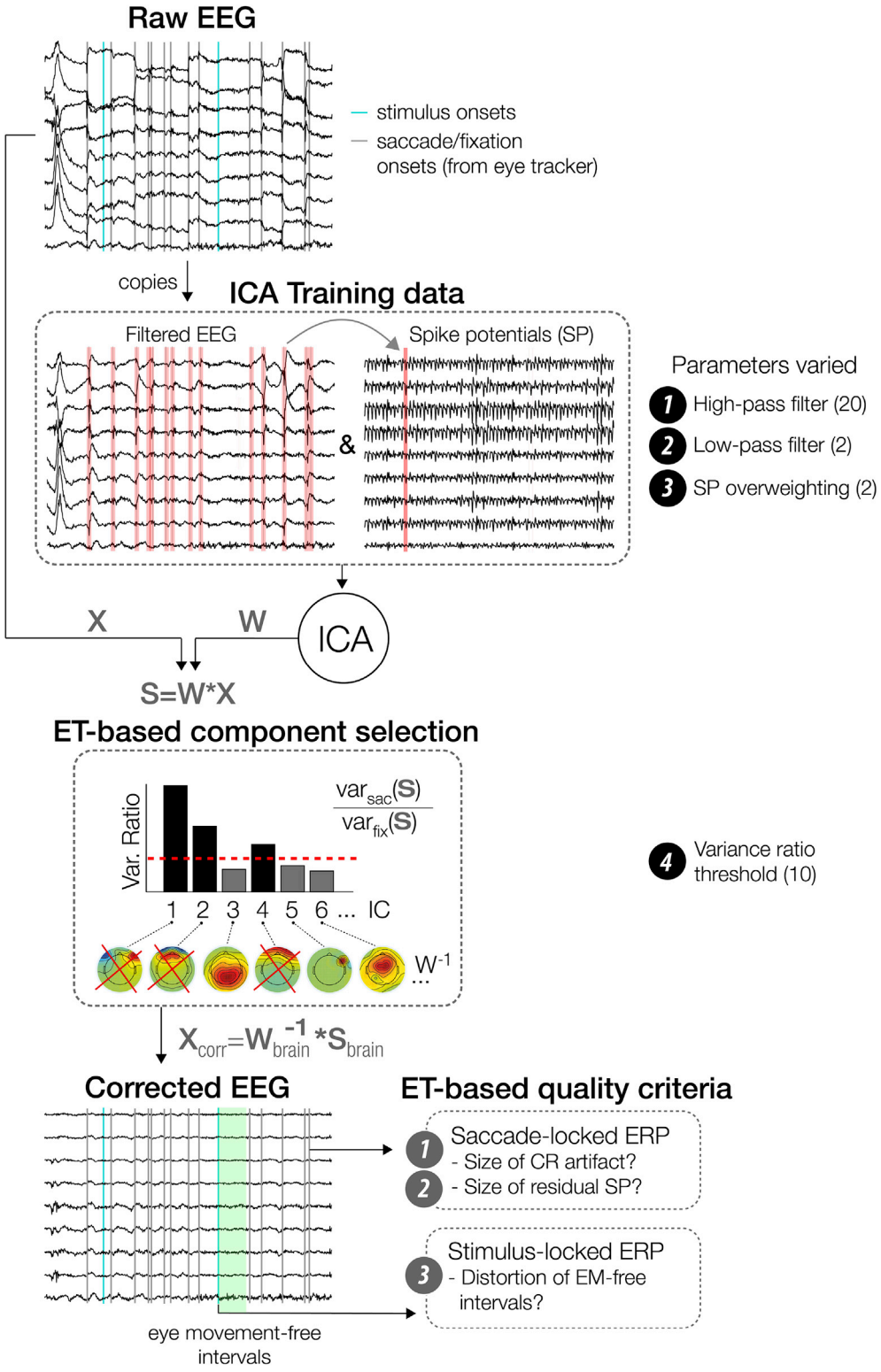

Fig. 1. Schematic workflow to determine parameters for optimized ICA training (OPTICAT) and component classification. Steps are described in the section "*Pipeline overview*".

frequency domain. Placing the baseline further away from saccade onset (e.g. −200 to −100 ms) is possible (e.g., Nikolaev et al., 2016; Simola et al., 2013), but will reduce the signal-to-noise ratio of the averaged waveforms. The spectral broadband artifact can also easily distort the EEG's beta and gamma band during free viewing, e.g. if it leaks into the post-saccadic interval in a time-frequency analysis. Second, the SP greatly complicates the study of phenomena occurring in temporal proximity to saccade onset, such as remapping (Kusunoki and Goldberg, 2003), saccadic suppression (Duffy and Lombroso, 1968), or saccade-related changes in cortical excitability (Ito et al., 2011). Finally, because a new saccade is executed every 200–400 ms during natural vision, late intervals of the SRP/FRP waveform are also contaminated by SPs from the following saccades on the stimulus (see also Fig. 6). It is therefore desirable to fully remove SPs. The following section reviews several ICA preprocessing parameters that can influence correction performance.

2. Explored parameters of the ICA pipeline

High-pass filter. The reliability of ICA decompositions has been shown to improve after signal offsets are removed by subtracting the mean voltages across each epoch (mean-centering, Groppe et al., 2009). In addition, practical experience suggests that decompositions improve after slow oscillations and drifts – for example caused by electrode potentials or skin potential fluctuations – are further suppressed by high-pass filtering (Debener et al., 2010; Winkler et al., 2015; Zakeri et al., 2014). The adverse effects of slow signals on ICA unmixing quality are not fully understood (Debener et al., 2010; Winkler et al., 2015), but one likely reason is that ICA is biased towards these high-amplitude signals since it tends to focus on data expressing the most power. Filtering may also help to satisfy ICA's stationarity assumption by removing signals that are spatially unstable.

Using data from an auditory oddball task, Winkler et al. (2015) systematically investigated effects of high-pass filtering on artifact reduction. Unmixing weights were computed on filtered data and then applied to the unfiltered recording (to preserve slow ERPs like P300; Debener et al., 2010). Filtering at 1 or 2 Hz produced the most dipolar ICs (Delorme et al., 2012), the best discrimination between targets/non-targets, and the least noisy ERPs. While these results underline the importance of data preprocessing, adequate filtering may be even more important for free viewing applications. The reason is that during normal vision – and in contrast to tasks with isolated saccades (e.g. Plöchl et al., 2012) – CR artifacts from multiple saccades frequently sum up to produce large deviations from baseline. This problem is most obvious in reading, where ~85% of all saccades point in the same direction (Rayner, 1998), creating DC offsets of ±250 μV at the end of each trial (Dimigen et al., 2011; their Fig. 1C). To explore this hypothesis, I high pass-filtered the ICA training data at 20 different frequencies.

Low-pass filter. Removing high frequencies can also improve decompositions since it attenuates electromagnetic noise and scalp-EMG. High cutoffs around 40–45 Hz are therefore commonly applied to ICA input data (e.g. Castellanos and Makarov, 2006; Gwin et al., 2010; Mannan et al., 2016; Ries et al., 2018; Winkler et al., 2015; Zakeri et al., 2014). However, EMG is not only produced by face, head, and neck muscles, but presumably also reflected in the SP, whose bandwidth extends to at least 90 Hz (Keren et al., 2010; Nativ et al., 1990). This implies that the SP may actually be modeled better if high frequencies remain in the data. Keeping high frequencies in the data also allows the researcher to isolate and remove other unwanted scalp-EMG sources from the data with ICA. To test this hypothesis, the training data was low pass-filtered at either 40 or 100 Hz.

Spike potential overweighting. The difficulty of removing the SP with ICA is likely due to the fact that it accounts for little energy in the signal, because the SP (1) is of moderate amplitude compared to CR artifacts, (2) possesses a different topography for different saccade directions, (3) lasts only a few samples (~20 ms, see Fig. 5), and (4) also changes its topography within this interval. A potential solution, suggested by Keren et al. (2010), is to overweight peri-saccadic samples in the training data (see also Mennes et al., 2010). In particular, both Keren et al. (2010) and Craddock et al. (2016) trained their ICAs on relatively short segments (81 ms and 200 ms long, respectively) centered on microsaccades. Similarly, Meyberg et al. (2017) found that only the inclusion of 15° saccades in the training data allowed for the removal of small CR artifacts produced by microsaccades. Finally, to aid component identification, Hassler et al. (2011) proposed an unconventional use of ICA for which "virtual" channels are added to the data, which contain a copy of the SP waveform (averaged across frontal channels) during microsaccade intervals, but zeros elsewhere. Taken together, these studies suggest that overweighting can improve ocular correction, particularly in case of the SP, but this approach has not been tested on free viewing data. I therefore aimed to improve the correction by massively overweighing peri-saccadic samples.

Eye tracker-guided IC classification. Unlike the EOG, eye-tracking provides an accurate gaze position signal that is electrically independent of the EEG and therefore potentially useful to improve ocular correction (Dimigen et al., 2011; Hironaga et al., 2004; Kierkels et al., 2007; Lourenço et al., 2016; Mannan et al., 2016; Noureddin et al., 2012; Plöchl et al., 2012).

One application is the objective classification of ICs (Dimigen et al., 2011; Plöchl et al., 2012). A simple but elegant criterion for this purpose was proposed by Plöchl et al. (2012), who validated it on data from a guided-saccade task. Basis for their classification is the variability of each IC's activity time course during saccades versus fixations: ICs showing relatively more variance during saccade intervals are likely to reflect ocular artifacts, whereas those showing more variance during fixations are likely neurogenic (because each fixation evokes lambda waves). Thus, if the ratio of both variances ($var_{saccade}/var_{fixation}$) exceeds 1, an IC is likely to reflect artifact; conversely, ratios below 1 indicate neural sources. Non-ocular artifacts such as cardiac activity show similar variance during both intervals and are not flagged; they can be detected with other techniques (Campos Viola et al., 2009; Chaumon et al., 2015; Mognon et al., 2011; Nolan et al., 2010; Winkler et al., 2011).

Obviously, the success of this procedure depends on an appropriate threshold. Although a value of 1 seems like a logical choice, Plöchl et al. (2012) proposed using a slightly higher threshold (1.1) to reduce misclassifications. Here I explored different thresholds to identify the lowest variance ratio threshold that removes ocular ICs while preserving neurogenic activity.

3. Quantifying overcorrection with eye-tracking

Correction can distort data in two ways. If ocular ICs are missed, artifacts will remain in the data (*undercorrection*). A question less often discussed is whether ICA removes brain activity observed in the task (overcorrection; e.g. Castellanos and Makarov, 2006; Mennes et al., 2010; Pontifex et al., 2016; see Ries et al., 2018 for a recent investigation with FRPs). Some overcorrection is to be expected since the number of active sources is likely much higher than the number of recording channels, meaning that ICA will never perfectly isolated any single source (Groppe et al., 2008). Thus, overcorrection can happen because the ICA produces mixed non-neural/neural ICs (e.g. Castellanos and Makarov, 2006; McMenamin et al., 2010) but also because the experimenter accidently removes neurogenic ICs. For example, inquiries among colleagues suggest that some labs tend to remove all ICs with a frontal topography that is focal or bipolar around the eyes. Presumably, however, at least some of these sources reflect (pre)frontal brain activity.

Under most circumstances, it is impossible to quantify overcorrection since the "true" (artifact-free) data is unavailable for comparison. Here I argue that high-resolution eye-tracking provides a unique opportunity in this regard, since it allows us to identify intervals *objectively free* of significant oculomotor activity (Dimigen et al., 2011). In particular, every experiment contains at least some short intervals during which the eye

lids remained open and the eyes were (virtually) motionless. Blinks can be objectively identified in video-based eye-tracking data because pupil and corneal reflex tracking is lost during a blink. Similarly, significant eye movements during fixation – in particular microsaccades – can be detected with suitable algorithms (Engbert and Kliegl, 2003). Although intervals without microsaccades are rare, they can be found immediately after stimulus onsets, since any sufficiently strong stimulation triggers saccadic inhibition, a transient decrease in the rate of saccades (Reingold and Stampe, 2002) and microsaccades (Engbert and Kliegl, 2003). Since these eye movement-free intervals should not be modified by the ocular correction, they provide a ground truth to quantify overcorrection.

4. A benchmark for ICA: Multiple-Source Eye Correction

To put the performance of Infomax ICA into context, I compared it to that obtained with an alternative algorithm, the surrogate variant of *Multiple-Source Eye Correction* (MSEC, Berg and Scherg, 1994). Like ICA, MSEC can be described as a spatial filter (Ille et al., 2002), which separates brain and artifact activities based on their topographical definitions (for details, see the Supplementary Materials, Supplement A). A main difference to ICA is that artifact topographies are not obtained by blind source separation, but empirically defined by averaging the artifacts of isolated calibration saccades. In addition, a set of generic brain topographies is defined by a dipole model that is the same for all participants. The purpose of this "surrogate" brain source model is not to directly model neural activity in the task, but to reduce the subtraction of brain activity spatially correlated to the artifact topographies (i.e. to reduce overcorrection). Because MSEC has been shown to produce good corneoretinal correction in natural reading (e.g. Dimigen et al., 2011; Kornrumpf et al., 2016) is was considered a suitable benchmark.

5. Current study

To summarize, this study aimed to evaluate and improve the ICA-based ocular correction of free viewing data. For this purpose, I analyzed combined eye movement/EEG recordings from two frequently studied free viewing paradigms with different oculomotor behavior: visual search in scenes and left-to-right sentence reading. The four described parameters of the ICA pipeline were orthogonally manipulated (see Fig. 1). The outcomes of each ICA variant were then compared using objective measures of under- and overcorrection and benchmarked against those obtained with MSEC.

6. Methods

6.1. Pipeline overview

I analyzed data from two eye-tracking/EEG experiments with unconstrained eye movements: scene viewing and reading (Fig. 2). Both were recorded in the same laboratory with identical hardware and largely identical recording settings (details below).

Fig. 1 summarizes the analysis pipeline. In a first step, I created for each participant 40 differently filtered copies of the original recording by crossing 20 high-pass filters (with passband edges between 0.016 and 30 Hz) with two low-pass filters (passband edge of 40 or 100 Hz). These differently filtered datasets were then cut into epochs serving as ICA training data. From each of these "basic" versions of the training data, I then created a second, "overweighted" version in which peri-saccadic samples were overrepresented. This was achieved by cutting 30 ms segments around saccade onsets and repeatedly re-appending them to the training data.

In the next step, ICAs were computed on each of the resulting 80 training datasets (20 high-pass filters × 2 low-pass filters × 2 versions). In a third step, the resulting unmixing matrices (**W**) were multiplied with the original, *unfiltered* recording (**X**), thereby producing activity waveforms (**S**) for the ICs in the original recording. In a fourth step, eye movement events provided by the eye-tracker were used to remove ICs whose activity waveforms showed more variance during saccades than fixations. To find the best threshold for this classification, 10 different thresholds were applied (0.6, 0.7, …1.5) and the corresponding artifact ICs were marked. The artifact-corrected data was then obtained by back-projecting only the neural sources ($X_{corr} = W^{-1}_{brain} * S_{brain}$). Since every threshold was applied to every ICA solution, this produced 800 versions of artifact-corrected data per participant (or 19,200 in total).

Finally, three eye tracker-based criteria were used to compare correction quality. To quantify *undercorrection*, I measured the residual amplitude of (1) CR and (2) SP artifacts in saccade onset-ERPs (SRPs). To quantify *overcorrection*, I identified short stimulus onset-aligned epochs

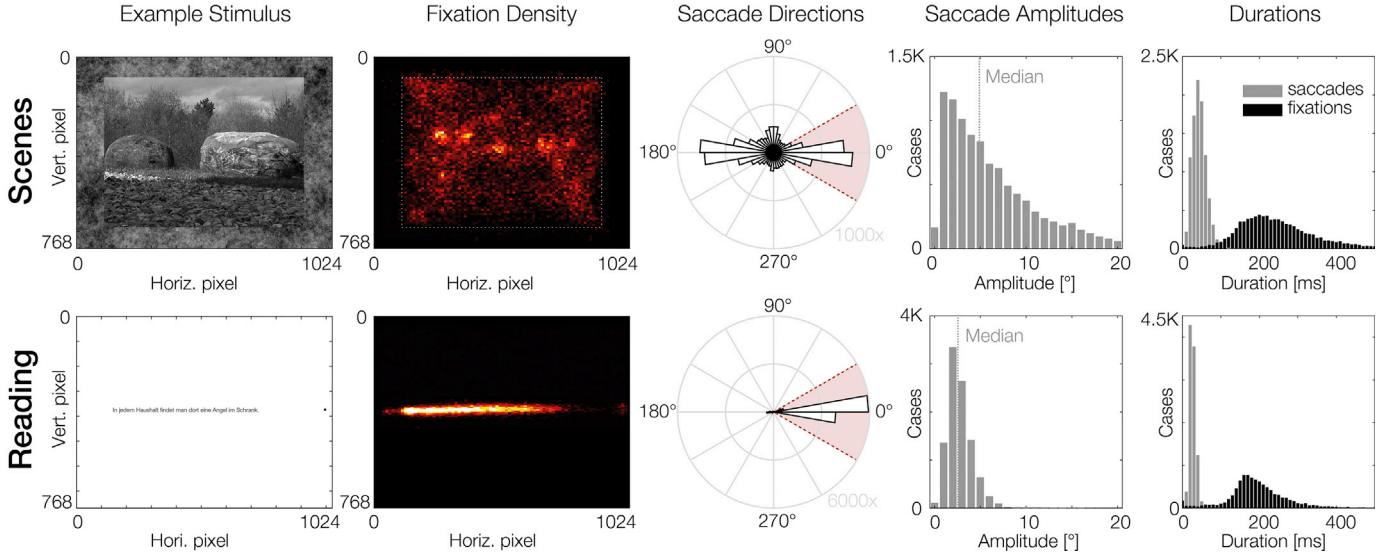

Fig. 2. Stimuli and eye movement behavior. In the Scenes experiment, participants searched for a small target stimulus hidden within natural images. In the Reading experiment, participants read short stories. For each experiment, plots show an example stimulus, a "heat map" of fixation locations (warmer colors indicate higher density), a directional histogram of saccade directions, and the distributions of saccade amplitude and saccade duration. Red shading in the saccade direction plots indicates saccades classified as "rightward" (±30°).

without detectable oculomotor activity. Because these intervals should not be affected by the ocular correction, overcorrection was quantified as (3) the degree to which these intervals were changed by ICA.

6.2. Participants

For the current analysis, I used the first 12 participants of each experiment. Most were students at Humboldt-University, with a mean age of 25.3 years in the *Scenes* (range 19–25 yrs, 7 female) and 21.3 years in the *Reading* experiment (18–33 yrs, 11 female). Procedures complied with the declaration of Helsinki and participants provided written informed consent.

6.3. Scene viewing experiment

In the *Scenes* experiment, participants searched for a target stimulus hidden within greyscale natural images. Most images were taken from the *Zurich Natural Image Database* (Einhäuser et al., 2006), a set of photographs shot in a forest. During the experiment, 99 images were presented at a resolution of 800 × 600 pixels (28.8° × 21.6°), centered on a 1024 × 768 pixels background showing 1/f noise. On each trial, a single scene was presented and the participant's task was to find a gray disc (0.4 cd/m^2) that appeared at a random location within the image 8–16 s after scene onset. This disc had an initial diameter of just 0.07° but then slowly increased in size. Once the participant found the target, he/she pressed a button, terminating the trial. The top row of Fig. 2 summarizes eye movement behavior in the task. Saccades had a median amplitude of 4.9°, mean duration of 44.2 ms and pointed in all directions, although most were horizontal. Fixation locations were distributed across the images, with some bias towards the image center and corners (heat map in Fig. 2). Average fixation duration was 264 ms.

6.4. Reading experiment

In the *Reading* experiment, participants read 152 short stories (including 8 practice trials) of the Potsdam Sentence Corpus III, a set of sentences used in previous ERP research (Dambacher et al., 2012). Each story consisted of two sentences successively presented as a single line of black text on a white background (0.45° per character). Purpose of the experiment was to study fixation-related N400 effects elicited by a semantically congruent/incongruent word in the second sentence. In addition, the experiment manipulated whether that target word was parafoveally visible or not while readers were looking at an earlier word in the sentence (n+2 boundary paradigm; Kliegl, Risse, & Laubrock, 2007). Only data for the second sentence was analyzed. As Fig. 2 shows, reading saccades had a median amplitude of 2.6° and mostly pointed rightward. Average fixation duration was 197 ms. Please note that the relatively short fixation durations in both experiments (cf. Rayner, 1998) are explained by the sensitive saccade detection algorithm used here, which also detects microsaccades.

6.5. Common methods: stimulation & eye-tracking

Recordings were made in an electromagnetically shielded room. Stimuli were presented on a 22 inch CRT monitor (Iiyama Vision Master Pro 510; vertical refresh: 160 Hz) at a viewing distance of 60 cm using *Presentation* software (Neurobehavioral Systems Inc.). Binocular eye movements were recorded at a rate of 500 Hz with a video-based eye-tracker (IView-X Hi-Speed, SMI GmbH). Offline, saccades and fixations were detected using Engbert and Kliegl's (2003) velocity-based algorithm as implemented in the EYE-EEG toolbox (Dimigen et al., 2011). Saccades were defined as intervals in which the velocity of both eyes exceeded for ≥10 ms a threshold set at 5 median-based SDs of all recorded eye velocities (excluding blink intervals). This low threshold (cf. Dimigen et al., 2009) was chosen to maximize detection sensitivity for microsaccades. In cases where multiple saccades were detected within <50 ms, I only kept the first to avoid detecting post-saccadic oscillations as separate saccades.

6.6. Common methods: electrophysiology

EEG and EOG were recorded from 45 (*Scenes*) or 63 (*Reading*) Ag/AgCl electrodes referenced against the left mastoid (A1). Electrodes were mounted in a cap at standard 10-10 system positions, except for four EOG electrodes that were affixed to the outer canthus and infraorbital ridge of each eye. Throughout all analyses, EOG electrodes remained in the data and were treated like EEG channels (which is crucial to detect residual artifacts and ocular ICs; Dimigen et al., 2011; Nikolaev et al., 2016). To make the montages in both experiments comparable, I reduced the *Reading* dataset also to 45 electrodes, which were largely identical to those in the *Scenes* experiment. Exact electrode locations are visualized in Supplementary Fig. S1. Signals were recorded with BrainAmp DC amplifiers (Brain Products GmbH) at a rate of 500 Hz with impedances kept < 5 kΩ. The *Scenes* data was initially acquired with a time constant of 10 s, whereas the *Reading* data was acquired as direct current data. To make the datasets directly comparable, the raw data of both experiments was high-pass filtered at 0.016 Hz (10 s time constant) using ERPLAB's Butterworth filter (*pop_basicfilter.m*, Lopez-Calderon and Luck, 2014). The EEG and eye-tracking data were synchronized offline based on shared trigger pulses with a synchronization error <2 ms.

6.7. Creating differently filtered ICA training data

For each participant, I trained the ICA on 80 differently preprocessed versions of the original data. In a first step, I created 20 copies of the continuous EEG recording and high-pass filtered them differently using EEGLAB's (Delorme and Makeig, 2004) standard filter function *pop_eegfiltnew.m* (a zero-phase Hamming-windowed sinc FIR filter, Widmann et al., 2015) with its default, frequency-dependent transition bandwidth. Importantly, this filter function expects as the input the edge of the filter's passband (the frequency at which the filter starts attenuating the signal) rather than the more commonly reported half-amplitude cutoff (the frequency at which the signal is attenuated by −6 dB). Therefore, for the present analysis, all filters are described by their passband edge. The corresponding cutoff values are given in Supplementary Table 1. Supplementary Fig. S2 visualizes the filter characteristics. For high-pass filtering, the 20 passband edges were set to: No filtering, 0.1, 0.25, 0.5, 0.75, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5, 7.5, 10, 12.5, 15, 20, 25, and 30 Hz. Afterwards, each dataset was also low-pass filtered with the (upper) edge of the passband set to either 40 Hz or 100 Hz.

Each filtered dataset was then cut into 3 s epochs (−200 to +2800 ms) around stimulus presentations. In the *Scene* dataset, I used the onset of the photograph as well as the onset of the search target, yielding 198 epochs. In the *Reading* data, I used the onset of the second sentence, yielding 152 epochs. Note that these epochs contain a representative sample of brain and artifact activity in these tasks, since they include both the stimulus onset and several subsequent saccades. From each epoch, I then removed the mean channel voltages across the epoch (mean-centering, Groppe et al., 2009). Any high-pass filtering was therefore done in addition to mean-centering.[1] To exclude epochs with large non-ocular artifacts, I rejected epochs containing extreme outliers (>±500 μV) in any channel. No additional pruning was performed (see also Winkler et al., 2015).

The resulting 40 differently filtered versions of training data will be called “basic” versions in the following, because they contain no

[1] Highly similar results were obtained without the mean-centering of the 3 s epochs. This step was only included here to illustrate that the effects of high-pass filtering go beyond those of the mean-centering for this epoch length. In contrast, the mean-centering step was beneficial in the case of the short, peri-saccadic epochs that were added during the overweighting procedure (see section *Overweighting spike potentials*).

overweighted artifacts. From each of these filtered and epoched datasets, I then used the first 162,000 points to train the ICA (corresponding to $k = 80$ points per weight in the 45×45 unmixing matrix).[2]

6.8. Overweighting spike potentials

From each basic versions of the training data, I then created an "overweighted" copy in which SP samples were overrepresented (Fig. 1). For this purpose, short 30 ms epochs (−20 to +10 ms) were extracted around all saccade onsets found in the basic version. These brief saccade-locked epochs were then again mean-centered, concatenated together, and repeatedly appended to the end of the training dataset until its total length was doubled to 324,000 points ($k = 160$), half of which now only consisted of parts of the SP waveform. This means that all of the appended samples were already contained in the basic version of the training data, that is, they were redundant except for the renewed mean-centering applied to the brief peri-saccadic epochs.

6.9. ICA decomposition

ICAs were computed on each of the resulting 80 training datasets per participant using EEGLAB's binary implementation of extended Infomax ICA (minimum change criterion: 10^{-7}). Infomax was chosen because it is widely used as the default option in EEGLAB and produces rather reliable decompositions (Groppe et al., 2009; Pontifex et al., 2016). The resulting ICA weights were then applied to the original, unfiltered recording. More precisely, source waveforms (S) for the original recording were computed by multiplying each of the 80 unmixing matrices (W) computed on the training data (this also includes the corresponding sphering matrices, stored separately in EEGLAB) with the matrix (X) containing the unfiltered recording (Fig. 1).

6.10. Eye tracker-guided component identification

The next step was to remove ocular ICs using the procedure by Plöchl et al. as implemented in EYE-EEG (function *pop_eyetrackerica.m*). I applied 10 thresholds between 0.6 and 1.5 (in steps of 0.1). To compute variance ratios, the saccade time window was defined as lasting from −10 ms before saccade onset until saccade offset. Conversely, fixations were defined as lasting from fixation onset until 10 ms before the saccade.[3] To obtain the corrected EEG (X_{corr}), only the activity of the neural sources was back-projected to the scalp ($X_{corr} = W^{-1}_{brain} * S_{brain}$).

6.11. Measuring correction quality

To quantify correction quality and visualize results, the corrected EEG was converted to average reference. I then extracted three measures of correction quality:

6.12. Undercorrection: CR artifact size

The size of residual CR artifacts was quantified by averaging all rightward saccades (tolerance of ±30°; see red shading in Fig. 2) and measuring the remaining EEG lateralization at frontal electrodes. Rightward saccades were selected because they occur frequently in both scene viewing and reading. However, for *Scenes*, comparable results were obtained for other saccade directions (Supplementary Fig. S4). Epochs were cut from −200 to +600 ms around saccade onsets, baseline-corrected from −50 to −10 ms (thereby excluding the SP from the baseline), and averaged. Residual artifact amplitude in these saccade-locked ERPs was then summarized as a single measure by subtracting the mean voltage at eight frontal left-hemispheric EEG and EOG electrodes (lateral EOG left, infraorbital EOG left, FP1, AF7, F7, F3, FT9, FC5) from that at their right-hemispheric counterparts (lateral EOG right, …, FC6) in the window +10 to +200 ms after saccade onset. Hemispheric lateralization within this window captures the CR artifact but excludes the SP, which was quantified separately.

A possible concern with this measure is that saccades are accompanied by genuine lateralized brain activity. However, most of the lateralized potentials related to attention shifts (Eimer, 2014) or oculomotor preparation (Becker et al., 1972; Berchicci et al., 2012; Everling et al., 1996; Moster and Goldberg, 1990) have been observed *before* rather than after the saccade and post-saccadic effects (Meyberg et al., 2015) have been observed at posterior sites. Thus, although some influence of neural sources on post-saccadic frontal lateralization cannot be ruled out, its influence was likely small in comparison to that of residual corneoretinal artifacts.

6.13. Undercorrection: SP size

The same rightward SRPs were used to measure residual SP amplitude. The SP is difficult to capture in a single measure, because it is biphasic and also has a diffuse parietal pole. As a suitable measure, I computed the average Global Field Power across the scalp (GFP, Lehmann and Skrandies, 1980) at the seven sampling points between −8 ms and +6 ms around saccade onset. Because the GFP of EEG data rarely drops to zero, the GFP around saccade onset was compared to that during a neutral and artifact-free reference interval before saccade onset (seven samples from −66 to −52 ms). The average GFP during this reference interval provides an estimate of the lowest possible GFP expected after a "perfect" SP correction. It is indicated by the horizontal dotted lines in Fig. 3B.

6.14. Overcorrection

Brain signal distortion.Overcorrection was quantified in intervals without significant oculomotor activity. As explained above, a suitable interval to find such intervals is immediately after the trial-initial onset of the scene/sentence-stimulus. Another set of epochs was therefore cut around stimulus onsets and baseline-corrected with a 100 ms pre-stimulus baseline. The eye-tracker was then used to identify a subset of these epochs in which no blink or eye movement was detected between −100 and + 200 ms after stimulus onset; this was the longest feasible interval that yielded some (virtually) eye movement-free epochs for each participant. Three criteria were used to identify these epochs: First, I excluded epochs in which a (micro)saccade was detected. Second, I removed epochs in which gaze position in any of the four eye-tracker channels (vertical and horizontal, left and right eye) changed by $> 0.2°$ between successive samples. Finally, to detect significant binocular drift, I removed epochs during which horizontal or vertical gaze position (averaged binocularly) varied by $> 0.5°$ (peak-to-peak) within the epoch.

These criteria identified $M = 38.0$ eye movement-free epochs per participant for *Scenes* (range: 5–66, SD = 21.1) and $M = 101.3$ (range: 32–129, SD = 26.3) for *Reading* which were then averaged to obtain a "clean" stimulus-ERP. To quantify overcorrection, this ERP was subtracted from the same ERP after ocular correction with each of the 800 ICA variants described above. The resulting difference waves (Fig. 7) therefore capture the distortion introduced by ICA and the average GFP of these difference waves across the 300 ms was used to quantify overcorrection.

[2] The length of the training data ($k = 80$ points per weight) was defined arbitrarily, but clearly exceeds minimum recommendations for mid-density EEG montages (e.g., $k = 20$ in Onton et al., 2006; $k = 30$ in Miyakoshi, 2019). Effects of data length on ICA reliability are explored in Groppe et al. (2009).

[3] The SP begins before saccade onset. A comparison of 20 different saccade window sizes (not reported here) showed that this definition (−10 ms before saccade until saccade offset) distinguished best between ocular and non-ocular ICs (i.e. maximized the distance between the two variance ratio distributions separated by a threshold of 1.1; cf. Fig. 8A).

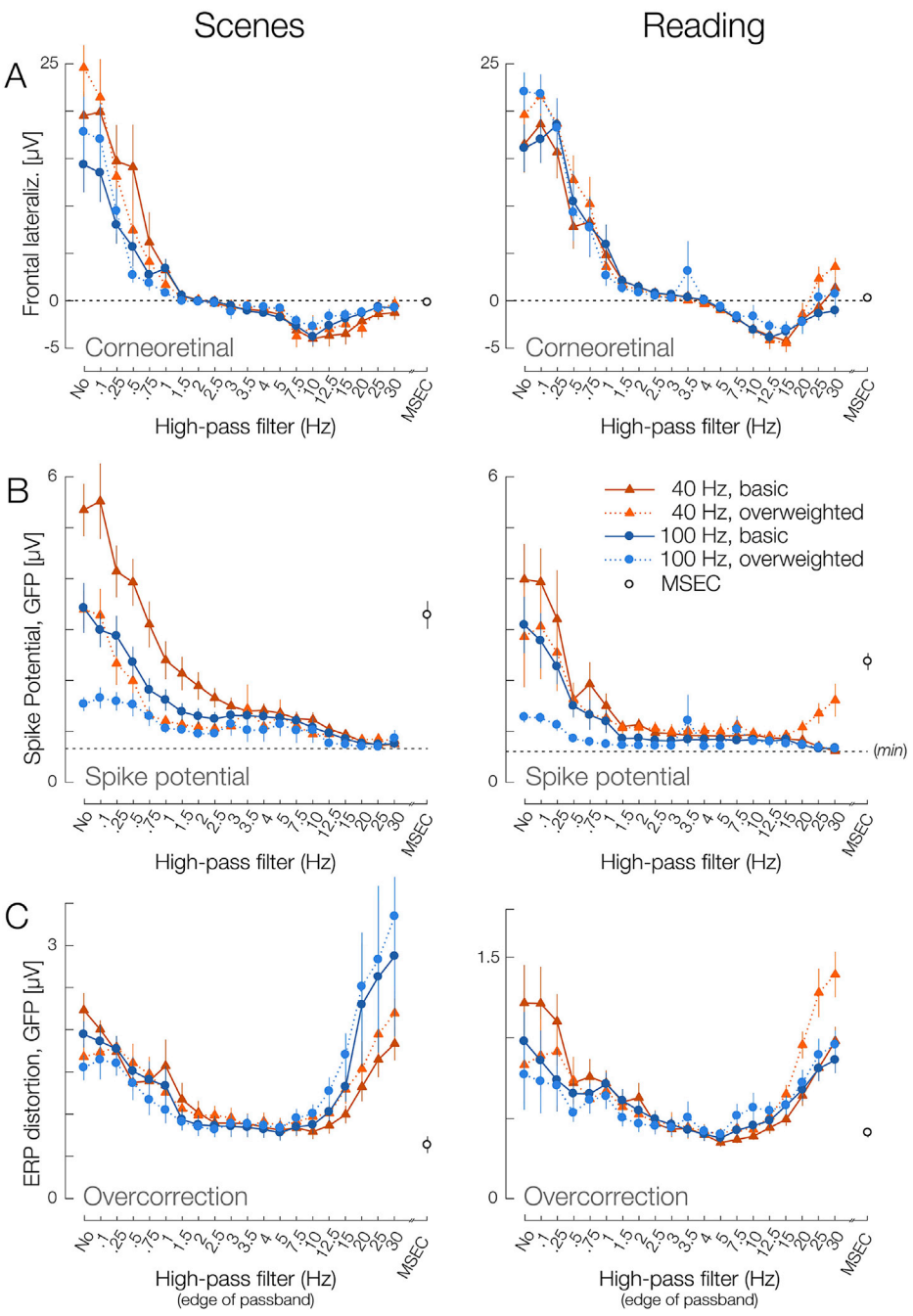

Fig. 3. Overview over main results for the Scenes (left) and Reading (right) experiment. Plots in the three rows depict the (A) size of the residual corneoretinal artifact, (B) size of the spike potential, and (C) distortion of eye movement-free intervals (overcorrection) after correction with ICA. Note that all measures of correction quality shown in this figure were computed on the original, unfiltered experimental data corrected with ICA. Results only differ in terms of the high-pass filter, low-pass filter, and overweighting applied to the training data. For this plot, ocular ICs were flagged using a variance ratio threshold of 1.1. Results for the MSEC algorithm are shown on the right of each plot (black open circles). In the panels in B, the horizontal dotted lines (min) mark the mean GFP during an artifact-free reference interval (i.e., the best possible correction). Error bars indicate ±1 SEM. The standard error for MSEC is too small to be visible in some plots.

6.5. Comparison to MSEC

Results were compared to those obtained with surrogate MSEC, as implemented in BESA (version 6.0, Besa GmbH, Gräfeling). Details on the algorithm and its implementation are provided in Supplement A.

6.16. Statistics

The three measures of correction quality were entered into separate mixed ANOVAs on the factors *Experiment* (2-level between-subject factor), *High-pass filter* (20), *Low-pass filter* (2), and *Overweighting* (2). For these analyses, the saccade/fixation variance ratio threshold was fixed at 1.1. ANOVAs were conducted using the "ez" package in *R* (Lawrence, 2013). To handle violations of the sphericity assumption, degrees of freedom were adjusted by multiplication with the *Greenhouser-Geisser epsilon*. Here I report the original degrees of freedom, the epsilon (ε), the adjusted *p*-values, and effect size as generalized eta-squared (ηG^2). Post-hoc paired *t*-test were used to assess the relationship between threshold and overcorrection and to compare MSEC with the best ICA solution.

6.17. Validation on new data

To test whether the specific best parameters determined in the present study (e.g. the most suitable high-pass filter) generalize to new data, all analyses described above were also repeated on new data. For *Reading*, analyses were re-run on the next twelve participants of the same experiment (mean age: 21.8 years, 7 female). For *Scenes,* analyses were repeated for the 9 participants of a follow-up experiment (mean age: 24.1 years, 7 female) who performed the same search task with different

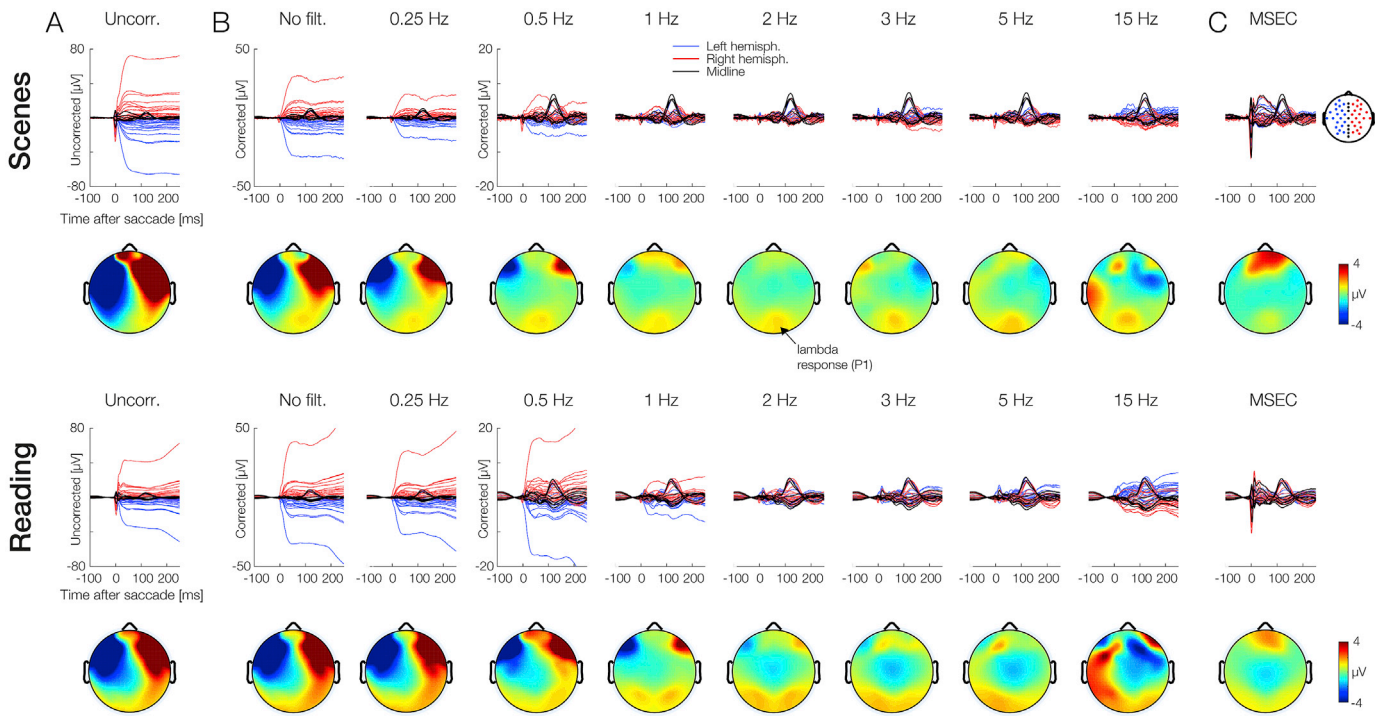

Fig. 4. Effect of high-pass filtering the training data on ICA correction of the original, unfiltered data. Plotted is the grand-average saccade-locked ERP (SRP) for rightward saccades in the Scenes (upper half) and Reading (lower half) experiment. Time zero marks saccade onset. (A) The leftmost column shows the SRP without ocular correction. Strong lateralized CR artifacts are evident. Maps depict the corresponding scalp distribution in the interval 10–200 ms after saccade onset. The positive deflection over midoccipital sites reflects the visual lambda response. Note that all plots also include EOG electrodes, which show the strongest residual artifacts. (B) Same data corrected with ICA. Best results were obtained with the passband edge set to about 2 Hz. Note the different y-axis scaling between plots. (C) Same data, corrected with MSEC. (Note: results in this figure are based on training data low-pass filtered at 100 Hz, with overweighted SPs). Frontal views of all topographies shown in this figure are provided in Supplementary Fig. S3.

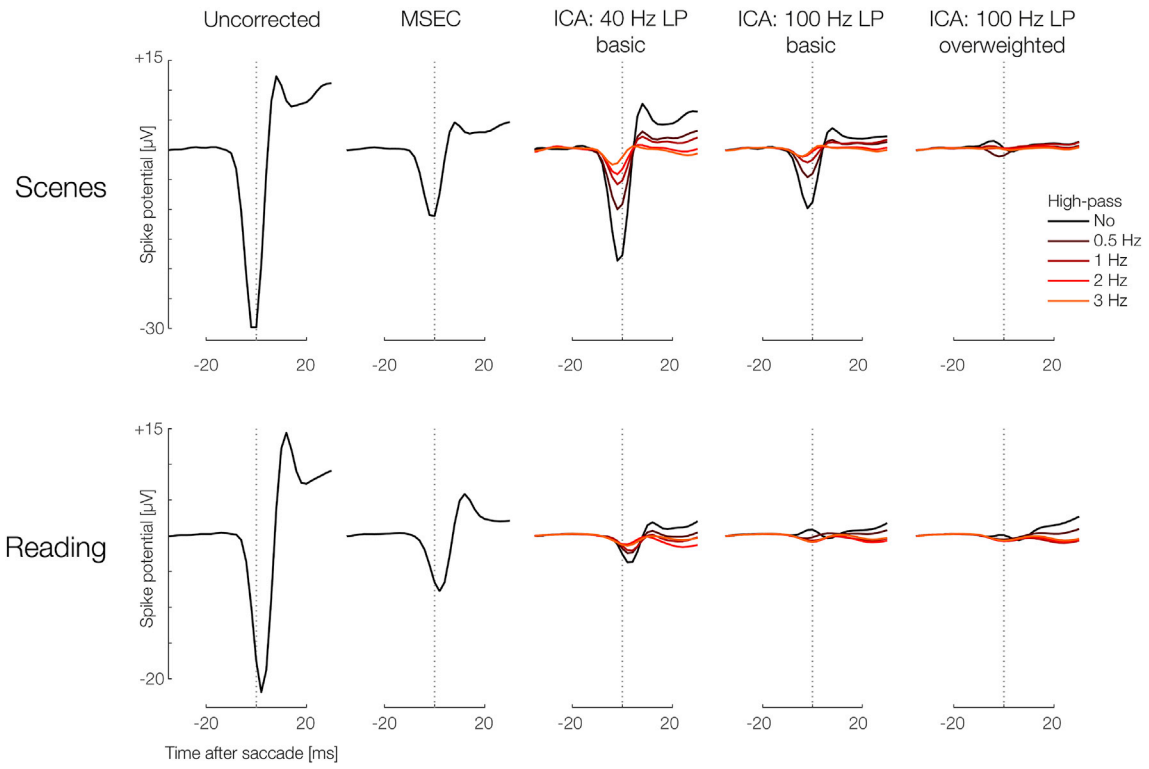

Fig. 5. Spike potential correction as a function of high-pass filter, low-pass filter, and overweighting. Waveforms show the SRP for rightward saccades at a radial EOG channel (mean of all facial EOG electrodes minus electrode POz). Note that the SP is almost entirely removed if the three parameters are combined in an optimal way (rightmost panels). In contrast, MSEC failed to fully correct the SP.

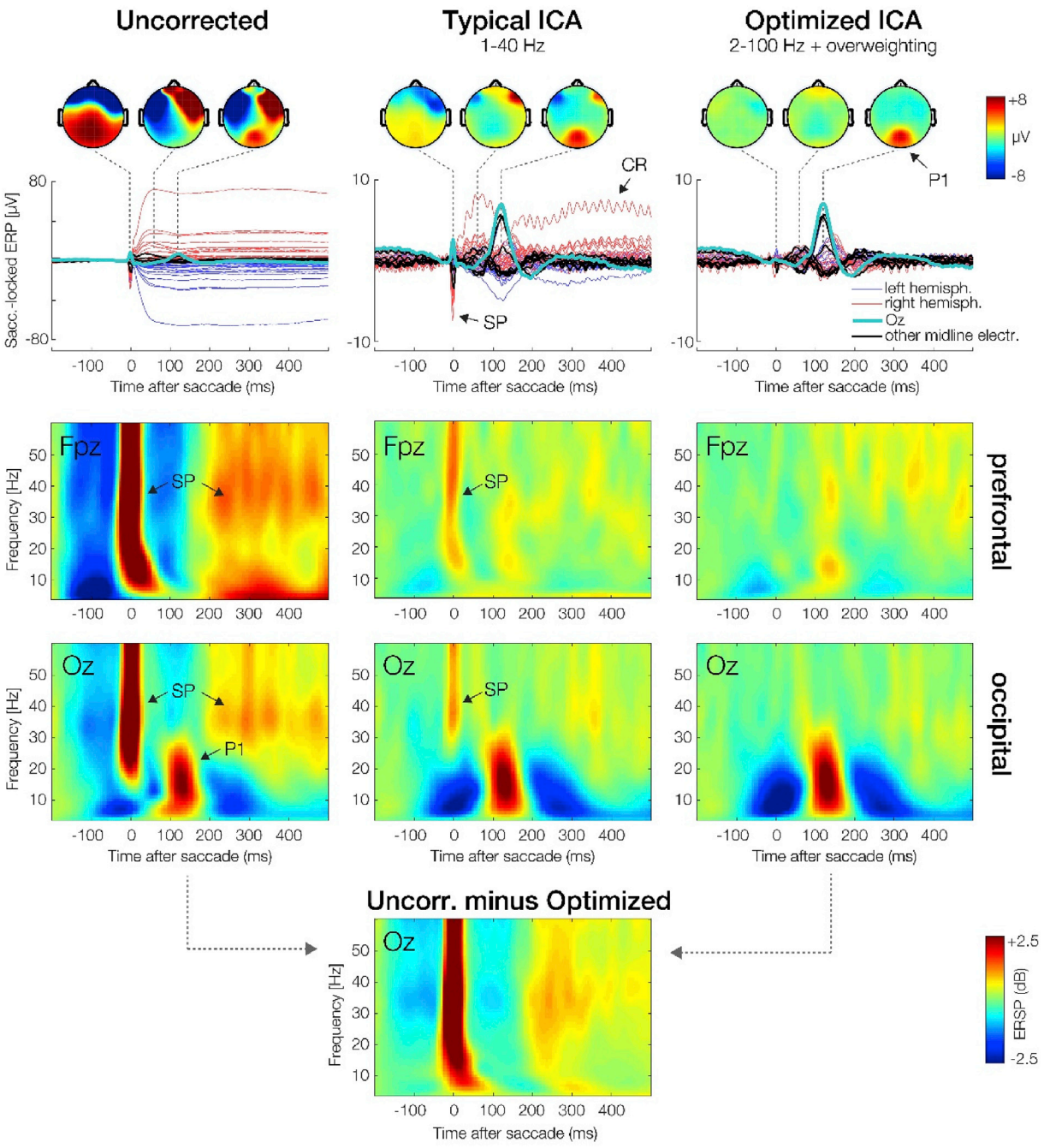

Figure 6. Correction with a "typical" versus "optimized" ICA. In all panels, time zero indicates the onset of rightward saccades during scene viewing. Left panels: Grand-average saccade-related activity without correction. Middle panels: Same data, corrected with a typically trained ICA (1–40 Hz bandwidth, no overweighting). Right panels: Same data, corrected with a well-trained ICA (2–100 Hz, with overweighting). Upper panels plot the grand-average saccade-locked ERPs with occipital electrode Oz highlighted by the turquoise line. Maps show the scalp distribution at −2 ms (SP artifact), 60 ms (CR artifact) and 120 ms (P1 brain potential), respectively. Lower panels present the same data at prefrontal electrode Fpz and occipital electrode Oz after a wavelet transform (ERSP). SP artifacts from the current saccade (around time zero) and the next saccade (beginning around 200 ms) are evident. Whereas the "typically"-trained ICA left some artifacts in the data, the optimized ICA removed them better. The bottom panel shows the pure artifact in the frequency domain at electrode Oz, i.e., the difference between the spectrograms for uncorrected data minus corrected data.

greyscale images. Everything else was identical.

7. Results

Below, I first present the results for CR and SP correction and for overcorrection. These are based on the removal of ICs with the saccade/fixation variance threshold fixed at 1.1, which proved to be suitable. Afterwards, I describe the effects of different thresholds. Finally, ICA is compared to MSEC.

7.1. Correction of CR artifacts

Fig. 3A summarizes the impact of high-pass filtering, low-pass filtering, and overweighting on the removal of CR artifacts. As evident from this figure, correction quality in both experiments differed drastically as a function of filtering. Results for CR artifacts are also depicted in

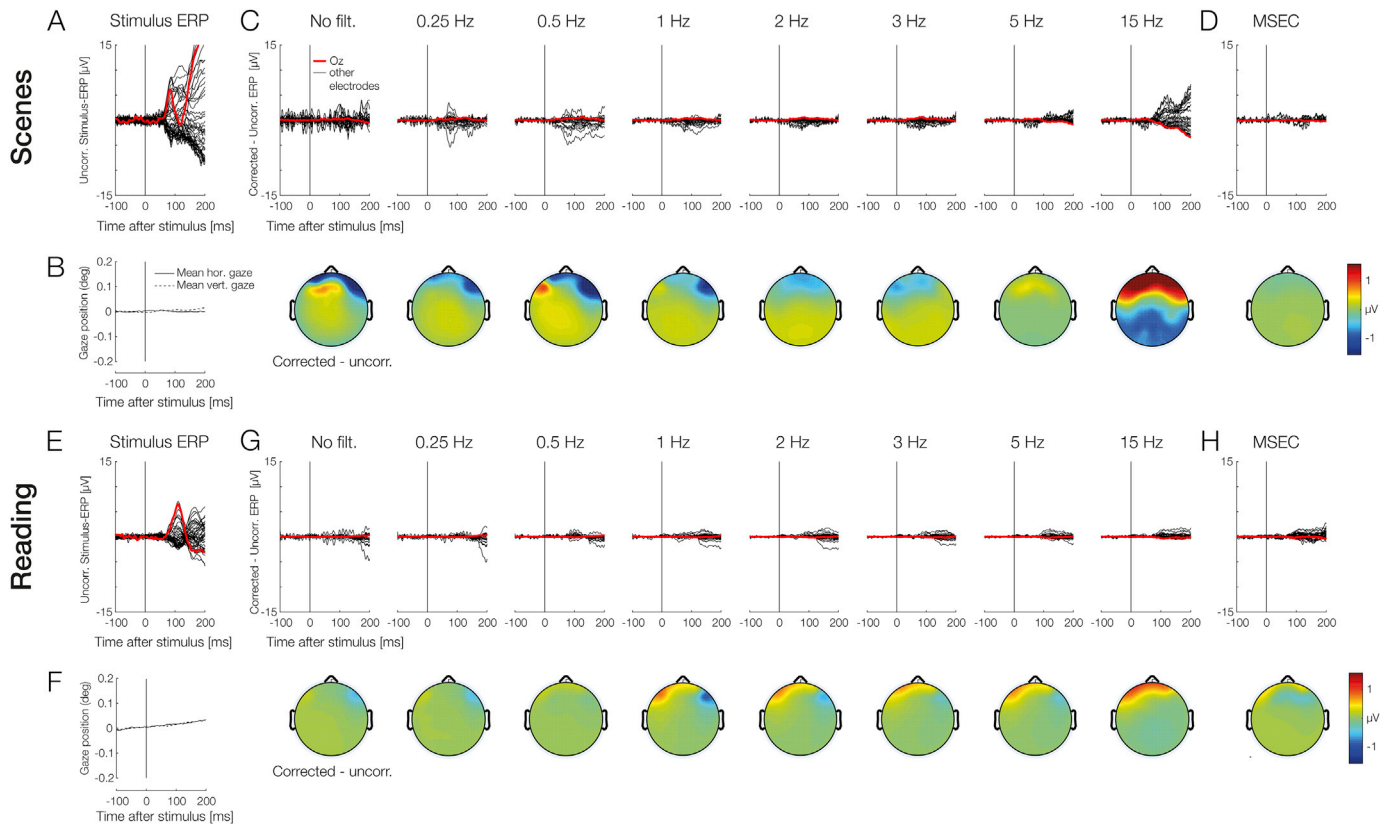

Fig. 7. Effect of high-pass filtering on overcorrection. Shown is the stimulus-locked ERP for (virtually) eye movement-free epochs. **(A, E)** Grand-average ERPs in the *Scene* (A) and *Reading* (E) experiment. Time zero marks image/sentence onset, respectively. **(B, F)** Averaged horizontal and vertical eye position. **(C, G)** Difference waves of the stimulus-ERP after ocular correction minus that without any correction for ICAs trained on differently filtered data. Scalp maps depict the corresponding difference topography between 0 and 200 ms. In many cases, ICA introduced overcorrection at frontal electrodes. (*Note*: ICAs shown here were computed on training data with overweighted SPs, low-pass filtered at 100 Hz). **(D, H)** Results for MSEC.

detail in Fig. 4, which shows the grand-average SRP locked to rightward saccades. Panel A in Fig. 4 displays the raw SRP without ocular correction; here, strong CR artifacts of up to ±70 µV are evident as differences between left- (blue lines) and right-hemisphere (red lines) channels. Black lines mark sagittal midline electrodes that are less affected by lateral saccades. At these channels, the visual lambda responses (P1 component) can be seen peaking over occipital cortex about 110 ms after saccade onset.

Fig. 4B shows the same data after correction with differently trained ICAs. Note that all averages shown here were extracted from the original, *unfiltered* data corrected with ICA; the results only differ in terms of the preprocessing of the training data. Although ICA reduced artifacts at all filter settings, visual inspection suggests that residual artifact size depended strongly on the high-pass filter, with filters <1.5 Hz producing suboptimal correction. Numerically, best results were obtained with passband edges between 2 and 2.5 Hz for *Scenes*, and 3–4 Hz for *Reading*. Interestingly, very strong low-pass filtering (>5 Hz) tended to invert the topography of residual CR artifacts in the corrected data, that is, right-hemisphere channels became more *negative* than left-hemisphere channels (see values below zero in Fig. 3A; see reversed lateralization in Fig. 4B). In other words, overly aggressive filtering distorted the data, whereas filtering at < 1.5 Hz left CR artifacts in the data. This visual impression was confirmed by a significant effect of *high-pass filter* on CR artifact amplitude, $F(19,418) = 82.917$, $\varepsilon = 0.155$, $p < 0.0001$, $\eta G^2 = 0.638$, which did not interact with *Experiment*. Factors *low-pass filter* and *overweighting* did not affect CR correction.

7.2. Correction of spike potential

Fig. 3B summarizes results for the SP. Waveforms are depicted in Fig. 5. High-pass filtering also strongly influenced SP correction. With increasingly stricter high-pass filtering, residual SP amplitude decreased in an almost monotonous fashion (main effect of *high-pass filter* on residual SP amplitude, $F(19,418) = 50.940$, $\varepsilon = 0.103$, $p < 0.0001$, $\eta G^2 = 0.380$). Importantly, however, the other two parameters also strongly modulated the SP: First, correction was improved by *overweighting*, but as Fig. 3B shows, this benefit was only present in datasets that were only moderately high-pass filtered (at frequencies of about 3 Hz or less), resulting in a significant *high-pass filter* × *overweighing* interaction, $F(19,418) = 10.943$, $\varepsilon = 0.124$, $p < 0.0001$, $\eta G^2 = 0.086$. Second, a similar pattern was seen for the *low-pass filter* (passband edge 40 vs. 100 Hz). Leaving the filter open up to 100 Hz clearly improved SP correction, but again only if high-pass filtering was moderate (interaction *high-pass filter* × *low-pass filter*, $F(19,418) = 9.799$, $\varepsilon = 0.111$, $p < 0.001$, $\eta G^2 = 0.068$).

Fig. 5 shows that if all three parameters were combined in a near-optimal manner, that is, high-pass filtering at 2 Hz, overweighting of SPs, and no additional low-pass filtering, SP artifacts could be almost entirely removed from both the *Scenes* and *Reading* data. Fig. 6 also shows the *Scenes* data in the frequency domain, that is, after saccade-locked epochs (cut from −600 to 1000 ms for this analysis) were individually wavelet-transformed using EEGLABs *newtimef.m* function, yielding event-related spectral pertubation (ERSP) values relative to a pre-saccadic (−200 to −100 ms) spectral baseline. Whereas a more "typically" trained ICA (bandwidth 1–40 Hz, no overweighting) left some spectral artifacts in the data, the near-optimally trained ICA removed the broadband artifact in the beta and gamma band (Craddock et al., 2016; Hassler et al., 2011; Hipp and Siegel, 2013; Yuval-Greenberg et al., 2008); not only from the current saccade (at time zero, see Fig. 6), but also from the following saccades on the stimulus.

7.3. Brain signal distortion (overcorrection)

Next, I tested whether ICA distorts neurogenic activity during stimulus-locked intervals (virtually) free of oculomotor activity. Aggregated results are again shown in Fig. 3C. Exemplary waveforms are provided in Fig. 7 (the data shown is based on ICAs trained on data low-pass filtered at 100 Hz with overweighted SPs, since these settings provided the best SP correction, see above). Panels B and F in Fig. 7 show the binocular gaze position in these clean epochs, measured by the eye-tracker. In both experiments, grand-average gaze position changed by less than 0.05° across the 300 ms intervals. Panels A and D in Fig. 7 depict the corresponding stimulus-ERPs for these clean epochs without correction. As expected, the trial-initial stimulation elicited a visual P1/N1 complex in both experiments. The following panels C and G only show the *difference waves* (at all channels) between this artifact-free ERP after ICA correction minus the same ERP without any ocular correction. In other words, these panels show the distortions introduced by ICA.

A first interesting finding is that all ICA variants produced some changes in eye movement-free intervals, regardless of filter settings. Crucially, however, overcorrection depended again strongly on details of the preprocessing of the training data, as confirmed by a main effect of *high-pass filter*, $F(19,418) = 13.213$, $\varepsilon = 0.098$, $p < 0.0001$, $\eta G^2 = 0.208$. More precisely, overcorrection and high-pass filtering displayed a U-shaped relationship such that with stricter filtering, distortions of the stimulus-ERP first decreased, then reached a plateau (from about 2 to 5 Hz), and then rebounded again. This effect of *high-pass filtering* also interacted with that of *overweighting*, $F(19,418) = 8.627$, $\varepsilon = 0.280$, $p < 0.0001$, $\eta G^2 = 0.015$.

Maps in Fig. 7 show the scalp distribution of overcorrection for some ICA solutions. Distortions occurred mostly at frontal channels and their topographies closely resembled those of typical CR and/or SP artifacts in most cases, suggesting that the activity time courses of some of the ocular ICs removed during the correction were non-zero during eye movement-free intervals.

7.4. Validation on new data

To validate the suitable values for high-pass filtering, low-pass filtering and overweighting, all main analyses were repeated on new data from different participants. Results are shown in Supplementary Fig. S5. As can be seen, the overall pattern of results and the numerically best parameter combinations largely generalized to new data. The most suitable high-pass filter for the correction of CR artifacts was again numerically higher for *Reading* than for *Scenes*.

7.5. Comparison to MSEC

Figs. 3–5 and 7 also include the corresponding results for MSEC. To assess its performance, MSEC was compared to a near-optimal ICA variant (2–100 Hz, with overweighting) using paired t-tests. In terms of CR artifact removal, there was no evidence that MSEC performed significantly better or worse than this optimized ICA ($p = 0.863$ for *Scenes*, $p = 0.123$ for *Reading*). However, MSEC produced less overcorrection than the optimized ICA variant for *Scenes*, $t(11) = 2.41$, $p < 0.05$, although the actual numerical difference between the methods was small (difference: −0.21 μV, 95% confidence interval of difference: −0.41 to 0.02 μV). For *Reading*, overcorrection did not differ between methods ($p = 0.274$). Importantly, for both *Scenes* and *Reading*, MSEC failed to remove the SP, which was only reduced to about a third of its original amplitude in the radial EOG (see Fig. 5). Residual SP artifacts therefore remained significantly larger with MSEC than optimized ICA. This difference in SP correction performance was highly significant for *Scenes* (difference between methods: 2.33 μV, 95% CI: 1.74–2.93 μV, $t(11) = 8.65$, $p < 0.0001$) and *Reading* (difference between methods: 1.65 μV, 95% CI: 1.27–2.03 μV, $t(11) = 9.6$, $p < 0.0001$).

7.6. Choice of threshold

All analyses above were based on a variance ratio threshold of 1.1 for the identification of ocular ICs (Plöchl et al., 2012). At this threshold, the criterion removed on average 5.3 components for *Scenes* (SD = 1.37) and 4.0 components for *Reading* (SD = 0.74), a difference that is likely explained by the absence of vertical saccades in reading. Fig. 8 summarizes effects of changing the threshold. For this analysis, the other parameters were fixed at near-optimal values (bandwidth 2–100 Hz, with overweighting). Ideally, we would want a bimodal distribution of variance ratios that clearly distinguishes between ocular and non-ocular ICs. As Fig. 8A shows, the observed distribution was indeed fairly bimodal, with relatively few ICs located in a "gray area" between 1.0 and 1.4. As expected, decreasing the threshold increased the number of rejected ICs (Fig. 8B). Overweighting had no effect on the number of removed ICs at any threshold (Fig. 8C).

Fig. 8D shows the relationship between threshold and overcorrection. Whereas overcorrection was not significantly different for thresholds between 1.5 and 1.1, it increased once the threshold was lowered further. Specifically, for *Scenes*, overcorrection started to increase (relative to a lenient threshold of 1.5) once the threshold was lowered to 1.0 or less (comparison threshold 1.0 vs. 1.5; $t(11) = 2.79$, $p < 0.05$). For *Reading*, a statistically significant increase happened at thresholds of 0.9 or less, $t(11) = 2.90$, $p < 0.05$; however, as the right panel of Fig. 8D shows, a numerical increase was already seen at 1.0. Together, these results indicate that 1.1 is indeed the lowest suitable threshold that avoids overcorrection, while also almost fully removing ocular artifacts (given optimized training data, see above).

8. Discussion

Human behavior is characterized by frequent saccades of varying amplitude and direction. Although independent component analysis (ICA) has been previously used to reduce the strong ocular EEG artifacts generated during natural vision (e.g. during visual search, reading, face perception, or whole-body motion), little is known about how to adapt and improve the procedures for this purpose. Here I used simultaneous eye-tracking to explore the parameter space for data preprocessing, to select ocular ICA components, and to objectively validate the outcomes of the correction. Specifically, the goal was to combine the concepts of optimized filtering (Winkler et al., 2015), artifact overweighting (Keren et al., 2010), eye tracker-guided component classification (Plöchl et al., 2012), and eye tracker-based quality control (Dimigen et al., 2011) into one coherent pipeline. Results show that with common preprocessing parameters, ICA left substantial artifacts in the data and also distorted neural activity during eye movement-free intervals. However, with optimized procedures, Infomax ICA removed artifacts almost entirely from natural scene viewing and reading with little overcorrection. Below, I first discuss the individual parameters and then summarize specific recommendations.

8.1. High-pass filter

Adverse effects of slow signals on unmixing quality are known among ICA practitioners (see also Miyakoshi, 2019) and many filter their input data at cutoffs between 0.1 and 1 Hz (e.g. Nieuwland et al., 2018). Similar settings have also been used in SRP/FRP research (e.g. 0.5 Hz; as used in Henderson et al., 2013; Nikolaev et al., 2011; and Ossandon et al., 2010). However, filter effects have rarely been formally investigated (Winkler et al., 2015) and not in tasks with multiple eye movements, where filtering was hypothesized to be crucial. Results confirm this hypothesis. Whereas the impact of suboptimal filtering on the resulting ERP waveforms was comparatively mild in the oddball task investigated by Winkler and colleagues (cf. their Fig. 5), effects were strong for both scene viewing and reading. Without filtering of the training data (i.e. 10 s time constant), ICA produced strong distortions of eye movement-free

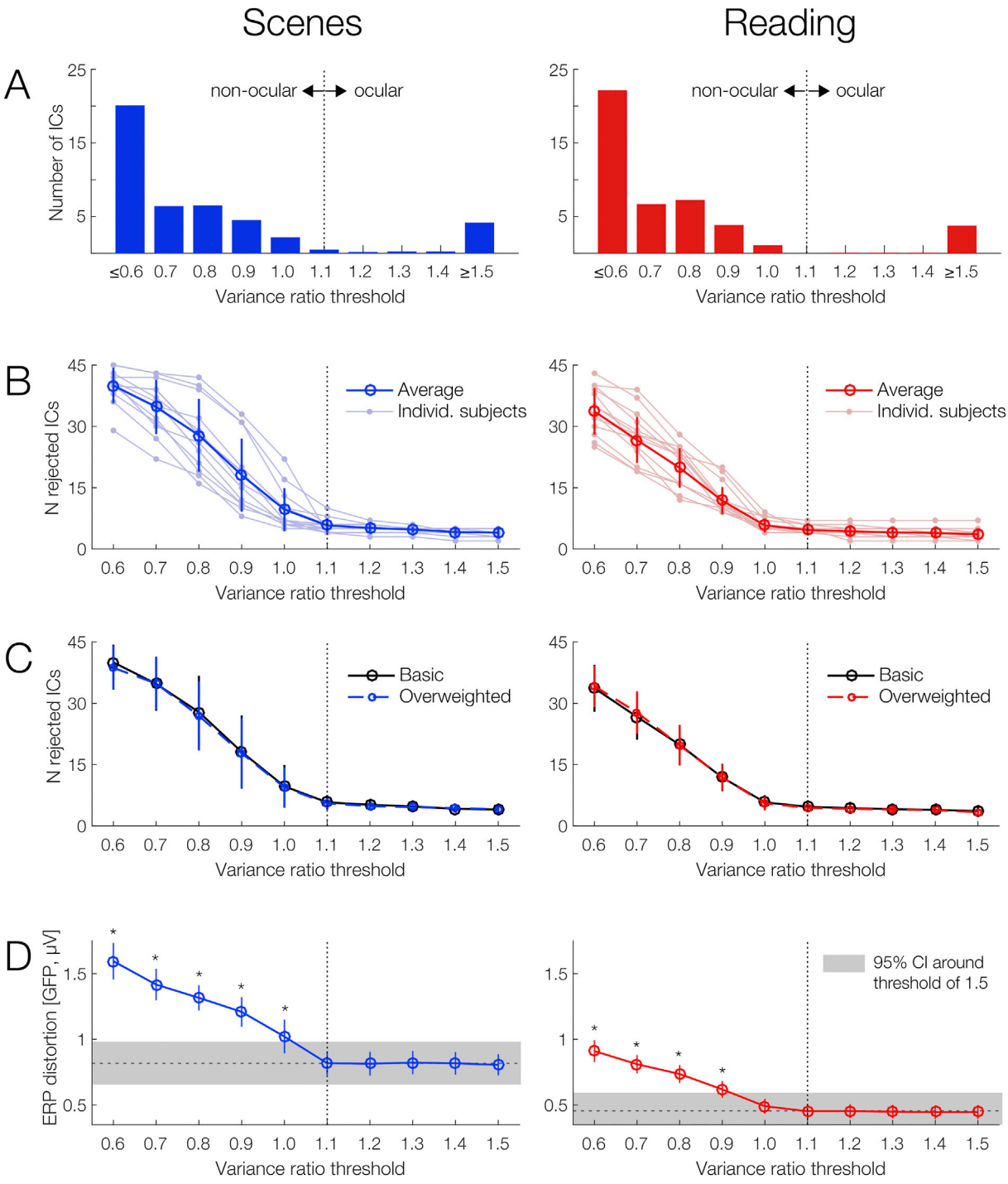

Fig. 8. Selecting a threshold for automatic component classification (Plöchl et al., 2012). (A) Grand-average distribution of variance ratios for all 45 ICs per participant, shown here for a near-optimal ICA variant (2–100 Hz, with overweighting). The threshold of 1.1, used in all analyses above is highlighted by the vertical line. For both Scenes (left) and Reading (right), there was a rather clear separation between ICs with a low vs. high variance ratio. (B) Number of ICs rejected as "ocular" as a function of threshold. (C) Overweighting SPs in the training data did not influence the number of rejected ICs. (D) Effect of threshold on overcorrection, with y-axis indicating the distortion of artifact-free epochs. Gray shading marks the 95% confidence interval (CI) around the threshold of 1.5. Compared to this lenient threshold, overcorrection increased significantly ($p < 0.05$, asterisks) once the threshold was lowered beyond 1.1 (for Scenes) or 1.0 (for Reading), indicating that 1.1 is a suitable threshold to avoid overcorrection (while ensuring good CR and SP correction).

intervals and left large CR artifacts in the saccade-related potentials; however, even if the edge of the passband was raised to 1 Hz (corresponding to a half-amplitude cutoff of 0.5 Hz), correction was still suboptimal. Instead, filtering at a passband edge of 2 or 2.5 Hz (corresponding to a cutoff of 1 and 1.5 Hz, respectively) produced the smallest residual CR artifacts and the least overcorrection for scene viewing. These results are in line with those of Winkler et al. (2015). For *Reading*, the numerically best results were obtained with even stronger filtering (passband edge settings as high as 4 Hz), but filtering at 2.5 Hz still produced good results. Although the interaction of *High-pass filter* with *Experiment* did not reach significance, the same pattern was observed when the analyses were repeated on a new data

(Supplementary Fig. S5). This tendency for a higher optimal cutoff in reading is likely explained by the stronger summation of CR artifacts from repeated rightward saccades during reading. It indicates that aggressive filtering may be especially important in paradigms with an assymetric distribution of saccade orientations. If the edge of the passband was raised much higher, beyond about 7.5 Hz, residual CR artifacts and overcorrection increased again markedly in both paradigms.

Importantly, it should be noted that the most suitable cutoff frequency also depends on the characteristics of the filter used (Widmann et al., 2015), in particular its steepness (transition bandwidth). In the present study, all filtering was done with EEGLAB's FIR filter with its default, frequency-dependent transition bandwidth setting (Supplementary Fig. S2). Alternatively, users could manually define a steeper filter. For example, in a control analysis with a much steeper filter setting, decent correction results for *Scenes* were already obtained with the passband edge set to 1 Hz and a cutoff at 0.75 Hz (for details see Supplementary Fig. S7).

8.2. Low-pass filter

Another parameter of interest was the low-pass filter which is often set to cutoffs around 30–40 Hz for traditional ERP analyses. Results show that SP correction was significantly better if filters were instead kept open to 100 Hz (Fig. 5), even though this meant that the input data contained more line noise and scalp EMG. The likely reason for this benefit is that the spectrum of the SP extends well beyond 40 Hz (see Fig. 6), meaning that this artifact is larger – and presumably better modeled – without low-pass filtering. Of course, without low-pass filtering, a large proportion of the produced ICs may reflect scalp EMG. This is not a problem if the goal is to remove ocular ICs and to back-project the data. If the experimenter plans to work with the neural ICs in source space instead (and wants to avoid problems associated with re-running ICA on rank-reduced data after removing EMG sources; Artoni et al., 2018), it may be better to keep the low-pass filter and only optimize the other two parameters (high-pass and overweighting), since this may yield more neural components (at the cost of a slightly less clearly modeled SP).

8.3. Overweighting spike potentials

In combination with suitable filtering, spike potential overweighing strongly improved SP correction and made it possible to almost fully remove this artifact – and its associated beta and gamma band distortions – from both paradigms (Figs. 5 and 6). These findings are consistent with the beneficial effects of overweighting for suppressing gamma-band artifacts from much smaller microsaccades in steady-fixation experiments (Keren et al., 2010; Craddock et al., 2016). Overweighting also did not influence the number of removed ICs, suggesting that it produced qualitatively better SP components. I also observed that overweighting was only effective if the brief peri-saccadic epochs taken from the basic training data were mean-centered again across their 30 ms duration before appending them, most likely because this further emphasizes signal variance in high frequency bands. In line with this observation, results in Fig. 3B show that both overweighting and low-pass filtering became irrelevant if the training data was already high-pass filtered very aggressively (e.g. at 10 Hz), presumably because in this data, the SP already accounts for sufficient variance. Importantly, however, such extreme filtering produced bad CR correction and strong overcorrection. The tradeoff to remove *both* the SP and the CR artifact is therefore to combine moderate high-pass filtering (here: passband edge set to 2–2.5 Hz for scenes) with overweighting and no low-pass filtering. All three measures increase the proportion of signal variance due to spike potential artifacts.

In the current study, the overweighting of peri-saccadic samples doubled the length of the training data. Since a longer training dataset increases ICA computation time, this raises the question of how many samples need to be added. In a supplementary analysis (Supplementary Fig. S6), the length of the appended samples was varied between 0% and 400% of the original training data length. In this analysis, SP correction generally improved with increasing overweighing proportions, but this was especially the case if the high-pass filter was chosen poorly (e.g. 0.1 Hz).

8.4. Overcorrection & threshold choice

One question motivating the current study was to what degree ICA distorts neurogenic signals. Results show that all tested ICA variants modified the EEG during intervals that were free of significant oculomotor activity. In scene viewing, ICA also produced stronger distortions that MSEC, which was explicitly designed to reduce overcorrection (Berg and Scherg, 1994). Interestingly, distortions mostly affected frontal sites and their scalp distributions resembled those of SP and CR artifacts (Fig. 7). Given this topographical resemblance, one might suspect that these changes in the stimulus-ERP were not really caused by "overcorrection", but simply reflect the removal of tiny, overlooked artifacts elicited by undetected microsaccades (Meyberg et al., 2017; Plöchl et al., 2012) or other fixational eye movements (Rolfs, 2009).[4] This seems unlikely for at least two reasons: First, it is unlikely that the sensitive detection algorithm (Engbert and Kliegl, 2003) missed a relevant number of microsaccades and significant binocular drift was also eliminated. Second, and more importantly, if this was the case, overcorrection should have been largest at filter settings that were also the most effective at removing artifacts. However, the opposite was true: "Bad" ICA solutions – those that failed to suppress all CR and SP artifacts – generated the *strongest* changes in eye movement-free intervals, whereas "good" ICA solutions – those that effectively removed ocular artifacts (see Fig. 3A and B) – produced the *least* overcorrection (Fig. 3C). The more likely explanation is therefore that with a suboptimal unmixing, some neural activity or noise was modeled in the activity time courses of the ocular ICs that were later removed from the data, thereby also affecting eye movement-free intervals.[5]

As expected, the distortion of stimulus-ERPs increased if the variance ratio threshold was set too low; that is, once it was lowered to 1.0 (for *Scenes*) or 0.9 (for *Reading*). Thus, a threshold of 1.1, as initially proposed by Plöchl et al. 2012 (see also Ries et al., 2018 for converging evidence) appears to be suitable.

8.5. Comparison to MSEC

Surrogate MSEC (Berg and Scherg, 1994) was used as a benchmark for Infomax ICA. In terms of CR correction and overcorrection, this alternative method performed as well or better than the best ICA solutions obtained here. Crucially, however, at least with its default procedures (Scherg, 2013; see Supplement A), MSEC failed to remove the SP. MSEC also has practical drawbacks: First, it requires the experimenter to record 5–10 min of isolated eye movements from each participant before or after the experiment. Second, to my knowledge, the method is

[4] Even during microsaccade-free intervals, the eyes are never entirely motionless but show slow (<0.5°/s, Rolfs, 2009) conjugate or non-conjugate *drift* movements of limited spatial extent (resembling a random walk) as well as a high-frequency *microtremor* of extremely small (~0.1–0.5 min-arc) amplitude. Significant conjugate drift was eliminated by the criteria applied here. Regarding microtremor, it has been putatively suggested by Onton and Makeig (2009) that it produces a tiny, high-frequency (>40 Hz) oscillation in peri-ocular EEG channels; however, such a small frontal oscillation would not have affected the overcorrection measure used here in a significant manner.

[5] The distortion of stimulus onset-ERPs occurring during artifact-free intervals was not larger if the ICA was trained on filtered data (e.g. at 2 Hz) as compared to unfiltered data. This indicates that it is permissible to transfer ICA weights trained on filtered data to the unfiltered version of the same data (e.g. Debener et al., 2010).

currently only implemented in proprietary software (see Berg, 2003 for an outdated open implementation). Finally, it provides a less flexible framework than ICA for removing non-ocular artifacts. Nevertheless, MSEC appears to be a viable alternative if the focus is on removing CR artifacts with little brain-signal distortion.

8.6. Other challenges when analyzing multi-saccadic EEG

Artifacts are only one of four challenges when analyzing EEG experiments with multiple saccades (Dimigen et al., 2011). The other problems relate to (1) the integration of eye-tracking and EEG data, (2) the temporally varying overlap between the neural responses generated by successive fixations, and (3) the complex influences of visual and oculomotor low-level variables (such as saccade size) on the morphology of the post-saccadic lambda waves. The first of these problem can now be solved with dedicated toolboxes (such as EYE-EEG, see also Baekgaard et al., 2014; Xue et al., 2017), whereas the latter two are effectively addressed by analyzing the artifact-corrected EEG with regression-based linear deconvolution models (Burns et al., 2013; Cornelissen et al., 2019; Dandekar et al., 2011; Dimigen and Ehinger, 2019; Ehinger & Dimigen, 2019; Kristensen et al., 2017; Smith and Kutas, 2015). If ocular correction is also optimized, there are now viable solutions to all four problems.

9. Conclusions & recommendations

Results motivate the following practical recommendations: First, ICAs of free viewing data should be trained on high-pass filtered data. Using EEGLAB's standard FIR filter with its default settings, best results were obtained at passband edges of around 2–2.5 Hz (corresponding to cutoffs of 1–1.5 Hz). For reading, optimal values tended to be higher. Second, frequencies higher than about 40 Hz should remain in the training data, since this improves SP correction. Third, correction is further improved if SPs are overweighted. Fourth, a threshold of 1.1 is suitable for component identification (in combination with a saccade window that begins −10 ms before saccade onset), with lower thresholds leading to overcorrection.

Importantly, Infomax ICA trained in this manner removed ocular artifacts almost fully with little distortion of neural activity and no need for any subjective classifications by the experimenter. Since the suitable parameters were overall similar for scene viewing and reading, they will likely also generalize to other free viewing taks, e.g. in virtual reality or during mobile brain/body imaging (Gwin et al., 2010). The fact that it is feasible to remove the spike potential artifact under free viewing conditions should also facilitate future studies of the attentional, oculomotor, and visual processes occurring around the time of saccade onset.

Finally, it should be noted that the explored parameters only cover some of the choices when designing the ICA pipeline. Parameters not investigated include the specific algorithm used (e.g. Infomax vs. AMICA), the number and locations of EEG channels in the montage, the overall number of data points submitted to ICA, or the question whether and how the variance ratio criterion should be combined with other flagging methods (Chaumon et al., 2015). The objective eye tracker-based quality measures proposed here may help to quantify the role of these other parameters in the future.

9.1. Implementation

A function to create training data with overweighted artifacts was added to the EYE-EEG extension for EEGLAB and a simple Matlab script implementing the current procedures is found at www.github.com/olafdimigen/opticat.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116117.

References

Ai, G., Sato, N., Singh, B., Wagatsuma, H., 2016. Direction and viewing area-sensitive influence of EOG artifacts revealed in the EEG topographic pattern analysis. Cognitive neurodynamics 10 (4), 301–314.

Armington, J.C., 1978. Potentials that precede small saccades. In: Armington, J.C. (Ed.), Visual Psychophysics and Physiology. Academic Press, New York, pp. 363–372.

Artoni, F., Delorme, A., Makeig, S., 2018. Applying dimension reduction to EEG data by Principal Component Analysis reduces the quality of its subsequent Independent Component decomposition. Neuroimage 175, 176–187.

Baccino, T., 2011. Eye movements and concurrent event-related potentials: eye fixation-related potential investigations in reading. In: Liversedge, S.P., Gilchrist, I., Everling, S. (Eds.), Oxford Handbook on Eye Movements. Oxford University Press, pp. 857–870.

Baekgaard, P., Petersen, M.K., Larsen, J.E., 2014. In the twinkling of an eye: synchronization of EEG and eye tracking based on blink signatures. In: Paper Presented at the 2014 4th International Workshop on Cognitive Information Processing.

Balaban, C.D., Weinstein, J.M., 1985. The human pre-saccadic spike potential: influences of a visual target, saccade direction, electrode laterality and instructions to perform saccades. Brain Res. 347 (1), 49–57.

Barry, W., Jones, G.M., 1965. Influence of eyelid movement upon electro-oculographic recording of vertical eye movements. Aero. Med. 36, 855–858.

Becker, W., Hoehne, O., Iwase, K., Kornhuber, H.H., 1972. Bereitschaftspotential, prämotorische Positivierung und andere Hirnpotentiale bei sakkadischen Augenbewegungen. Vis. Res. 12, 421–436.

Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7 (6), 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129.

Berchicci, M., Stella, A., Pitzalis, S., Spinelli, D., Di Russo, F., 2012. Spatio-temporal mapping of motor preparation for self-paced saccades. Biol. Psychol. 90 (1), 10–17. https://doi.org/10.1016/j.biopsycho.2012.02.014.

Berg, P., 2003. Konstanz Raw Data Format Handbook - EEG and Psychophysiology Data Analysis Programs.

Berg, P., Scherg, M., 1991. Dipole models of eye movements and blinks. Electroencephalogr. Clin. Neurophysiol. 79 (1), 36–44.

Berg, P., Scherg, M., 1994. A multiple source approach to the correction of eye artifacts. Electroencephalogr. Clin. Neurophysiol. 90 (3), 229–241.

Blinn, K.A., 1955. Focal anterior temporal spikes from external rectus muscle. Electroencephalogr. Clin. Neurophysiol. 7 (2), 299–302.

Boylan, C., Ross Doig, H., 1989. Effect of saccade size on presaccadic spike potential amplitude. Investig. Ophthalmol. Vis. Sci. 30 (12), 2521–2527.

Brouwer, A.-M., Reuderink, B., Vincent, J., van Gerven, M.A., van Erp, J.B., 2013. Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. J. Vis. 13 (3), 17-17.

Brunia, C.H.M., Möcks, J., Van den Berg-Lenssen, M.M.C., Coelho, M., Coles, M.G.H., Elbert, T., Roth, W.T., 1989. Correcting ocular artifacts in the EEG: a comparison of several methods. J. Psychophysiol. 3, 1–50.

Burns, M.D., Bigdely-Shamlo, N., Smith, N.J., Kreutz-Delgado, K., Makeig, S., 2013. Comparison of averaging and regression techniques for estimating event related potentials. In: Paper Presented at the Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE.

Campos Viola, F., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., Debener, S., 2009. Semi-automatic identification of independent components representing EEG artifact. Clin. Neurophysiol. 120 (5), 868–877. https://doi.org/10.1016/j.clinph.2009.01.015.

Carl, C., Acik, A., Konig, P., Engel, A.K., Hipp, J.F., 2012. The saccadic spike artifact in MEG. Neuroimage 59 (2), 1657–1667. https://doi.org/10.1016/j.neuroimage.2011.09.020.

Castellanos, N.P., Makarov, V.A., 2006. Recovering EEG brain signals: artifact suppression with wavelet enhanced independent component analysis. J. Neurosci. Methods 158 (2), 300–312.

Chaumon, M., Bishop, D.V.M., Busch, N.A., 2015. A practical guide to the selection of independent components of the electroencephalogram for artifact correction. J. Neurosci. Methods 250, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025.

Cornelissen, T., Sassenhagen, J., Võ, M.L.-H., 2019. Improving free-viewing fixation-related EEG potentials with continuous-time regression. J. Neurosci. Methods 313, 77–94.

Craddock, M., Martinovic, J., Müller, M.M., 2016. Accounting for microsaccadic artifacts in the EEG using independent component analysis and beamforming. Psychophysiology 53 (4), 553–565.

Dambacher, M., Dimigen, O., Braun, M., Wille, K., Jacobs, A.M., Kliegl, R., 2012. Stimulus onset asynchrony and the timeline of word recognition: event-related potentials during sentence reading. Neuropsychologia 50 (8), 1852–1870. https://doi.org/10.1016/j.neuropsychologia.2012.04.011.

Dandekar, S., Ding, J., Privitera, C., Carney, T., Klein, S.A., 2012. The fixation and saccade p3. PLoS One 7 (11), e48761.

Dandekar, S., Privitera, C., Carney, T., Klein, S.A., 2011. Neural saccadic response estimation during natural viewing. J. Neurophysiol. 107 (6), 1776–1790.

Debener, S., Thorne, J., Schneider, T.R., Viola, F.C., 2010. Using ICA for the Analysis of Multi-Channel EEG Data. Simultaneous EEG and fMRI: Recording, Analysis, and Application: Recording, Analysis, and Application, pp. 121–133.

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. https://doi.org/10.1371/journal.pone.0030135.
Delorme, A., Sejnowski, T., Makeig, S., 2007. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. Neuroimage 34 (4), 1443–1449.
Dimigen, O., Ehinger, B.V., 2019. Analyzing combined eye-tracking/EEG experiments with (non) linear deconvolution models. bioRxiv. https://doi.org/10.1101/735530.
Dimigen, O., Kliegl, R., Sommer, W., 2012. Trans-saccadic parafoveal preview benefits in fluent reading: a study with fixation-related brain potentials. Neuroimage 62 (1), 381–393. https://doi.org/10.1016/j.neuroimage.2012.04.006.
Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A.M., Kliegl, R., 2011. Coregistration of eye movement and EEG in natural reading: analyses and review. J. Exp. Psychol. Gen. 140 (4), 552–572.
Dimigen, O., Valsecchi, M., Sommer, W., Kliegl, R., 2009. Human microsaccade-related visual brain responses. J. Neurosci. 29 (39), 12321–12331. https://doi.org/10.1523/JNEUROSCI.0911-09.2009.
Duffy, F.H., Lombroso, C.T., 1968. Electrophysiological evidence for visual suppression prior to the onset of a voluntary saccadic eye movement. Nature 218, 1074–1075.
Ehinger, B.V., Dimigen, O., 2019. Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. PeerJ 7, e7838. https://doi.org/10.7717/peerj.7838.
Eimer, M., 2014. The time course of spatial attention. In: Oxford Handbooks Online, vol. 1. Oxford University Press, pp. 289–317.
Einhäuser, W., Kruse, W., Hoffmann, K.P., König, P., 2006. Differences of monkey and human overt attention under natural conditions. Vis. Res. 46 (8–9), 1194–1209. https://doi.org/10.1016/j.visres.2005.08.032.
Engbert, R., Kliegl, R., 2003. Microsaccades uncover the orientation of covert attention. Vis. Res. 43 (9), 1035–1045.
Evans, C.C., 1953. Spontaneous excitation of the visual cortex and association areas; lambda waves. Electroencephalogr. Clin. Neurophysiol. 5 (1), 69–74.
Everling, S., Krappmann, P., Flohr, H., 1996. Cortical potentials preceding pro- and antisaccades in man. Electroencephalogr. Clin. Neurophysiol. 102 (4), 356–362.
Gaarder, K., Krauskopf, J., Graf, V., Kropfl, W., Armington, J.C., 1964. Averaged brain activity following saccadic eye movement. Science 146 (3650), 1481–1483.
Gratton, G., 1998. Dealing with artifacts: the EOG contamination of the event-related brain potential. Behav. Res. Methods Instrum. Comput. 30 (1), 44–53.
Groppe, D.M., Makeig, S., Kutas, M., 2008. Independent component analysis of event-related potentials. Cognitive science online 6 (1), 1–44.
Groppe, D.M., Makeig, S., Kutas, M., 2009. Identifying reliable independent components via split-half comparisons. Neuroimage 45 (4), 1199–1211. https://doi.org/10.1016/j.neuroimage.2008.12.038.
Gwin, J.T., Gramann, K., Makeig, S., Ferris, D.P., 2010. Removal of movement artifact from high-density EEG recorded during walking and running. J. Neurophysiol. 103 (6), 3526–3534.
Hassler, U., Barreto, N.T., Gruber, T., 2011. Induced gamma band responses in human EEG after the control of miniature saccadic artifacts. Neuroimage 57 (4), 1411–1421. https://doi.org/10.1016/j.neuroimage.2011.05.062.
Henderson, J.M., Luke, S.G., Schmidt, J., Richards, J.E., 2013. Co-registration of eye movements and event-related potentials in connected-text paragraph reading. Front. Syst. Neurosci. 7, 28.
Hipp, J.F., Siegel, M., 2013. Dissociating neuronal gamma-band activity from cranial and ocular muscle activity in EEG. Front. Hum. Neurosci. 7, 338.
Hironaga, N., Haruhana, K., Liu, L., Fenwick, P., Ioannides, A., 2004. Monitoring of eye movement and its use for artifact elimination. Int. Congr. Ser. 1270, 134–137.
Ille, N., Berg, P., Scherg, M., 2002. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. J. Clin. Neurophysiol. 19 (2), 113–124.
Ito, J., Maldonado, P., Singer, W., Grün, S., 2011. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. Cereb Cortex 21 (11), 2482–2497. https://doi.org/10.1093/cercor/bhr020.
Jerbi, K., Freyermuth, S., Dalal, S., Kahane, P., Bertrand, O., Berthoz, A., Lachaux, J.-P., 2009. Saccade related gamma-band activity in intracerebral EEG: dissociating neural from ocular muscle activity. Brain Topogr. 22 (1), 18–23.
Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.
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), 4. https://doi.org/10.1167/12.7.4.
Kazai, K., Yagi, A., 2003. Comparison between the lambda response of eye-fixation-related potentials and the P100 component of pattern-reversal visual evoked potentials. Cognit. Affect Behav. Neurosci. 3 (1), 46–56.
Keren, A.S., Yuval-Greenberg, S., Deouell, L.Y., 2010. Saccadic spike potentials in gamma-band EEG: characterization, detection and suppression. Neuroimage 49 (3), 2248–2263.
Kierkels, J.J., Riani, J., Bergmans, J.W., van Boxtel, G.J., 2007. Using an eye tracker for accurate eye movement artifact correction. Biomedical Engineering, IEEE Transactions on 54 (7), 1256–1267.
Kliegl, R., Risse, S., Laubrock, J., 2007. Preview benefit and parafoveal-on-foveal effects from word n+ 2. *Journal of Experimental Psychology: Human Perception and Performance* 33 (5), 1250.
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.
Kornrumpf, B., Niefind, F., Sommer, W., Dimigen, O., 2016. Neural correlates of word recognition: a systematic comparison of natural reading and rapid serial visual presentation. J. Cogn. Neurosci. 28 (9), 1374–1391.
Kovach, C.K., Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M.A., Adolphs, R., 2011. Manifestation of ocular-muscle EMG contamination in human intracranial recordings. Neuroimage 54 (1), 213–233. https://doi.org/10.1016/j.neuroimage.2010.08.002.
Kristensen, E., Rivet, B., Guérin-Dugué, A., 2017. Estimation of overlapped eye fixation related potentials: the general linear model, a more flexible framework than the ADJAR algorithm. Journal of Eye Movement Research 10 (1), 1–27.
Kusunoki, M., Goldberg, M.E., 2003. The time course of perisaccadic receptive field shifts in the lateral intraparietal area of the monkey. J. Neurophysiol. 89 (3), 1519–1527.
Lawrence, M.A., 2013. ez: easy analysis and visualization of factorial experiments. R package version 4, 2-2.
Lee, T.-W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput. 11 (2), 417–441.
Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. Electroencephalogr. Clin. Neurophysiol. 48 (6), 609–621.
Lins, O.G., Picton, T.W., Berg, P., Scherg, M., 1993. Ocular artifacts in recording EEGs and event-related potentials II: source dipoles and source components. Brain Topogr. 6 (1), 65–78.
Lopez-Calderon, J., Luck, S.J., 2014. ERPLAB: an open-source toolbox for the analysis of event-related potentials. Front. Hum. Neurosci. 8 https://doi.org/10.3389/fnhum.2014.00213.
Lourenço, P.R., Abbott, W., Faisal, A.A., 2016. Supervised EEG ocular artefact correction through eye-tracking. In: Advances in Neurotechnology, Electronics and Informatics. Springer, pp. 99–113.
Mannan, M.M.N., Kim, S., Jeong, M.Y., Kamran, M.A., 2016. Hybrid EEG—eye tracker: automatic identification and removal of eye movement and blink artifacts from electroencephalographic signal. Sensors 16 (2), 241.
Marmor, M.F., Zrenner, E., 1993. Standard for clinical electro-oculography. Doc. Ophthalmol. 85 (2), 115–124.
Marton, M., Szirtes, J., Breuer, P., 1985. Electrocortical signs of word categorization in saccade-related brain potentials and visual evoked potentials. Int. J. Psychophysiol. 3 (2), 131–144.
Matsuo, F., Peters, J.F., Reilly, E., 1975. Electrical phenomena associated with movements of the eyelid. Electroencephalogr. Clin. Neurophysiol. 38, 507–511.
McMenamin, B.W., Shackman, A.J., Maxwell, J.S., Bachhuber, D.R., Koppenhaver, A.M., Greischar, L.L., Davidson, R.J., 2010. Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG. Neuroimage 49 (3), 2416–2432.
Mennes, M., Wouters, H., Vanrumste, B., Lagae, L., Stiers, P., 2010. Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP. Psychophysiology 47 (6), 1142–1150.
Meyberg, S., Sommer, W., Dimigen, O., 2017. How microsaccades relate to lateralized ERP components of spatial attention: a co-registration study. Neuropsychologia 99, 64–80.
Meyberg, S., Werkle-Bergner, M., Sommer, W., Dimigen, O., 2015. Microsaccade-related brain potentials signal the focus of visuospatial attention. Neuroimage 104, 79–88.
Miyakoshi, M., 2019. Makoto's preprocessing pipeline. Retrieved from. https://sccn.ucsd.edu/wiki/Makoto%27s_preprocessing_pipeline.
Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48 (2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x.
Moster, M.L., Goldberg, G., 1990. Topography of scalp potentials preceding self-initiated saccades. Neurology 40, 644–648.
Nativ, A., Weinstein, J.M., Rosas-Ramos, R., 1990. Human presaccadic spike potentials. Of central or peripheral origin? Investig. Ophthalmol. Vis. Sci. 31 (9), 1923–1928.
Nieuwland, M.S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Ito, A., 2018. Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. eLife 7, e33468.
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.
Nikolaev, A.R., Nakatani, C., Plomp, G., Jurica, P., van Leeuwen, C., 2011. Eye fixation-related potentials in free viewing identify encoding failures in change detection. Neuroimage 56 (3), 1598–1607. https://doi.org/10.1016/j.neuroimage.2011.03.021.
Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: fully automated statistical thresholding for EEG artifact rejection. J. Neurosci. Methods 192 (1), 152–162. https://doi.org/10.1016/j.jneumeth.2010.07.015.
Noureddin, B., Lawrence, P.D., Birch, G.E., 2012. Online removal of eye movement and blink EEG artifacts using a high-speed eye tracker. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 59 (8), 2103–2110.
Onton, J., Makeig, S., 2009. High-frequency broadband modulations of electroencephalographic spectra. Front. Hum. Neurosci. 3 https://doi.org/10.3389/Neuro.09.061.2009.
Onton, J., Westerfield, M., Townsend, J., Makeig, S., 2006. Imaging human EEG dynamics using independent component analysis. Neurosci. Biobehav. Rev 30, 808–822.
Ossandon, J.P., Helo, A.V., Montefusco-Siegmund, R., Maldonado, P.E., 2010. Superposition model predicts EEG occipital activity during free viewing of natural scenes. J. Neurosci. 30 (13), 4787–4795. https://doi.org/10.1523/JNEUROSCI.5769-09.2010.

Parks, N.A., Corballis, P.M., 2008. Electrophysiological correlates of presaccadic remapping in humans. Psychophysiology 45, 776–783.
Picton, T.W., Van Roon, P., Armilio, M.L., Berg, P., Ille, N., Scherg, M., 2000. The correction of ocular artifacts: a topographic perspective. Clin. Neurophysiol. 111, 53–65.
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 https://doi.org/10.3389/fnhum.2012.00278.
Pontifex, M.B., Gwizdala, K.L., Parks, A.C., Billinger, M., Brunner, C., 2016. Variability of ICA decomposition may impact EEG signals when used to remove eyeblink artifacts. Psychophysiology 54 (3), 386–398. https://doi.org/10.1111/psyp.12804.
Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychol. Bull. 124 (3), 372–422.
Reingold, E.M., Stampe, D.M., 2002. Saccadic inhibition in voluntary and reflexive saccades. J. Cogn. Neurosci. 14 (3), 371–388. https://doi.org/10.1162/089892902317361903.
Reva, N.V., Aftanas, L.I., 2004. The coincidence between late non-phase-locked gamma synchronization response and saccadic eye movements. Int. J. Psychophysiol. 51 (3), 215–222. https://doi.org/10.1016/j.ijpsycho.2003.09.005.
Ries, A.J., Slayback, D., Touryan, J., 2018. The fixation-related lambda response: effects of saccade magnitude, spatial frequency, and ocular artifact removal. Int. J. Psychophysiol. 134, 1–8.
Riggs, L., Merton, P., Morton, H., 1974. Suppression of visual phosphenes during saccadic eye movements. Vis. Res. 14 (10), 997–1011.
Rolfs, M., 2009. Microsaccades: small steps on a long way. Vis. Res. 49, 2415–2441.
Sakamoto, A., Lüders, H., Burgess, R., 1991. Intracranial recordings of movement-related potentials to voluntary saccades. J. Clin. Neurophysiol. 8 (2), 223–233.
Scherg, M., 2013. Tutorial: Using Calibration Data to Generate Artifact Coefficients. Documentation for BESA (Brain-Electrical Source Analysis) (Version 6.0). Megis Software GmbH, Munich, Germany.
Shinomiya, K., Itsuki, N., Kubo, M., Shiota, H., 2008. Analyses of the characteristics of potential and cross-talk at each electrode in electro-oculogram. J. Med. Investig. 55, 120–126.
Simola, J., Torniainen, J., Moisala, M., Kivikangas, M., Krause, C.M., 2013. Eye movement related brain responses to emotional scenes during free viewing. Front. Syst. Neurosci. 7, 41.
Smith, N.J., Kutas, M., 2015. Regression-based estimation of ERP waveforms: II. Nonlinear effects, overlap correction, and practical considerations. Psychophysiology 52 (2), 169–181.
Soto, V., Tyson-Carr, J., Kokmotou, K., Roberts, H., Cook, S., Fallon, N., Stancak, A., 2018. Brain responses to emotional faces in natural settings: a wireless mobile EEG recording study. Front. Psychol. 9, 2003.
Stone, J.V., 2004. Independent Component Analysis: A Tutorial Introduction. MIT Press, Cambridge.
Thickbroom, G.W., Mastaglia, F.L., 1985. Presaccadic 'spike' potential: investigation of topography and source. Brain Res. 339 (2), 271–280.
Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data – a practical approach. J. Neurosci. Methods 250, 34–46. https://doi.org/10.1016/j.jneumeth.2014.08.002.
Winkler, I., Debener, S., Muller, K.-R., Tangermann, M., 2015. On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. In: Paper Presented at the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).
Winkler, I., Haufe, S., Tangermann, M., 2011. Automatic classification of artifactual ICA-components for artifact removal in EEG signals. Behav. Brain Funct. 7 (1), 30. https://doi.org/10.1186/1744-9081-7-30.
Xue, J., Quan, C., Li, C., Yue, J., Zhang, C., 2017. A crucial temporal accuracy test of combining EEG and Tobii eye tracker. Medicine 96 (13).
Yagi, A., 1979. Lambda-waves associated with offset of saccades - subject with large lambda-waves. Biol. Psychol. 8 (3), 235–238. https://doi.org/10.1016/0301-0511(79)90051-6.
Yamazaki, A., 1968. Electrophysiological study on "flick" eye movements during fixation. Nippon. Ganka Gakkai Zasshi 72 (12), 2446–2459.
Young, L.R., Sheena, D., 1988. Eye movement measurement techniques. In: Webster, J.G. (Ed.), Encyclopedia of Medical Devices and Instrumentation. Wiley, New York, pp. 1259–1269.
Yuval-Greenberg, S., Tomer, O., Keren, A.S., Nelken, I., Deouell, L.Y., 2008. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron 58 (3), 429–441.
Zakeri, Z., Assecondi, S., Bagshaw, A.P., Arvanitis, T.N., 2014. Influence of signal preprocessing on ICA-based EEG decomposition. In: IFMBE Proceedings. Springer International Publishing, pp. 734–737.