NeuroImage 49 (2010) 2416–2432

Contents lists available at ScienceDirect

NeuroImage

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

Validation of ICA-based myogenic artifact correction for scalp and source-localized EEG

Brenton W. McMenamin [a,*,1], Alexander J. Shackman [b,*,1], Jeffrey S. Maxwell [c], David R.W. Bachhuber [b], Adam M. Koppenhaver [b], Lawrence L. Greischar [b], Richard J. Davidson [b,*]

[a] Department of Psychology, Center for Cognitive Science, University of Minnesota, Twin Cities, Minneapolis, MN, USA
[b] Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin, Madison, WI, USA
[c] U.S. Army Research Laboratory, Aberdeen, MD, Laboratory for Affective Neuroscience, University of Wisconsin, Madison, WI, USA

ARTICLE INFO

Article history:
Received 8 July 2009
Revised 5 October 2009
Accepted 6 October 2009
Available online 13 October 2009

ABSTRACT

Muscle electrical activity, or "electromyogenic" (EMG) artifact, poses a serious threat to the validity of electroencephalography (EEG) investigations in the frequency domain. EMG is sensitive to a variety of psychological processes and can mask genuine effects or masquerade as legitimate neurogenic effects across the scalp in frequencies at least as low as the alpha band (8–13 Hz). Although several techniques for correcting myogenic activity have been described, most are subjected to only limited validation attempts. Attempts to gauge the impact of EMG correction on intracerebral source models (source "localization" analyses) are rarer still. Accordingly, we assessed the sensitivity and specificity of one prominent correction tool, independent component analysis (ICA), on the scalp and in the source-space using high-resolution EEG. Data were collected from 17 participants while neurogenic and myogenic activity was independently varied. Several protocols for classifying and discarding components classified as myogenic and non-myogenic artifact (e.g., ocular) were systematically assessed, leading to the exclusion of one-third to as much as three-quarters of the variance in the EEG. Some, but not all, of these protocols showed adequate performance on the scalp. Indeed, performance was superior to previously validated regression-based techniques. Nevertheless, ICA-based EMG correction exhibited low validity in the intracerebral source-space, likely owing to incomplete separation of neurogenic from myogenic sources. Taken with prior work, this indicates that EMG artifact can substantially distort estimates of intracerebral spectral activity. Neither regression- nor ICA-based EMG correction techniques provide complete safeguards against such distortions. In light of these results, several practical suggestions and recommendations are made for intelligently using ICA to minimize EMG and other common artifacts.

Peri-cranial muscle or *myogenic* activity is distinguished by its relatively high amplitude, broad spectral and often anatomical distributions, and exquisite sensitivity to a variety of psychologically interesting processes. Consequently, it poses a serious inferential hazard for any electroencephalography (EEG) investigation in the frequency-domain. This artifact can compromise sensitivity by masking effects of interest or diminish specificity by masquerading as a neurogenic effect. Although several techniques have been developed to correct myogenic activity (Shackman et al., 2009), many have been subjected to only limited attempts at validation, rendering their utility questionable. In particular, the sensitivity and specificity of one prominent electromyography (EMG) correction tool, independent component analysis (ICA), remains unclear.

Several properties of cranial EMG are collectively responsible for its pernicious effects. First, EMG is sufficiently sizable to perturb *all* classic EEG bands. Goncharova, McFarland, Vaughan, and Wolpaw (2003) report myogenic artifact reliably as low as 2 Hz, making even the widely used alpha band (8–13 Hz) vulnerable to muscle artifacts (Lee and Buchsbaum, 1987; Willis et al., 1993; Van Boxtel, 2001). Second, EMG can often be detected across the entire scalp (Goncharova et al., 2003) due to volume conduction of myogenic activity independently generated by muscles across the head, face and neck. Anterior electrodes are sensitive to facial muscles, such as the *corrugator supercilii* and *frontalis*; lateral electrodes are sensitive to the muscles of mastication, *masseter* and *temporalis*; and posterior electrodes are sensitive to muscles at the intersection of the cranium, spine, and torso, such as *occipitalis* (Supplementary Fig. 1). Third, EMG is temporally confounded with a variety of experimental manipulations. Facial EMG, in particular, is sensitive to numerous cognitive and affective processes, including cognitive load (Cohen et al., 1992; Waterink and van Boxtel, 1994), facial mimicry (Dimberg et al., 2000), vocalization (Brooker and Donald, 1980), and induced emotional states (Borden et al., 1991; Coan and Allen, 2003; Bradley et al., 2001).

* Corresponding authors. Laboratory for Affective Neuroscience, University of Wisconsin, Madison, 1202 West Johnson Street, Madison, WI 53706, USA.
E-mail addresses: mcmen020@umn.edu (B.W. McMenamin), shackman@wisc.edu (A.J. Shackman), rjdavids@wisc.edu (R.J. Davidson).
[1] Brenton McMenamin and Alexander Shackman are co-first authors on this manuscript.

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.10.010

EMG also exhibits less stereotypy than other biological artifacts. Ocular and cardiac artifacts, for example, arise from fixed sources and do not qualitatively differ across individuals. EMG, however, arises from the activity of spatially distributed, functionally independent muscle groups, with distinct topographic and spectral signatures (Goncharova et al., 2003). For instance, *frontalis* activity peaks around 25 Hz, whereas *temporalis* generates a low peak around 20 Hz and broad plateau centered around 40–80 Hz (Goncharova et al., 2003). The spectral composition of myogenic activity also varies as a function of contraction intensity (Goncharova et al., 2003) and fatigue (Chung et al., 2002). This is compounded by the fact that the relative contributions of each muscle group to the cranial EMG vary substantially across elicitors and individuals (Tassinary et al., 2007) and may differ somewhat between spontaneous and voluntary contractions (Davidson et al., 2004; Morecraft and Tanji, 2009).

Given the inferential hazards posed by EMG, there is substantial interest in developing tools to remedy myogenic artifact. Generally, EEG artifacts can be addressed in one of two ways, *rejecting* contaminated epochs of data or *filtering* artifact from neurogenic activity. Rejection-based techniques are most appropriate for transient artifacts, such as blinks, that influence a small portion of the data record. The protracted time-course of EMG makes such a solution impractical—the high data rejection rate would markedly erode the signal to noise ratio (Jung et al., 2000b; Talsma, 2008). Moreover, because EMG covaries with cognitive and affective processes of interest, rejecting data laden with EMG artifact would likely entail discarding some of the most interesting, discriminative periods of neural activity (Davidson et al., 1990). For these reasons, EMG mandates the use of filtering techniques capable of separating myogenic from neurogenic activity. Given marked individual differences in the spectral and anatomical profile of myogenic activity, spatial or spectral filters or templates that are fixed across subjects cannot be fruitfully applied to the correction of EMG artifact (cf. Frank and Frishkoff, 2006; Ille et al., 2002; Koskinen and Vartiainen, 2009). Instead, a more flexible approach is required.

One class of techniques for correcting EMG artifact employs variants of the general linear model (GLM), such as multiple regression and ANCOVA, to identify and discard variance in a neurogenic band of interest (e.g., alpha) that is predicted by activity in an *a priori* EMG band (e.g., 70–80 Hz). The advantage of this technique is that it does not require dedicated EMG channels or manual intervention, and, by performing separate corrections at each site, can accommodate individual differences in artifact topography. These GLM-based techniques have proven quite popular (Allen et al., 2004; Davidson et al., 2000), and McMenamin et al. (2009) have shown that at least one variant of this technique displays adequate sensitivity and specificity on the scalp.

Despite these strengths, GLM-based EMG correction techniques suffer from two key limitations. First, they do not permit reconstruction of the EEG time-series. Thus, while useful for investigations of tonic ("resting") and induced changes in the EEG spectra (e.g., Lutz et al., 2004), GLM-based tools cannot be applied to studies relying on event-related spectral perturbation (ERSP) measures (Onton et al., 2006). Second, McMenamin et al (2009) reported that applying GLM-based techniques to source-estimated EEG in a voxelwise manner is not appropriate if the data has been corrupted by EMG prior to localization. Source-estimation ("localization") is a technique that estimates neurogenic signals from scalp EEG recordings. This is achieved by developing a forward-model that uses the biophysics of the EEG (e.g. the spatial filtering imposed on neurogenic signals by the cerebrospinal fluid, skull and scalp) to predict signals on the scalp given a particular neural generator. Source-estimation occurs when this model is inverted and used to estimate a probable neural generator given scalp-recorded signals (Pizzagalli, 2007). McMenamin et al (2009) speculated that the EMG-contaminated data cannot be properly localized because a solution space that only allows intra-cranial dipoles cannot account for a scalp-recording that contain both intra-cranial (neurogenic) and extra-cranial (myogenic) sources. The resulting attempt at localization will be corrupted and the true neurogenic solution rendered unrecoverable. Removing extra-cranial source activity from the data prior to source-estimation may circumvent this problem. Unfortunately, this is not possible using GLM-based techniques because they cannot reconstruct the artifact-free EEG time-series.[2]

A second class of EMG correction methods employs ICA to decompose the EEG time-series into a set of temporally independent components (Delorme, 2007a; James and Hesse, 2005; Onton et al., 2006; Onton and Makeig, 2006; Makeig et al., 2004). Components are inspected visually for the presence of artifact and those classified as predominantly artifactual (e.g. EMG or blinks) are discarded. Like GLM-based correction techniques, ICA does not require dedicated EMG channels and can accommodate variation across the scalp. More importantly, unlike GLM-based techniques, ICA allows reconstruction of the artifact-filtered time-series, which can then be used for analyses employing averaging, spectral decomposition, or source modeling.

Although ICA shows great promise as a tool for correcting EMG and other kinds of biological artifact (e.g., Jung et al., 2000a), attempts to assess its validity have been limited. Many validation studies have relied on small samples of *ad hoc* data (Delorme et al., 2007b; Jung et al., 2000a,b; Wallstrom et al., 2004; Flexer et al., 2005; Ting et al., 2006; Frank and Frishkoff, 2006). While others have used simulations (e.g., Crespo-Garcia et al., 2008; De Clercq et al., 2005; Delorme et al., 2007a,b; Fitzgibbon et al., 2007; Frank and Frishkoff, 2006; Romero et al., 2008). In simulations, real or artificial EMG activity is mathematically "injected" into otherwise artifact-free EEG. The potential problem with this strategy is that the assumptions underlying injection (e.g., the degree of temporal and spatial correlation with neurogenic signals) may not characterize real EMG contamination, potentially limiting external validity and biasing the results in favor of correction techniques founded on similar assumptions (Grouiller et al., 2007; Hoffmann and Falkenstein, 2009).

Accordingly, the major aim of the present study was to quantitatively assess the quality of EMG artifact correction afforded by ICA. Ideally, validation would quantitatively establish that a technique possesses a high degree of sensitivity (i.e., attenuates myogenic artifact) and specificity (i.e., preserves neurogenic signals) in a reasonably large and varied dataset. This requires data in which the presence or absence of EMG ("ground truth") is definitive or can be reasonably assumed. To this end, the dataset previously employed by McMenamin et al. (2009) for testing the validity of GLM-based correction techniques was reanalyzed using ICA. This had the advantage of facilitating direct comparisons across correction techniques. In this dataset, 128-channel EEG was acquired while neurogenic and myogenic activity were independently varied. Alpha band neurogenic activity was selectively increased or decreased by instructing participants to close or open their eyes, a procedure sometimes termed the "Berger maneuver" (cf. Berger, 1929/1969). Myogenic activity was manipulated by instructing participants to alternately tense and relax their cranial muscles. The sensitivity and specificity of ICA-based EMG correction were then quantitatively assessed in the alpha band using methods similar to those described in our prior report (McMenamin et al., 2009). Several considerations led us to focus on the alpha band. First, it is relatively easy to manipulate neurogenic activity in this frequency. To our knowledge, comparably robust manipulations do not exist for the other classical EEG bands (Niedermeyer, 2005). Second, alpha activity has been among the most widely used spectral indices of neural activity, from the earliest EEG research (Berger, 1929/1969), to contemporary

[2] Source modeling in the frequency-domain requires phase information in the form of the cross-spectra. Extant GLM-correction techniques operate on estimates of spectral power (squared amplitude) and discard information about the phase of EEG oscillations required to compute the cross-spectra.

investigations of memory (Freuenberger et al., 2009; Gevins and Smith, 2000; Hamidi et al., 2009), perception and attention (Romei et al., 2008; Thut and Miniussi, 2009), emotion (Coan and Allen, 2003; Davidson et al., 1990), temperament and individual differences (Carver and Harmon-Jones, 2009; Shackman et al., in press), and psychopathology (Thibodeau et al., 2006; DeRubeis et al., 2008).

The other major aim of this study was to test whether ICA-based techniques constitute a valid EMG correction technique for distributed intracerebral source modeling. Source modeling is an increasingly popular technique for maximizing the anatomical information yielded by scalp-recorded EEG (Pizzagalli, 2007) and the dissemination of commercial and freely available software for performing distributed source localization, such as Cartool (http://brainmapping.unige.ch/Cartool.htm), EMSE (http://www.sourcesignal.com), LORETA-KEY (http://www.unizh.ch/keyinst/) and SPM5 (http://www.fil.ion.ucl.ac.uk/spm/), is likely to accelerate this trend. Furthermore, prior work indicates that EMG correction techniques deemed valid on the scalp do not necessarily confer validity in the intracerebral source-space (McMenamin et al., 2009). Accordingly, ICA-based procedures that proved valid on the scalp were also tested with source solutions estimated using the low-resolution electromagnetic tomography (LORETA) algorithm.

A minor aim of this study was to evaluate the degree to which variation in the protocol for filtering *non-myogenic* artifacts, such as eye movements, impacts the quality of EMG correction. To date, existing methodological and empirical reports employing ICA provide little direct guidance on the question of which components ought to be discarded (Shackman et al., 2009). Furthermore, despite the fact that ICA requires trained raters to inspect hundreds or even thousands of components for a single high-resolution EEG study (number of components ≈ channels × participants; see Supplementary Method), the reliability of component classification has only rarely been reported (Viola et al., 2009). Without such evidence, poor validity might simply reflect inadequate training or an ambiguous classification protocol. Accordingly, the inter-rater reliability was computed.

Methods

Participants

The dataset consisted of seventeen individuals recruited from the University of Wisconsin–Madison campus (16 female; $M = 24.1$ years, $SD = 7.1$) and described in an earlier report assessing the validity of GLM-based EMG correction techniques (McMenamin et al., 2009). Each received US$20 for their participation. Participants provided informed consent in accord with guidelines prescribed by the local Institutional Review Board.

Design

In order to independently manipulate neurogenic and myogenic activity in the alpha band (8–13 Hz), the experiment took the form of a 2 (Eyes Open/Closed) × 2 (Muscles Tense/Relaxed) repeated-measures design. We anticipated that participants would generate greater broad-spectrum power, including increases in alpha power, indicative of *reduced* neural activity (Allen et al., 2004; Oakes et al., 2004), during the eyes-closed condition. We further expected participants to generate greater alpha power, indicative of *increased* muscle activity, during the muscles-tense condition. Hereafter, these four conditions are referred to using the following acronyms: Open-Relaxed (OR), Open-Tense (OT), Closed-Relaxed (CR), and Closed-Tense (CT).

Procedure

Procedures were identical to those detailed by McMenamin et al (2009; see also Bonnett and Arand, 2001; Freeman et al., 2003). In brief, participants were instructed how to properly tense facial muscles at the outset of the session. Frontalis and corrugator muscles were contracted by lifting and squeezing the eyebrows together; masseter and temporalis were contracted by lightly clenching the jaw. EEG was acquired during sixteen 32-second blocks (order counterbalanced; 4 blocks/condition). Participants were continuously monitored via a closed-circuit audio-video circuit and real-time EEG.

EEG acquisition and preliminary reduction

EEG were collected using a 128-channel Geodesic Sensor Net (GSN128; Electrical Geodesics Inc., Eugene, OR) referenced to vertex (Cz) and sampled at 500-Hz (analog anti-aliasing: 0.1–250 Hz).[3] Data reduction used a combination of EEGLAB (Delorme and Makeig, 2004; http://www.sccn.ucsd.edu/eeglab) and in-house code written for MATLAB (http://www.mathworks.com). A zero-phase 60-Hz notch filter removed line noise from calibrated (μV) data, and bad channels (± 100 μV for >20 s) or gross artifacts (± 100 μV for >4 channels) were manually identified and rejected. A 0.5-Hz high-pass filter was used to attenuate channel drift and better satisfy ICA's stationarity assumption (Onton et al., 2006).[4] Such artifacts were rejected to better approximate the subtle contamination of signal that can occur when EMG covaries with an experimental treatment. Removal of non-stereotyped artifact also maximizes the quality of the ICA (Onton et al., 2006).

ICA

Overview

Consistent with other high-resolution EEG studies (Delorme, 2007b),a spatial Principal Components Analysis (PCA) was used to reduce the dimensionality of the EEG from 128 channels to 64 principal components (PCs) prior to performing ICA.[5] This was done in a single step as part of the ICA using the EEGLAB *runica* command, implementing the extended Infomax algorithm (Bell and Sejnowski, 1995; Lee et al., 1999).

The primary aim of this study was to assess the validity of ICA for EMG artifact correction. Accordingly, three protocols for the correction of EMG artifact, described below, were investigated. A secondary aim of this study was to investigate the degree to which the quality of EMG correction was dependent on the protocol for removing *non-myogenic* sources of variance (e.g., ocular artifact, noise components). Consequently, three ICA-based protocols for the correction of non-neurogenic/non-myogenic (NNNM) components, described below,

[3] In hindsight, the acquisition parameters were not optimal for measuring high-frequency EMG effects. We would recommend that future studies of myogenic artifact use a higher sampling rate (e.g., 1000 Hz) and more conservative anti-aliasing filter (e.g., 250 Hz) to compensate for non-zero filter roll-off.

[4] Some investigators (e.g., Huang et al., 2008; Milne et al., 2009) employ a more stringent 1–2 Hz highpass filter for this purpose. As noted by several reviewers, for researchers interested in low-frequency ERP components or frequency bands (e.g., delta, 1–4 Hz), it is possible to train ICA on highpass filtered data and then apply the weights to the unfiltered dataset.

[5] There were two reasons for doing so, aside from computational and classification efficiency. First, preliminary inspection of the 128 components extracted from the native electrode array indicated over-fitting, evidenced by fragmentation of artifacts across components (Li et al., 2007; Lawrence and Hancock, 1999). By contrast, exploratory analyses (not reported) showed that reduction to 48 or fewer PCs prior to ICA led to under-fitting, evidenced by cross-contamination of EEG, physiological artifacts, and noise (Fava and Velicer, 1996). Second, it has been suggested (Onton et al., 2006; Romero et al., 2008) that Infomax ICA requires a minimum of $20 \times c^2$ samples, where c is the number of channels or, equivalently, PCs. For the native electrode array, this would require $20 \times 128^2 = 327680$ samples, whereas we had at most 16 blocks × 32-s × 500 Hz = 256000 samples. Reducing the model order by half allowed us to satisfy this criterion ($20 \times 64^2 = 81920$ samples). Quantitative estimates of "model order," the number of components required to adequately but parsimoniously describe the data, suggested that this was sufficient (see Supplementary Method and Results). A viable alternative to PCA-based dimension reduction is to simply prune the number channels submitted to ICA (Milne et al., 2009), at the potential expense of spatial resolution (Srinivasan et al., 1998; Michel et al., 2004).

were also examined. The quality of EMG artifact correction was evaluated for all nine factorial combinations of the EMG and NNNM protocols. Following removal of the relevant components, the filtered 128-channel time-series were reconstructed. Subsequent analyses used only the 107 cephalic electrodes. Exploratory analyses (not reported) using the complete 128-channel array indicated worsened performance when the peri-cephalic electrodes on the face and along the posterior edge of the array were retained (Supplementary Fig. 1).

Component classification

Using in-house code, the variance accounted for by each of the ICs was assessed. By default, components that individually accounted for <0.2% of the variance were categorized as *Low-Variance*[6]. In cases where the determination was unambiguous, exceptions were made. As described in the Supplementary Method and Results (Supplementary Figs. 2–13), the remaining components were classified as neurogenic (*Neuro*), myogenic (*Myo*), a combination of the two sources (*Neuro-Dominant* or *Myo-Dominant*), or artifact (residual *Gross* or *Ocular*). Components classified as Gross included reference, ground, electrocardiographic, and alternating current artifacts. Components that met the minimum variance criterion, but proved impossible to unambiguously categorize were classified as *Noise*. Classifications were made by two raters based on inspection of the component's time-series, power spectrum, and topography. When disagreements occurred, final classification was by consensus. Inter-rater reliability, assessed prior to consensus using Krippendorff's alpha (Hayes and Krippendorff, 2007), was excellent, $\alpha = 0.98$ (for details, see Supplementary Method). We urge investigators with a practical interest in using ICA for artifact reduction to examine our detailed classification protocol (see Supplementary Method and Results).

Correction of EMG artifact

Three different ICA-based protocols for removing myogenic artifact were assessed. The Minimal-EMG protocol discarded only those components that contained clear EMG activity in the absence of any identifiable neurogenic activity (i.e., rejected Myo components). The Intermediate-EMG protocol expanded this definition to include mixed components in which myogenic activity was more prominent than neurogenic activity (i.e. rejected both Myo and Myo-Dominant components). The Maximal-EMG protocol rejected any component containing myogenic signal, even if myogenic activity was less prominent than neurogenic activity (i.e., rejected Myo, Myo-Dominant, and Neuro-Dominant components). Thus, the Maximal-EMG protocol performs the strictest filtering of the data, at the potential expense of discarding neurogenic signals of interest.

Filtering of non-neurogenic/non-myogenic (NNNM) signals

To provide a specific test of ICA's utility for removing EMG artifact, it is necessary to first filter signals that are not clearly neurogenic or myogenic (cf. McMenamin et al., 2009). However, the choice of which components to remove is subjective and has a marked impact on the number of components and percentage of variance retained (see Results). Accordingly, three different ICA-based protocols for filtering non-neurogenic/non-myogenic signals were used. The Minimal-NNNM protocol made the fewest assumptions, filtering only those components that were explicitly classified as Gross or Ocular artifact, similar to the method used in McMenamin et al. (2009). The Intermediate-NNNM protocol made the additional assumption that components categorized as Noise do not contain meaningful neurogenic signal and filtered them as well. The Maximal-NNNM protocol further assumed that Low-Variance components do not contain significant neurogenic signal and filters them as well.

Scalp spectral power density estimation

Following reconstruction of the filtered time-series, epochs with residual artifact (i.e., deviations exceeding ± 200 μV for more than half an epoch or variance exceeding 1000 μV^2) or flat channels (epoch variance less than 0.25 μV^2) were automatically rejected (Delorme et al., 2007a,b). After residual artifact-rejection, the rejected channels were interpolated with a spherical spline when at least one neighboring electrode was usable (Greischar et al., 2004). Data were re-referenced to an average montage (Davidson et al., 2000; Dien, 1998) and spectral power density (μV^2/Hz) estimated for the alpha (8–13 Hz) band using Welch's (1967) method on sliding Hanning-windowed epochs (50% overlap). Estimates were $\log_{10}$ transformed to normalize the distribution (Allen et al., 2004; Gasser et al., 1982).

LORETA distributed source current density modeling

The modeling of distributed sources from scalp-recorded electrical activity was performed using previously published procedures (McMenamin et al., 2009; Shackman et al., in press) via in-house MATLAB code implementing the LORETA algorithm (Pascual-Marqui et al., 1994) to estimate intracerebral current density. LORETA has undergone extensive cross-modal validation (reviewed in Shackman et al., in press; Pizzagalli, 2007).

An inverse operator distributed with the LORETA-Key software suite (Pascual-Marqui, 1999; http://www.unizh.ch/keyinst/; $\lambda = 10^{-5}$) was used to generate three-dimensional intracerebral current density estimates (A/m^2) from cross-spectra calculated using the artifact-free Hanning-windowed epochs from the scalp analyses. The forward-model is a 3-shell spherical head model using 107 cephalic EEG electrodes (Shackman et al., in press). The source-space is normalized to the Montreal Neurological Institute's probabilistic MRI anatomical template (i.e., MNI305; Evans et al., 1993; Collins et al., 1994), restricted to the cerebral gray matter, hippocampi, and amygdalae on a 7-mm^3 isotropic lattice. LORETA source-estimates were $\log_{10}$-transformed prior to analysis (Thatcher et al., 2005). Results are displayed on the rendered canonical brain distributed with LORETA-Key.

Analytic strategy

Overview

A valid correction technique should render EMG-contaminated data statistically equivalent to data collected under the same conditions in the absence of myogenic artifact (Frank and Frishkoff, 2006; Debener et al., 2007; Freyer et al., 2009). Accordingly, each combination of the EMG correction and NNNM filtering protocols was evaluated in terms of its (i) sensitivity, the attenuation of myogenic artifact (i.e., Tense vs. Relaxed) in the alpha band, (ii) specificity, the preservation of neurogenic effects (i.e., alpha-blocking: Eyes-Closed vs. Eyes-Open) in the alpha band, and (iii) the degree to which each protocol introduced correction artifacts, artificial effects generated by the correction. Sensitivity and specificity were assessed using regions of interest (ROIs) defined by the areas of peak myogenic and neurogenic activation, respectively. An ROI approach was used to constrain the number of comparisons in both scalp and LORETA source-space analyses. Only those filtering protocols that proved sufficiently valid on the scalp were assessed with LORETA. To permit a direct comparison of ICA- and GLM-based EMG correction techniques, key analyses reported in McMenamin et al. (2009) were recomputed using the identical validation techniques used here. These analyses are detailed in the Supplementary Method and Results.

[6] Preliminary inspection of the ICA results indicated that such low-variance (<0.2%) components were dominated by noise, making them difficult to reliably classify and leading raters to devote an undue amount of time to their consideration. The threshold of 0.2% was arbitrarily chosen to minimize the cumulative amount of variance that was automatically classified as "noise" (i.e., remained unclassified).

Sensitivity

On the scalp, a myogenic ROI was created for each of the three NNNM filters using electrodes exhibiting a significant ($p<0.05$) myogenic effect (OR-OT). Channels that were situated at the edge of the 107-channel electrode-array, were spatially discontiguous (i.e., lacked at least one nearest neighbor meeting the significance criterion), or also met the inclusion criteria for the neurogenic (i.e., specificity) ROI were excluded (range: 8–15 electrodes; many located at the posterior base of the array). The effect of the spatial contiguity criterion was minor, resulting in a single electrode being dropped. In the LORETA source-space, myogenic ROIs were created by identifying voxels in the OR-OT contrast with $p<0.001$, using a cluster-extent threshold to correct for multiple comparisons (Nichols and Holmes, 2002; Shackman et al., in press).

Using the resulting ROIs, the degree to which each EMG correction protocol attenuated the myogenic contrast (i.e., EMG-corrected OR-OT vs. 0) was tested. Additional contrasts tested the degree to which each EMG-correction protocol removed myogenic effects using double differences that compared three EMG-corrected contrasts of interest and their uncorrected, artifact-free analogs. ICA's ability to correct EMG artifact that negatively covaried with neurogenic signals was tested using the (EMG-corrected OT-CR) − (uncorrected OR-CR) contrast, and the ability to correct artifact that was positively covaried with neurogenic signals was tested using the (EMG-corrected OR-CT) − (uncorrected OR-CR) contrast. The amount of EMG artifact surviving each correction was indexed using median and peak ROI *t*-values, viewed as indices of typical and "worst-case" correction, respectively. Significant *t*-tests for these contrasts indicate that the EMG-corrected EEG signals deviate from their artifact-free analogs, evidence of poor sensitivity.

Conversely, failure to reject the null hypothesis does not indicate the absence of residual myogenic activity. In order to rigorously test whether the EMG-corrected contrasts were significantly equivalent to artifact-free data, the Westlake–Schuirmann test (Seaman and Serlin, 1998) was employed as a follow-up test to non-significant contrasts. Sometimes termed the two one-sided tests (TOST) method, a number of fields (e.g., the US Food and Drug Agency; Department of Health and Human Services, 2001) consider TOST the gold standard for testing statistical equivalence. The null hypothesis for TOST is that the mean difference lies outside of the range $[-\varepsilon, \varepsilon]$, where ε is an *a priori* error tolerance. To reject the null (i.e., demonstrate significant equivalence) for $\alpha = 0.05$, one must demonstrate that the 90th-percentile confidence interval of the mean difference between the artifact-free and EMG-corrected data lies completely within the interval $[-\varepsilon, \varepsilon]$. Following our prior report (McMenamin et al., 2009), ε was set to 0.5 standard deviations of the artifact-free contrast (i.e., OR for the OR-OT contrast, OR-CR for positively/negatively covarying contrasts).

Specificity

Neurogenic ROIs were generated by thresholding the neurogenic contrast (OR-CR). Owing to the large size of this effect ($p<0.001$ at all electrodes), it proved useful to threshold the contrast using a percentile approach. This had the advantage of creating ROIs that were similar in size to the myogenic ROIs used to interrogate sensitivity. On the scalp, this entailed selecting the upper tercile of channels. As before, channels that were situated on the edge of the array, were not spatially contiguous, or met the inclusion criteria for the myogenic (i.e., sensitivity) ROI were excluded. The latter two criteria led us to drop one electrode. In the source-space, voxels with absolute *t*-values in the upper tercile for the OR-CR contrast were selected.

As with the sensitivity analysis, *t*-tests and follow-up TOSTs were used to test whether neurogenic effects were preserved following EMG correction. First, to test the impact of correction *per se* on neurogenic activation, (EMG-corrected OR-CR) was compared against (uncorrected OR-CR). Second, to test correction's impact on negatively covarying neurogenic and myogenic signals, (EMG-corrected OT-CR) was compared to (uncorrected OR-CR). Third, to test correction's impact on positively covarying signals, (EMG-corrected OR-CT) was compared to (uncorrected OR-CR). The ε error tolerance for TOST follow-ups was defined using the uncorrected OR-CR contrast.

Correction artifacts

To investigate the degree to which protocols generated artificial results, two kinds of tests were conducted. The first test determined whether correction of the EMG-contaminated myogenic contrast produced artificial effects in the neurogenic ROI (i.e., EMG-corrected OR-OT). The second test assessed whether correction of the EMG-free neurogenic contrast yielded artificial effects in the myogenic ROI (i.e., [EMG-corrected OR-CR] – [uncorrected OR-CR]). The presence of artifactual effects was assessed using the same logic as the sensitivity and specificity tests.

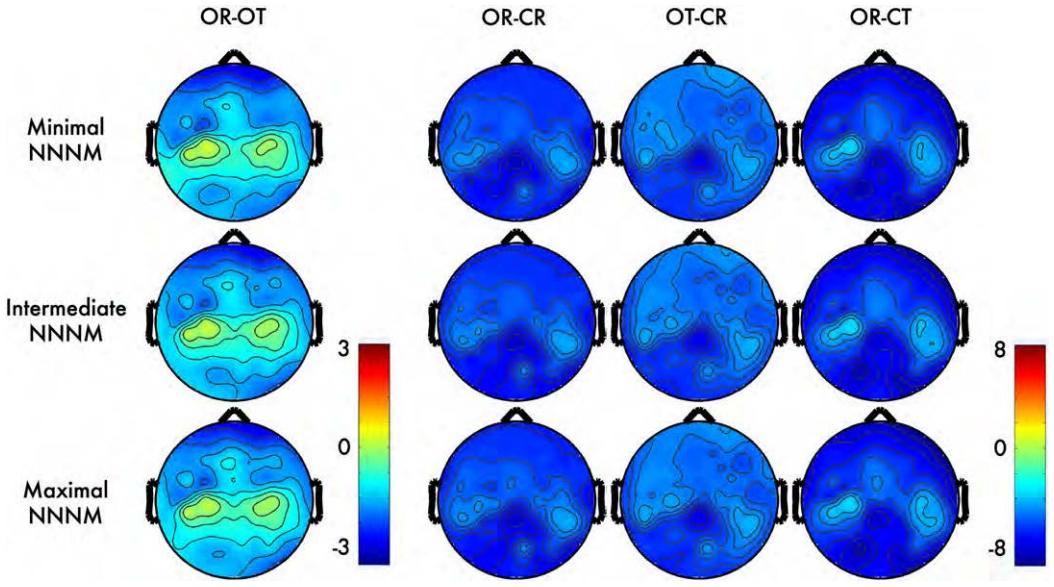


Fig. 1. Alpha-band contrasts prior to correction. Topographic maps depict spline-interpolated *t*-maps for each condition-contrast (columns) and non-neurogenic/non-myogenic (NNNM) artifact filter (rows). There were four conditions, reflecting the factorial manipulation of myogenic (Muscles: **R**elaxed, **T**ensed) and neurogenic activity (Eyes: **O**pen, **C**losed). Contrasts were computed to isolate *myogenic* (OR–OT), *neurogenic* (OR–CR), *positively-covarying* (OT–CR), and *negatively-covarying* (OR–CT) activity. Note the less extreme values for the myogenic contrast (OR–OT; first column).

Performance ratings

For each contrast and correction protocol, sensitivity was rated as: *poor* ($p_{\text{Median } t\text{-test}}<0.05$ or median $p_{TOST}>0.05$), *questionable* ($0.10>p_{\text{Median } t\text{-test}}>0.05$ or median $p_{TOST}<0.05$), *adequate* ($p_{\text{Median } t\text{-test}}>0.10$ and median $p_{TOST}<0.05$) or *excellent* ($p_{\text{Peak } t\text{-test}}>0.05$ and all $p_{TOST}<0.05$). Thus, the significance of the peak *t*-test was only considered in cases where a particular EMG-correction protocol showed evidence of adequate or excellent sensitivity using the median-based tests.

Results

Scalp

Effects prior to EMG correction

Visual inspection indicated that the topography of the four alpha-band contrasts was similar across the three NNNM protocols (Fig. 1).

Myogenic ROI. Consistent with expectation, scripted muscle tensing increased alpha power near facial muscles at midline-, left- and right-frontal electrodes. The myogenic contrast (OR-OT) was significant at 34–39 anterior electrodes (Fig. 2) with qualitatively similar peak locations across the three NNNM filters. The median *t*-scores ($ts=-2.30$ to -2.44, $ps<0.04$; $\eta_p^2s=0.24$ to 0.27) and extreme *t*-scores ($ts=-2.82$ to -3.21, $ps<0.01$; $\eta_p^2s=0.33$ to 0.39) were similar across the three protocols, indicating that the *degree* of EMG contamination was also similar. This contrast was used to form the Myogenic ROIs (Fig. 3), resulting in clusters of 23–25 contiguous anterior electrodes, extending to mid-frontal and fronto-central leads (e.g., AF7, F2, F3, F5, F7, FC2, FC3, FT7, T7).

Neurogenic ROI. Consistent with expectation, the Berger maneuver (OR-CR) altered power at all electrodes ($ts<-4.37$, $ps<0.001$). The peak difference occurred at midline parietal sites (Fig. 1), and defined the neurogenic ROIs (Fig. 3). The three neurogenic ROIs contained 21–26 contiguous electrodes (e.g., Pz, P1, P2, P3, P5, P7, P9, POz, PO3, PO4, PO8, Oz, CPz) with comparable median ($ts=-7.35$ to -7.49, $p<0.001$; $\eta_p^2s=0.77$ to 0.78) and extreme *t*-scores ($ts=-8.85$ to -9.12, $p<0.001$; $\eta_p^2s=0.83$ to 0.84), indicating that the strength of neurogenic effect was minimally affected by the choice of NNNM protocol.

Covarying effects. In the absence of EMG correction, myogenic activity distorted the magnitude of neurogenic effects in the alpha band. For instance, when changes in EMG and EEG negatively covaried (OT-CR), the effect remained significant at all electrodes; but, the magnitude of the neurogenic effect was significantly *attenuated* at 34–39 electrodes relative to the uncontaminated effect (OR-CR). This resulted in a slightly shifted topography that deemphasizes activation at anterior sites. Notably, significant attenuation was present at posterior electrodes far removed from the area of peak myogenic artifact (Fig. 4). A parallel, albeit non-significant, pattern of *amplification* occurred at anterior sites (not shown) when changes in EMG and EEG positively covaried (OR-CT). In this case, significant attenuation was only observed at a small number of parietal sites (Fig. 5).

Descriptive statistics for ICA

Classification. Fig. 6 depicts the relative frequency and percentage of scalp variance predicted by each class of components. A similar pattern was found using means. Visual inspection indicates that the most frequent classification, comprising about one-fifth of the total, was Noise (Fig. 6A). The frequency of the other classifications was somewhat smaller, but similar to one another (11–15%). Neuro-Dominant and Gross Artifact components were infrequent (2–3%). The frequency of the Myogenic and Myogenic-Dominant components

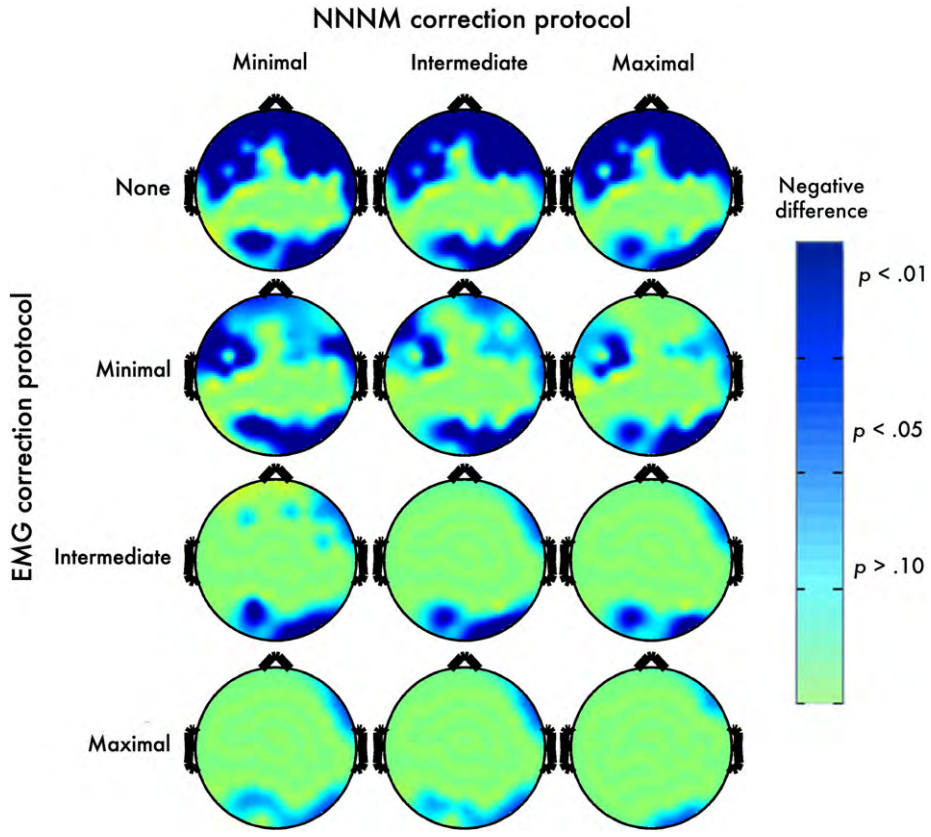


Fig. 2. Myogenic contrast (OR–OT) after EMG correction. Topographic maps depict thresholded *p*-values at each electrode after applying each method of non-neurogenic/non-myogenic (NNNM) artifact filtering and ICA-based EMG correction. Negative values are depicted in blue (dark-blue: $p<0.05$; light-blue: $p<0.10$; green: $p>0.10$). Note that the row labeled "None" depicts the thresholded OR-OT contrast from Fig. 1.

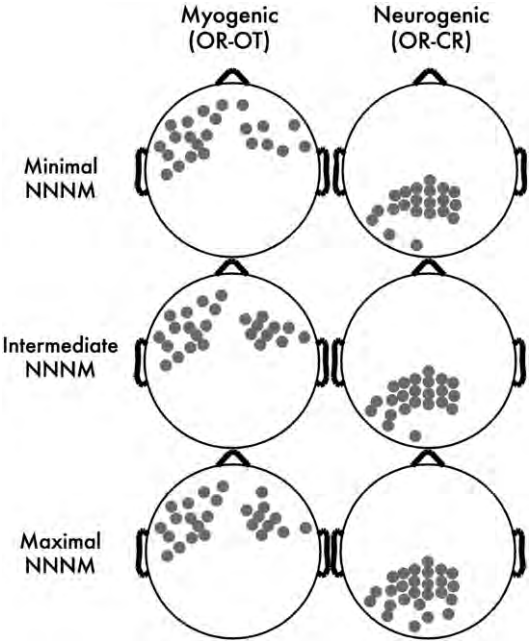


Fig. 3. Scalp regions of interest (ROIs). Gray circles depict electrodes included in the ROIs.

showed marked variability and was positively skewed. That is, a few individuals displayed many more of these two EMG-related components than the remainder of the group.

Collectively, the Neurogenic (25%) and Ocular components (21%) accounted for nearly half of the variance in scalp electrical activity (Fig. 6B). Myogenic, Myogenic-Dominant, and Noise components each accounted for another 8–10%, while the Neurogenic-Dominant, Low Variance, and Gross Artifact components each accounted for less than 2% of the variance. There was substantial variability and positive skew in the percentage of variance accounted for by several artifact components, particularly Ocular and Myogenic-Dominant. It is worth emphasizing that a sizable proportion of the variance—equal to that accounted for by the purely myogenic component—was predicted by the more heterogeneous Myogenic-Dominant component.

Fig. 7 depicts the median percentage of variance that was discarded or retained following each combination of NNNM filter and EMG correction protocol. The choice of NNNM filter and EMG correction protocol markedly affected the percentage of variance discarded. Application of the NNNM filters removed from one-quarter to slightly more than one-third of the variance in scalp electrical activity (Fig. 7). There was little difference between Intermediate- and Maximal-NNNM filters (center and right columns), presumably owing to the small net contribution of the Low Variance components. Application of the Minimal-EMG protocol removed another 10% of the variance (second row). When paired with the different NNNM filters, application of either the Intermediate- or Maximal-EMG protocols removed two-thirds to three-quarters of the variance (bottom two rows). Put another way, artifact accounted for about twice as much variance as neurogenic activity in this sample. The small difference between the Intermediate- and Maximal-EMG protocols (bottom two rows), presumably stems from the infrequency of Neuro-Dominant components.

Validity of ICA-based EMG correction

As summarized in Table 1, only four protocols showed questionable or better performance across all tests of sensitivity (Figs. 2, 4, and 5), specificity (Figs. 4, 5, and 8), and correction-induced artifact (Figs. 2 and 8): the Minimal-EMG protocol paired with Minimal- or Intermediate-NNNM filtering and the Maximal-EMG protocol paired with

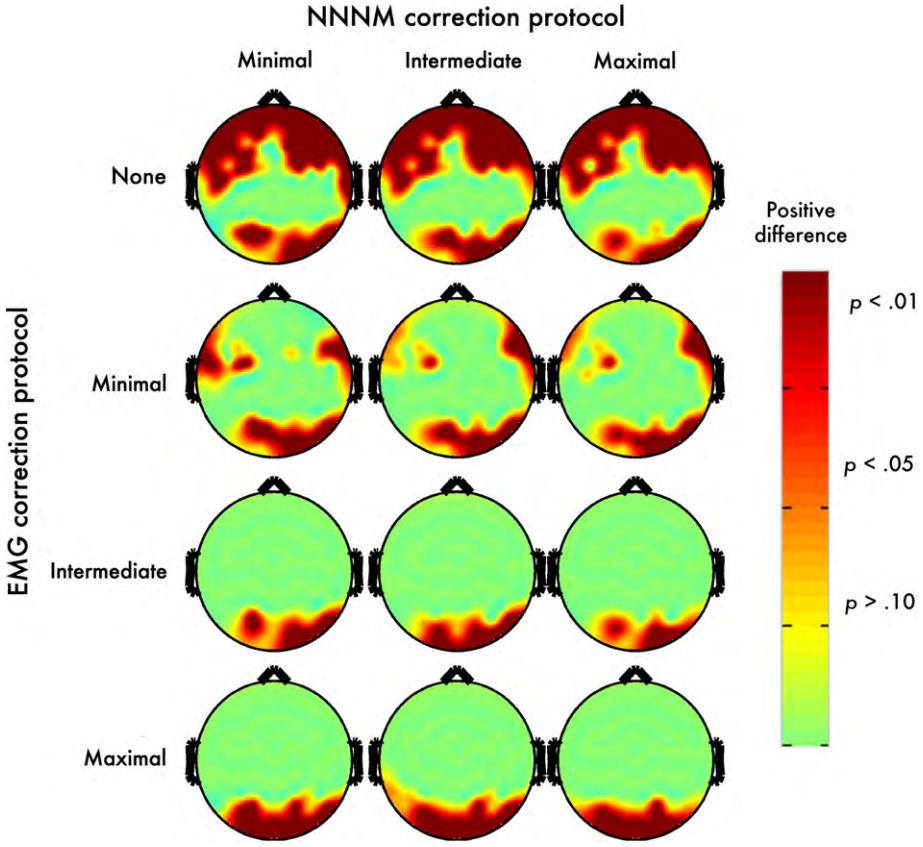


Fig. 4. Negatively-covarying contrast: post-correction error. Topographic maps depict *p*-values corresponding to the corrected OT-CR minus uncorrected OR-CR contrast. Positive values, indicating an increase in magnitude for the contrast, are shown in red (dark-red: $p<0.05$; light-red: $p<0.10$; yellow: $p>0.10$).

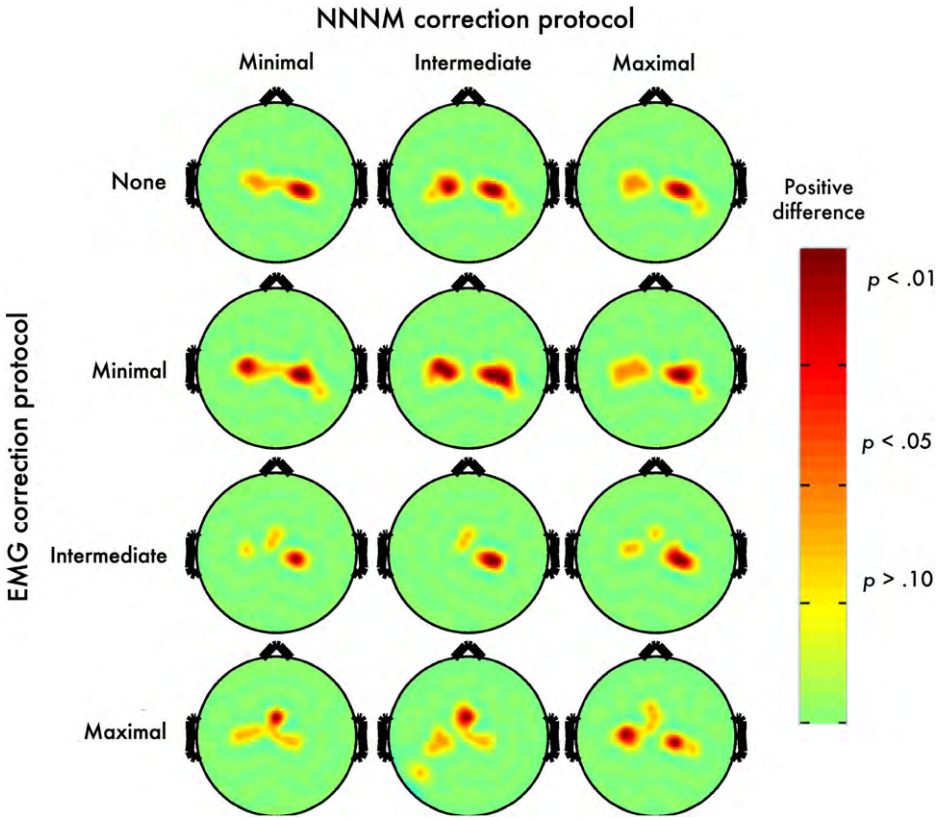


Fig. 5. Positively covarying contrast: post-correction error. Topographic maps depict *p*-values corresponding to the corrected OR-CT minus uncorrected OR-CR contrast. Positive values, indicating an increase in magnitude for the contrast, are shown in red (dark-red: $p<0.05$; light-red: $p<0.10$; yellow: $p>0.10$).

Minimal- or Maximal-NNNM filtering.[7] For detailed results, see Supplementary Tables 1–3.

Moreover, inspection of Table 1 indicates that among these four, the Minimal-EMG/Intermediate-NNNM protocol invariably equaled or exceeded the performance of the Minimal-EMG/Minimal-NNNM protocol; likewise, the Maximal-EMG/Maximal–NNNM protocol always outperformed the Maximal-EMG/Minimal-NNNM protocol. Accordingly, these two combinations of EMG correction and NNNM filtering were subjected to additional testing in the intracerebral source-space using LORETA.

Intracerebral source modeling

Effects prior to EMG correction

Myogenic ROI. Muscle tensing increased current density in a large number of frontopolar and ventral prefrontal voxels near the facial muscles (Fig. 9). Across correction protocols, the myogenic contrast (OR-OT) was significant at 717–789 voxels (~30% of the source-space). Myogenic ROIs (715–749 voxels) were created from this contrast after applying a cluster-extent threshold (Supplementary Fig. 14). The median *t*-scores ($ts=-3.75$ to -3.83, $ps<0.01$, $\eta_p^2=0.47$ to 0.48) and extreme *t*-scores ($ts=-6.19$ to -6.33, $ps<0.001$, $\eta_p^2=0.71$) were similar across the two filters.

Neurogenic ROI. The Berger maneuver was associated with widespread attenuation of current density across the posterior cortex (Fig. 9). Using the neurogenic (OR-CR) contrast, ROIs were generated for the Intermediate- and Maximal-NNNM filters (765–766 voxels) across posterior voxels (Supplementary Fig. 15). Median ($ts=-3.92$ to -3.96, $ps<0.01$, $\eta_p^2=0.49$) and extreme *t*-scores ($ts=-7.26$ to -7.68, $ps<0.001$, $\eta_p^2=0.77$ to 0.79) were similar across the two filters.

Covarying effects. The presence of uncorrected EMG artifact altered the magnitude of the neurogenic effects produced by the Berger maneuver (Fig. 9). In particular, when changes in neurogenic and myogenic activity negatively covaried (OT-CR) there was a substantial *attenuation* of effects in the myogenic ROI for both median ($ts>3.75$, $ps<0.01$, $\eta_p^2s>0.47$) and extreme *t*-scores ($ts>6.19$, $ps<0.001$, $\eta_p^2s>0.71$). The attenuation was reduced, but still observable, in the posterior neurogenic ROI (median $ps<0.08$, $\eta_p^2s>0.18$; extreme $ts>4.53$, $ps<0.01$, $\eta_p^2s>0.56$). Conversely, when changes in neurogenic and myogenic activity positively covaried (OR-CT), effects in the myogenic ROI were *amplified* as indexed by both the median ($ts<-2.30$, $ps<0.04$, $\eta_p^2s>0.25$) and extreme *t*-scores ($ts<-4.32$, $ps<0.001$, $\eta_p^2s>0.54$). These deleterious effects were weaker in the neurogenic ROI, reaching significance for the extreme ($ts<-3.81$, $ps<0.01$, $\eta_p^2s>0.48$) but not the median *t*-scores ($ps>0.72$, $\eta_p^2s<0.01$).

Sensitivity

Myogenic contrast (OR-OT). Although both protocols quantitatively reduced EMG contamination in the myogenic ROI (Fig. 10, red points), their sensitivity was poor (Supplementary Table 4, Supplementary Figs. 16–17).

Negatively covarying contrast (OT-CR). The Maximal-EMG/Maximal-NNNM pairing showed adequate sensitivity (Supplementary Table 4), whereas the Minimal-EMG/Intermediate-NNNM pairing evidenced poor sensitivity. Across protocols, voxels in the myogenic ROI showing weaker EMG contamination were more resistant to correction (Fig. 10).

[7] Exploratory analyses indicated dramatically worse performance in the "EMG" band (70–80 Hz). Prior to EMG correction, the myogenic contrast (OR-OT) in this frequency range was significant at 89–90 electrodes across NNNM filters ($p<.01$ threshold). Following correction, it remained significant at 68–86 electrodes.

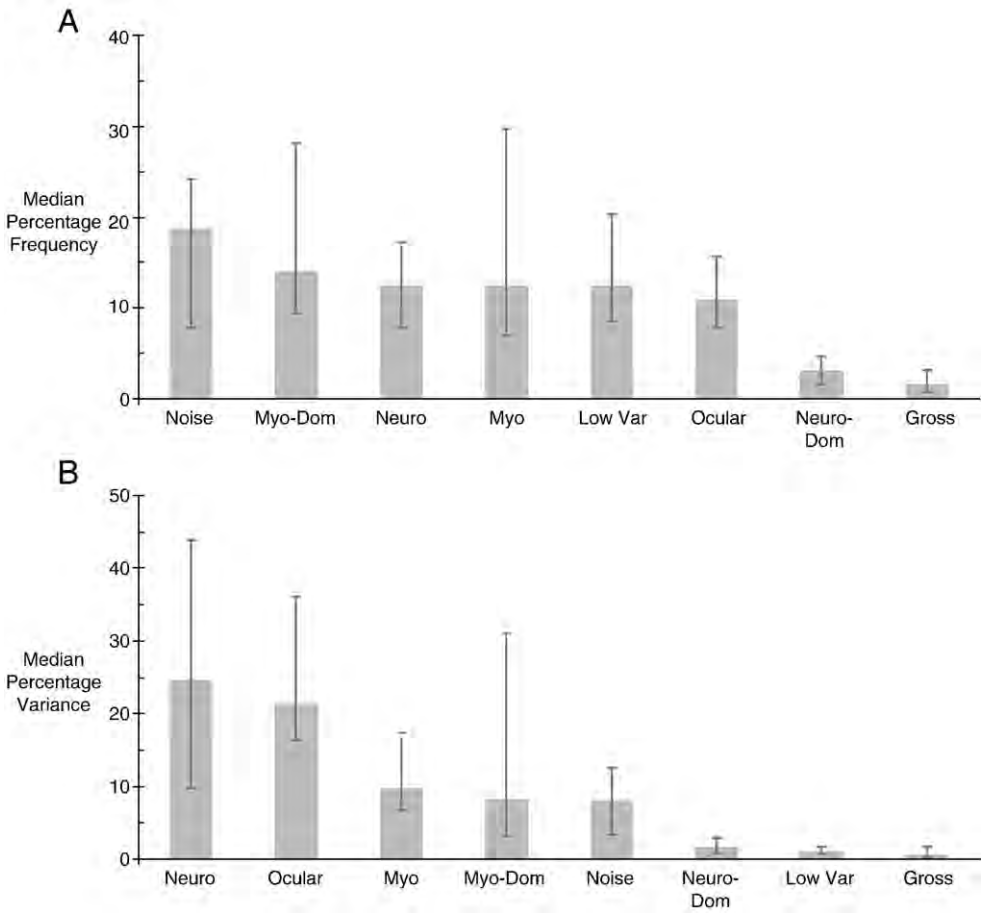


Fig. 6. (A) Median frequency and (B) variance (in percent) accounted for by each class of components. Frequencies and percentages were computed within participants. Medians were then computed across participants. Error bars indicate the 25th and 75th percentiles.

Positively covarying contrast (OR-CT). Again, the Maximal-EMG/Maximal-NNNM pairing demonstrated adequate sensitivity, whereas the Minimal-EMG/Intermediate-NNNM pairing showed poor sensitivity (Supplementary Table 4). Inspection of the myogenic ROI voxels that yielded peak errors indicated that the Minimal-EMG/Intermediate-NNNM pairing tended to undercorrect the data (negative *t*-scores), whereas the Maximal-EMG/Maximal-NNNM pairing tended to overcorrect the data (positive *t*-scores), albeit to a lesser degree. This is depicted in Figs. 8 and 10.

Specificity

Neurogenic contrast (OR-CR contrast). The Maximal-EMG/Maximal-NNNM pairing showed questionable specificity, whereas the Minimal-EMG/Minimal-NNNM pairing showed adequate specificity (Fig. 10, blue points, Supplementary Table 5, Supplementary Figs. 18–19).[8] Inspection of the neurogenic ROI voxels that yielded peak errors indicated that in the absence of EMG artifact both protocols attenuated changes in activation associated with the Berger maneuver. This was particularly evident for the Maximum-EMG protocol paired with Minimal-NNNM filtering.

Negatively covarying contrast (OT-CR). Both protocols tended to attenuate neurogenic activity, yielding questionable or worse specificity (Supplementary Table 5 and Fig. 10).

Positively covarying contrast (OR-CT). Both pairings exhibited acceptable or better specificity (Supplementary Table 5). Inspection of the neurogenic ROI voxels that yielded peak errors indicated that the Minimal-EMG/Intermediate-NNNM pairing led to undercorrection (negative *t*-scores), whereas the Maximal-EMG/Maximal-NNNM pairing led to overcorrection (positive *t*-scores). This is depicted in Fig. 10.

Correction artifact

Myogenic contrast (OR-OT) in the neurogenic ROI. The Maximal-EMG/Maximal-NNNM pairing yielded an acceptable amount of correction-induced artifact when EMG was present (Supplementary Table 6 and Fig. 10, blue points), whereas the Minimal-EMG/Intermediate-NNNM pairing showed poor performance.

Neurogenic contrast (OR-CR) in the myogenic ROI. Both pairings showed an acceptable level of correction artifact when EMG was absent (Supplementary Table 6 and Fig. 10, red points).

Discussion

Given the substantial inferential threat posed by EMG contamination, there is a pressing need for valid correction tools. In recent years, ICA has rapidly become a popular tool for correcting EMG artifact, despite limited work assessing its validity for this purpose.

[8] We interpreted differences in the size of the neurogenic contrast as evidence of non-specificity. Nonetheless, such differences might instead reflect amplification of the neurogenic statistical effect ("signal-to-noise") following EMG correction. This would occur if EMG correction led to reductions in error ("noise") that were substantially larger than reductions in the mean difference across conditions ("signal"). If this were so, the magnitude of the *t*-test following EMG correction should increase. Contrary to this alternative account, the *t*-test for the neurogenic contrast was decreased following Maximal-EMG correction (Maximal-NNNM: $t = -3.92$; Maximal-EMG/Maximal-NNNM: $t = -3.70$).

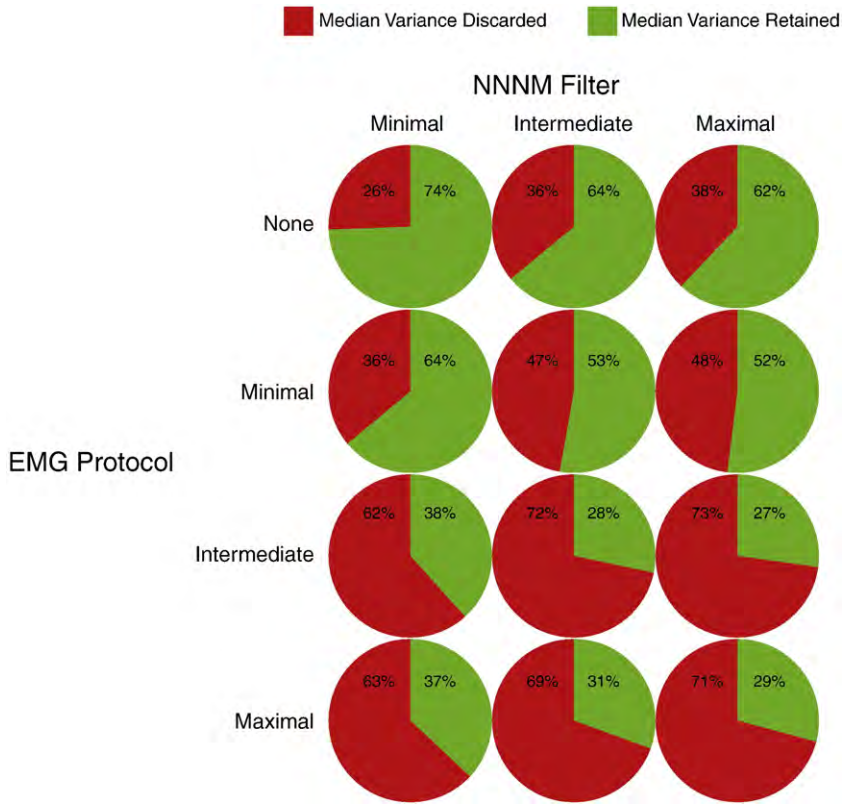


Fig. 7. Median variance retained and discarded for each combination of NNNM filter and EMG correction protocol.

Accordingly, we quantitatively tested the sensitivity and specificity of ICA-based EMG correction using a naturalistic dataset in which neurogenic (opening and closing the eyes, i.e. the "Berger maneuver") and myogenic activity (scripted muscle tensing and quiescence) were independently manipulated. This design allowed investigation of ICA-based EMG correction under conditions in which changes in neurogenic and myogenic activity covaried, as they do in many experimental settings.

Low-intensity clenching of the face increased EEG spectral power across much of the scalp, extending as far as the fronto-central electrodes. Smaller pockets of myogenic activity were also present at the posterior edge of the electrode array (Fig. 1). In contrast, the Berger maneuver altered power across the entire array, peaking at midline parietal sites. And while the effect-size of the anterior myogenic effect was only one-third that associated with the posterior neurogenic effect, it was sufficient to alter the magnitude of neurogenic effects. For instance, when changes in EMG and EEG negatively covaried, as they would be expected to do in many experiments, neurogenic effects associated with the Berger maneuver were attenuated at electrodes across the array, including posterior sites far removed from the area of peak myogenic activity (Figs. 1 and 4).

Consistent with these findings, uncorrected EMG artifact increased alpha-band current density across large regions (~30%) of the intracerebral source-space located near the facial muscles (i.e., frontopolar and ventral prefrontal cortex, insula, temporal poles; Fig. 9). This effect reflects the fact that LORETA-Key and many other distributed modeling packages restrict the source space to the cerebral cortex. Consequently, activity generated in the cranial muscles tends to be explained by dipoles fitted to the proximal region of cortex. The Berger maneuver was associated with attenuated current density across the posterior cortex (Fig. 9). Furthermore, covariation in neurogenic and myogenic activity significantly altered the magnitude of neurogenic effects in anterior regions of the brain (Fig. 9). Similar, albeit less dramatic, distortions were found in posterior regions, particularly for negatively covarying activity.

Table 1
Validity ratings on the scalp.

	EMG correction	NNNM filtering	Myogenic[a]	Neurogenic[b]	Negatively-covarying[c]	Positively-covarying[d]	Lowest rating
Myogenic/neurogenic ROI[e]	Minimal	Minimal	?/+	+/+	?/+	++/++	?
		Intermediate	?/+	+/++	+/+	++/++	?
		Maximal	?/+	+/+	+/−	++/++	−
	Intermediate	Minimal	+/+	−/?	++/−	++/++	−
		Intermediate	−/+	?/+	++/−	++/+	−
		Maximal	+/−	−/+	++/−	++/++	−
	Maximal	Minimal	+/+	+/+	++/+	++/++	+
		Intermediate	+/+	+/+	++/−	++/+	−
		Maximal	++/++	+/+	++/+	++/++	+

[a] Results from the myogenic contrast (corrected OR–OT) in the myogenic and neurogenic ROI.
[b] The corrected OR–CR vs. Uncorrected OR–CR contrast.
[c] Corrected OT–CR vs. uncorrected OR–CR.
[d] Corrected OR-CT vs. uncorrected OR-CR.
[e] Ratings from each ROI: − *poor*, ? *questionable*, + *adequate*, ++ *excellent* (see Methods).

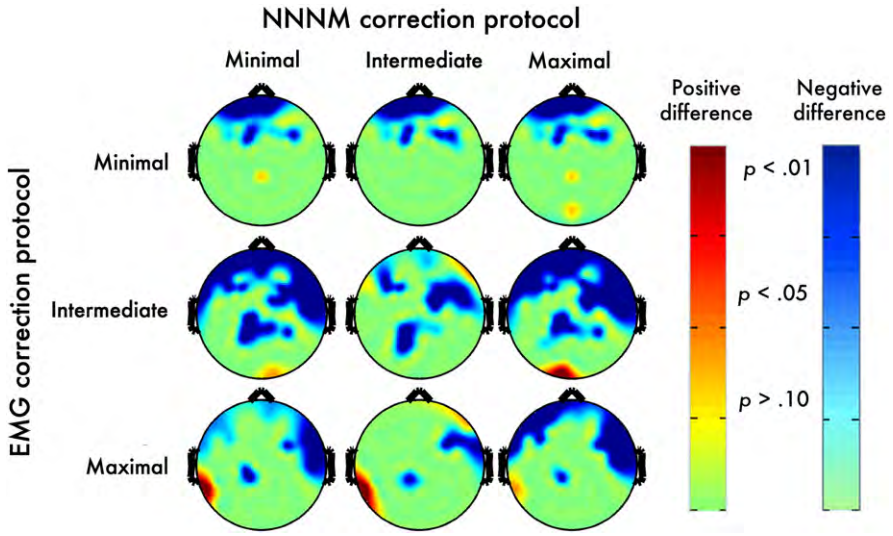


Fig. 8. Neurogenic contrast: post-correction error. Topographic maps depict thresholded *p*-values corresponding to the corrected OR–CR minus uncorrected OR–CR contrast. Negative values, indicating an increase in magnitude for the neurogenic contrast, are shown in blue (dark-blue: $p<0.05$; light-blue: $p<0.10$; green: $p>0.10$). Positive values, indicating correction-induced magnitude reduction of the neurogenic contrast, are shown in red (dark-red: $p<0.05$; light-red: $p<0.10$; yellow: $p>0.10$).

From these data, independent components were extracted and classified for each participant. Neurogenic components typically accounted for the most variance in scalp electrical activity (median: 25%), whereas those classified as entirely or predominantly myogenic each accounted for another 8–10%. Participants showed pronounced variability in the amount of variance accounted for by components exhibiting characteristics of both myogenic and neurogenic activity (interquartile range for "myogenic-dominant" components: 3–33%; see Fig. 6). Descriptively, the choice of which components to exclude had a marked impact on the amount of variance retained for analysis (Fig. 7). Nine different protocols for determining which components to discard were examined, reflecting the factorial pairing of three for removing non-neurogenic/non-myogenic (NNNM) components with three for removing EMG. NNNM filtering removed between one-quarter and one-third of the total variance and EMG correction removed between a tenth and one-third of the total. Together, these two filters led to the exclusion of as little as one-third to as much as three-quarters of the variance in scalp electrical activity.

Not surprisingly, the validity of these protocols was also quite variable. In the present study, *sensitivity* (i.e., attenuation of myogenic effects), *specificity* (i.e., preservation of neurogenic effects), and *correction artifacts* (i.e., generation of effects in the absence of artifact) were each quantitatively assessed using ROIs corresponding to areas of peak myogenic and neurogenic activation. Results indicated that most of the nine protocols did a reasonable job removing EMG artifact, evidenced by adequate or excellent sensitivity (Table 1). Unfortunately, many of these pairings also altered neurogenic activity, evidenced by inadequate specificity or excessive correction-induced artifact (Table 1). In fact, only two pairings—Maximal-EMG correction combined with Minimal- or Maximal-NNNM filtering—showed adequate or excellent performance across all three measures. Furthermore, the Maximal-EMG/Maximal-NNNM pairing tended to outperform the Maximal-EMG/Minimal-NNNM pairing. None of the nine procedures consistently displayed excellent performance.

On the scalp, the most sensitive and specific procedure for removing EMG artifact from the alpha band was among the strictest, indexed by the amount of variance discarded (median: 71%). That is, the Maximal-EMG/Maximal-NNNM pairing entailed rejecting any component containing myogenic signal—including those where myogenic activity was *less* prominent than neurogenic activity—in addition to those indexing gross or ocular artifacts, noise, or unclassifiable low-variance signals. The only other procedure that consistently showed adequate performance—the Maximal-EMG/Minimal-NNNM pairing—was similarly strict (median variance discarded: 63%), differing only in the retention of noise and unclassifiable components.

In contrast to the scalp analyses, Maximal-EMG correction paired with Maximal-NNNM filtering was associated with inadequate performance in the intracerebral source-space (Fig. 10 and Supplementary Tables 4–6). In particular, it failed to adequately remove EMG when neural activity was fixed and overcorrected neurogenic activity when EMG was absent. Poor sensitivity was also evident for the other procedures examined in the source-space (i.e., Minimal-EMG/Intermediate-NNNM).

Prospects for ICA-based EMG correction

As noted in the Introduction, many studies have documented, with varying degrees of rigor, the utility of infomax ICA for attenuating various physical and biological artifacts. Nevertheless, several studies have found ICA to exhibit worse performance than alternative source separation algorithms for ocular (Wallstrom et al., 2004; Romero et al., 2008) and EMG artifacts (Crespo-Garcia et al., 2008; Fitzgibbon et al., 2007). More recently, Debener et al. (2007) showed that ICA displays low specificity for ballistocardiogram artifacts, evidenced by attenuation of event-related neurogenic activity, under some circumstances (Debener et al., 2008). Paralleling Debener and colleagues' research, the present findings suggest that ICA is a valid means of correcting EMG in some, but not all, cases. On the scalp, some of the ICA-based protocols we tested displayed adequate sensitivity and specificity, whereas other did not. Furthermore, in the intracerebral source-space, even those protocols that showed the most promising performance on the scalp failed.

The inadequate performance of ICA in the source-space likely reflects two factors. First, source modeling, implemented here using the LORETA algorithm and a three-shell head model, makes use of data from *all* electrodes in the array, not just those in the scalp ROIs. Less-than-perfect EMG correction at even a modest number of electrodes in- or outside of the scalp ROIs could, therefore, exert a substantial influence on the EEG source model. As such, source modeling can be viewed as providing a "global" check on the quality of EMG correction performed on the scalp, complimenting the more "local" ROI analyses.

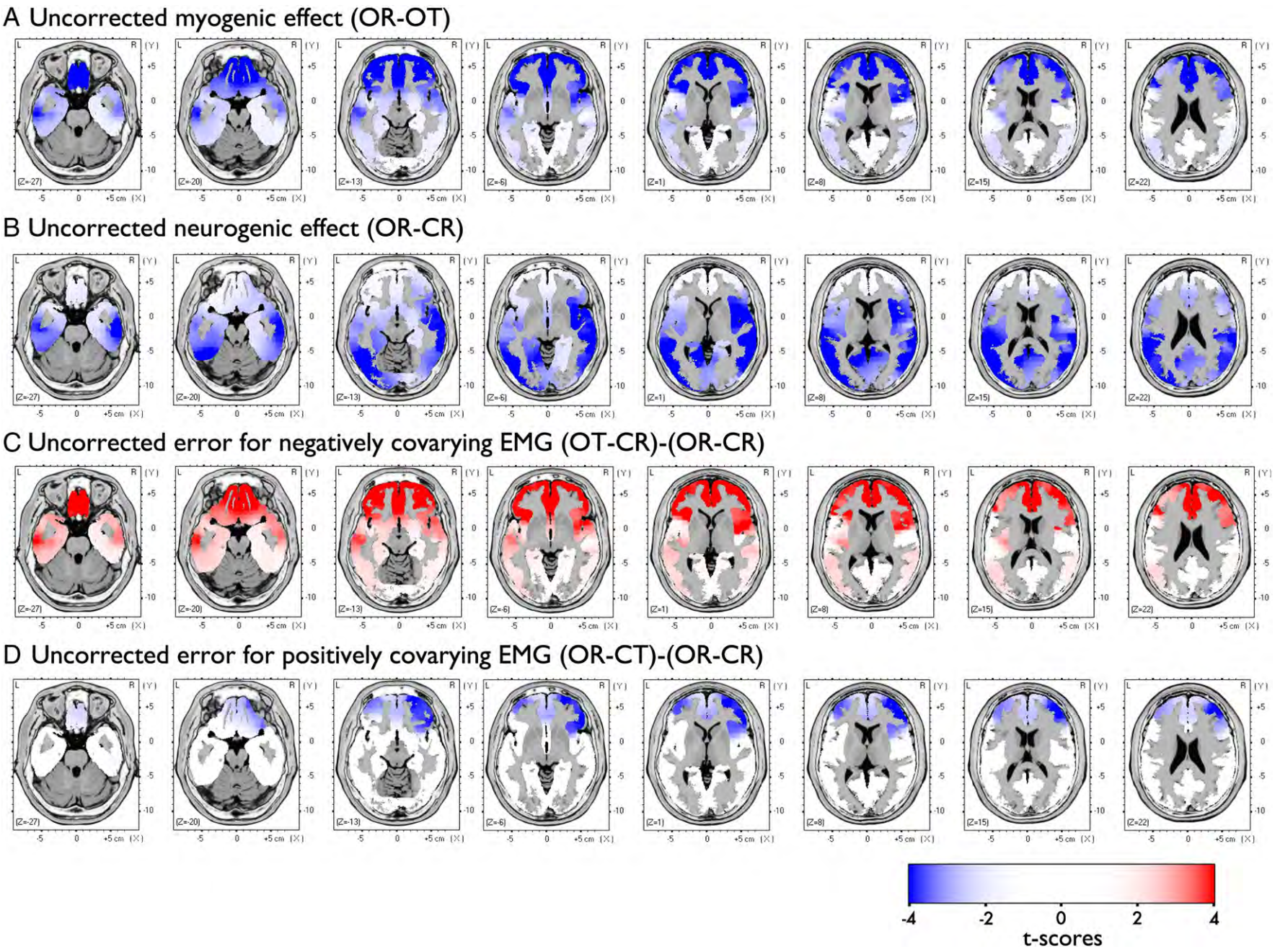


Fig. 9. Contrasts of interest in the source-space after applying the Intermediate-NNNM filter. (A) Myogenic contrast (OR–OT), (B) neurogenic contrast (OR–CR), (C) error induced by negatively-covarying artifact prior to EMG correction ([OT–CR] − [OR–CR]), and (D) error induced by positively-covarying artifact ([OR–CT] − [OR–CR]) prior to EMG correction.

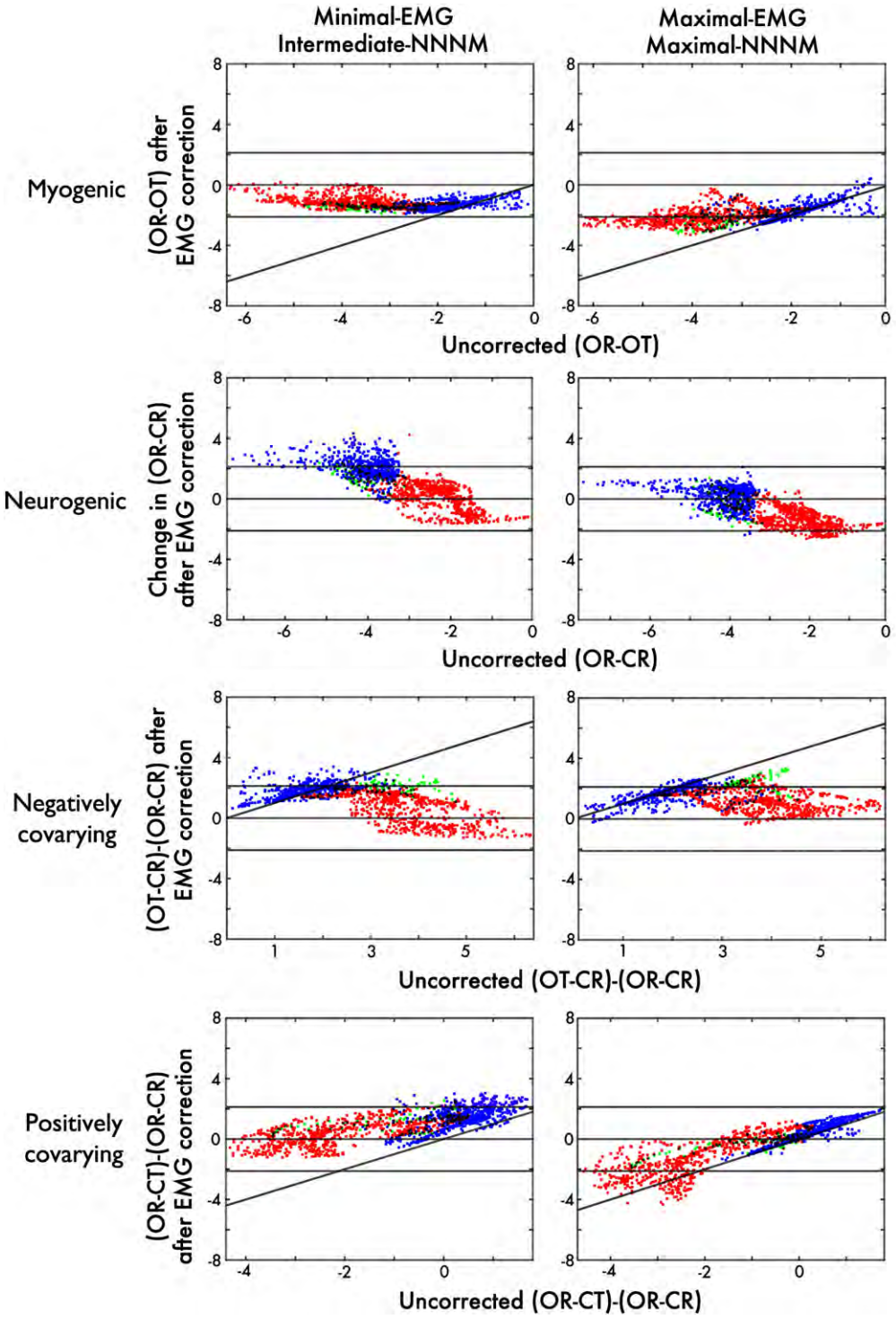


Fig. 10. Source-space effects of EMG correction. Uncorrected (*x*-axis) compared to ICA-corrected *t*-values (*y*-axis) for voxels in the source-space, color-coded by ROI (blue: neurogenic ROI; red: myogenic ROI; green: both ROIs). Points close to the solid horizontal line are voxels where *t* approached zero after correction, indicating high sensitivity for EMG-contaminated contrasts and low specificity for the EMG-free neurogenic contrast. Conversely, points lying beyond the broken horizontal lines ($p = 0.05$) were significantly altered by correction. Points along the diagonal were unchanged by correction, indicating low sensitivity for EMG-contaminated contrasts and high specificity for the neurogenic contrast. Note that the second row plots the uncorrected OR–CR contrast (*x*-axis) against the change in OR–CR after correction (i.e. correction induced error; *y*-axis).

Second, several lines of evidence suggest that ICA failed to adequately separate myogenic from neurogenic sources on the scalp, causing under-correction (low sensitivity) and over-correction (low specificity) in the source-space. In particular, the amount of variance accounted for by "mixed" components, those displaying characteristics of both neurogenic and myogenic activity, was equivalent (e.g., myogenic-dominant vs. pure myogenic components, Fig. 6). If ICA had cleanly separated the two sources, one would instead expect the mixed components to be infrequent and to account for little variance.

Quantitatively, none of the nine protocols consistently showed excellent performance on the scalp. That is, a reasonably small number of "worst-case" electrodes located *inside* of the scalp ROIs (Fig. 3) evinced under- or overcorrection (Supplementary Tables 1–3).

Qualitatively, visual inspection of the topographic maps created following EMG correction indicated significant distortions *outside* of the scalp ROIs. For instance, significant residual artifact was present at the edge of the array for the corrected myogenic contrast depicted in Fig. 2. Likewise, evidence of overcorrection (blue regions) and, to a lesser degree, undercorrection (red regions) was present outside of the defined ROIs when neurogenic and myogenic activity covaried (Figs. 4, 5).

It is worth emphasizing that these observations cannot be attributed to limitations in the procedure for classifying or rejecting components. The manual classification protocol was detailed (see Supplementary Method), and raters were extensively trained and highly reliable in its application ($\alpha = 0.98$). In contrast to prior studies, the impact of systematically varying criteria for rejecting both non-myogenic and myogenic components was examined. Of the nine protocols examined, only four consistently showed questionable or better performance on the scalp, and none did so in the source-space. Systematic biases in the classification or rejection of components cannot explain the combination of low sensitivity and specificity in the source-space, whereas a failure to cleanly separate myogenic and neurogenic sources does.

Two factors could plausibly account for the inability of infomax ICA to fully separate myogenic from neurogenic sources. Systematically testing these hypotheses represents a key challenge for future research. First, inadequate separation might reflect non-optimal specification of the number of components to extract ("model order"). Like most other high-resolution studies, a combination of PCA and ICA was used to extract fewer components, 64, than the limit imposed by the number of electrodes in the array, 128. As detailed in Footnote 3, the number of components was subjectively optimized using a trial-and-error approach. Model order was also constrained to be identical across participants. Notably, specifying too many components or too few (i.e., over- or under-fitting) might explain poor source separation (Naeem et al., 2009; Ryali et al., 2009). This possibility could be tested using an information theoretic approach to objectively identify the optimal model order for each participant, as is typical in the functional magnetic resonance imaging literature (Beckmann and Smith, 2004; Calhoun et al., 2001; Li et al., 2007) and has occasionally been done in the EEG literature (Moraux and Iannetti, 2009; see also Supplementary Method and Results). Alternatively, the deflationary approach implemented in the *fastica* package (http://www.cis.hut.fi/projects/ica/fastica) could be used (Mantini et al., 2008). Stepwise ICA fitting, in which model order and sources are estimated simultaneously, is another possible approach (Hesse and James, 2004). Procedures for optimizing model order on an individual basis would also facilitate the development of criteria for discarding problematic participants (e.g., those with an unusually large number of components). And, from a more practical perspective, application of such procedures would reduce the need to classify large numbers of closely related or uninformative components.

Second, inadequate separation could reflect EMG violating key assumptions of the infomax ICA algorithm (Bell and Sejnowski, 1995; James and Hesse, 2005). For instance, infomax assumes that the activation of each component is sparse and non-Gaussian. The lengthy blocks of neurogenic and myogenic activation in the present study—and in many studies of emotion (e.g., Coan and Allen, 2003; Davidson et al., 1990; Shackman et al., 2006)—may not adequately satisfy this assumption. Infomax also assumes that sources are mutually temporally independent. To the degree that neurogenic and myogenic activity are too closely coupled in the time-domain they would violate this assumption (but cf. Ohla et al., 2009). Future studies could test the degree to which second-order blind source separation algorithms, that do not require strong assumptions, such as Second Order Blind Identification (SOBI) or Algorithm for Multiple Unknown Signals Extraction (AMUSE), produce better separation (Joyce et al., 2004; Romero et al., 2008; Tang et al., 2005). It might also be fruitful to investigate the utility of signal-space projection methods (Nolte and Curio, 1999; Tesche et al., 1995; Uusitalo and Ilmoniemi, 1997) or ICA performed in the frequency-domain, which may improve the separation of sources with overlapping spectral profiles ((Annemuller et al., 2003, 2004; Lee et al., 2008). Finally, it will be important to examine the degree to which these conclusions generalize to paradigms characterized by punctuate bursts of time-locked activation, as is typical of ERSP studies.

GLM-based EMG correction

The present results would seem to partially contradict prior reports (McMenamin et al., 2009; Shackman et al., 2009) in which the use of one form of GLM-based EMG correction, "epoch-wise regression," was recommended for scalp analyses. Epoch-wise regression removes epoch-to-epoch variance in alpha-band activity predicted by contemporaneous EMG-band activity (e.g., 70–80 Hz) separately for each electrode and participant. To facilitate a more direct comparison of GLM- and ICA-based correction techniques, the validity of epoch-wise regression was re-examined in the present report using the identical methods used for testing ICA (see Supplementary Method and Results). Consistent with McMenamin et al (2009), epoch-wise regression exhibited adequate performance across nearly all validation tests on the scalp (Supplementary Tables 7–9), with the lone exception of poor specificity in the presence of positively covarying neurogenic activity. The performance discrepancy between these reports is likely to be a consequence of using much larger neurogenic ROIs for the present analyses (i.e., 23 vs. 7 electrodes). Larger ROIs were employed with the aim of indexing the impact of EMG artifact and EMG correction in regions characterized by less extreme signals (myogenic and neurogenic) with the idea that such signals would be more representative of real-world changes in spectral activity. By contrast, the small, highly focused neurogenic ROI used by McMenamin et al (2009) was blind to distortions outside of areas of peak neurogenic activity.

Future challenges

Four limitations of the present study represent additional avenues for future research. First, the impact of EMG artifact and EMG correction on *individual differences* in state or trait brain electrical activity remains unknown. In particular, the degree to which either ICA- or GLM-based correction techniques preserve hemispheric asymmetries in tonic ("resting") frontal activity, a neural marker of individual differences in emotional reactivity (Shackman et al., in press; Coan et al., 2006) and affective disorders (Thibodeau et al., 2006) remains untested.

Second, our conclusions derive from analyses of the alpha band (8–13 Hz) during extended blocks of myogenic activity. While this represents a reasonable analog to studies using blocked manipulations of emotion (e.g., threat of shock, emotional films), the degree to which these conclusions generalize to event-related designs or other frequency bands is unclear. Still, it seems likely that the quality of performance for the neighboring theta band (4–8 Hz) would be similar to that observed for alpha. In contrast, we anticipate that the quality of correction would be lower for bands, such as gamma (>30 Hz), that lie closer to peak EMG activity. Indeed, exploratory analyses indicated poor performance for ICA in the 70–80 Hz range (Footnote 5). Finally, the degree to which the present conclusions generalize to stronger or weaker EMG contamination is unknown (Fitzgibbon et al., 2007).

Third, source modeling was not incorporated into the component classification protocol. Classifications were instead based on visual inspection of time-series, power spectrum, and topography (see Supplementary Method and Results). While this is a conventional

approach, it is possible that inspection of component dipoles would have facilitated more accurate classifications, particularly in the case of "mixed" components (e.g., Myogenic dominant). More specifically, dipoles could be modeled using the *dipfit2* or *besafit* plug-ins for EEGLAB. Components characterized by dipoles at the edge of or outside of the brain could then be classified as artifactual (myogenic or otherwise; Onton and Makeig, 2006; Milne et al., 2009).

Fourth, like nearly all prior validation studies, the tests of specificity are founded on the assumption that myogenic activity was absent when participants were instructed to relax (Shackman et al., 2009). If, in fact, modest amounts of EMG were present during this condition, estimates of specificity would be artifactually reduced. We consider this a minor concern; the location of the neurogenic ROI (~Pz) for testing specificity on the scalp (Fig. 3) should minimize contributions from the anterior, lateral, and posterior muscle groups (Supplementary Fig. 1) when myogenic activity is weak. Nevertheless, it would be informative to validate ICA and other EMG correction techniques using data obtained during neuromuscular blockade (Whitham et al., 2007).

Recommendations and conclusions

Consideration of these observations and the extant literature yields several recommendations. First, if widespread EMG artifacts are suspected, Maximal-EMG correction protocol with Maximal-NNNM filtering should be employed. This method exhibited the best combination of sensitivity and specificity across all tests (Table 1). It also outperformed GLM-based EMG correction, which never demonstrated excellent performance in any of our validation tests (Supplementary Tables 7–9). Second, given its merely adequate sensitivity and inconsistent specificity, we no longer recommend the use of GLM-based EMG correction techniques (cf. McMenamin et al., 2009) for studies characterized by widespread EMG artifact. Third, for investigations where specificity is a smaller concern than sensitivity, any of the Intermediate- or Maximal-EMG correction ICA-based protocols represent reasonable choices. Fourth, the use of distributed modeling techniques, such as LORETA, to estimate the intracerebral sources of spectral EEG in studies with prominent myogenic activity is not recommended. It remains to be seen whether it is reasonable to do so using other approaches, such as dipoles or beamformers (Michel et al., 2004; Nazarpour et al., 2008). Fifth, the results of the present study and several others (e.g., Whitham et al., 2007; Yuval-Greenberg et al., 2009a,b; Shackman et al., 2009) indicate that findings in the upper frequency bands (i.e., beta: 14–13 Hz; gamma: >30 Hz), should be interpreted with extreme caution, particularly when they occur in the vicinity of scalp muscles. At minimum, plots depicting the scalp topography and spectral character of the results should be presented in sufficient detail to allow readers to independently assess whether the phenomenon in question is neurogenic (Shackman, in press). And while ICA cannot be viewed as a panacea for EMG contamination, careful application of ICA or related techniques for attenuating EMG artifact represents a useful means of rejecting the most dubious results.

Recent years have witnessed a resurgence of interest in using scalp-recorded and source-modeled EEG to answer fundamental questions about the neural implementation of cognitive and affective processes (Makeig et al., 2004; Pizzagalli, 2007). The continued development and careful validation of more sophisticated techniques, such as frequency-domain ICA, for separating myogenic from neurogenic signals will have substantial benefits for this important endeavor.

Acknowledgments

The first two authors (BWM and AJS) contributed equally to this research. We thank Alanna Clare, Donna Cole, Isa Dolski, Andre Mouraux, and Aaron Teche for assistance and three anonymous reviewers for thoughtful comments. This work was supported by the NIMH (P50-MH069315 and R37/R01-MH43454 to RJD; BWM was supported by T32-HD007151).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.10.010.

References

Allen, J.J.B., Coan, J.A., Nazarian, M., 2004. Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. Biol. Psychol. 67, 183–218.
Anemuller, J., Sejnowski, T.J., Makeig, S., 2003. Complex independent component analysis of frequency-domain electroencephalographic data. Neural Netw. 16, 1311–1323.
Anemuller, J., Sejnowski, T.J., Makeig, S., 2004. Reliable measurement of cortical flow patterns using complex independent component analysis of electroencephalographic signals. Lect. Notes Comput. Sci. 3195, 1009–1016.
Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans. Med. Imaging 23, 137–152.
Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 7, 1004–1034.
Berger, H., (1929/1969). On the electroencephalogram of man. The fourteen original reports on the human electroencephalogram (P. Gloor, Trans.). NY: Elsevier.
Bonnett, M.H., Arand, D.L., 2001. Impact of activity and arousal upon spectral EEG parameters. Physiol. Behav. 74, 291–298.
Borden, J.W., Peterson, D.R., Jackson, E.A., 1991. The Beck anxiety inventory in nonclinical samples: Initial psychometric properties. J. Psychopathol. Behav. Assess. 13, 345–356.
Bradley, M.M., Codispoti, M., Cuthbert, B.N., Lang, P.J., 2001. Emotion and motivation I: Defensive and appetitive reactions in picture processing. Emotion 1, 276–298.
Brooker, B.H., Donald, M.W., 1980. Contribution of the speech musculature to apparent human EEG asymmetries prior to vocalization. Brain Lang. 9, 226–245.
Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp. 14, 140–151.
Carver, C.S., Harmon-Jones, E., 2009. Anger is an approach-related affect: Evidence and implications. Psychol. Bull. 135, 183–204.
Chung, J.W., Kim, C., McCall Jr, W.D., 2002. Effect of sustained contraction on motor unit action potentials and EMG power spectrum of human masticatory muscles. J. Dent. Res. 81, 646–649.
Coan, J.A., Allen, J.J.B., 2003. The state and trait nature of frontal EEG asymmetry in emotion. The asymmetrical brain. MIT Press, Cambridge, MA, pp. 565–615.
Coan, J.A., Allen, J.J.B., McKnight, P.E., 2006. A capability model of individual differences in frontal EEG asymmetry. Biol. Psychol. 72, 198–207.
Cohen, B.H., Davidson, R.J., Senulis, J.A., Saron, C.D., Weisman, D.R., 1992. Muscle tension patterns during auditory attention. Biol. Psychol. 33, 133–156.
Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18, 192–205.
Crespo-Garcia, M., Atienza, M., Cantero, J.L., 2008. Muscle artifact removal from human sleep EEG by using independent component analysis. Ann. Biomed. Eng. 36, 467–475.
Davidson, R.J., Ekman, P., Saron, C.D., Senulis, J.A., Friesen, W.V., 1990. Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. J. Pers. Soc. Psychol. 58, 330–341.
Davidson, R.J., Jackson, D.C., Larson, C.L., 2000. Human electroencephalography. Handbook of psychophysiology, 2nd ed. Cambridge University Press, NY, pp. 27–52.
Davidson, R.J., Shackman, A.J., Maxwell, J.S., 2004. Asymmetries in face and brain related to emotion. Trends Cogn. Sci. 8, 389–391.
Debener, S., Strobel, A., Sorger, B., Peters, J., Kranczioch, C., Engel, A.K., Goebel, R., 2007. Improved quality of auditory event-related potential recorded simultaneously with 3T fMRI: Removal of the ballistocardiogram artefact. NeuroImage 34, 587–597.
Debener, S., Mullinger, K.J., Niazy, R.K., Bowtell, R.W., 2008. Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 Tesla. Int. J. Psychophysiol. 67, 189–199.
De Clercq, W., Vergult, A., Vanrumste, B., Van Hees, J., Palmini, A., Van Paesschen, W., et al., (2005). A new muscle artifact removal technique to improve the interpretation of the ictal scalp electroencephalogram. 27th Annual International Conference of the Engineering in Medicine and Biology Society (IEEE-EMBS 2005), 944-947.
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, 9–21.
Delorme, A., Sejnowski, T., Makeig, S., 2007a. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 34, 1443–1449.
Delorme, A., Westerfield, M., Makeig, S., 2007b. Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. J. Neurosci. 27, 11949–11959.
DeRubeis, R.J., Siegle, G.J., Hollon, S.D., 2008. Cognitive therapy versus medication for depression: Treatment outcomes and neural mechanisms. Nat. Rev. Neurosci. 9, 788–795.

Dien, J., 1998. Issues in the application of the average reference: Review, critiques, and recommendations. Behav. Res. Methods Instrum. Comput. 30, 34–43.
Dimberg, U., Thunberg, M., Elmehed, K., 2000. Unconscious facial reactions to emotional facial expressions. Psychol. Sci. 11, 86–89.
Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M., 1993. 3D statistical neuroanatomical models from 305 MRI volumes. Nuclear Science Symposium and Medical Imaging Conference 1813–1817.
Fava, J.L., Velicer, W.F., 1996. The effects of underextraction in factor and component analyses. Educ. Psychol. Meas. 56, 907–929.
Fitzgibbon, S.P., Powers, D.M., Pope, K.J., Clark, C.R., 2007. Removal of EEG noise and artifact using blind source separation. J. Clin. Neurophysiol. 24, 232–243.
Flexer, A., Bauer, H., Pripfl, J., Dorffner, G., 2005. Using ICA for removal of ocular artifacts in EEG recorded from blind subjects. Neural Netw. 18, 998–1005.
Frank, R.M., Frishkoff, G.A., 2006. Automated protocol for evaluation of electromagnetic component separation (APECS): Application of a framework for evaluating statistical methods of blink extraction from multichannel EEG. Clin. Neurophysiol. 118, 80–97.
Freeman, W.J., Holmes, M.D., Burke, B.C., Vanhatalo, S., 2003. Spatial spectra of scalp EEG and EMG from awake humans. Clin. Neurophysiol. 114, 1053–1068.
Freunberger, R., Fellinger, R., Sauseng, P., Gruber, W., Klimesch, W., 2009. Dissociation between phase-locked and nonphase-locked alpha oscillations in a working memory task. Hum. Brain Mapp., Epub ahead of print.
Freyer, F., Becker, R., Anami, K., Curio, G., Villringer, A., Ritter, P., 2009. Ultrahigh-frequency EEG during fMRI: Pushing the limits of imaging-artifact correction. NeuroImage. Epub ahead of print.
Gasser, T., Bacher, P., Mocks, J., 1982. Transformations towards the normal distribution of broad band spectral parameters of the EEG. Electroencephalogr. Clin. Neurophysiol. 53, 119–124.
Gevins, A., Smith, M.E., 2000. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cereb. Cortex 10, 829–839.
Goncharova, I.I., McFarland, D.J., Vaughan, T.M., Wolpaw, J.R., 2003. EMG contamination of EEG: Spectral and topographical characteristics. Clin. Neurophysiol. 114, 1580–1593.
Greischar, L.L., Burghy, C.A., van Reekum, C.M., Jackson, D.C., Pizzagalli, D.A., Mueller, C., et al., 2004. Effects of electrode density and electrolyte spreading in dense array electroencephalographic recording. Clin. Neurophysiol. 115, 710–720.
Grouiller, F., Vercueil, L., Krainik, A., Segebarth, C., Kahane, P., David, O., 2007. A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI. NeuroImage 38, 124–137.
Hamidi, M., Slagter, H.A., Tononi, G., Postle, B.R., 2009. Repetitive transcranial magnetic stimulation affects behavior by biasing endogenous cortical oscillations. Frontiers in Integrative Neuroscience 3, 1–12.
Hayes, A.F., Krippendorff, K., 2007. Answering the call for a standard reliability measure for coding data. Communication Methods and Measures 1, 77–89.
Hesse, C.W., James, C.J., 2004. Stepwise model order estimation in blind source separation applied to ictal EEG. Engineering in Medicine and Biology Society 26th Annual International Conference of the IEEE (IEMBS 2004), 1, 986-989.
Hoffmann, S., Falkenstein, M., 2009. The correction of eye blink artefacts in the EEG: A comparison of two prominent methods. PLoS ONE 3, e3004.
Huang, R.S., Jung, T.P., Delorme, A., Makeig, S., 2008. Tonic and phasic electroencephalographic dynamics during continuous compensatory tracking. NeuroImage 39, 1896–1909.
Ille, N., Berg, B., Scherg, M., 2002. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. J. Clin. Neurophysiol. 19, 113–124.
James, C.J., Hesse, C.W., 2005. Independent component analysis for biomedical signals. Physiol. Meas. 26, R15–R39.
Joyce, C.A., Gorodnitsky, I.F., Kutas, M., 2004. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 41, 313–325.
Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., et al., 2000a. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 37, 163–178.
Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000b. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.
Koskinen, M., Vartiainen, N., 2009. Removal of imaging artifacts in EEG during EEG/fMRI recording: Reconstruction of a high-precision artifact template. NeuroImage 46, 160–167.
Lawrence, F.R., Hancock, G.R., 1999. Conditions affecting integrity of a factor solution under varying degrees of overextraction. Educ. Psychol. Meas. 59, 549–579.
Lee, J.H., Lee, T.W., Jolesz, F.A., Yoo, S.S., 2008. Independent vector analysis (IVA): Mulitvariate approach for fMRI group study. NeuroImage 40, 86–109.
Lee, S., Buchsbaum, M.S., 1987. Topographic mapping of EEG artifacts. Clin. Electroencephalogr. 18, 61–67.
Lee, T.W., Lewicki, M.S., Girolami, M., Sejnowski, T.J., 1999. Blind source separation of more sources than mixtures using overcomplete representations. IEEE Signal Process. Lett. 6, 87–90.
Li, Y.O., Adali, T., Calhoun, V.D., 2007. Estimating the number of independent components for functional magnetic resonance imaging data. Hum. Brain Mapp. 28, 1251–1266.
Lutz, A., Greischar, L.L., Rawlings, N.B., Ricard, M., Davidson, R.J., 2004. Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. Proc. Natl. Acad. Sci. U. S. A. 101, 16369–16373.
Makeig, S., Debener, S., Onton, J., Delorme, A., 2004. Mining event-related brain dynamics. Trends Cogn. Sci. 8, 204–210.
Mantini, D., Franciotti, R., Romani, G.L., Pizzella, V., 2008. Improving MEG source localizations: an automated method for complete artifact removal based on independent component analysis. NeuroImage 40, 160–173.
McMenamin, B.W., Shackman, A.J., Maxwell, J.S., Greischar, L.L., Davidson, R.J., 2009. Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. Psychophysiology 46, 578–592.
Michel, C., Murray, M., Lantz, G., Gonzalez, S., Spinelli, L., Grave de Peralta, R., 2004. EEG source imaging. Clin. Neurophysiol. 115, 2195–2222.
Milne, E., Scope, A., Pascalis, O., Buckley, D., Makeig, S., 2009. Independent component analysis reveals atypical electroencephalographic activity during visual perception in individuals with autism. Biol. Psychiatry 65, 22–30.
Morecraft, R.J., Tanji, J., 2009. Cingulofrontal interactions and the cingulate motor areas. In: Vogt, B.A. (Ed.), Cingulate neurobiology and disease: Infrastructure, diagnosis, treatment. Oxford University Press, NY, pp. 113–144.
Moraux, A., Ianetti, G.D., 2009. Nociceptive laser-evoked brain potentials do not reflect nociceptive-specific neural activity. J. Neurophysiol. 101, 3258–3269.
Naeem, M., Brunner, C., Pfurtscheller, G., 2009. Dimensionality reduction and channel selection of motor imagery electroencephalographic data. Comput. Intell. Neurosci. Epub ahead of print.
Nazarpour, K., Wongsawat, Y., Sanei, S., Chambers, J.A., Oraintara, S., 2008. Removal of the eye-blink artifacts for EEGs via STF-TS modeling and robust minimum variance beamformers. IEEE Trans. Biomed. Eng. 55, 2221–2231.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
Niedermeyer, E., 2005. The normal EEG of the waking adult. In: Niedermeyer, E., Lopes da Silva, F.H. (Eds.), Electroencephalography: basic principles, clinical applications, and related fields. Lippincott, Williams & Wilkins, Philadelphia, PA, pp. 167–192.
Nolte, G., Curio, G., 1999. The effect of artifact rejection by signal-space projection on source localization accuracy in MEG measurements. IEEE Trans. Biomed. Eng. 46, 400–408.
Oakes, T.R., Pizzagalli, D.A., Hendrick, A.M., Horras, K.A., Larson, C.L., Abercrombie, H.C., et al., 2004. Functional coupling of simultaneous electrical and metabolic activity in the human brain. Hum. Brain Mapp. 21, 257–270.
Ohla, K., Hudry, J., le Coutre, J., 2009. The cortical chronometry of electrogustatory event-related potentials. Brain Topogr. 22, 73–82.
Onton, J., Makeig, S., 2006. Information-based modeling of event-related brain dynamics. Prog. Brain Res. 159, 99–120.
Onton, J., Westerfield, M., Townsend, J., Makeig, S., 2006. Imaging human EEG dynamics using independent component analysis. Neurosci. Biobehav. Rev. 30, 808–822.
Pascual-Marqui, R.D., 1999. Review of methods for solving the EEG inverse problem. Int. J. Bioelectromagn. 1, 75–86.
Pascual-Marqui, R.D., Michel, C.M., Lehmann, D., 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. Int. J. Psychophysiol. 18, 49–65.
Pizzagalli, D.A., 2007. Electroencephalography and high density electrophysiological source localization, In: Cacioppo, J.T., Tassinaru, L.G., Berntson, G.G. (Eds.), Handbook of psychophysiology, 3rd ed. Cambridge University Press, NY, pp. 56–84.
Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., Thut, G., 2008. Spontaneous fluctuations in posterior α-band EEG activity reflect variability in excitability of human visual areas. Cereb. Cortex 18, 2010–2018.
Romero, S., Mananas, M.A., Barbanoj, M.J., 2008. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: a simulation case. Comp. Biol. Med. 38, 348–360.
Ryali, S., Glover, G.H., Chang, C., Menon, V., 2009. Development, validation, and comparison of ICA-based gradient artifact reduction algorithms for simultaneous EEG-spiral in/out echo-planar fMRI recordings. NeuroImage. Epub ahead of print.
Seaman, M.A., Serlin, R.C., 1998. Equivalence confidence intervals for two-group comparisons of means. Psychol. Methods 3, 403–411.
Shackman, A.J., in press. *The potentially deleterious impact of muscle activity on gamma band inferences*. Neuropsychopharmacology.
Shackman, A.J., Sarinopoulos, I., Maxwell, J.S., Pizzagalli, D.A., Lavric, A., Davidson, R.J., 2006. Anxiety selectively disrupts visuospatial working memory. Emotion 6, 40–61.
Shackman, A.J., McMenamin, B.W., Slagter, H.A., Maxwell, J.S., Greischar, L.L., Davidson, R.J., 2009. Electromyogenic artifacts and electroencephalographic inferences. Brain Topogr. 21, 7–12.
Shackman, A.J., McMenamin, B.W., Maxwell, J.S., Greischar, L.L., Davidson, R.J., 2009. Right dorsolateral prefrontal cortical activity and behavioral inhibition. Psychol. Sci. 20, 1500–1506.
Srinivasan, R., Tucker, D.M., Murias, M., 1998. Estimating the spatial Nyquist of the human EEG. Behav. Res. Methods Instrum. Comput. 30, 8–19.
Talsma, D., 2008. Auto-adaptive averaging: detecting artifacts in event-related potential data using a fully automated procedure. Psychophysiology 45, 216–228.
Tang, A.C., Liu, J.Y., Sutherland, M.T., 2005. Recovery of correlated neuronal sources from EEG: the good and bad ways of using SOBI. NeuroImage 28, 507–519.
Tassinary, L.G., Cacioppo, J.T., Vanman, E.J., 2007. The skeletomotor system: Surface electromyography, In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of psychophysiology, 3rd ed. Cambridge University Press, NY, pp. 267–302.
Tesche, C.D., Uusitalo, M.A., Ilmoniemi, R.J., Huotilainen, M., Kajola, M., Salonen, O., 1995. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr. Clin. Neurophysiol. 95, 189–200.

Thatcher, R.W., North, D., Biver, C., 2005. Parametric vs. non-parametric statistics of low resolution electromagnetic tomography (LORETA). Clin. Electroencephalogr. Neurosci. 36, 1–8.
Thibodeau, R., Jorgensen, R.S., Kim, S., 2006. Depression, anxiety, and resting frontal EEG asymmetry: a meta-analytic review. J. Abnorm. Psychology 115, 715–729.
Thut, G., Miniussi, C., 2009. New insights into rhythmic brain activity from TMS-EEG studies. Trends Cogn. Sci. 13, 182–189.
Ting, K.H., Fung, P.C.W., Chang, C.Q., Chan, F.H.Y., 2006. Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. Med. Eng. Phys. 28, 780–794.
U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research, 2001. Guidance for industry: Statistical approaches to establishing bioequivalence. Retrieved November 11th, 2007, from http://www.fda.gov/cder/guidance/index.htm.
Uusitalo, M.A., Ilmoniemi, R.J., 1997. Signal-space projection method for separating MEG or EEG into components. Med. Biol. Eng. Comput. 35, 135–140.
Van Boxtel, A., 2001. Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. Psychophysiology 38, 22–34.
Viola, F.C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., Debener, S., 2009. Semi-automatic identification of independent components representing EEG artifact. Clin. Neurophysiol. 120, 868–877.
Wallstrom, G.L., Kass, R.E., Miller, A., Cohn, J.F., Fox, N.A., 2004. Automatic correction of ocular artifacts in the EEG A comparison of regression-based and component-based methods. Int. J. Psychophysiol. 53, 105–119.
Waterink, W., van Boxtel, A., 1994. Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. Biol. Psychol. 37, 183–198.
Welch, P.D., 1967. The use of Fast Fourier Transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. 15, 70–73.
Whitham, E.M., Pope, K.J., Fitzgibbon, S.P., Lewis, T., Clark, C.R., Loveless, S., Broberg, M., Wallace, A., DeLosAngeles, D., Lillie, P., Hardy, A., Fronsko, R., Pulbrook, A., Willoughby, J.O., 2007. Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clin. Neurophysiol. 118, 1877–1888.
Willis, J., Nelson, A., Rice, J., Black, F.W., 1993. The topography of muscle activity in quantitative EEG. Clin. Electroencephalogr. 24, 123–126.
Yuval-Greenberg, S., Tomer, O., Keren, A.S., Nelken, I., Deouell, L.Y., 2009a. Transient induced gamma-band responses in EEG as a manifestation of miniature saccades. Neuron 58, 429–441.
Yuval-Greenberg, S., Keren, A.S., Tomer, O., Nelken, I., Deouell, L.Y., 2009b. Response to letter. Neuron 62, 10–12.

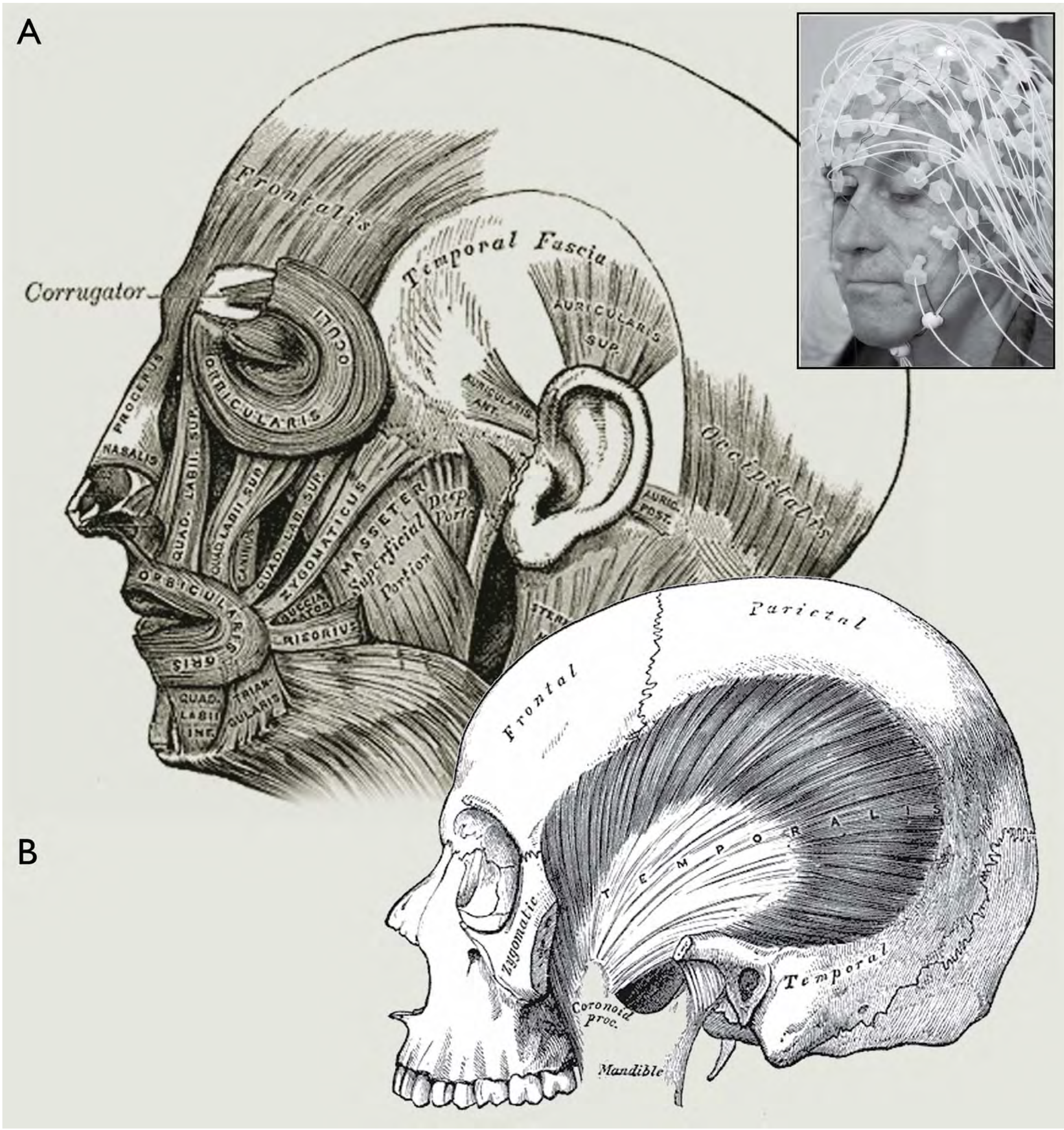
Supplementary Method and Results for

Validation of ICA-Based Myogenic Artifact Correction for Scalp and Source-Localized EEG

McMenamin, Shackman, Maxwell, Bachhuber, Koppenhaver, Greischar & Davidson

Contents

Cranial Musculature



Supplementary Figure 1. *Cranial Musculature*. **A)** From front to back, key generators of myogenic artifact include the corrugator along the brow, orbicularis oculi around the eyes, frontalis above the brow, masseter on the jaw, the peri-auricular muscles surrounding the ear, and occipitalis at the base of the skull. **Inset:** High-resolution (128-channel) EEG array. **B)** Temporalis. Both engravings adapted from *Gray's Anatomy* (Gray, 1918/2000).

Component Classification

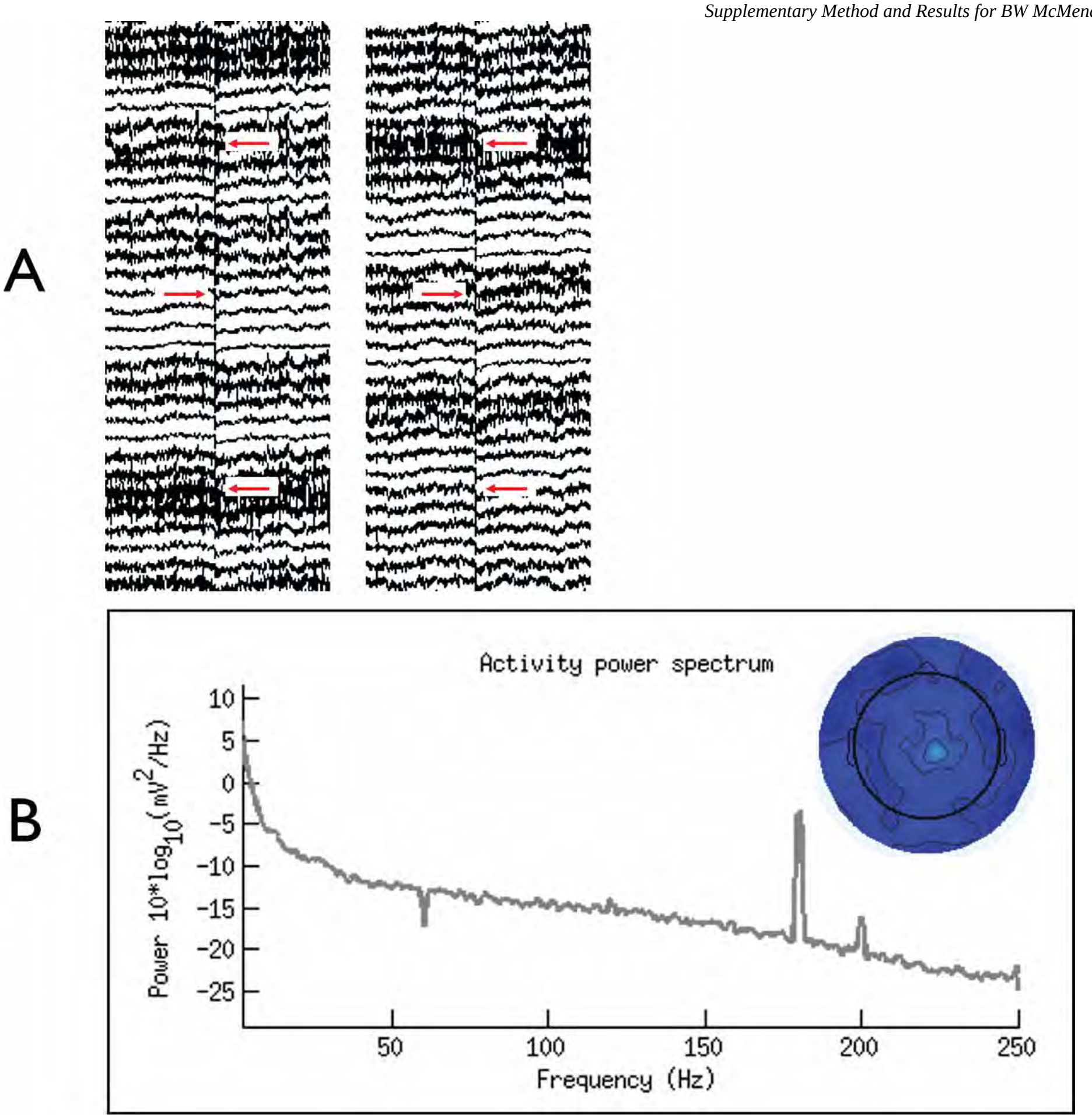
Components were classified by two independent, well-trained baccalaureate-level raters (DRWB and AMK) based on inspection of the time-series, power spectrum, and absolute topography. In general, our criteria were in close accord with classical descriptions (Gibbs & Gibbs, 1950). As detailed below, the relative importance of each measure to classification varied somewhat across categories. Where necessary, the raw time-series was also inspected. Preliminary classifications by each rater were used for reliability estimates. Final classification was by consensus. Classification required approximately two hours per participant per rater once the raters were calibrated to the protocol (see below).[1] In the present study, the time invested in calibration approximately equaled that devoted to "final" classifications.

It is worth noting that this could be reduced somewhat by using a MATLAB script to create low-resolution image files depicting the topography, PSD, snippets of the raw and component time-series, and/or dipole fits for each component. These could be rapidly reviewed by rater(s) using any ordinary image viewer software. We are currently in the process of developing the code necessary to do so. Future studies might also consider using the method of Groppe, Makeig and Kutas (2009) to reduce the number of components requiring classification to those evidencing adequate split-half reliability. The application of clustering algorithms prior to classification might also prove useful in this regard (see http://sccn.ucsd.edu/eeglab/clusttut/clustertut.html).

Gross Artifact. Several kinds of residual physiological and electromechanical artifacts were collectively classified as *Gross*. These included reference (Cz) and ground (nasion) sensor artifacts, electrocardiographic (ECG) artifacts, and alternating current (AC) artifacts.

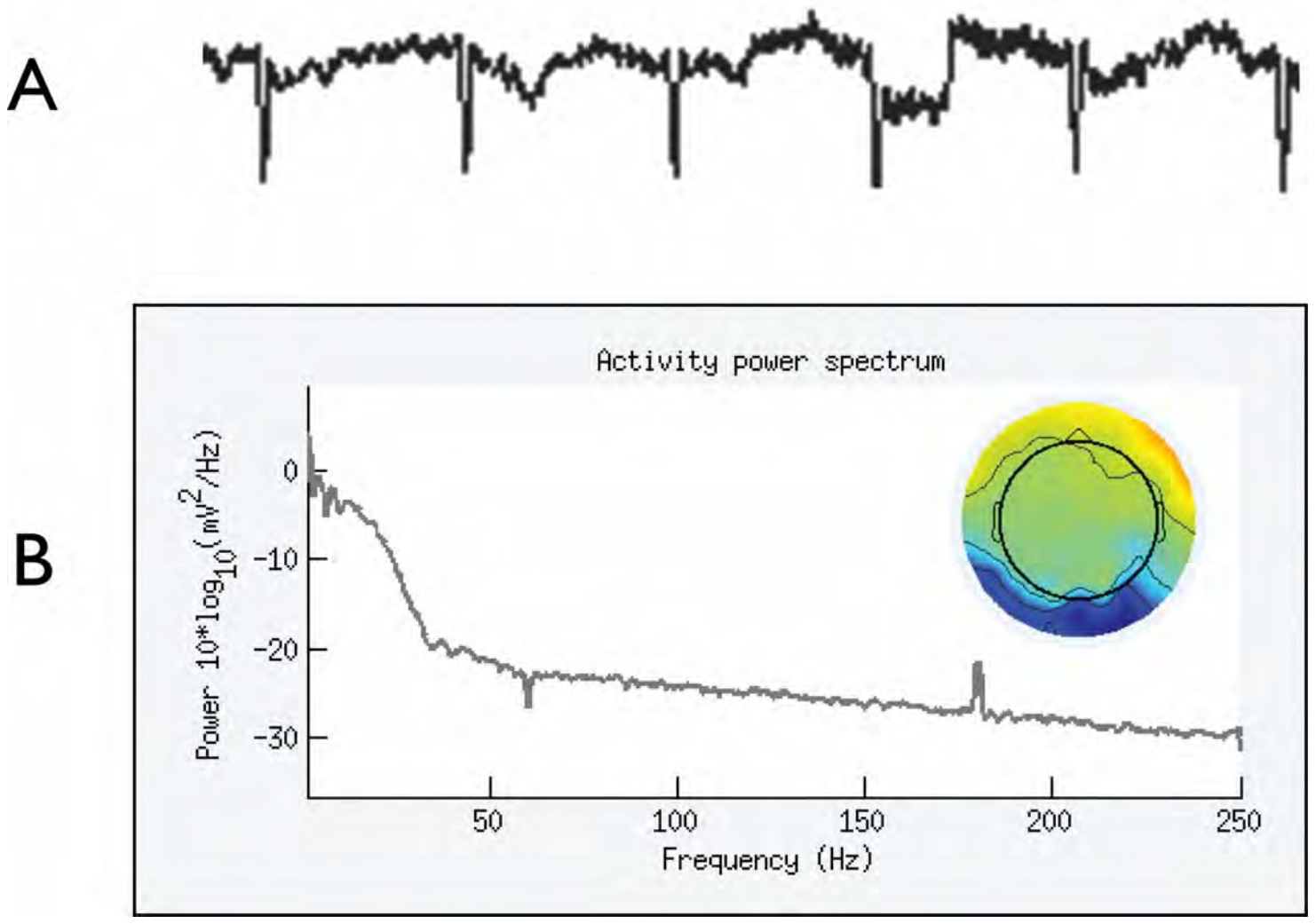
Reference and ground artifacts showed widespread, synchronous deflections in the raw time-series corresponding to periods of apparent "activation" in the component time-series, combined with a characteristically uniform topography (Supplementary Figure 2).

[1] This could be further reduced in future studies by using a MATLAB script to create low-resolution image files depicting the topography, PSD, snippets of the raw and component time-series, and/or dipole fits for each component. These could be rapidly reviewed by rater(s) using any ordinary image viewer software.



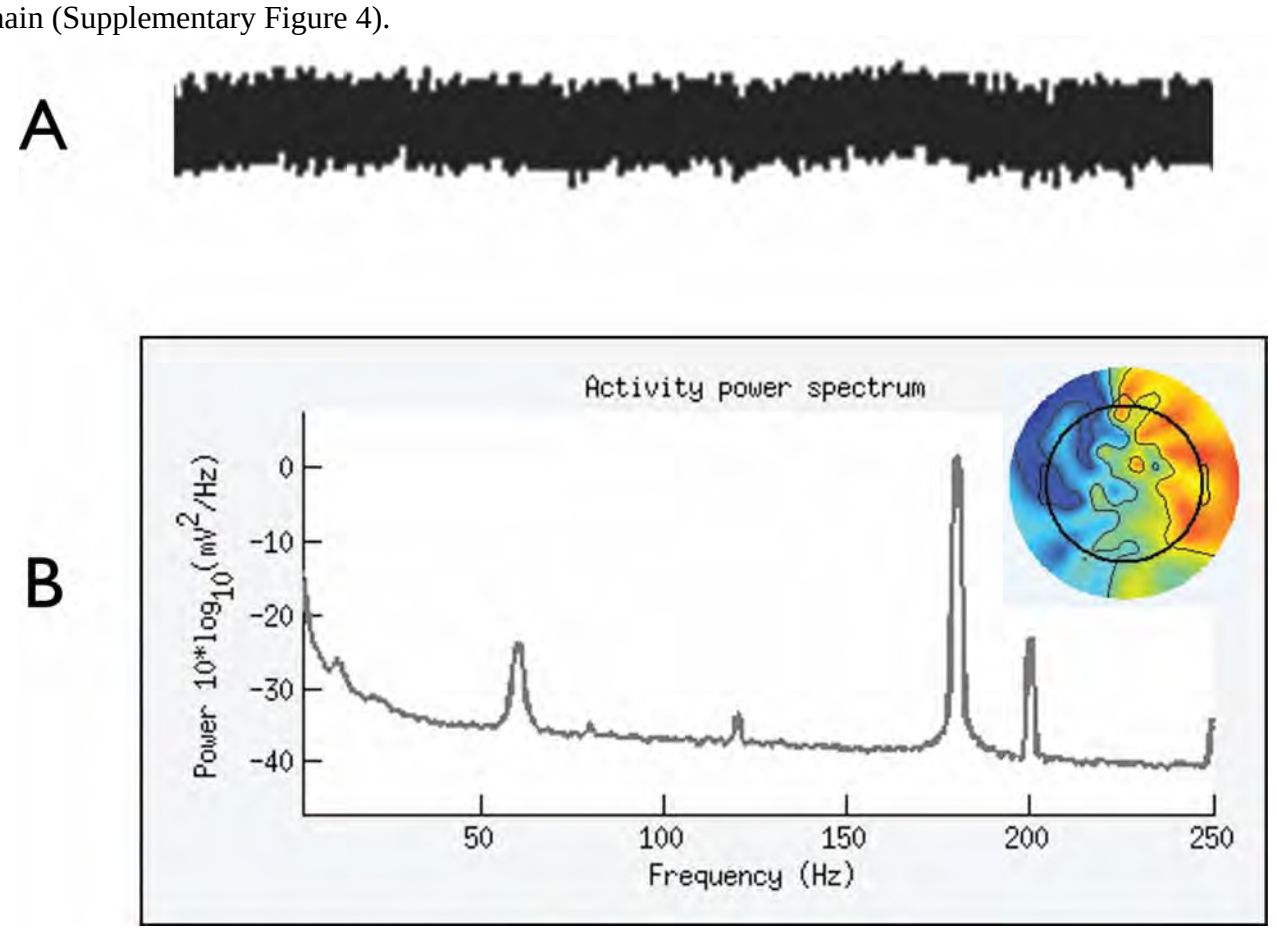
Supplementary Figure 2. *Reference Artifact.* **A)** Representative multi-channel slices of the raw time-series. Red arrows indicate location of transient reference artifacts. In both cases, all 128 channels (not displayed) were similarly affected. **B)** Back-projected and interpolated topographic map of whole-head (128-channel) component weights. The polarity of the map is arbitrary with scaling proportional to μV (Zeman, Till, Livingston, Tanaka, & Driessen, 2007).

ECG artifacts showed a characteristic pattern of deflections in the component time-series, typically persisting throughout the recording; a unilateral or, more rarely, bilateral posterior topography; and a low frequency (<3Hz) peak in the frequency-domain (Supplementary Figure 3). In contrast to one prior report (Viola et al., 2009), the raters anecdotally found ECG to be the easiest artifact to classify.



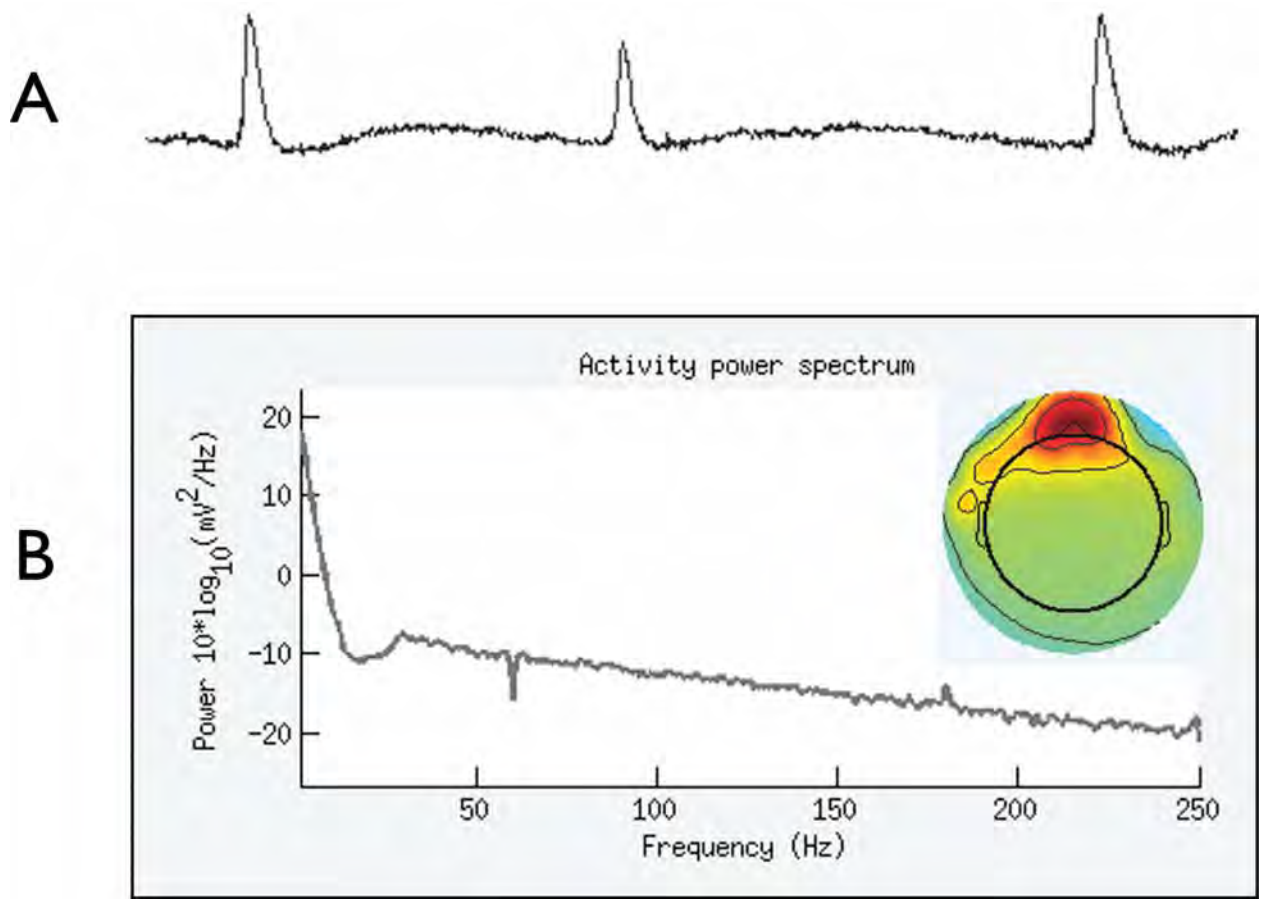
Supplementary Figure 3. *ECG Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

AC artifacts were chiefly identified by a 60Hz peak, reflecting residual signal following notch-filtering, and harmonics in the frequency-domain along with sustained activation in the time-domain (Supplementary Figure 4).

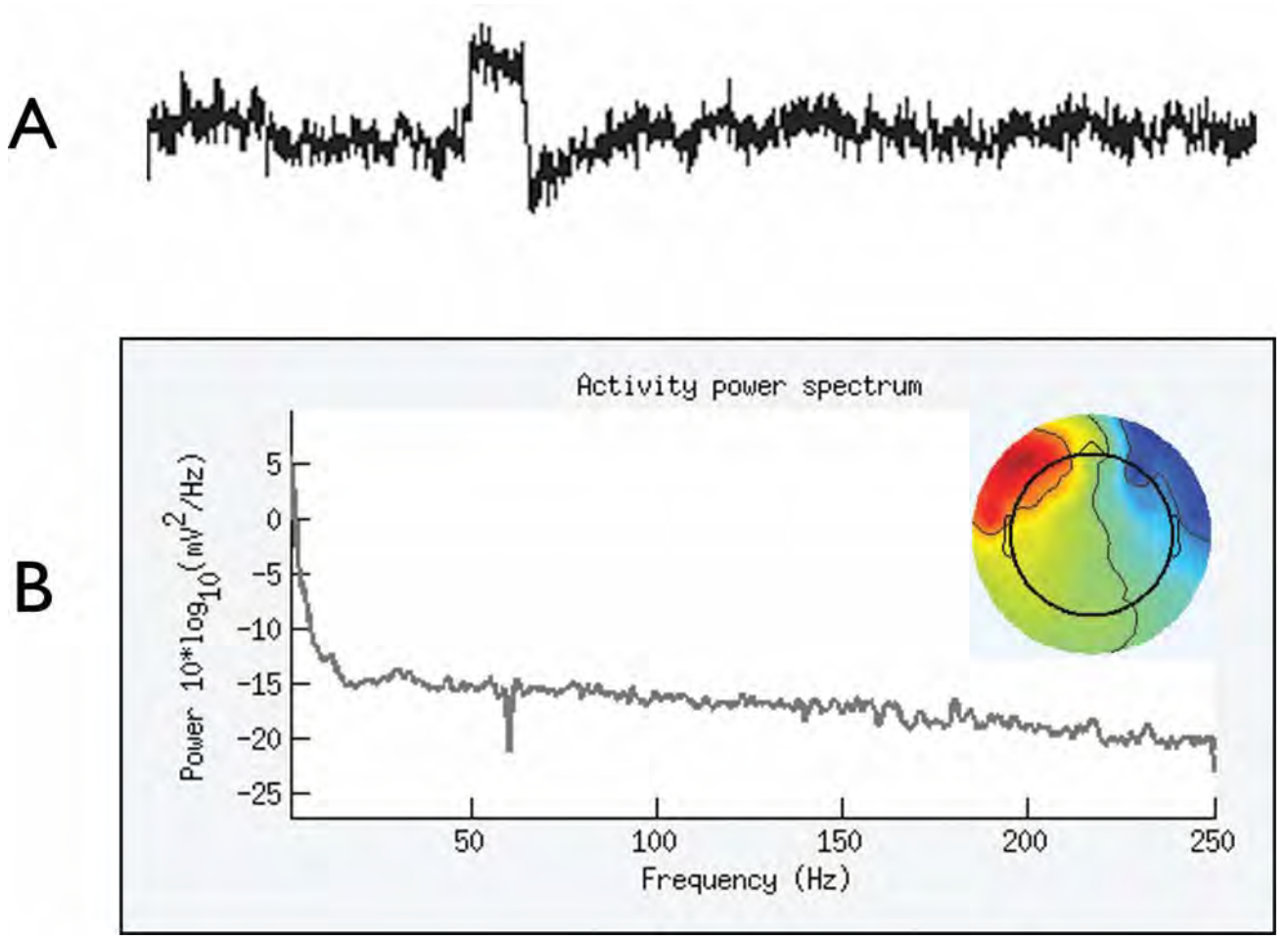


Supplementary Figure 4. *AC Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

Ocular Artifact. Eye movements were classified as *Ocular*. Blinks showed high loadings at the most anterior sites (Supplementary Figure 5), whereas saccades manifested as an anterior dipole (Supplementary Figure 6). Both kinds of ocular artifact exhibited transient, high amplitude deflections in the time-domain and low-frequency peaks in the frequency-domain. The raters noted that large-variance ocular artifacts were relatively easy to classify, whereas small-variance ocular artifacts were easy, but time-consuming.

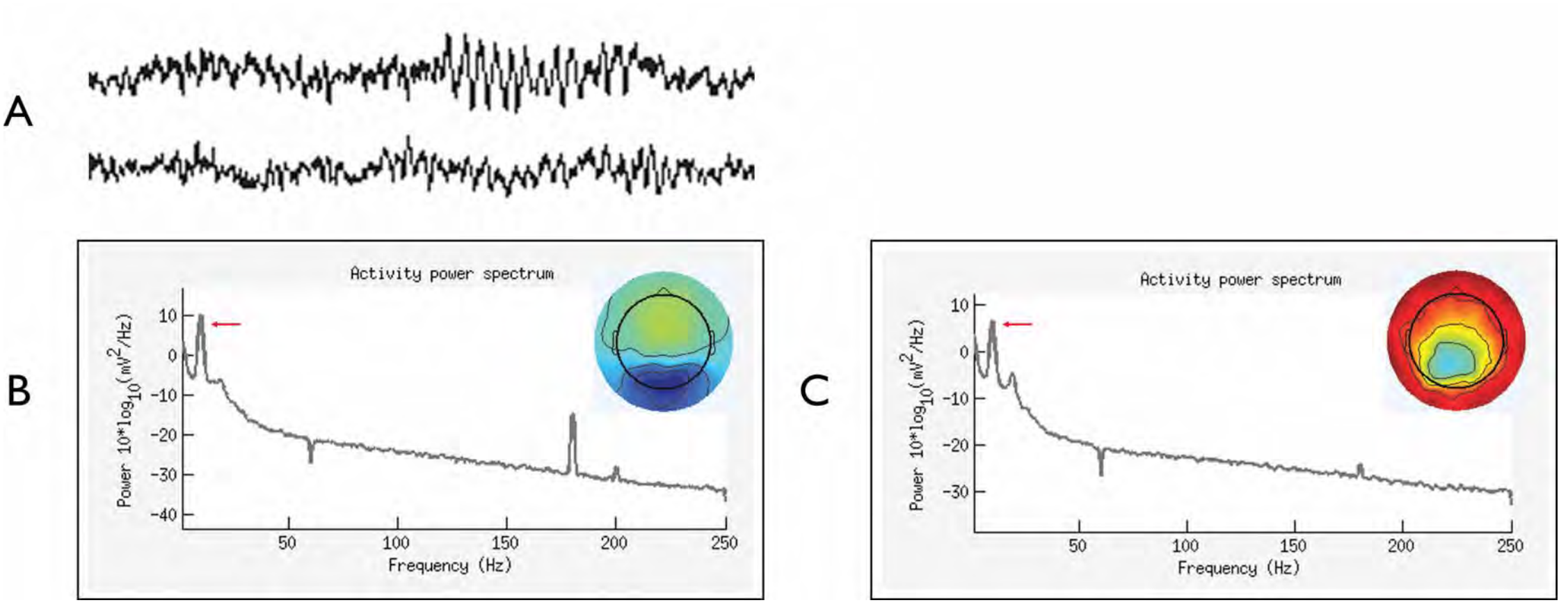


Supplementary Figure 5. *Blink Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

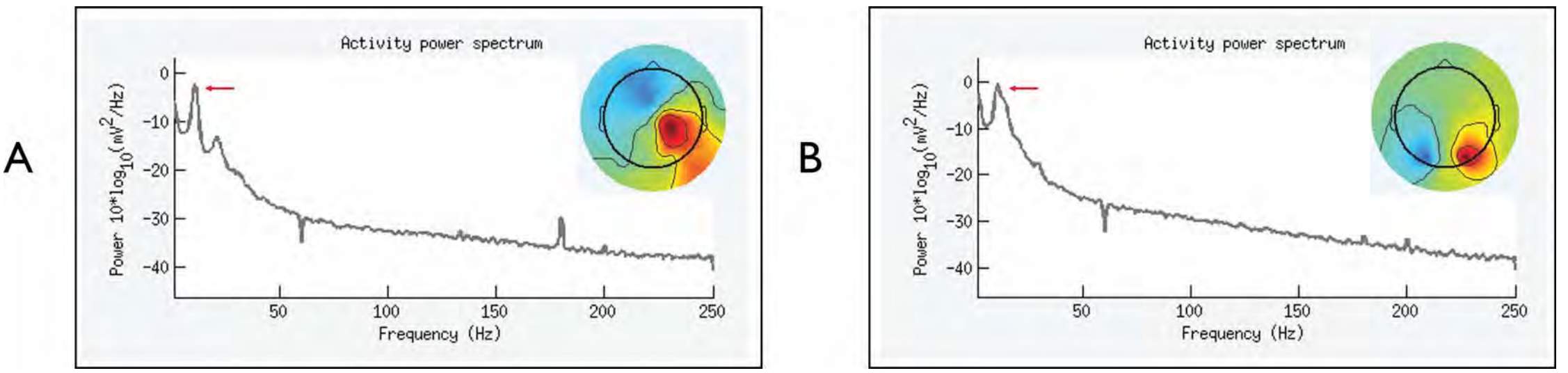


Supplementary Figure 6. *Saccadic Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

Neurogenic. Components containing frank neurogenic activity in the absence of any artifactual activity were classified as *Neuro* (Supplementary Figures 7-8). These components were characterized by broad, smooth topographies, often with a clearly dipolar pattern, with peak loadings well away from the edge of the montage. In the frequency-domain, they exhibited a clear 1/*f* pattern, often with a peak in the alpha band (8-13Hz). In the time-domain, they displayed sustained periods of activation with low-frequency oscillations. The raters noted that large-variance neurogenic components were relatively easy to classify

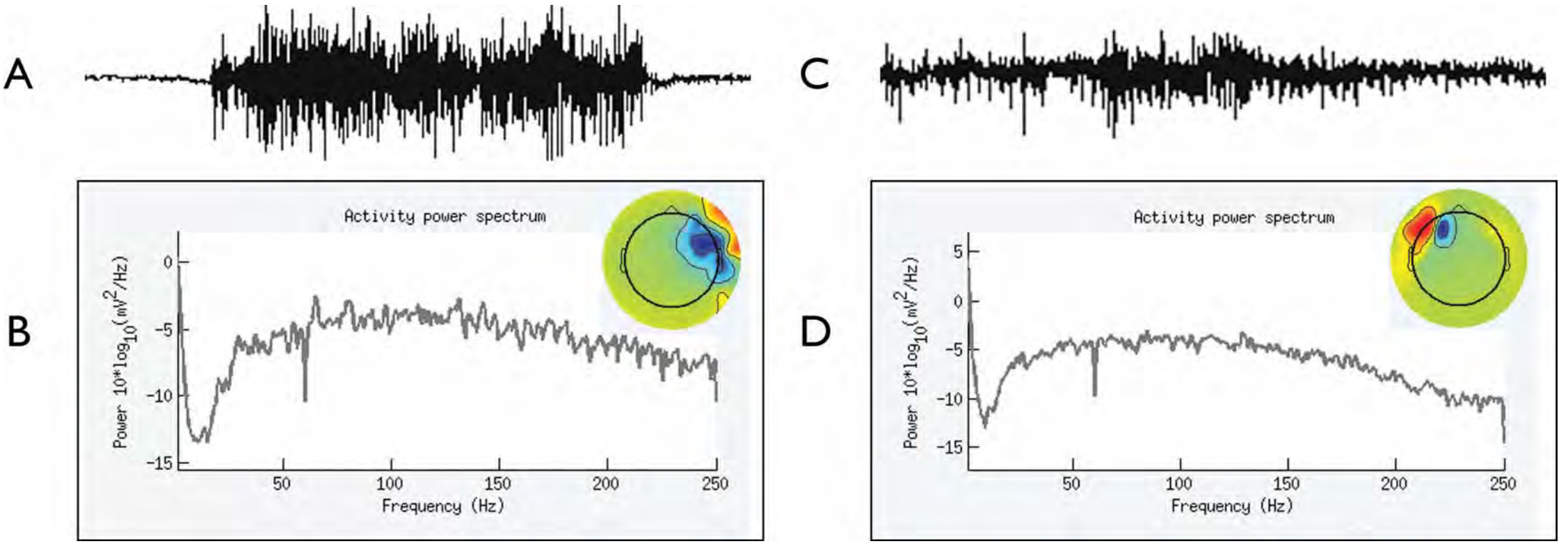


Supplementary Figure 7. *High-Variance Neurogenic Components*. **A)** Component time-series. **B) and C)** Component power spectra for the fourth and fifth components. Sixty-four components were extracted and ranked in descending order according to the amount of variance predicted. Red arrows indicate the location of the alpha peak (~10Hz). Insets show corresponding topographic maps.



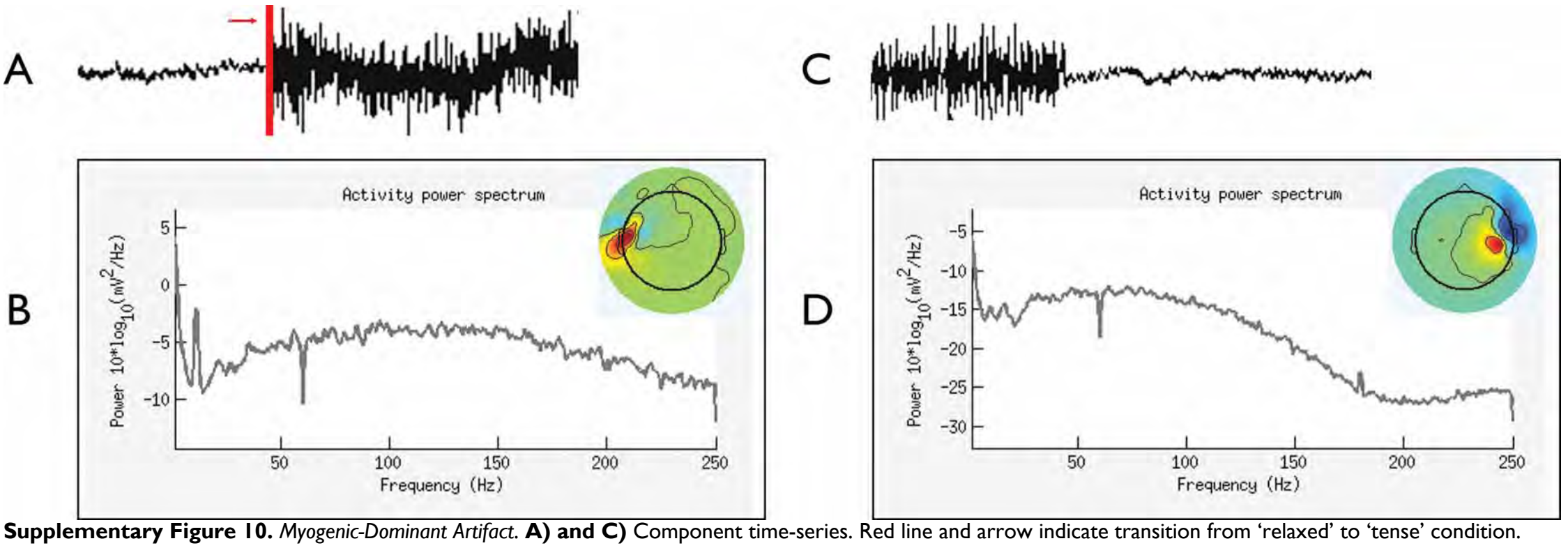
Supplementary Figure 8. *Low-Variance Neurogenic Components*. **A)** and **B)** Component power spectra for the sixteenth and eighteenth components. Red arrows indicate the location of the alpha peak (~10Hz). Insets show corresponding topographic maps.

Myogenic. Components containing frank EMG activity in the absence of any identifiable neurogenic activity were classified as *Myo* (Supplementary Figure 9). These components were chiefly distinguished based on spectra with broad peaks around either 40Hz or greater than 70Hz. On the scalp, they showed one of two topographies: a moderately broad distribution that mimicked the underlying scalp musculature and peaked along the edge of the montage, or small cluster(s) of cephalic or extracephalic electrodes. In the time-domain, they exhibited periods of high-frequency activation.

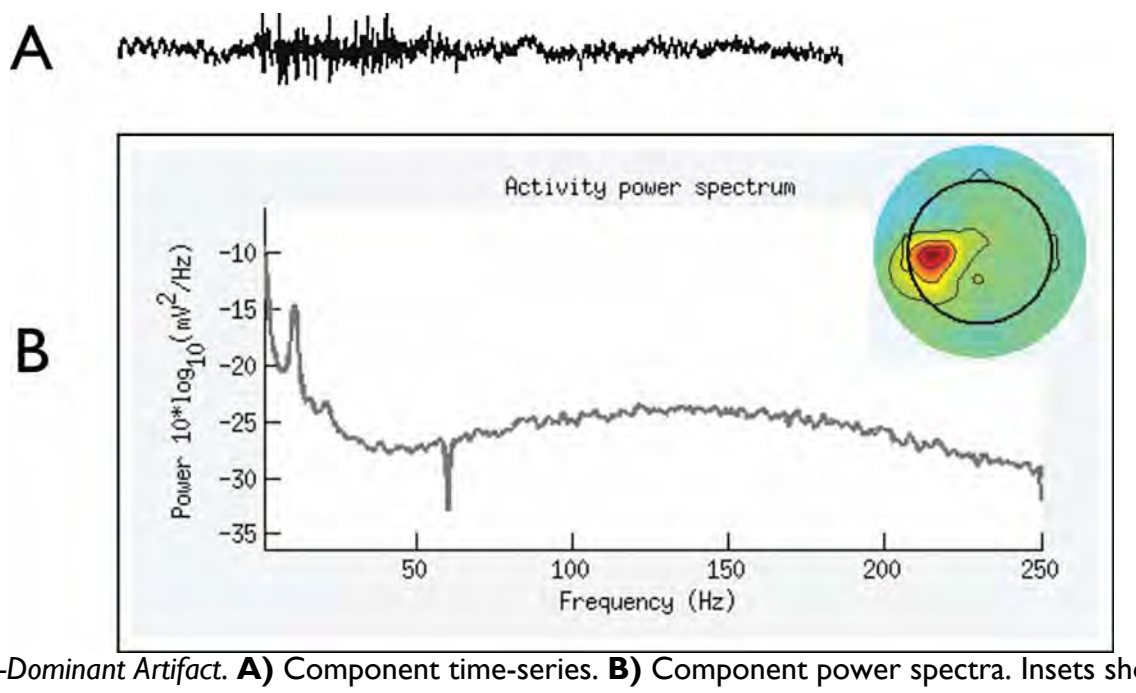


Supplementary Figure 9. *Myogenic Artifact.* **A) and C)** Component time-series. **B) and D)** Component power spectra. Insets show corresponding topographic maps.

Myo-Dominant and Neuro-Dominant. Based on the criteria detailed above, many components contained mixtures of myogenic and neurogenic activity. Depending on the relative dominance of the two sources, they were classified as either *Myo-Dominant* or *Neuro-Dominant*. Dominance was determined using the power spectrum and component time-series. If the amplitude of the apparently myogenic peak (>35Hz) was greater than the neurogenic peak, the component was classified as Myo-Dominant (Supplementary Figure 10). If the reverse pattern was observed *and* the component time-series was *not* dominated by high-frequency activity, the component was classified as Neuro-Dominant (Supplementary Figure 11). More typically, the time-series was dominated by high-frequency activity and the component was instead classified as Myo-Dominant. The raters noted that mixed-dominance components were among the more difficult to classify. In light of these results (see below), future studies might consider treating them as a single category.

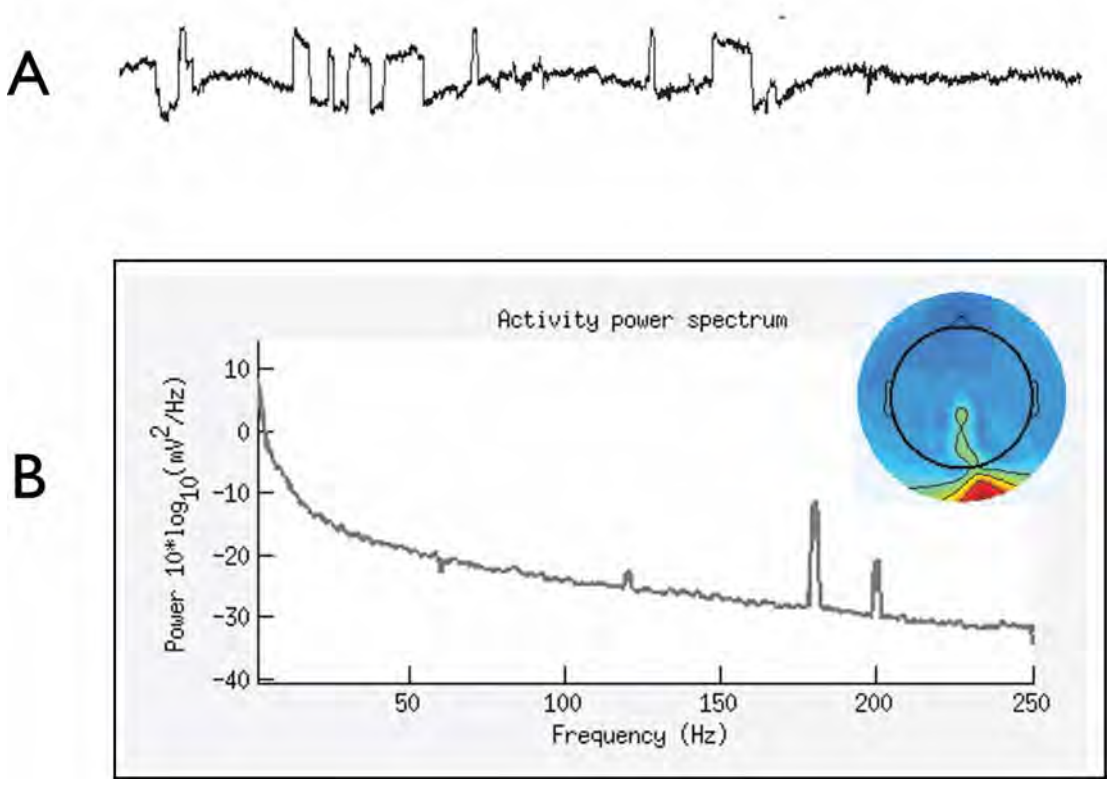


Supplementary Figure 10. *Myogenic-Dominant Artifact.* **A) and C)** Component time-series. Red line and arrow indicate transition from 'relaxed' to 'tense' condition. **B) and D)** Component power spectra. Insets show corresponding topographic maps.



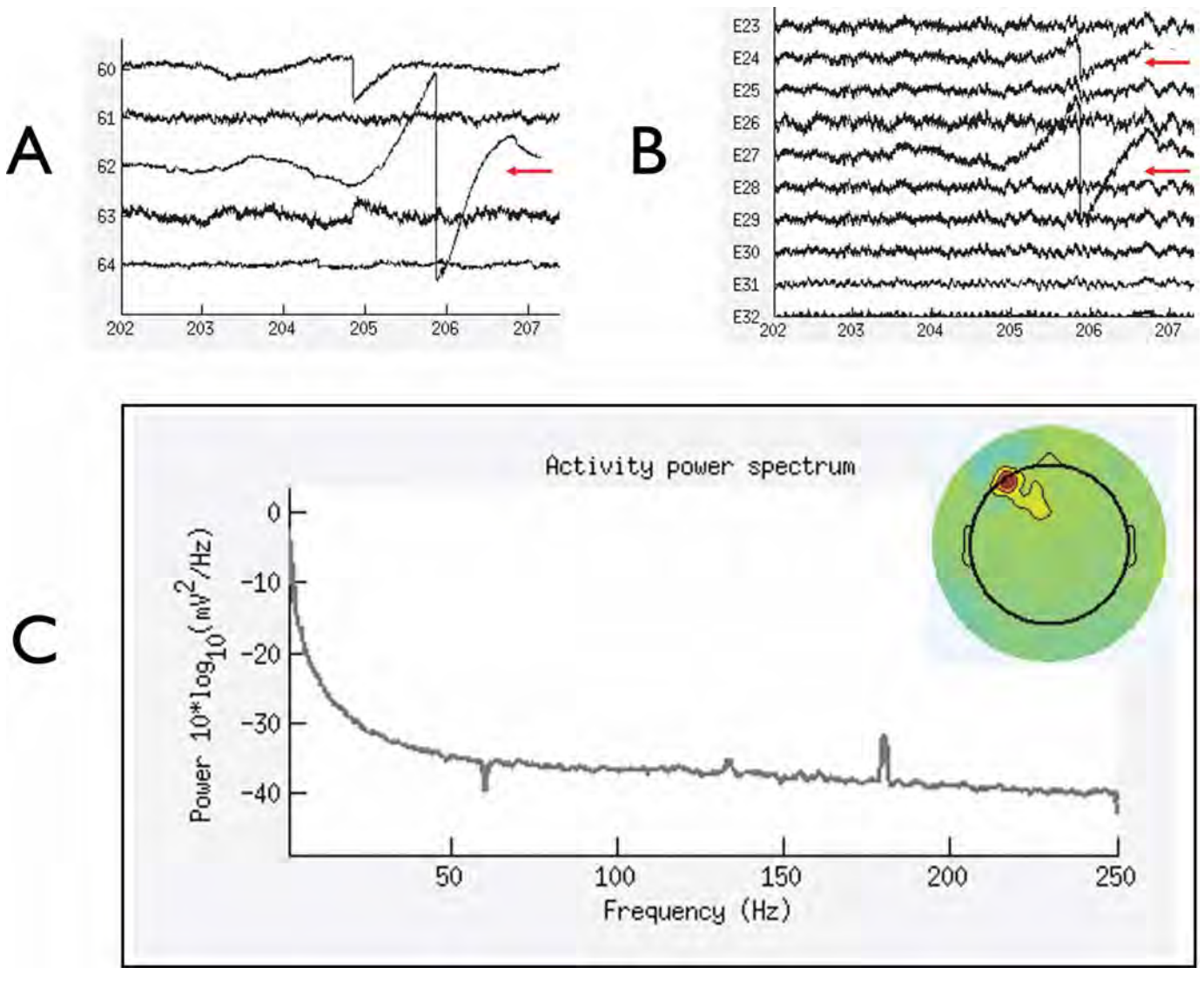
Supplementary Figure 11. *Neurogenic–Dominant Artifact.* **A)** Component time-series. **B)** Component power spectra. Insets show corresponding topographic maps.

***Noise*.** A small number of components failed to clearly meet the criteria detailed above, and were classified as *Noise* (Supplementary Figure 12). Noise components tended to show strongly 1/*f* –shaped spectra.



Supplementary Figure 12. *Noise*. **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

Low-Variance. As noted in the main report, components were, by default, classified only if they accounted for at least 0.2% of the variance in the EEG for a particular participant.[2] In cases where the determination was unambiguous (approximately one-third of such components), exceptions were made. Otherwise, they were categorized as *Low-Variance* by default (Supplementary Figure 13).



Supplementary Figure 13. *Unclassified Low-Variance Component.* **A)** Component time-series for ICs 60-64. Note the deviation centered around 206-s (red arrow). **B)** Raw time-series for electrodes 23-32. Note the corresponding deviation at 206-s on frontal electrodes 24 and 27 (red arrows). **C)** Component power spectrum. Inset shows corresponding topographic map.

[2] The use of a slightly higher threshold might prove a useful means of reducing the classification burden. For instance, using a threshold of 0.39% (i.e., 1/256th of the total variance) would have reduced the number of components requiring classification by ~25% while ensuring that ~95% of the variance was manually classified.

Inter-Rater Reliability

Although both raters were highly experienced, additional study-specific calibration was performed on a subset of five participants. Inter-rater reliability was estimated from the remaining participants using Krippendorff's alpha (*KA*; Hayes & Krippendorff, 2007). *KA*, which ranges from 0 (0% agreement) to 1 (100% agreement), exhibits superior performance compared to more familiar metrics, such as kappa or percentage agreement (Hayes & Krippendorff, 2007). Using a bootstrap program to estimate *KA* (http://www.comm.ohio-state.edu/ahayes; 5,000 bootstraps), two kinds of inter-rater reliability were computed. First, global agreement was computed based on the eight mutually exclusive, nominal categories that raters used to code the components (myogenic, myogenic-dominant, neurogenic-dominant, neurogenic, ocular, gross, noise, low-variance). This yielded excellent overall agreement: *KA* = .98 (95% CI: .97-.99). Second, in order to identify systematic disagreements between the raters, indicative of more difficult classifications or residual ambiguities in the classification protocol, agreement was computed separately for each of the eight categories. In contrast to one prior report in which raters used only topographic maps to classify components (Viola et al., 2009), agreement was excellent for all categories, with the lowest levels exhibited for neurogenic-dominant components, *KA* = .93 (95% CI: .85-.1.0); agreement for the remaining categories was better still, *KA*s > .96. It is worth emphasizing that these high levels of agreement reflect both the classification protocol and the raters' training—*reliability is not guaranteed by the mere adoption of this protocol*. Future investigations must provide adequate training and should report observed agreement.

Supplementary Tables for Scalp Analyses

Supplementary Table 1.
Sensitivity on the Scalp

			ROI Median Contrast			ROI Peak Contrast		
	EMG Correction	NNNM Filtering	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	None	Minimal	-2.44*	2.44*	.03	-3.21**	3.21**	-1.04
		Intermediate	-2.30*	2.30*	.21	-2.82**	2.82**	-1.04
		Maximal	-2.35*	2.35*	.19	-3.05**	3.05**	-1.03
	Minimal	Minimal	-2.05†	1.79†	.27	-2.90**	2.58*	1.06
		Intermediate	-1.91†	1.66	.47	-2.73**	2.41*	1.05
		Maximal	-1.83†	1.60	.50	-2.75**	2.30*	1.04
	Intermediate	Minimal	-1.66	.98	.07	-1.89†	1.46	-1.24
		Intermediate	-1.59	1.01	.21	-1.89†	1.31	1.12
		Maximal	-1.43	.81	.35	-1.82†	1.31	1.19
	Maximal	Minimal	-1.59	1.10	.25	-1.80†	1.63	-1.61
		Intermediate	-1.49	1.15	.51	-1.80†	1.49	1.39
		Maximal	-1.37	.90	.54	-1.72†	1.52	1.48
Equivalence[e]	None	Minimal	.06 (-.07 -.01)	.07 (.01 .07)*	.06 (-.03 .03)*	.10 (-.14 -.03)	.09 (.03 .14)	.07 (-.05 .02)*
		Intermediate	.06 (-.07 .00)	.08 (.00 .07)*	.08 (-.03 .04)*	.10 (-.13 -.02)	.09 (.02 .13)	.07 (-.05 .02)*
		Maximal	.07 (-.08 -.01)	.08 (.01 .08)*	.07 (-.03 .03)*	.10 (-.06 -.03)	.09 (.03 .13)	.07 (-.05 .02)*
	Minimal	Minimal	.08 (-.08 .00)*	.06 (.00 .04)*	.08 (-.03 .04)*	.07 (-.05 -.01)*	.06 (.01 .06)*	.06 (-.02 .05)*
		Intermediate	.08 (-.08 .00)*	.09 (-.01 .09)*	.07 (-.02 .04)*	.07 (-.05 -.01)*	.06 (.00 .06)*	.06 (-.02 .05)*
		Maximal	.08 (-.08 .00)*	.07 (-.01 .07)*	.09 (-.03 .05)*	.07 (-.05 -.01)*	.06 (.00 .06)*	.06 (-.02 .05)*
	Intermediate	Minimal	.08 (-.08 .01)*	.10 (-.03 .08)*	.09 (-.04 .04)*	.08 (-.07 .00)*	.06 (-.01 .05)*	.07 (-.05 .01)*
		Intermediate	.09 (-.10 .01)	.09 (-.03 .09)*	.08 (-.03 .04)*	.08 (-.07 .00)*	.09 (-.02 .09)*	.06 (-.02 .05)*
		Maximal	.11 (-.10 .02)*	.07 (-.03 .06)*	.09 (-.03 .05)*	.08 (-.08 .00)*	.09 (-.02 .09)*	.06 (-.01 .05)*
	Maximal	Minimal	.10 (-.10 .01)*	.07 (-.02 .06)*	.09 (-.03 .04)*	.09 (-.08 .01)*	.06 (-.01 .06)*	.07 (-.05 .01)*
		Intermediate	.08 (-.06 .01)*	.07 (-.02 .06)*	.09 (-.03 .05)*	.09 (-.08 .01)*	.06 (-.01 .05)*	.06 (-.01 .06)*
		Maximal	.07 (-.05 .01)*	.10 (-.03 .08)*	.05 (-.02 .04)*	.09 (-.08 .01)*	.06 (-.01 .05)*	.06 (-.01 .06)*

Note: [a] Corrected OR-OT vs. 0. [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test. [e] Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > p > .05, * p < .05, ** p < .01.

Supplementary Table 2.
Specificity on the Scalp

			ROI Median Contrast			ROI Peak Contrast		
	EMG Correction	NNNM Filtering	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	None	Minimal		1.25	.48		2.10†	1.57
		Intermediate		1.31	.52		2.06†	1.72
		Maximal		1.33	.54		2.12*	1.64
	Minimal	Minimal	.23	1.30	.55	2.06†	2.08†	1.57
		Intermediate	.27	1.28	.59	1.58	2.04†	1.72
		Maximal	.92	1.30	.58	1.93†	2.14*	1.68
	Intermediate	Minimal	-1.93†	1.12	.36	-3.24**	1.87†	1.28
		Intermediate	-1.20	1.28	.26	-3.18**	2.13*	1.64
		Maximal	-1.04	1.29	.43	-2.93**	2.15*	1.36
	Maximal	Minimal	.35	1.39	.80	-2.89**	2.21*	1.51
		Intermediate	.12	1.38	.79	-2.50*	2.31*	1.90†
		Maximal	.69	1.40	.97	-2.50*	2.57*	1.69
Equivalence[e]	None	Minimal		.13 (-.03 .11)*	.11 (-.04 .06)*		.11 (.00 .15)	.09 (-.01 .08)*
		Intermediate		.14 (-.03 .13)*	.13 (-.04 .07)*		.13 (.00 .16)	.09 (-.01 .09)*
		Maximal		.12 (-.03 .13)	.13 (-.04 .07)*		.12 (.00 .12)*	.09 (-.01 .09)*
	Minimal	Minimal	.11 (.00 .00)*	.13 (-.02 .11)*	.09 (-.04 .06)*	.09 (.00 .00)*	.11 (.00 .15)	.09 (-.01 .08)*
		Intermediate	.09 (.00 .00)*	.13 (-.03 .11)*	.13 (-.04 .07)*	.09 (.00 .00)*	.13 (.00 .16)	.09 (-.01 .08)*
		Maximal	.12 (.00 .00)*	.12 (-.03 .13)	.13 (-.04 .07)*	.09 (.00 .00)*	.12 (.00 .11)*	.09 (-.01 .08)*
	Intermediate	Minimal	.09 (-.01 .00)*	.08 (-.03 .09)	.13 (-.05 .07)*	.14 (-.01 .00)*	.15 (-.01 .15)*	.09 (-.01 .06)*
		Intermediate	.11 (-.01 .00)*	.12 (-.03 .13)	.11 (-.05 .06)*	.14 (-.01 .00)*	.12 (.00 .13)	.09 (-.01 .10)
		Maximal	.13 (.00 .00)*	.09 (-.02 .10)	.09 (-.04 .06)*	.09 (-.02 .00)*	.12 (.00 .11)*	.09 (-.01 .07)*
	Maximal	Minimal	.14 (-.01 .02)*	.14 (-.03 .14)*	.15 (-.05 .11)*	.09 (-.02 .00)*	.11 (.01 .16)	.09 (-.01 .06)*
		Intermediate	.13 (-.01 .01)*	.13 (-.03 .14)	.09 (-.02 .05)*	.09 (-.03 .00)*	.15 (.01 .16)	.09 (.00 .12)
		Maximal	.14 (-.01 .02)*	.14 (-.02 .13)*	.08 (-.02 .07)*	.09 (-.02 .00)*	.12 (.01 .13)	.09 (-.01 .06)*

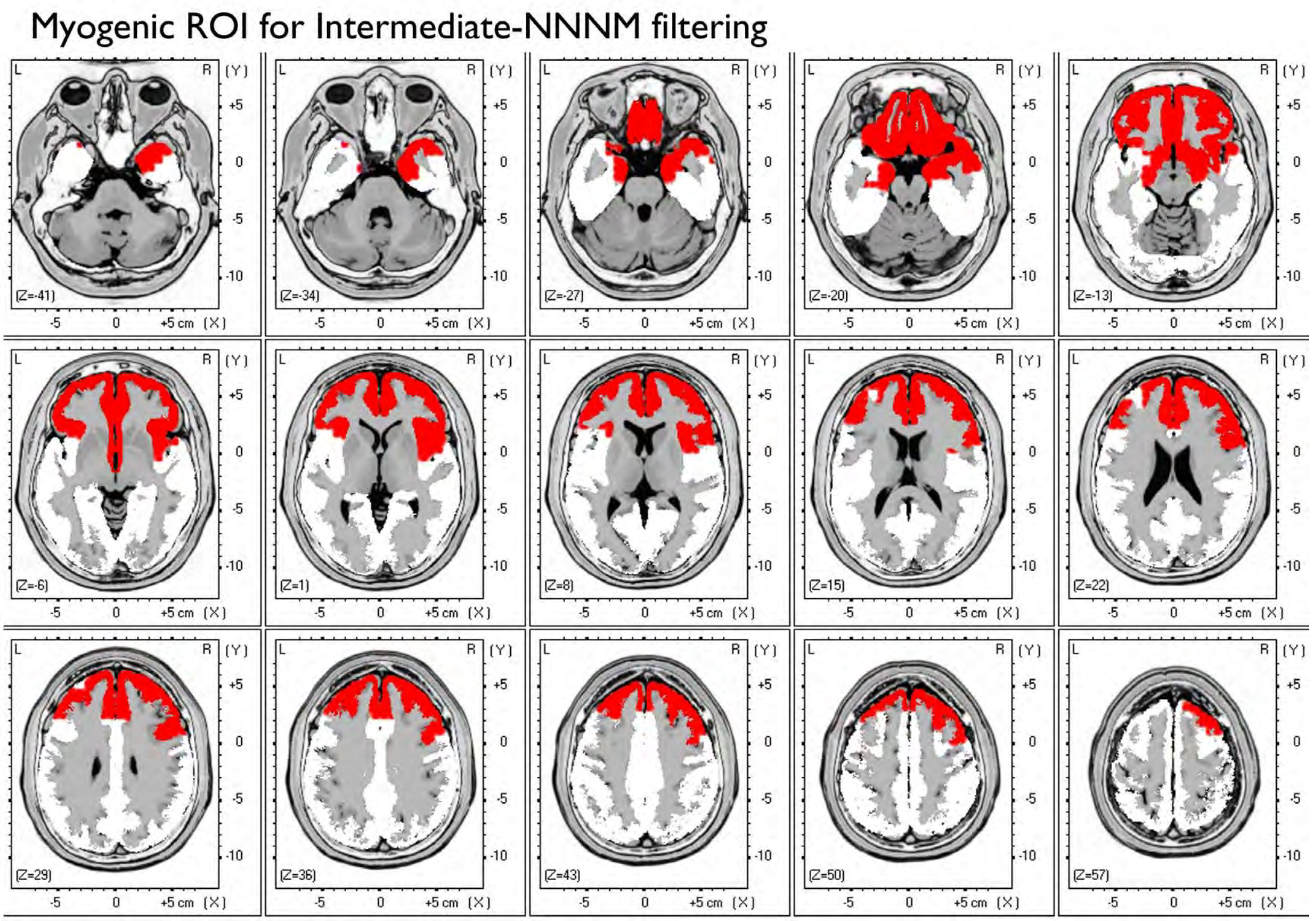
Note: [a] Corrected OR-CR vs. uncorrected OR-CR [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test.. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > p > .05, * p < .05, ** p < .01.

Supplementary Table 3.
Correction Artifact on the Scalp

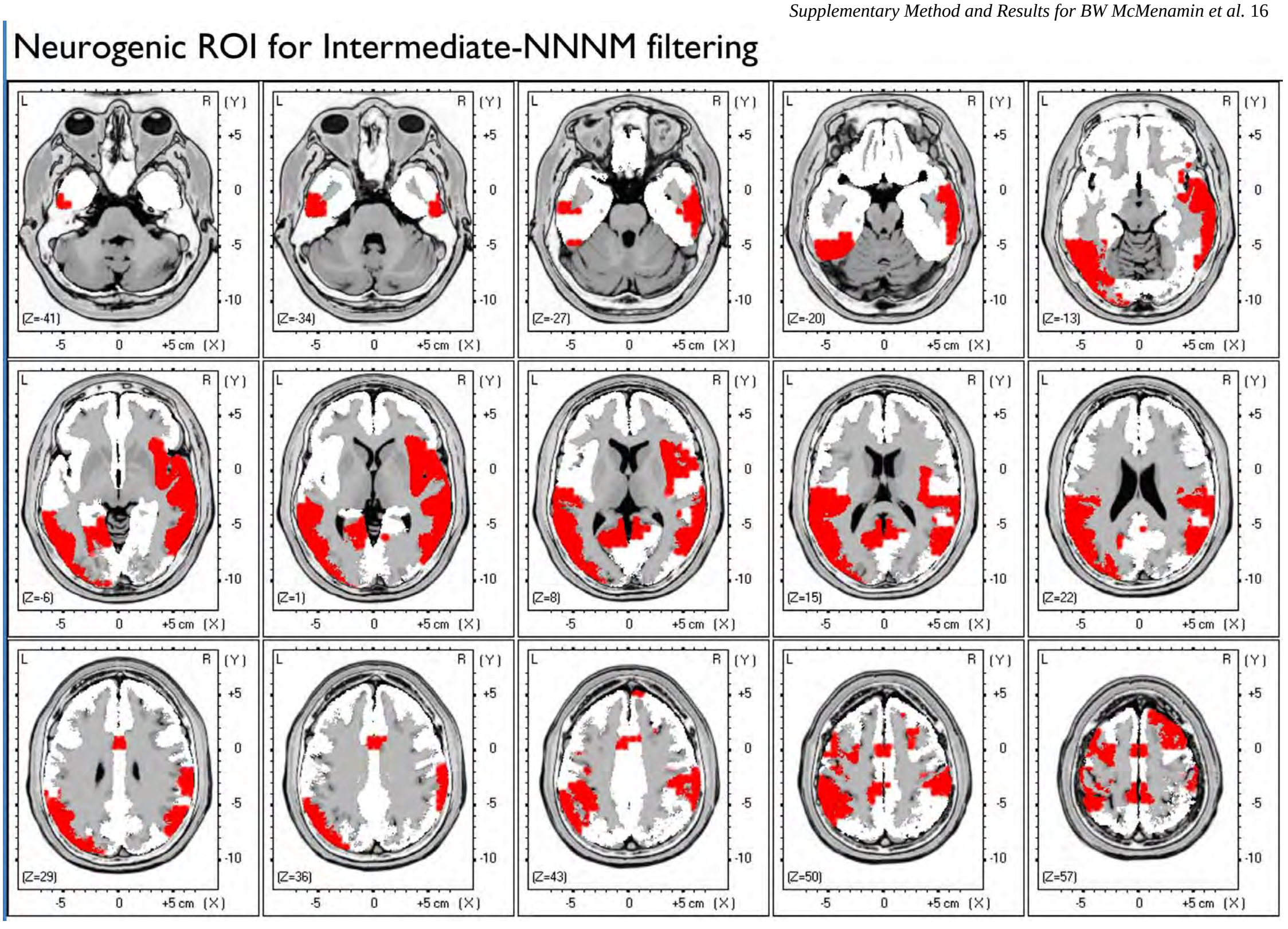
			ROI Median Contrast		ROI Peak Contrast	
	EMG Correction	NNNM Filtering	Myogenic[a]	Neurogenic[b]	Myogenic[a]	Neurogenic[b]
Corrected Contaminated vs. Uncorrected Analog[c]	None	Minimal	-1.25		-2.10†	
		Intermediate	-1.31		-2.06†	
		Maximal	-1.33		-2.12*	
	Minimal	Minimal	-1.29	-1.53	-2.08†	-2.41*
		Intermediate	-1.27	-1.57	-2.03†	-2.70*
		Maximal	-1.30	-1.61	-2.15*	-2.54*
	Intermediate	Minimal	-1.24	-2.46*	-1.95†	-4.01**
		Intermediate	-1.24	-1.93†	-2.11†	-3.74**
		Maximal	-1.34	-2.89**	-2.05†	-4.55**
	Maximal	Minimal	-1.21	-1.27	-1.84†	-3.62**
		Intermediate	-1.22	-0.64	-1.89†	-3.30**
		Maximal	-1.25	-1.51	-1.68	-3.39**
Equivalence[d]	None	Minimal	.15 (-0.11 .03)*		.09 (-.15 .00)	
		Intermediate	.14 (-0.13 .03)*		.12 (-.16 .00)	
		Maximal	.13 (-0.13 .03)*		.12 (-.12 .00)*	
	Minimal	Minimal	.15 (-0.11 .02)*	.09 (-.01 .00)*	.09 (-.15 .00)	.09 (-.02 .00)*
		Intermediate	.14 (-0.13 .03)*	.09 (-.03 .00)*	.12 (-.16 .00)	.08 (-.02 .00)*
		Maximal	.13 (-0.12 .03)*	.09 (-.01 .00)*	.12 (-.11 .00)*	.08 (-.02 .00)*
	Intermediate	Minimal	.14 (-0.11 .03)*	.09 (-.03 .00)*	.10 (-.14 .00)	.08 (-.02 -.01)*
		Intermediate	.14 (-0.11 .03)*	.09 (-.03 .00)*	.10 (-.14 .00)	.08 (-.02 -.01)*
		Maximal	.07 (-0.10 .02)	.09 (-.02 .00)*	.13 (-.15 .00)	.08 (-.03 -.01)*
	Maximal	Minimal	.14 (-0.11 .03)*	.09 (-.03 .01)*	.10 (-.14 .01)	.07 (-.03 -.01)*
		Intermediate	.15 (-0.12 .03)*	.08 (-.02 .01)*	.10 (-.14 .01)	.08 (-.03 -.01)*
		Maximal	.16 (-0.13 .03)*	.08 (-.03 .01)*	.13 (-.12 .01)*	.07 (-.03 -.01)*

Note: [a] Corrected OR-OT evaluated in the *neurogenic* ROI. [b] Corrected OR-CR vs. uncorrected OR-CR evaluated in the *myogenic* ROI. [c] Cells show the *t*-test.. [d] Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > p > .05, * p < .05, ** p < .01.

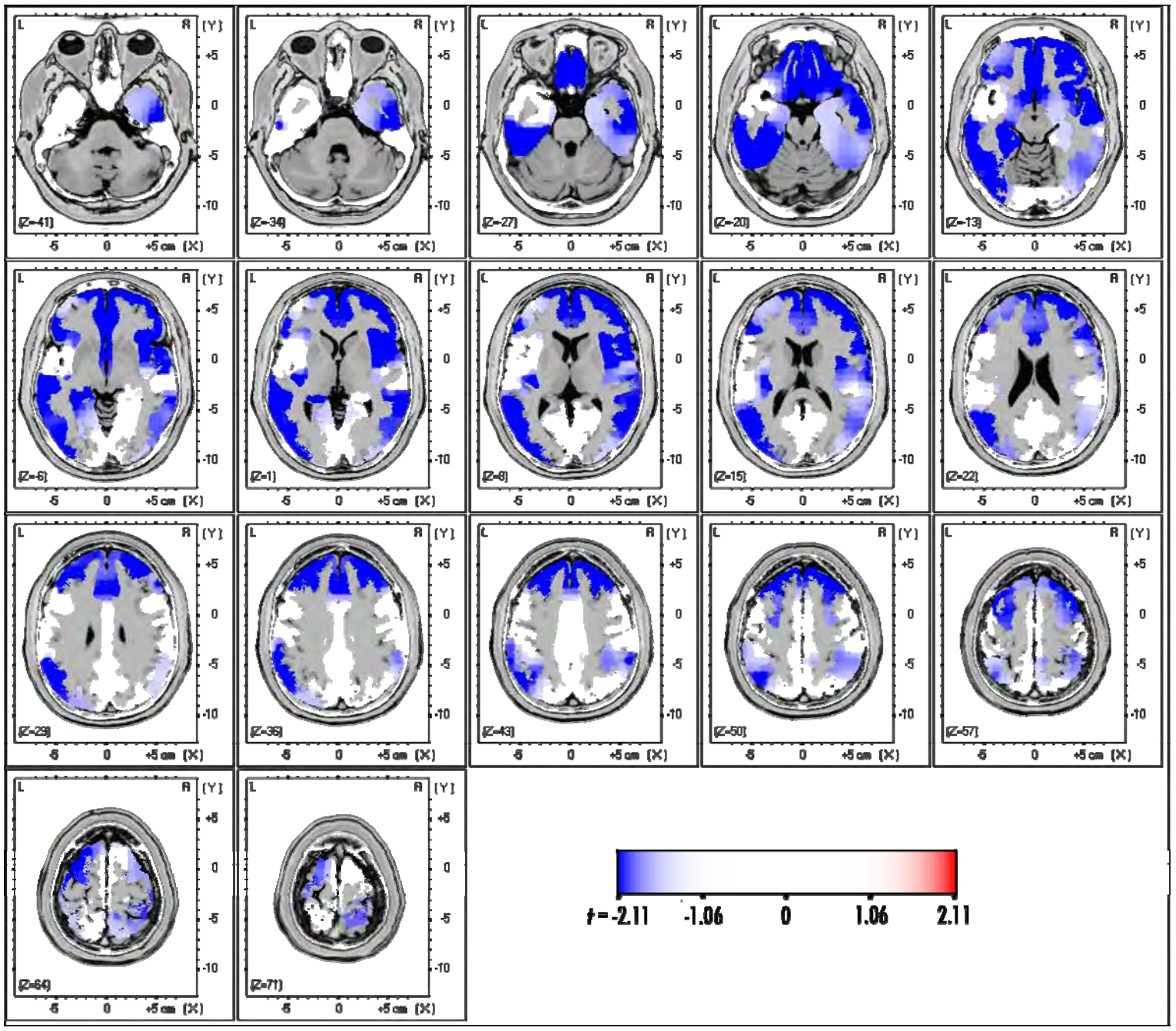
Supplementary Figures for Source-Modeling Analyses



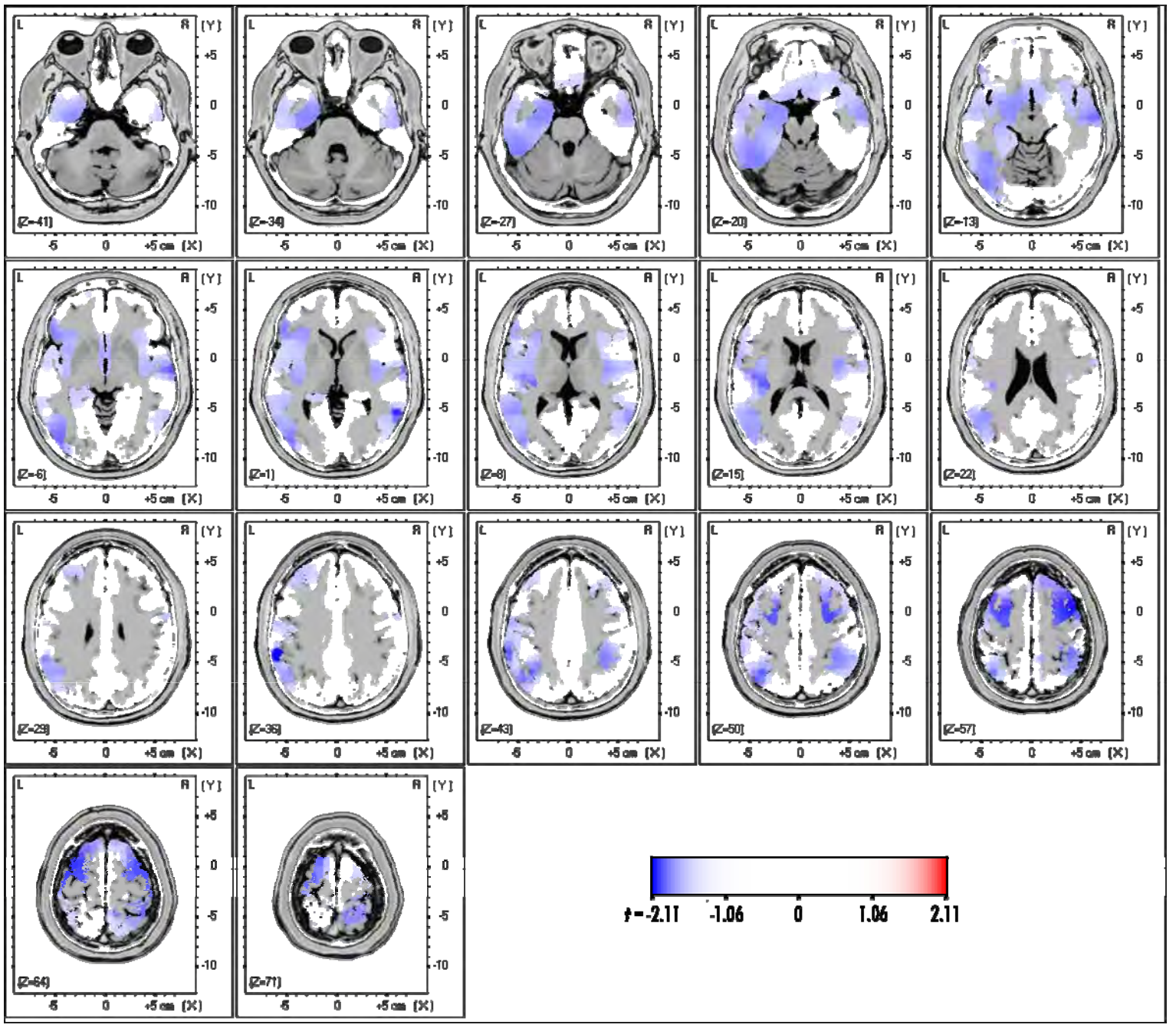
Supplementary Figure 14. *LORETA regions of interest (ROIs) for the myogenic contrast.*



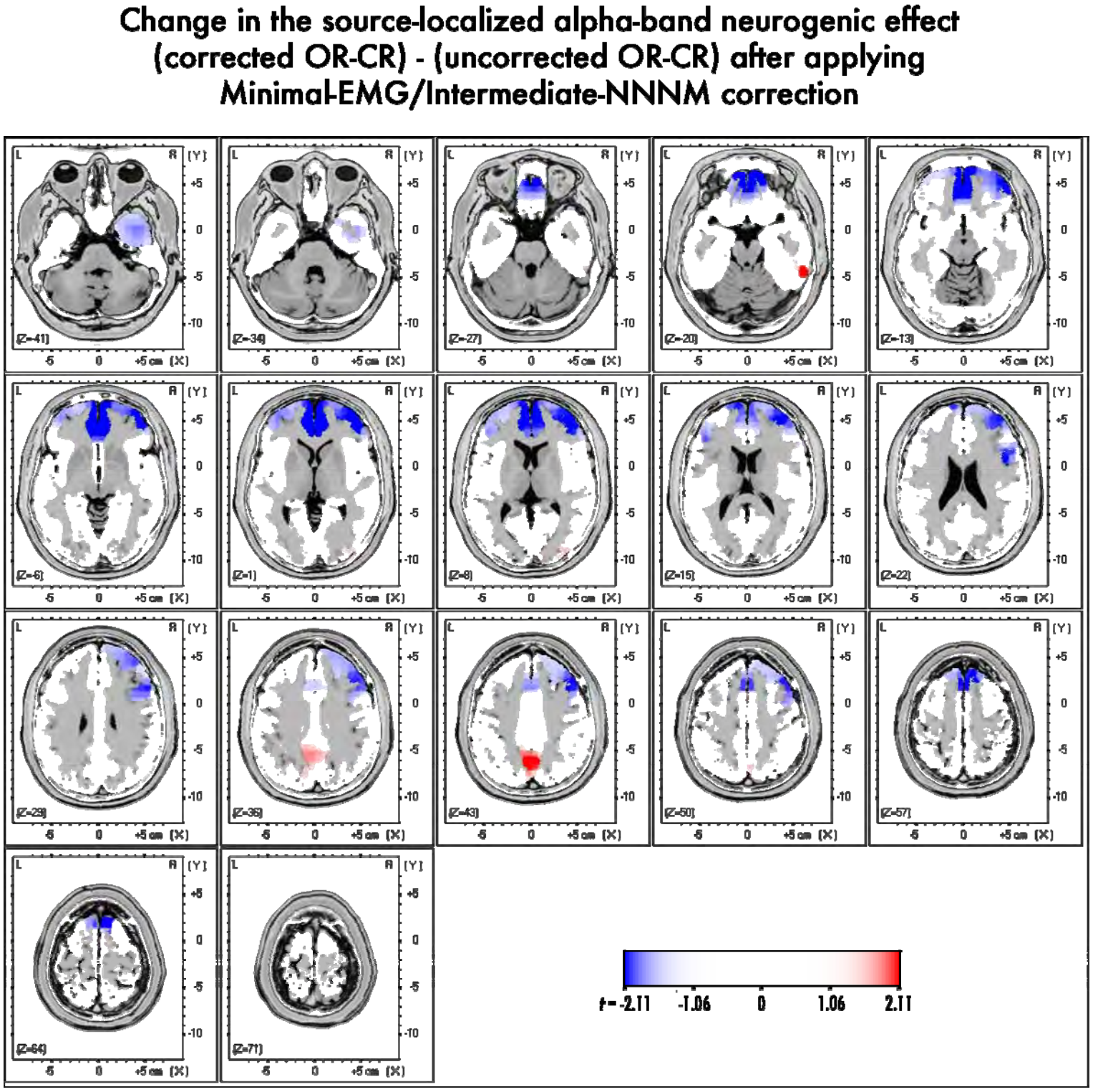
Supplementary Figure 15. *LORETA regions of interest (ROIs) for the neurogenic contrast.*



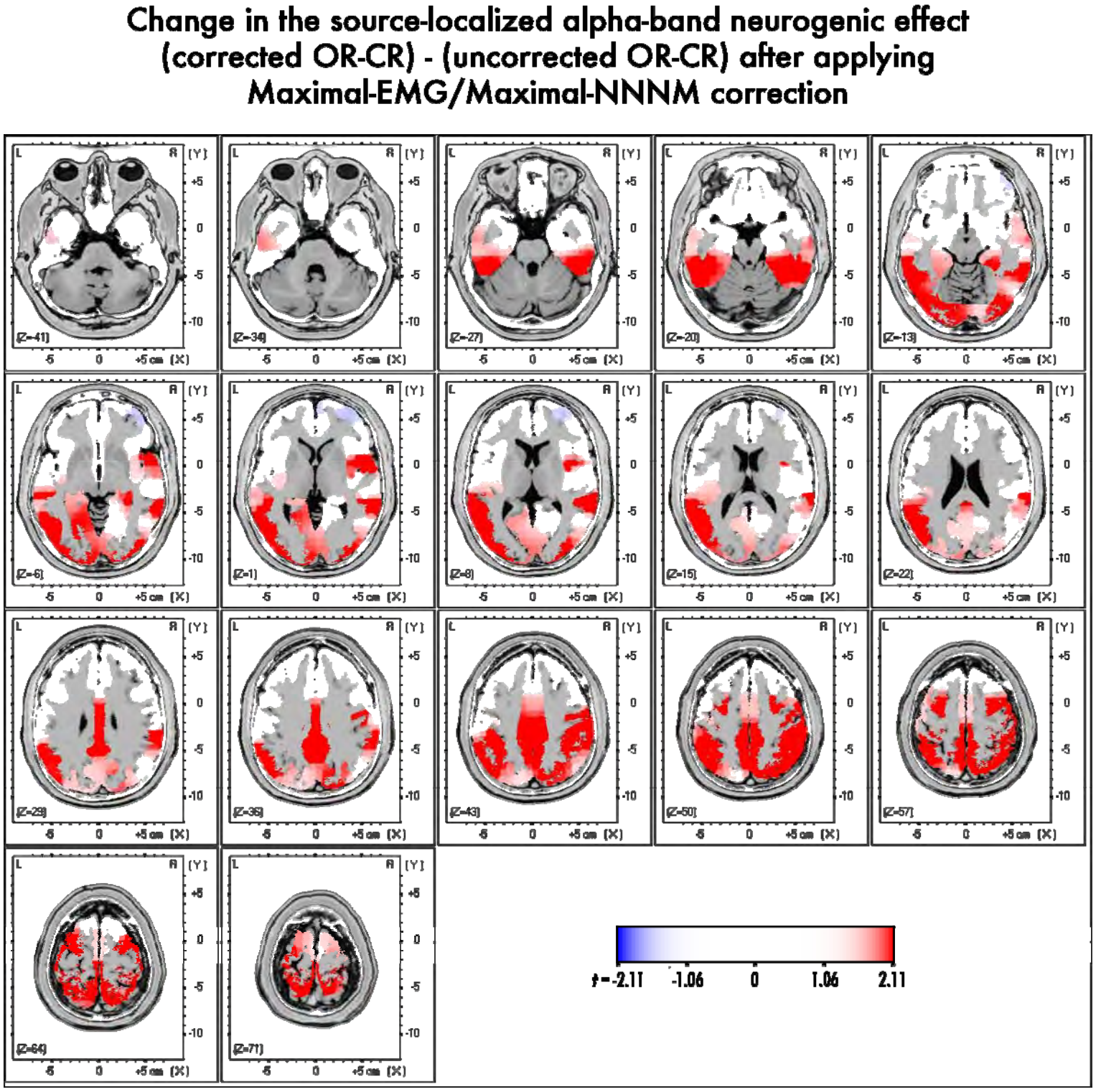
Supplementary Figure 16. *LORETA solution for the myogenic effect (OR-OT) after applying Minimal-EMG/Intermediate-NNNM correction*



Supplementary Figure 17. *LORETA solution for the myogenic effect (OR-OT) after applying Maximal-EMG/Maximal-NNNM correction*



Supplementary Figure 18. *Error in the LORETA solution for the neurogenic effect (OR-OT) after applying Minimal-EMG/Intermediate-NNNM correction*



Supplementary Figure 19. *Error in the LORETA solution for the neurogenic effect (OR-OT) after applying Maximal-EMG/Maximal-NNNM correction*

Supplementary Tables for Source-Modeling Analyses

Supplementary Table 4.
Sensitivity in the Source-space

			ROI Median Contrast			ROI Peak Contrast		
	EMG Correction	NNNM Filtering	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	None	Intermediate	-3.83**	3.83**	-2.30*	-6.19**	6.19**	-4.57**
	None	Maximal	-3.75**	3.75**	-2.30*	-6.33**	6.33**	-4.32**
	Minimal	Intermediate	-2.19**	1.22	-1.12	-3.30**	3.32**	-4.32**
	Maximal	Maximal	-1.23	1.09	0.48	-1.81†	2.91*	2.52*
Equivalence[e]	None	Intermediate	.60 (-1.79 -.54)	.35 (-2.72 -0.15)	.35 (.54 1.79)	.50 (-0.88 -.44)	.38 (0.44 .88)	.38 (-1.09 -0.41)
	None	Maximal	.66 (-1.06 .31)	.62 (-3.05 0.17)	.60 (.31 1.06)	.92 (-1.58 .03)	.75 (0.80 1.58)	.35 (-1.22 -0.43)
	Minimal	Intermediate	.97 (-0.72 -.02)*	2.58 (-1.11 0.33)	.62 (-.22 .84)	2.07 (-1.00 .03)	.24 (0.07 .29)*	.18 (-0.27 -0.10)
	Maximal	Maximal	.62 (-1.27 .32)	.18 (-0.11 0.17)*	.46 (-.07 .23)*	.97 (-0.47 .03)	.31 (0.07 .38)	1.64 (0.11 1.11)*

Note: [a] Corrected OR-OT vs. 0. [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Supplementary Table 5.
Specificity in the Source-space

			ROI Median Contrast			ROI Peak Contrast		
	EMG Correction	NNNM Filtering	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	None	Intermediate		2.01†	0.23		4.53**	-3.81**
	None	Maximal		1.90†	0.36		4.77**	-3.75**
	Minimal	Intermediate	.26	1.82†	0.38	2.70*	3.32**	-1.96†
	Maximal	Maximal	1.86†	1.98†	1.41	4.33**	3.39**	3.04**
Equivalence[e]	None	Intermediate		.93 (-0.01 0.69)*	10.20 (-4.83 6.04)*		.39 (.29 .78)	.29 (-0.87 -0.26)
	None	Maximal		2.86 (-0.14 3.40)	4.58 (-2.19 3.11)*		.39 (.30 .76)	.42 (-1.27 -0.37)
	Minimal	Intermediate	1.75 (-0.58 0.74)*	.73 (-0.04 0.70)*	.60 (-0.55 0.81)*	6.42 (.01 .09)*	.24 (.07 .29)	.19 (-0.35 0.01)
	Maximal	Maximal	3.07 (-0.15 3.07)*	1.86 (-0.04 2.34)	.60 (-0.50 2.70)*	1.59 (1.59 .44)*	.48 (.11 .45)*	1.90 (0.34 1.78)*

Note: [a] Corrected OR-CR vs. uncorrected OR-CR. [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Supplementary Table 6.
Correction Artifact in the Source-Space

			ROI Median Contrast		ROI Peak Contrast	
	EMG Correction	NNNM Filtering	Myogenic effect in Neurogenic ROI[a]	Neurogenic effect in Myogenic ROI[b]	Myogenic effect in Neurogenic ROI[a]	Neurogenic effect in Myogenic ROI[b]
Corrected Contaminated vs. Uncorrected Analog[c]	None	Intermediate	-2.01†		-4.53**	
	None	Maximal	-1.90†		-4.77**	
	Minimal	Intermediate	-1.89†	1.42	-3.23*	-2.62*
	Maximal	Maximal	-1.38	.31	-1.92	2.46*
Equivalence[d]	None	Intermediate	.70 (-0.69 .01)*		.55 (-.78 -.29)	
	None	Maximal	1.75 (-3.40 .14)		.54 (-.76 -.30)	
	Minimal	Intermediate	1.75 (-1.67 .07)	.46 (-0.33 .06)*	.27 (-.27 .06)*	.65 (-.52 -.06)*
	Maximal	Maximal	1.75 (-3.88 .76)*	2.40 (-0.42 .57)*	1.50 (-.99 .03)*	.68 (.05 .51)*

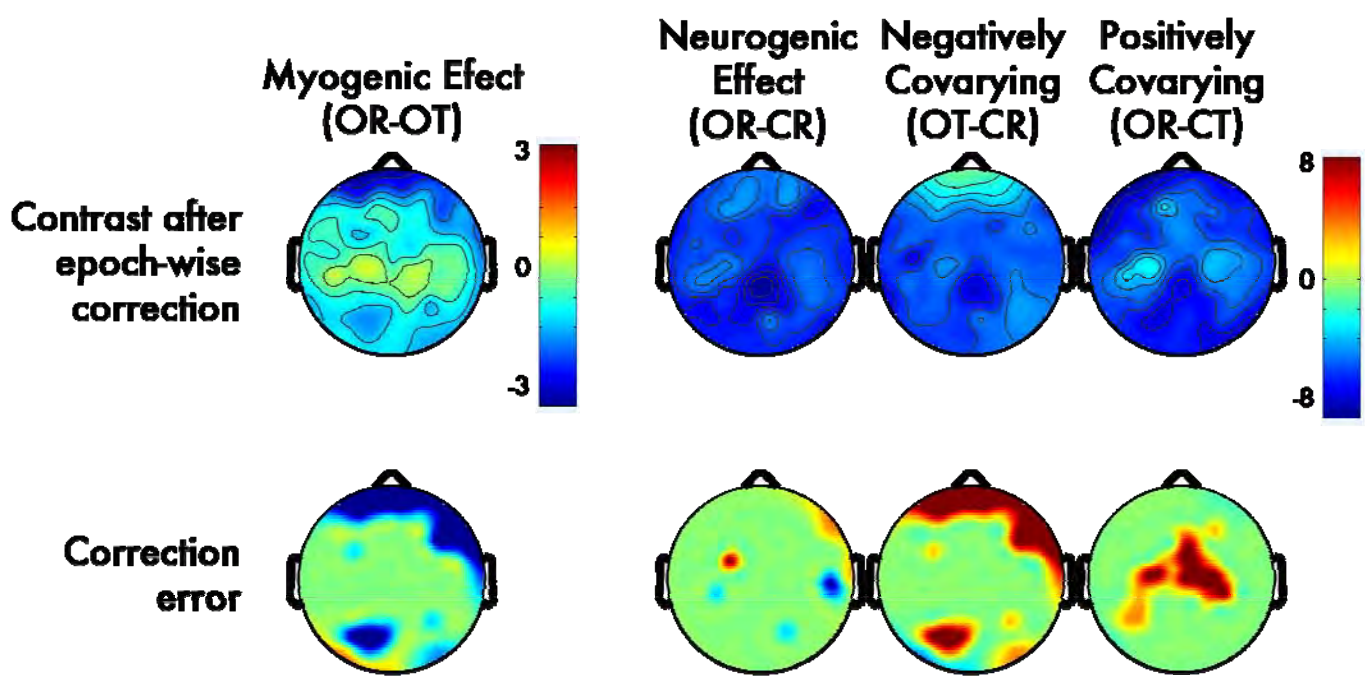
Note: [a] Corrected OR-OT contrast in the neurogenic ROI [b] Corrected OR-CR vs. uncorrected OR-CR in the myogenic ROI. [c]Cells show the *t*-test. [d]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Regression-Based EMG Correction

McMenamin *et al* (McMenamin, Shackman, Maxwell, Greischar, & Davidson, 2009) tested the sensitivity and specificity of a variety of GLM-based myogenic correction techniques using the dataset described in the present report. Only intra-individual correction ("epoch-wise regression")—which removes epoch-by-epoch variance in alpha-band activity predicted by EMG-band activity (e.g., 70-80Hz) separately for each participant and electrode—showed consistently adequate performance on the scalp. No method proved valid in the source-space. Although the analytic pathway employed by McMenamin et al. was similar to the present report, it differs in two ways that makes it difficult to compare regression- and ICA-based EMG correction techniques. First, McMenamin et al. (2009) used relatively small ROIs to quantify myogenic and neurogenic effects. Here, larger ROIs were used with the aim of understanding the impact of EMG artifact and correction in regions characterized by less extreme contamination, in the case of the myogenic ROI, and more modest signals of interest, in the case of the neurogenic ROI. Second, only the median *t*-scores were examined in the prior report. Here, the extreme *t*-scores were also examined as an index of "worst case" performance.

To permit direct comparison of the two techniques, the validity of intra-individual EMG correction was re-assessed on the scalp using the methods employed in the present report. Data were first pre-processed to remove a minimal degree of non-myogenic artifact (i.e., "Minimal NNNM" protocol), as in McMenamin et al (2009). Log-transformed power in the alpha (8-13 Hz) and EMG (70-80 Hz) bands was then calculated for each 1.024s epoch. Next, general linear models were computed for each combination of channel and participant. The resulting EMG-residualized estimates of alpha power were then averaged across epochs for each combination of condition, channel and participant.

This approach showed adequate sensitivity for all contrasts (Supplementary Figure 20 and Table 7, see below). In contrast to the previous report, specificity was deemed poor for the positively-covarying contrast. Specificity was adequate for the remaining contrasts (Supplementary Table 8). The amount of correction-induced artifact was also adequate (Supplementary Table 9).



Supplementary Figure 20. ***Effects of GLM-Based EMG Correction*****. First Row:** Topographic maps depict spline-interpolated *t*-scores for contrasts of interest after applying regression-based EMG correction. **Second Row:** Topographic maps depict *t*-scores corresponding to the correction error for each contrast of interest, measured as the difference between the corrected version of a contaminated and it's uncorrected artifact-free analogue (e.g. corrected OT-CR compared to uncorrected OR-CR). Negative values are shown in blue (dark-blue: $p < .05$; light-blue: $p < .10$; green: $p > .10$); Positive values are shown in red (dark-red: $p < .05$; light-red: $p < .10$; green: $p > .10$).

Supplementary Table 7.
Sensitivity of GLM-Based EMG Correction.

	ROI Median Contrast			ROI Peak Contrast		
	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Myogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	-.80	1.52	.99	-2.27*	2.34*	2.74*
Equivalence[e]	.15 (-.10 .04)*	0.11 (-.02 .11)*	.13 (-.04 .12)*	.14 (-.14 -.01)*	.09 (0.01 .14)	.13 (.02 .16)

[a] Corrected OR-OT. [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test.. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Supplementary Table 8.
Specificity of GLM-Based EMG Correction.

	ROI Median Contrast			ROI Peak Contrast		
	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]	Neurogenic[a]	Negatively-Covarying[b]	Positively-Covarying[c]
Corrected Contaminated vs. Uncorrected Analog[d]	1.39	1.48	1.50	2.35*	2.59*	2.25*
Equivalence[e]	.11 (.00 .02)*	.16 (-.02 .10)*	.19 (-.04 .23)	.11 (.00 .03)*	.11 (.02 .13)	.11 (.01 .18)

[a] Corrected OR-CR vs. uncorrected OR-CR [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test.. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Supplementary Table 9.
Correction-Induced Artifact for GLM-Based EMG Correction.

	ROI Median Contrast		ROI Peak Contrast	
	Myogenic[a]	Neurogenic[b]	Myogenic[a]	Neurogenic[b]
Corrected Contaminated vs. Uncorrected Analog[c]	1.47	-1.36	2.86*	-1.96†
Equivalence[d]	**.11 (-.01 .03)***	**0.21 (-.20 .04)***	**.09 (.01 .04)***	**-0.17 (-.21 .00)***

[a] Corrected OR-OT in the Neurogenic ROI. [b]Corrected OR-CR vs. uncorrected OR-CR in the myogenic ROI. [c]Cells show the *t*-test.. [d]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically equivalent to the uncontaminated data for the analogous contrast. † .10 > *p* > .05, * *p* < .05, ** *p* < .01.

Post Hoc Model Order Estimation

Just prior to submitting this report, we became aware of a recent ERP study (Mouraux & Iannetti, 2009) exploiting the Bayesian model order estimation procedure used by the FSL Melodic software package (Beckmann & Smith, 2002, 2004; Rajan & Rayner, 1997). Using Matlab code kindly provided by the lead author, Andre Mouraux, a post hoc analysis was performed to determine the number of dimensions ("model order") characterizing the (128-channel) dataset. This indicated that the median number of dimensions was 39.5 (*SD*: 6.8) with a range of 23-53, which is broadly consistent with prior reports (e.g., Naeem, Brunner & Pfurtscheller, 2009; Onton, Westerfield, Townsend & Makeig, 2006). This finding suggests that the 64-component extraction used in the present report was sufficient to avoid underfitting, but moderately overfitted most participants. Nevertheless, given the potential limitations of Bayesian methods (e.g., Hesse & James, 2004), it would be helpful for future investigations to examine the utility of other model order estimation algorithms (e.g., Cordes & Nandy, 2006; Hesse, 2008).

Supplementary References

Beckmann, C. F., & Smith, S. A. (2002). Probabilistic independent component analysis for functional magnetic resonance imaging. *FMRIB Technical Report TR02CB1*. *http://www.fmrib.ox.ac.uk/analysis/techrep/*.

Beckmann, C. F., & Smith, S. A. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Tranactions in Medical Imaging, 23*, 137-152.

Cordes, D. & Nandy, R. R. (2006). Estimation of the intrinsic dimensionality of fMRI data. *NeuroImage, 29*, 145-154.

Gibbs, F. A. & Gibbs, E. L. (1950). *Atlas of electroencephalography* (2nd ed.). Cambridge, MA: Addison-Wesley Press.

Gray, H. (1918/2000). *Anatomy of the human body*: http://www.bartleby.com/107.

Groppe, D. M. , Makeig, S. & Kutas, M. (2009) Identifying reliable independent components via split-half comparisons. *NeuroImage, 45*, 1199-1211.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*, 77-89.

Hesse, C. W. (2008) Model order estimation for blind source separation of multichannel magnetoencephalogram and electroencephalogram signals. *Engineering in Medicine and Biology Society 30th Annual International Conference of the IEEE (EMBS 2008), 1*, 3348-3351.

Hesse, C. W. & James, C. J. (2004). Stepwise model order estimation in blind source separation applied to ictal EEG. *Engineering in Medicine and Biology Society 26th Annual International Conference of the IEEE (IEMBS 2004)*, 1, 986-989.

McMenamin, B. W., Shackman, A. J., Maxwell, J. S., Greischar, L. L., & Davidson, R. J. (2009). Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. *Psychophysiology, 46*, 578-592.

Mouraux, A., & Iannetti, G. D. (2009). Nociceptive laser-evoked brain potentials do not reflect nociceptive-specific neural activity. *Journal of Neurophysiology, 101*, 3258-3269.

Naeem, M., Brunner, C. & Pfurtscheller, G. (2009). Dimensionality reduction and channel selection of motor imagery electroencephalographic data. Computational Intelligence and Neuroscience. Epub ahead of print.

Onton, J., Westerfield, M., Townsend, J., & Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. Neuroscience and Biobehavioral Reviews, 30, 808-822.

Rajan, J. J., & Rayner, P. J. W. (1997). Model order selection for the singular value decomposition and the discrete Karhunen-Loeve transform using a Bayesian approach. *Vision Image Signal Processing IEEE Proc, 144*, 116-123.

Viola, F. C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T. & Debener, S. (2009) Semi-automatic identification of independent components representing EEG artifact. *Clinical Neurophysiology, 120*, 868-877.

Zeman, P. M., Till, B. C., Livingston, N. J., Tanaka, J. W., & Driessen, P. F. (2007). Independent component analysis and clustering improve signal-to-noise ratio for statistical analysis of event-related potentials. *Clinical Neurophysiology, 118*, 2591-2604.