Published in final edited form as:
Neuroimage. 2010 February 1; 49(3): 2416. doi:10.1016/j.neuroimage.2009.10.010.

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

Brenton W. McMenamin[*],
Department of Psychology, Center for Cognitive Science, University of Minnesota – Twin Cities

Alexander J. Shackman[*],
Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison

Jeffrey S. Maxwell,
U.S. Army Research Laboratory, Aberdeen, MD, Laboratory for Affective Neuroscience, University of Wisconsin – Madison

David R. W. Bachhuber,
Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison

Adam M. Koppenhaver,
Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison

Lawrence L. Greischar, and
Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison

Richard J. Davidson
Laboratory for Affective Neuroscience, Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin – Madison

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 (8–13Hz). 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,

MS Correspondance: Brenton W. McMenamin, Alexander J. Shackman or Richard J. Davidson, mcmen020@umn.edu, shackman@wisc.edu or rjdavids@wisc.edu, Laboratory for Affective Neuroscience, University of Wisconsin—Madison, 1202 West Johnson Street, Madison, Wisconsin 53706.
[*]Brenton McMenamin and Alexander Shackman are co-first authors on this manuscript.

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, McMenamin, Slagter, Maxwell, Greischar & Davidson, 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 2Hz, making even the widely used alpha band (8–13Hz) vulnerable to muscle artifacts (Lee & Buchsbaum, 1987; Willis, Nelson, Rice & Black, 1993; Van Boxtel, 2001). Second, EMG can often be detected across the entire scalp (Goncharova, McFarland, Vaughan & Wolpaw, 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 Figure 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, Davidson, Seulis, Saron & Weisman, 1992; Waterink & Van Boxtel, 1994), facial mimicry (Dimberg, Thunberg & Elmehed, 2000), vocalization (Brooker & Donald, 1980), and induced emotional states (Borden, Peterson & Jackson, 1991; Coan & Allen, 2003; Bradley, Codispoti, Cuthbert & Lang 2001).

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 25Hz, 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, Kim & McCall, 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, Cacioppo & Vanman, 2007) and may differ somewhat between spontaneous and voluntary contractions (Davidson, Shackman & Maxwell, 2004; Morecraft & 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, Makeig, Westerfield, Townsend, Courchesne & Sejnowski, 2000; 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, Ekman, Saron, Senulis & Friesen, 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 & Frishkoff, 2006; Ille, Berg & Scherg, 2002; Koskinen & 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., 70–80Hz). 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, Coan & Nazarian, 2004; Davidson, Jackson & Larson, 2000), and McMenamin, Shackman, Maxwell, Greischar and Davidson (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, Greischar, Rawlings, Ricard & Davidson 2004), GLM-based tools cannot be applied to studies relying on event-related spectral perturbation (ERSP) measures (Onton, Westerfield, Townsend & Makeig, 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.[1]

A second class of EMG correction methods employs ICA to decompose the EEG time-series into a set of temporally independent components (Delorme, Sejnowski & Makeig, 2007; James & Hesse, 2005; Onton et al., 2006; Onton & Makeig, 2006; Makeig, Debener, Onton, & Delorme, 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

[1]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.

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, Makeig, Humphries, Lee, McKeown, Iragui & Sejnowski, 2000), attempts to assess its validity have been limited. Many validation studies have relied on small samples of *ad hoc* data (Delorme, Westerfield & Makeig, 2007; Jung et al., 2000; Wallstrom, Kass, Miller, Cohn & Fox et al., 2004; Flexer, Bauer, Pripfl, & Dorffner, 2005; Ting, Fung, Chang, & Chan, 2006; Frank & Frishkoff, 2006). While others have used simulations (e.g., Crespo-Garcia, Atienza & Cantero, 2008; De Clercq, Vergult, Vanrumste, Van Hees, Palmini, Van Paesschen & Van Huffel, 2005; Delorme et al., 2007; Fitzgibbon, Powers, Pope & Clark, 2007; Frank & Frishkoff, 2006; Romero, Mananas & Barbanoj, 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, Vercueil, Krainik, Segebarth, Kahane & David, 2007; Hoffmann & 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 investigations of memory (Freuenberger, Fellinger, Sauseng, Gruber & Klimesch, 2009; Gevins & Smith, 2000; Hamidi, Slagter, Tononi & Postle, 2009), perception and attention (Romei Brodbeck, Michel, Amedi, Pascual-Leone & Thut, 2008; Thut & Miniussi, 2009), emotion (Coan & Allen, 2003; Davidson et al., 1990), temperament and individual differences (Carver & Harmon-Jones, 2009; Shackman, McMenamin Maxwell, Greischar & Davidson in press), and psychopathology (Thibodeau, Jorgensen & Kim, 2006; DeRubeis, Siegle & Hollon, 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, Thorne, Edmonds, Schneider, Eichele & Debener, 2009). Without such evidence, poor validity might simply reflect inadequate training or an ambiguous classification protocol. Accordingly, the inter-rater reliability was computed.

Method

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 (8–13Hz), 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 Bonnet & Arand, 2001; Freeman, Holmes, Burke & Vanhatalo, 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).[2] Data reduction used a combination of EEGLAB (Delorme & 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 (±100μV for >20s) or gross artifacts (±100μV for >4 channels) were manually identified and rejected. A 0.5Hz high-pass filter was used to attenuate channel drift and better satisfy ICA's stationarity assumption (Onton et al., 2006).[3] 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, Westerfield & Makeig, 2007), 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.[4] This was done in a single step as part of the ICA using the EEGLAB *runica* command, implementing the extended Infomax algorithm (Bell & Sejnowski, 1995; Lee, Lewicki, Girolami & Sejnowski, 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, 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 Figure 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

[2]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., 1000Hz) and more conservative anti-aliasing filter (e.g., 250Hz) to compensate for non-zero filter roll-off.

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

[4]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, Adali & Calhoun, 2007; Lawrence & 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 & Velicer, 1996). Second, it has been suggested (Onton et al., 2006; Romero, Mananas & Barbanoj, 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 × 500Hz = 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, Scope, Pascalis, Buckley & Makeig, 2009), at the potential expense of spatial resolution (Srinivasan, Tucker & Murias, 1998; Michel, Murray, Lantz, Gonzalez, Spinelli & Grave de Peralta, 2004).

variance were categorized as *Low-Variance*[5]. In cases where the determination was unambiguous, exceptions were made. As described in the Supplementary Method and Results (Supplementary Figures 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 & Krippendorff, 2007), was excellent, α=.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., 2007). 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 (8–13Hz) band using Welch's (1967) method

[5]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).

on sliding Hanning-windowed epochs (50% overlap). Estimates were $\log_{10}$ transformed to normalize the distribution (Allen et al., 2004; Gasser, Bacher & Mocks, 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, Michel & Lehmann, 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, Neelin, Peters & Evans, 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, North & Biver, 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 & Frishkoff, 2006; Debener, Strobel, Sorger, Peters, Kranczioch, Engel & Goebel, 2007; Freyer, Becker, Anami, Curio, Villringer & Ritter, 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) s*pecificity*, 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.

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 $p < .001$, using a cluster-extent threshold to correct for multiple comparisons (Nichols & 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 & 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 $\alpha = .05$, one must demonstrate that the 90^{th}-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 ($p < .001$ 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.

Performance ratings—For each contrast and correction protocol, sensitivity was rated as: *poor* ($p_{Median\ t\text{-test}} < .05$ or median $p_{TOST} > .05$), *questionable* ($.10 > p_{Median\ t\text{-test}} > .05$ or median $p_{TOST} < .05$), *adequate* ($p_{Median\ t\text{-test}} > .10$ and median $p_{TOST} < .05$) or *excellent* ($p_{Peak\ t\text{-test}} > .05$ and all $p_{TOST} < .05$). 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 (Figure 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 (Figure 2) with qualitatively similar peak locations across the three NNNM filters. The median *t*-scores (*t*s = −2.30 to −2.44, *p*s < .04; η_p^2s = 0.24 to 0.27) and extreme *t*-scores (*t*s = −2.82 to −3.21, *p*s < .01; η_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 (Figure 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 (*t*s < −4.37, *p*s < 0.001). The peak difference occurred at midline parietal sites (Figure 1), and defined the neurogenic ROIs (Figure 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 (*t*s = −7.35 to −7.49, $p < .001$; η_p^2s = 0.77 to 0.78) and extreme *t*-scores (*t*s = −8.85 to −9.12, $p < .001$; η_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 (Figure 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 (Figure 5).

Descriptive Statistics for ICA

Classification: Figure 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 (Figure 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 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 (Figure 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.

Figure 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 (Figure 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 (Figure 2, Figure 4, and Figure 5), specificity (Figure 4, Figure 5, and Figure 8), and correction-induced artifact (Figure 2 and Figure 8): the Minimal-EMG protocol paired with Minimal- or Intermediate-NNNM filtering and the Maximal-EMG protocol paired with Minimal- or Maximal-NNNM filtering.[6] 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 (Figure 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 Figure 14). The median *t*-scores (*t*s = −3.75 to −3.83, *p*s < .01, η_p^2 = 0.47 to 0.48) and extreme *t*-scores (*t*s = −6.19 to −6.33, *p*s < .001, η_p^2 = 0.71) were similar across the two filters.

[6]Exploratory analyses indicated dramatically worse performance in the “EMG” band (70–80Hz). 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.

Neurogenic ROI—The Berger maneuver was associated with widespread attenuation of current density across the posterior cortex (Figure 9). Using the neurogenic (OR-CR) contrast, ROIs were generated for the Intermediate- and Maximal-NNNM filters (765–766 voxels) across posterior voxels (Supplementary Figure 15). Median (*t*s = −3.92 to −3.96, *p*s < .01, η_p^2 = .49) and extreme *t*-scores (*t*s = −7.26 to −7.68, *p*s < .001, η_p^2 = .77 to .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 (Figure 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 (*t*s > 3.75, *p*s < .01, η_p^2s > 0.47) and extreme *t*-scores (*t*s > 6.19, *p*s < .001, η_p^2s > 0.71). The attenuation was reduced, but still observable, in the posterior neurogenic ROI (median *p*s < 0.08, η_p^2s > 0.18; extreme *t*s > 4.53, *p*s < 0.01, η_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 (*t*s < −2.30, *p*s < .04, η_p^2s > 0.25) and extreme *t*-scores (*t*s < −4.32, *p*s < .001, η_p^2s > 0.54). These deleterious effects were weaker in the neurogenic ROI, reaching significance for the extreme (*t*s < −3.81, *p*s < .01, η_p^2s > 0.48) but not the median *t*-scores (*p*s > .72, η_p^2s < 0.01).

Sensitivity

Myogenic contrast (OR-OT)—Although both protocols quantitatively reduced EMG contamination in the myogenic ROI (Figure 10a, red points), their sensitivity was poor (Supplementary Table 4, Supplementary Figures 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 (Figure 10c).

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 Figure 8c and Figure 10d.

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 (Figure 10b, blue points, Supplementary Table 5, Supplementary Figures 18–19).[7] 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

[7]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).

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 Figure 10c).

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 Figure 10d.

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 Figure 10a, 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 Figure 10b, 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. 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 (Figure 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 (Figure 1 and Figure 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; Figure 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 (Figure 9). Furthermore, covariation in neurogenic and myogenic activity significantly altered the

magnitude of neurogenic effects in anterior regions of the brain (Figure 9). Similar, albeit less dramatic, distortions were found in posterior regions, particularly for negatively covarying activity.

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 Figure 6). Descriptively, the choice of which components to exclude had a marked impact on the amount of variance retained for analysis (Figure 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 (Figures 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, Mullinger, Niazy & Bowtell, 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.

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,

1. 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, Figure 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.
2. 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 (Figure 3) evinced under- or overcorrection (Supplementary Tables 1–3).
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 Figure 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 (Figure 4–Figure 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 (α=.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, Brunner & Pfurtscheller, 2009;Ryali, Glover, Chang & Menon, 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 & Smith, 2004;Calhoun, Adali, Pearlson, & Pekar, 2001;Li, Adali & Calhoun, 2007) and has occasionally been done in the EEG literature (Mouraux & 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, Franciotti, Romani, & Pizzella, 2008). Stepwise ICA fitting, in which model order and sources are estimated simultaneously, is another possible approach (Hesse & 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 & Sejnowski, 1995; James & 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 & Allen, 2003; Davidson et al., 1990; Shackman, Sarinopoulos, Maxwell, Pizzagalli, Lavric & Davidson, 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, Hudry & le Coutre, 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, Gorodnitsky & Kutas, 2004; Romero et al., 2008; Tang, Liu & Sutherland, 2005). It might also be fruitful to investigate the utility of signal-space projection methods (Nolte & Curio, 1999; Tesche, Uusitalo, Ilmoniemi, Huotilainen, Kajola & Salonen, 1995; Uusitalo & Ilmoniemi, 1997) or ICA performed in the frequency-domain, which may improve the separation of sources with overlapping spectral profiles (Annemuller, Sejnowski & Makeig, 2003, 2004; Lee, Lee, Jolesz & Yoo, 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., 70–80Hz) 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 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, Allen & McKnight, 2006) and affective disorders (Thibodeau et al., 2006) remains untested.

Second, our conclusions derive from analyses of the alpha band (8–13Hz) 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 (4–8Hz) 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 (>30Hz), that lie closer to peak EMG activity. Indeed, exploratory analyses indicated poor performance for ICA in the 70–80Hz 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 & 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 (Figure 3) should minimize contributions from the anterior, lateral, and posterior muscle groups (Supplementary Figure 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, Murray, Lantz, Gonzalez, Spinelli & Grave de Peralta, 2004;Nazarpour, Wongsawat, Sanei, Chambers & Oraintara, 2008). Fifth, the results of the present study and several others (e.g., Whitham et al., 2007; Yuval-Greenberg, Tomer, Keren, Nelken & Deouell, 2008; Yuval-Greenberg, Keren, Tomer, Nelken & Deouell, 2009;Shackman et al., 2009) indicate that findings in the upper frequency bands (i.e., beta: 14–13Hz; gamma: >30Hz), 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

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).

References

Allen JJB, Coan JA, Nazarian M. Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. Biological Psychology 2004;67:183–218. [PubMed: 15130531]

Anemuller J, Sejnowski TJ, Makeig S. Complex independent component analysis of frequency-domain electroencephalographic data. Neural Networks 2003;16:1311–1323. [PubMed: 14622887]

Anemuller J, Sejnowski TJ, Makeig S. Reliable measurement of cortical flow patterns using complex independent component analysis of electroencephalographic signals. Lecture Notes in Computer Science 2004;3195:1009–1016.

Beckmann CF, Smith SM. Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Transactions on Medical Imaging 2004;23:137–152. [PubMed: 14964560]

Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. Neural Computation 1995;7:1004–1034.

Berger, H. On the electroencephalogram of man. In: Gloor, P., translator. The fourteen original reports on the human electroencephalogram. NY: Elsevier; 19291969.

Bonnett MH, Arand DL. Impact of activity and arousal upon spectral EEG parameters. Physiology & Behavior 2001;74:291–298. [PubMed: 11714491]

Borden JW, Peterson DR, Jackson EA. The Beck anxiety inventory in nonclinical samples: Initial psychometric properties. Journal of Psychopathology and Behavioral Assessment 1991;13:345–356.

Bradley MM, Codispoti M, Cuthbert BN, Lang PJ. Emotion and motivation I: Defensive and appetitive reactions in picture processing. Emotion 2001;1:276–298. [PubMed: 12934687]

Brooker BH, Donald MW. Contribution of the speech musculature to apparent human EEG asymmetries prior to vocalization. Brain and Language 1980;9:226–245. [PubMed: 7363067]

Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. Human Brain Mapping 2001;14:140–151. [PubMed: 11559959]

Carver CS, Harmon-Jones E. Anger is an approach-related affect: Evidence and implications. Psychological Bulletin 2009;135:183–204. [PubMed: 19254075]

Chung JW, Kim C, McCall WD Jr. Effect of sustained contraction on motor unit action potentials and EMG power spectrum of human masticatory muscles. Journal of Dental Research 2002;81:646–649. [PubMed: 12202649]

Coan, JA.; Allen, JJB. The asymmetrical brain. Cambridge, MA: MIT Press; 2003. The state and trait nature of frontal EEG asymmetry in emotion; p. 565-615.

Coan JA, Allen JJB, McKnight PE. A capability model of individual differences in frontal EEG asymmetry. Biological Psychology 2006;72:198–207. [PubMed: 16316717]

Cohen BH, Davidson RJ, Senulis JA, Saron CD, Weisman DR. Muscle tension patterns during auditory attention. Biological Psychology 1992;33:133–156. [PubMed: 1525291]

Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. Journal of Computer Assisted Tomography 1994;18:192–205. [PubMed: 8126267]

Crespo-Garcia M, Atienza M, Cantero JL. Muscle artifact removal from human sleep EEG by using independent component analysis. Annals of Biomedical Engineering 2008;36:467–475. [PubMed: 18228142]

Davidson RJ, Ekman P, Saron CD, Senulis JA, Friesen WV. Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. Journal of Personality and Social Psychology 1990;58:330–341. [PubMed: 2319445]

Davidson, RJ.; Jackson, DC.; Larson, CL. Handbook of psychophysiology. 2nd ed.. NY: Cambridge University Press; 2000. Human electroencephalography; p. 27-52.

Davidson RJ, Shackman AJ, Maxwell JS. Asymmetries in face and brain related to emotion. Trends in Cognitive Sciences 2004;8:389–391. [PubMed: 15350238]

Debener S, Strobel A, Sorger B, Peters J, Kranczioch C, Engel AK, Goebel R. Improved quality of auditory event-related potential recorded simultaneously with 3T fMRI: Removal of the ballistocardiogram artefact. NeuroImage 2007;34:587–597. [PubMed: 17112746]

Debener S, Mullinger KJ, Niazy RK, Bowtell RW. Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 Tesla. International Journal of Psychophysiology 2008;67:189–199. [PubMed: 17683819]

De Clercq, W.; Vergult, A.; Vanrumste, B.; Van Hees, J.; Palmini, A.; Van Paesschen, W., et al. 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); 2005. p. 944-947.

Delorme A, Sejnowski T, Makeig S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. NeuroImage 2007;34:1443–1449. [PubMed: 17188898]

Delorme A, Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 2004;134:9–21. [PubMed: 15102499]

Delorme A, Westerfield M, Makeig S. Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. Journal of Neuroscience 2007;27:11949–11959. [PubMed: 17978035]

DeRubeis RJ, Siegle GJ, Hollon SD. Cognitive therapy versus medication for depression: Treatment outcomes and neural mechanisms. Nature Reviews Neuroscience 2008;9:788–795.

Dien J. Issues in the application of the average reference: Review, critiques, and recommendations. Behavior Research Methods, Instruments, & Computers 1998;30:34–43.

Dimberg U, Thunberg M, Elmehed K. Unconscious facial reactions to emotional facial expressions. Psychological Science 2000;11:86–89. [PubMed: 11228851]

Evans AC, Collins DL, Mills SR, Brown ED, Kelly RL, Peters TM. 3D statistical neuroanatomical models from 305 MRI volumes. Nuclear Science Symposium and Medical Imaging Conference 1993:1813–1817.

Fava JL, Velicer WF. The effects of underextraction in factor and component analyses. Educational and Psychological Measurement 1996;56:907–929.

Fitzgibbon SP, Powers DM, Pope KJ, Clark CR. Removal of EEG noise and artifact using blind source separation. Journal of Clinical Neurophysiology 2007;24:232–243. [PubMed: 17545826]

Flexer A, Bauer H, Pripfl J, Dorffner G. Using ICA for removal of ocular artifacts in EEG recorded from blind subjects. Neural Networks 2005;18:998–1005. [PubMed: 15990276]

Frank RM, Frishkoff GA. Automated protocol for evaluation of electromagnetic component separation (APECS): Application of a framework for evaluating statistical methods of blink extraction from multichannel EEG. Clinical Neurophysiology 2006;118:80–97. [PubMed: 17064960]

Freeman WJ, Holmes MD, Burke BC, Vanhatalo S. Spatial spectra of scalp EEG and EMG from awake humans. Clinical Neurophysiology 2003;114:1053–1068. [PubMed: 12804674]

Freunberger R, Fellinger R, Sauseng P, Gruber W, Klimesch W. Dissociation between phase-locked and nonphase-locked alpha oscillations in a working memory task. Human Brain Mapping. 2009 Epub ahead of print.

Freyer F, Becker R, Anami K, Curio G, Villringer A, Ritter P. Ultrahigh-frequency EEG during fMRI: Pushing the limits of imaging-artifact correction. NeuroImage. 2009 Epub ahead of print.

Gasser T, Bacher P, Mocks J. Transformations towards the normal distribution of broad band spectral parameters of the EEG. Electroencephalography and Clinical Neurophysiology 1982;53:119–124. [PubMed: 6173196]

Gevins A, Smith ME. Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. Cerebral Cortex 2000;10:829–839. [PubMed: 10982744]

Goncharova II, McFarland DJ, Vaughan TM, Wolpaw JR. EMG contamination of EEG: Spectral and topographical characteristics. Clinical Neurophysiology 2003;114:1580–1593. [PubMed: 12948787]

Greischar LL, Burghy CA, van Reekum CM, Jackson DC, Pizzagalli DA, Mueller C, et al. Effects of electrode density and electrolyte spreading in dense array electroencephalographic recording. Clinical Neurophysiology 2004;115:710–720. [PubMed: 15036067]

Grouiller F, Vercueil L, Krainik A, Segebarth C, Kahane P, David O. A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI. NeuroImage 2007;38:124–137. [PubMed: 17766149]

Hamidi M, Slagter HA, Tononi G, Postle BR. Repetitive transcranial magnetic stimulation affects behavior by biasing endogenous cortical oscillations. Frontiers in Integrative Neuroscience 2009;3:1–12. [PubMed: 19225578]

Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. Communication Methods and Measures 2007;1:77–89.

Hesse, CW.; James, CJ. 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); 2004. p. 986-989.

Hoffmann S, Falkenstein M. The correction of eye blink artefacts in the EEG: A comparison of two prominent methods. PLoS ONE 2009;3:e3004. [PubMed: 18714341]

Huang R-S, Jung T-P, Delorme A, Makeig S. Tonic and phasic electroencephalographic dynamics during continuous compensatory tracking. NeuroImage 2008;39:1896–1909. [PubMed: 18083601]

Ille N, Berg B, Scherg M. Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. Journal of Clinical Neurophysiology 2002;19:113–124. [PubMed: 11997722]

James CJ, Hesse CW. Independent component analysis for biomedical signals. Physiological Measurement 2005;26:R15–R39. [PubMed: 15742873]

Joyce CA, Gorodnitsky IF, Kutas M. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 2004;41:313–325. [PubMed: 15032997]

Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, et al. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000;37:163–178. [PubMed: 10731767]

Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clinical Neurophysiology 2000;111:1745–1758. [PubMed: 11018488]

Koskinen M, Vartiainen N. Removal of imaging artifacts in EEG during EEG/fMRI recording: Reconstruction of a high-precision artifact template. NeuroImage 2009;46:160–167. [PubMed: 19457365]

Lawrence FR, Hancock GR. Conditions affecting integrity of a factor solution under varying degrees of overextraction. Educational and Psychological Measurement 1999;59:549–579.

Lee J-H, Lee T-W, Jolesz FA, Yoo S-S. Independent vector analysis (IVA): Mulitvariate approach for fMRI group study. NeuroImage 2008;40:86–109. [PubMed: 18165105]

Lee S, Buchsbaum MS. Topographic mapping of EEG artifacts. Clinical Electroencephalography 1987;18:61–67. [PubMed: 3594923]

Lee TW, Lewicki MS, Girolami M, Sejnowski TJ. Blind source separation of more sources than mixtures using overcomplete representations. IEEE Signal Processing Letters 1999;6:87–90.

Li YO, Adali T, Calhoun VD. Estimating the number of independent components for functional magnetic resonance imaging data. Human Brain Mapping 2007;28:1251–1266. [PubMed: 17274023]

Lutz A, Greischar LL, Rawlings NB, Ricard M, Davidson RJ. Long-term meditators self-induce high-amplitude gamma synchrony during mental practice. Proceedings of the National Academy of Sciences USA 2004;101:16369–16373.

Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. Trends in Cognitive Sciences 2004;8:204–210. [PubMed: 15120678]

Mantini D, Franciotti R, Romani GL, Pizzella V. Improving MEG source localizations: An automated method for complete artifact removal based on independent component analysis. NeuroImage 2008;40:160–173. [PubMed: 18155928]

McMenamin BW, Shackman AJ, Maxwell JS, Greischar LL, Davidson RJ. Validation of regression-based myogenic correction techniques for scalp and source-localized EEG. Psychophysiology 2009;46:578–592. [PubMed: 19298626]

Michel C, Murray M, Lantz G, Gonzalez S, Spinelli L, Grave de Peralta R. EEG source imaging. Clinical Neurophysiology 2004;115:2195–2222. [PubMed: 15351361]

Milne E, Scope A, Pascalis O, Buckley D, Makeig S. Independent component analysis reveals atypical electroencephalographic activity during visual perception in individuals with autism. Biological Psychiatry 2009;65:22–30. [PubMed: 18774554]

Morecraft, RJ.; Tanji, J. Cingulofrontal interactions and the cingulate motor areas. In: Vogt, BA., editor. Cingulate neurobiology and disease: Infrastructure, diagnosis, treatment. NY: Oxford University Press; 2009. p. 113-144.

Moraux A, Ianetti GD. Nociceptive laser-evoked brain potentials do not reflect nociceptive-specific neural activity. Journal of Neurophysiology 2009;101:3258–3269. [PubMed: 19339457]

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

Nazarpour K, Wongsawat Y, Sanei S, Chambers JA, Oraintara S. Removal of the eye-blink artifacts for EEGs via STF-TS modeling and robust minimum variance beamformers. IEEE Transactions on Biomedical Engineering 2008;55:2221–2231. [PubMed: 18713691]

Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: A primer with examples. Human Brain Mapping 2002;15:1–25. [PubMed: 11747097]

Nolte G, Curio G. The effect of artifact rejection by signal-space projection on source localization accuracy in MEG measurements. IEEE Transactions on Biomedical Engineering 1999;46:400–408. [PubMed: 10217878]

Oakes TR, Pizzagalli DA, Hendrick AM, Horras KA, Larson CL, Abercrombie HC, et al. Functional coupling of simultaneous electrical and metabolic activity in the human brain. Human Brain Mapping 2004;21:257–270. [PubMed: 15038007]

Ohla K, Hudry J, le Coutre J. The cortical chronometry of electrogustatory event-related potentials. Brain Topography 2009;22:73–82. [PubMed: 19199019]

Onton J, Makeig S. Information-based modeling of event-related brain dynamics. Progress in Brain Research 2006;159:99–120. [PubMed: 17071226]

Onton J, Westerfield M, Townsend J, Makeig S. Imaging human EEG dynamics using independent component analysis. Neuroscience and Biobehavioral Reviews 2006;30:808–822. [PubMed: 16904745]

Pascual-Marqui RD. Review of methods for solving the EEG inverse problem. International Journal of Bioelectromagnetism 1999;1:75–86.

Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. International Journal of Psychophysiology 1994;18:49–65. [PubMed: 7876038]

Pizzagalli, DA. Electroencephalography and high density electrophysiological source localization. In: Cacioppo, JT.; Tassinaru, LG.; Berntson, GG., editors. Handbook of psychophysiology. 3rd ed.. NY: Cambridge University Press; 2007. p. 56-84.

Romei V, Brodbeck V, Michel C, Amedi A, Pascual-Leone A, Thut G. Spontaneous fluctuations in posterior α-band EEG activity reflect variability in excitability of human visual areas. Cerebral Cortex 2008;18:2010–2018. [PubMed: 18093905]

Romero S, Mananas MA, Barbanoj MJ. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case. Computers in Biology and Medicine 2008;38:348–360. [PubMed: 18222418]

Ryali S, Glover GH, Chang C, Menon V. Development, validation, and comparison of ICA-based gradient artifact reduction algorithms for simultaneous EEG-spiral in/out echo-planar fMRI recordings. NeuroImage. 2009 Epub ahead of print.

Seaman MA, Serlin RC. Equivalence confidence intervals for two-group comparisons of means. Psychological Methods 1998;3:403–411.

Shackman AJ. The potentially deleterious impact of muscle activity on gamma band inferences. Neuropsychopharmacology. (in press).

Shackman AJ, McMenamin BW, Maxwell JS, Greischar LL, Davidson RJ. Right dorsolateral prefrontal cortical activity and behavioral inhibition. Psychological Science. (in press).

Shackman AJ, McMenamin BW, Slagter HA, Maxwell JS, Greischar LL, Davidson RJ. Electromyogenic artifacts and electroencephalographic inferences. Brain Topography 2009;21:7–12. [PubMed: 19214730]

Shackman AJ, Sarinopoulos I, Maxwell JS, Pizzagalli DA, Lavric A, Davidson RJ. Anxiety selectively disrupts visuospatial working memory. Emotion 2006;6:40–61. [PubMed: 16637749]

Srinivasan R, Tucker DM, Murias M. Estimating the spatial Nyquist of the human EEG. Behavior Research Methods, Instruments & Computers 1998;30:8–19.

Talsma D. Auto-adaptive averaging: Detecting artifacts in event-related potential data using a fully automated procedure. Psychophysiology 2008;45:216–228. [PubMed: 17971060]

Tang AC, Liu JY, Sutherland MT. Recovery of correlated neuronal sources from EEG: The good and bad ways of using SOBI. NeuroImage 2005;28:507–519. [PubMed: 16139528]

Tassinary, LG.; Cacioppo, JT.; Vanman, EJ. The skeletomotor system: Surface electromyography. In: Cacioppo, JT.; Tassinary, LG.; Berntson, GG., editors. Handbook of psychophysiology. 3rd ed.. NY: Cambridge University Press; 2007. p. 267-302.

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalography and Clinical Neurophysiology 1995;95:189–200. [PubMed: 7555909]

Thatcher RW, North D, Biver C. Parametric vs. non-parametric statistics of low resolution electromagnetic tomography (LORETA). Clinical Electroencephalography and Neuroscience 2005;36:1–8. [PubMed: 15683191]

Thibodeau R, Jorgensen RS, Kim S. Depression, anxiety, and resting frontal EEG asymmetry: A meta-analytic review. Journal of Abnormal Psychology 2006;115:715–729. [PubMed: 17100529]

Thut G, Miniussi C. New insights into rhythmic brain activity from TMS-EEG studies. Trends in Cognitive Sciences 2009;13:182–189. [PubMed: 19286414]

Ting KH, Fung PCW, Chang CQ, Chan FHY. Automatic correction of artifact from single-trial event-related potentials by blind source separation using second order statistics only. Medical Engineering & Physics 2006;28:780–794. [PubMed: 16406675]

U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research. Guidance for industry: Statistical approaches to establishing bioequivalence. 2001. Retrieved November 11th, 2007, from http://www.fda.gov/cder/guidance/index.htm

Uusitalo MA, Ilmoniemi RJ. Signal-space projection method for separating MEG or EEG into components. Medical and Biological Engineering and Computing 1997;35:135–140. [PubMed: 9136207]

Van Boxtel A. Optimal signal bandwidth for the recording of surface EMG activity of facial, jaw, oral, and neck muscles. Psychophysiology 2001;38:22–34. [PubMed: 11321618]

Viola FC, Thorne J, Edmonds B, Schneider T, Eichele T, Debener S. Semi-automatic identification of independent components representing EEG artifact. Clinical Neurophysiology 2009;120:868–877. [PubMed: 19345611]

Wallstrom GL, Kass RE, Miller A, Cohn JF, Fox NA. Automatic correction of ocular artifacts in the EEG A comparison of regression-based and component-based methods. International Journal of Psychophysiology 2004;53:105–119. [PubMed: 15210288]

Waterink W, van Boxtel A. Facial and jaw-elevator EMG activity in relation to changes in performance level during a sustained information processing task. Biological Psychology 1994;37:183–198. [PubMed: 7948464]

Welch PD. The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics 1967;15:70–73.

Whitham EM, Pope KJ, Fitzgibbon SP, Lewis T, Clark CR, Loveless S, Broberg M, Wallace A, DeLosAngeles D, Lillie P, Hardy A, Fronsko R, Pulbrook A, Willoughby JO. Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. Clinical Neurophysiology 2007;118:1877–1888. [PubMed: 17574912]

Willis J, Nelson A, Rice J, Black FW. The topography of muscle activity in quantitative EEG. Clinical Electroencephalography 1993;24:123–126. [PubMed: 8403444]

Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. Transient induced gamma-band responses in EEG as a manifestation of miniature saccades. Neuron 2009;58:429–441. [PubMed: 18466752]

Yuval-Greenberg S, Keren AS, Tomer O, Nelken I, Deouell LY. Response to letter. Neuron 2009;62:10–12.

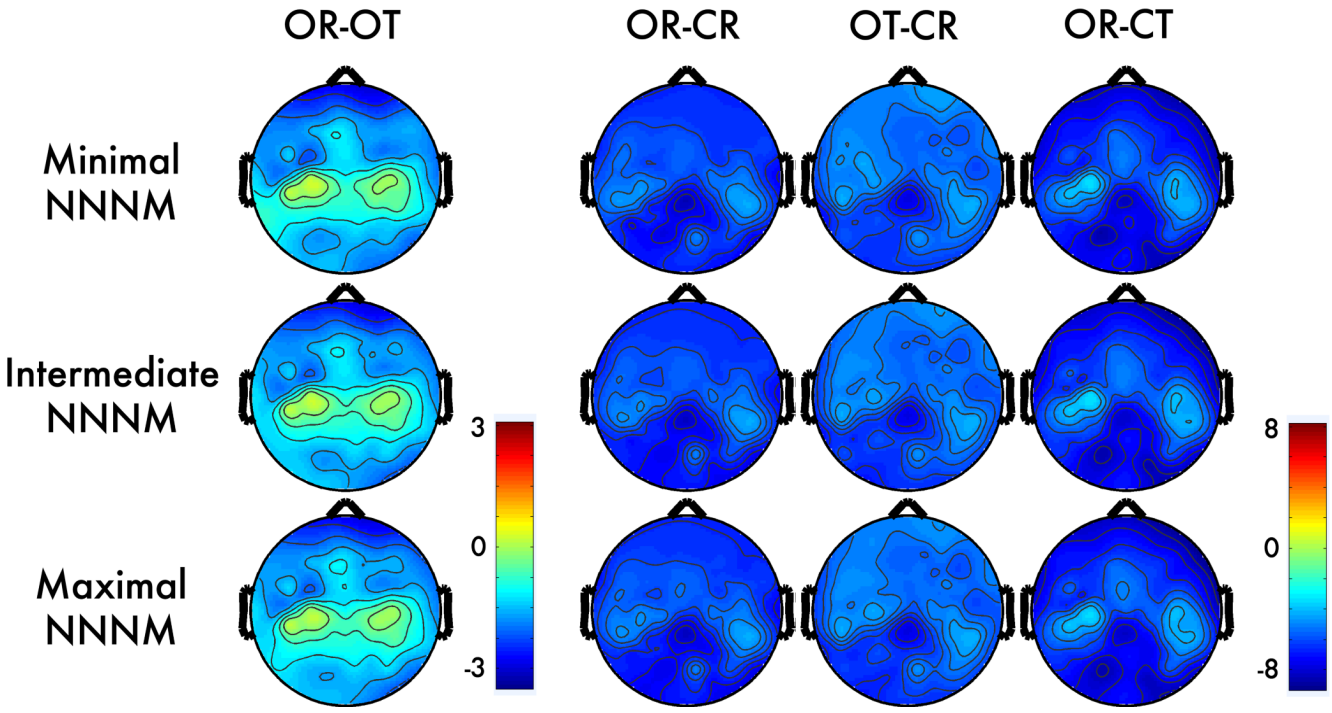

Figure 1. Alpha-band contrasts prior to correction

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).

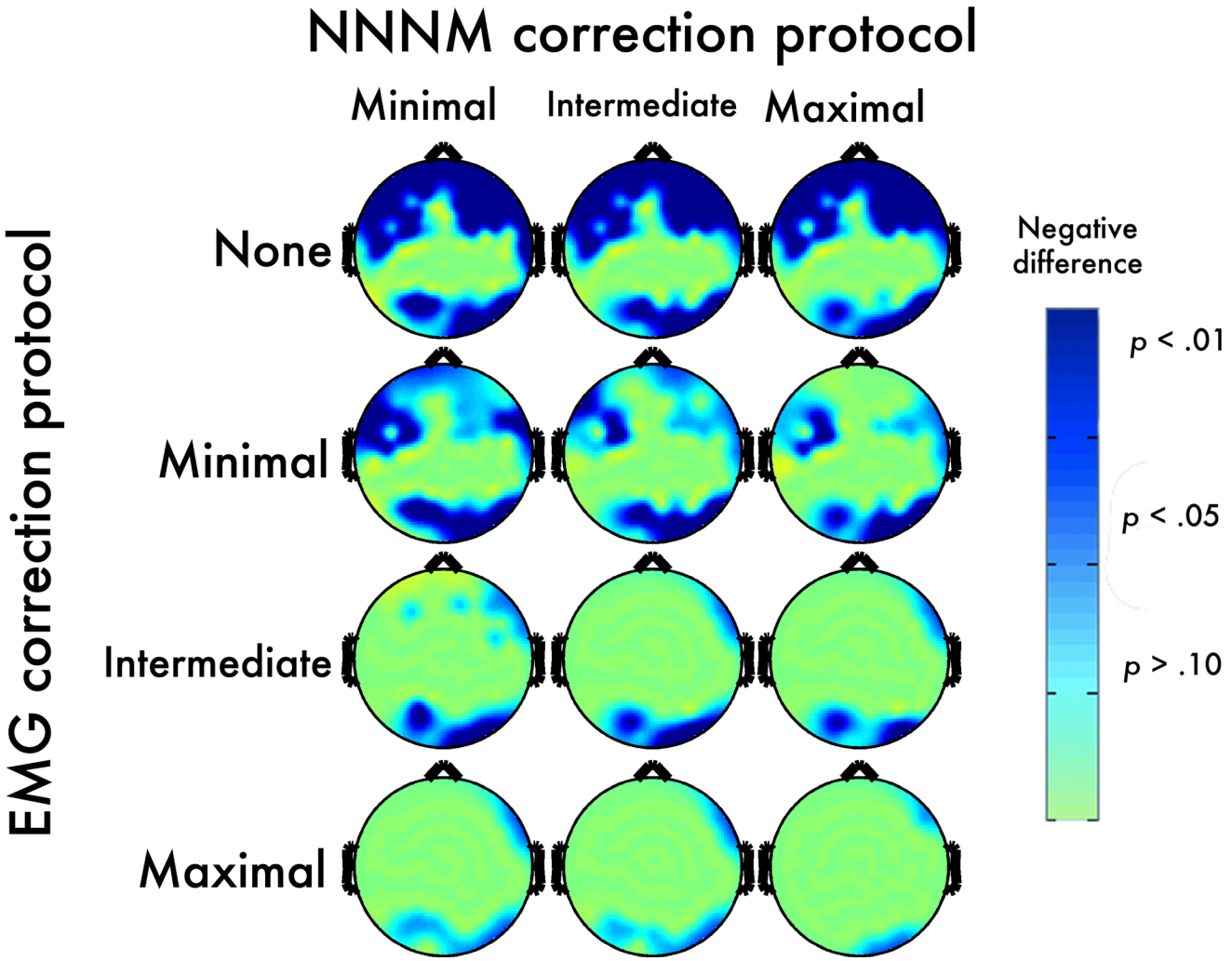

Figure 2. Myogenic contrast (OR-OT) after EMG correction

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: $p < .05$; light-blue: $p < .10$; green: $p > .10$). Note that the row labeled “None” depicts the thresholded OR-OT contrast from Figure 1.

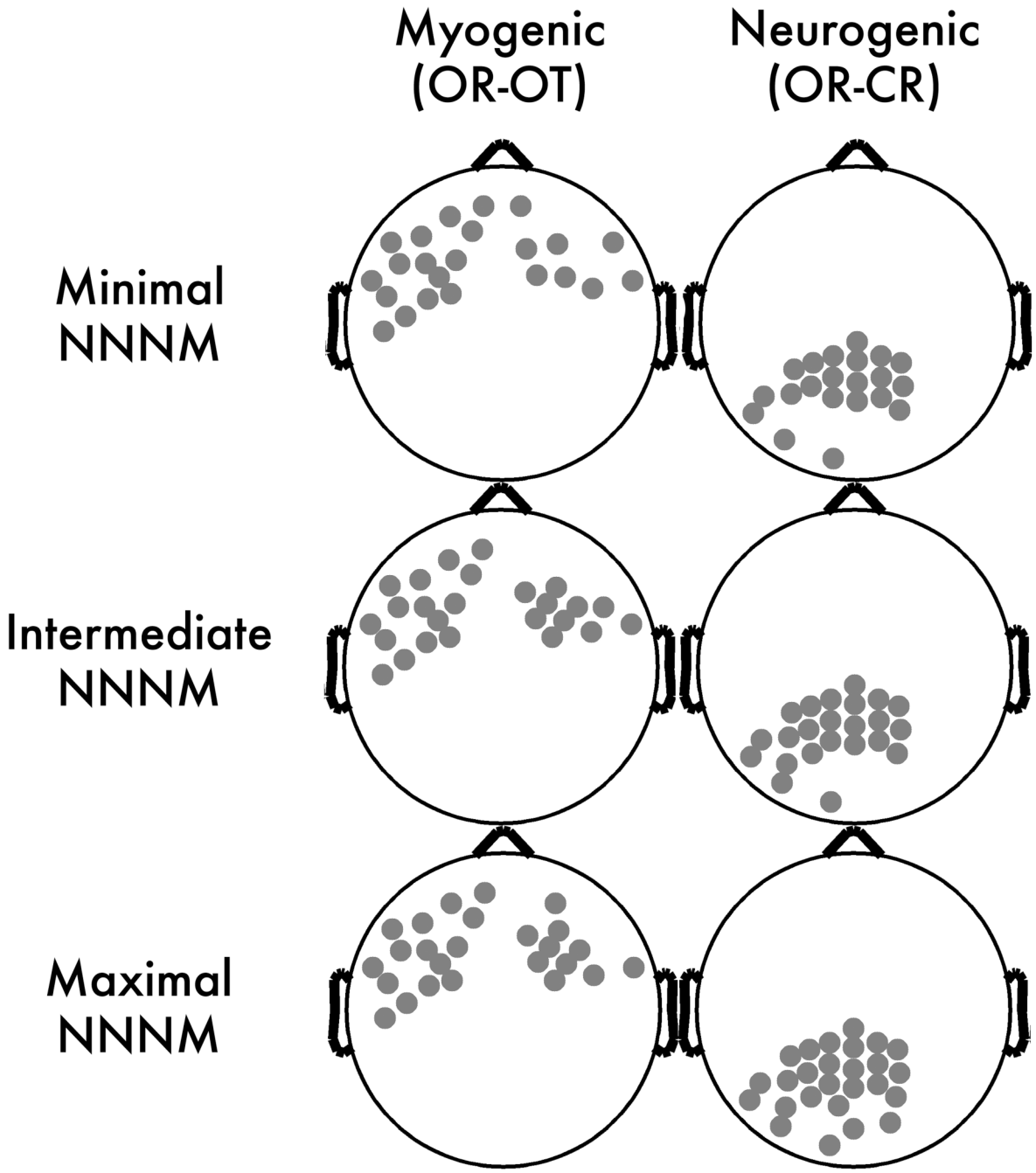

Figure 3. Scalp regions of interest (ROIs)
Gray circles depict electrodes included in the ROIs.

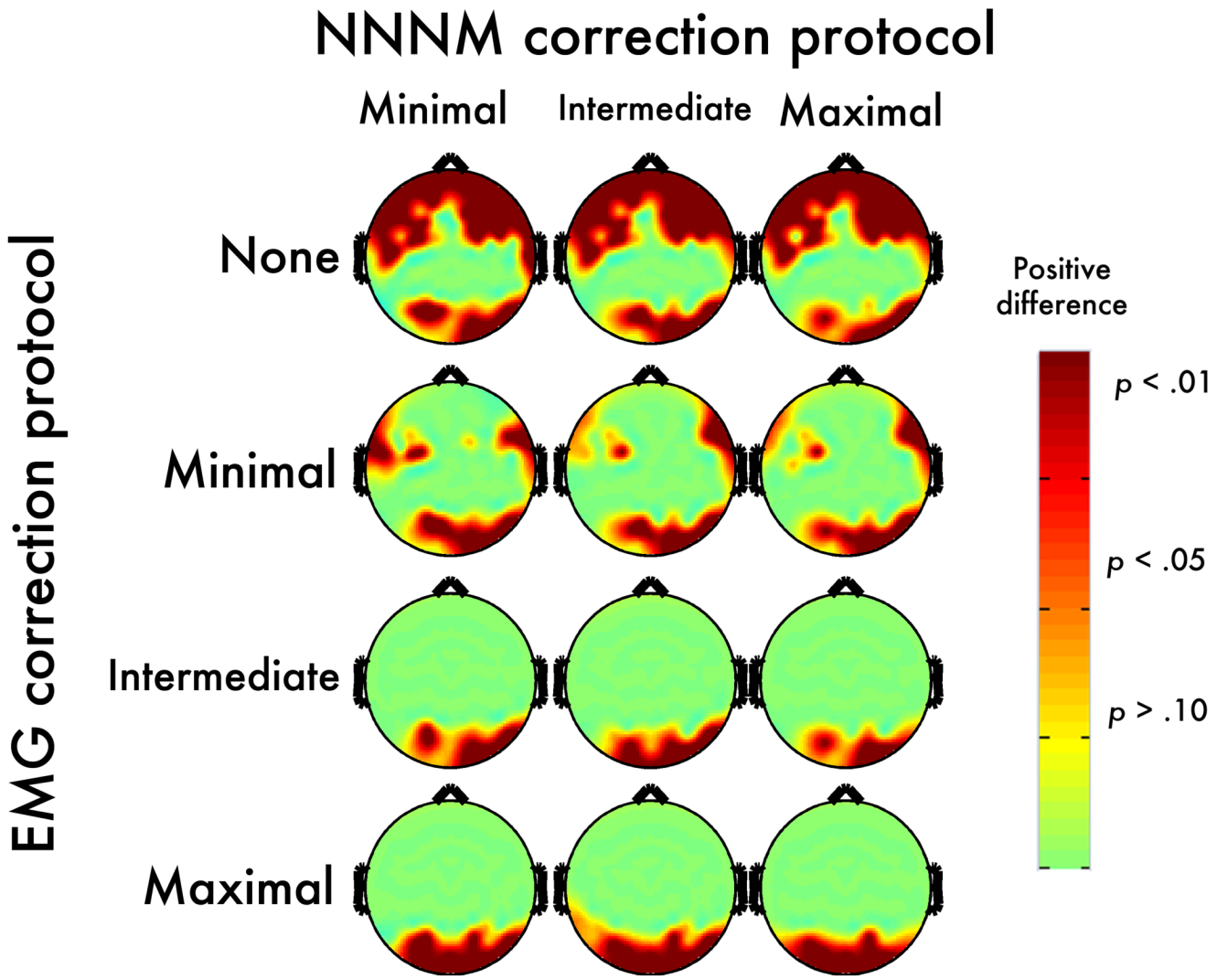

Figure 4. Negatively-covarying contrast: Post-correction error

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: $p < .05$; light-red: $p < .10$; yellow: $p > .10$).

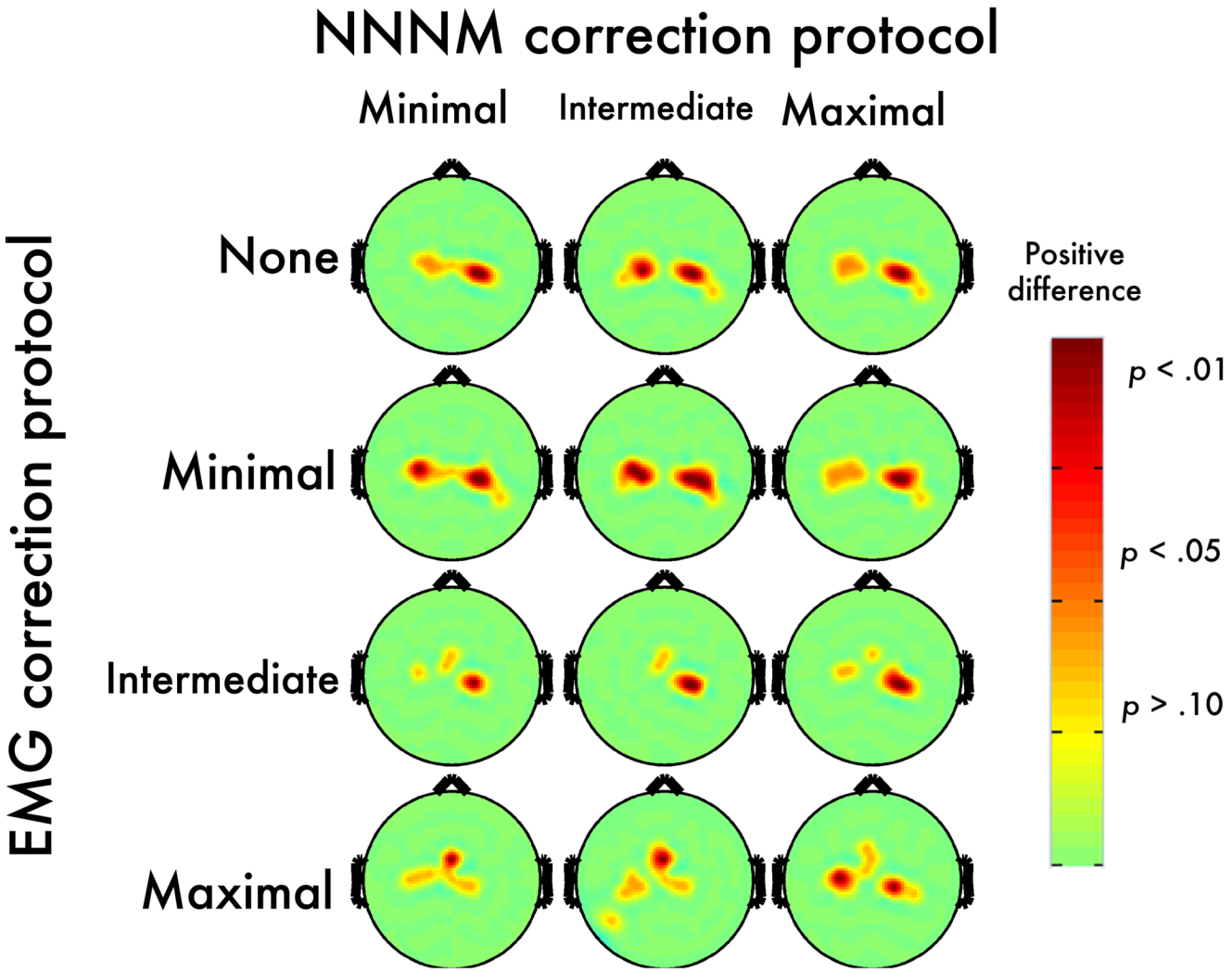

Figure 5. Positively covarying contrast: Post-correction error

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: $p < .05$; light-red: $p < .10$; yellow: $p > .10$).

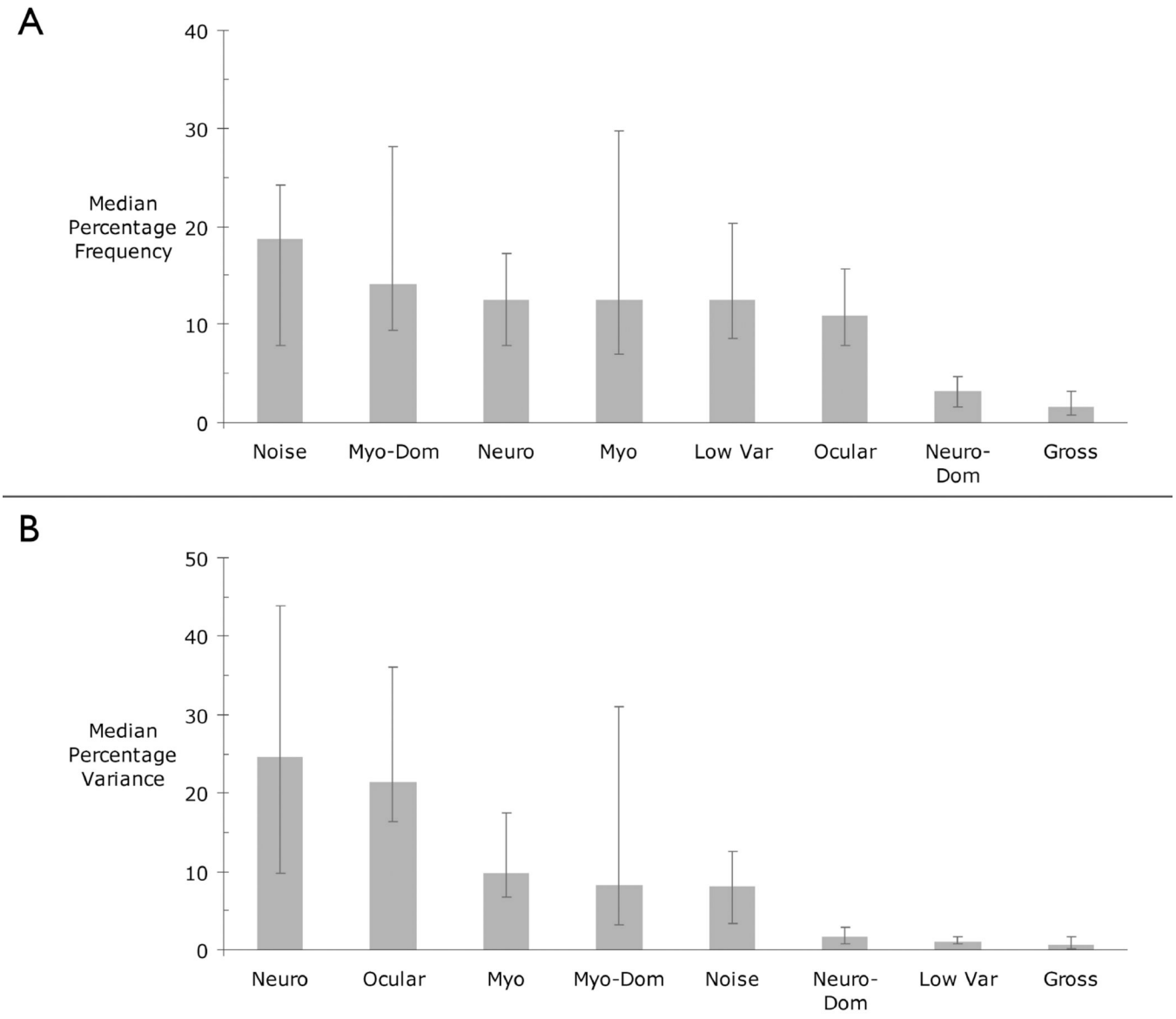

Figure 6. *A*) Median frequency and *B*) variance (in percent) accounted for by each class of components

Frequencies and percentages were computed within participants. Medians were then computed across participants. Error bars indicate the 25th and 75th percentiles.

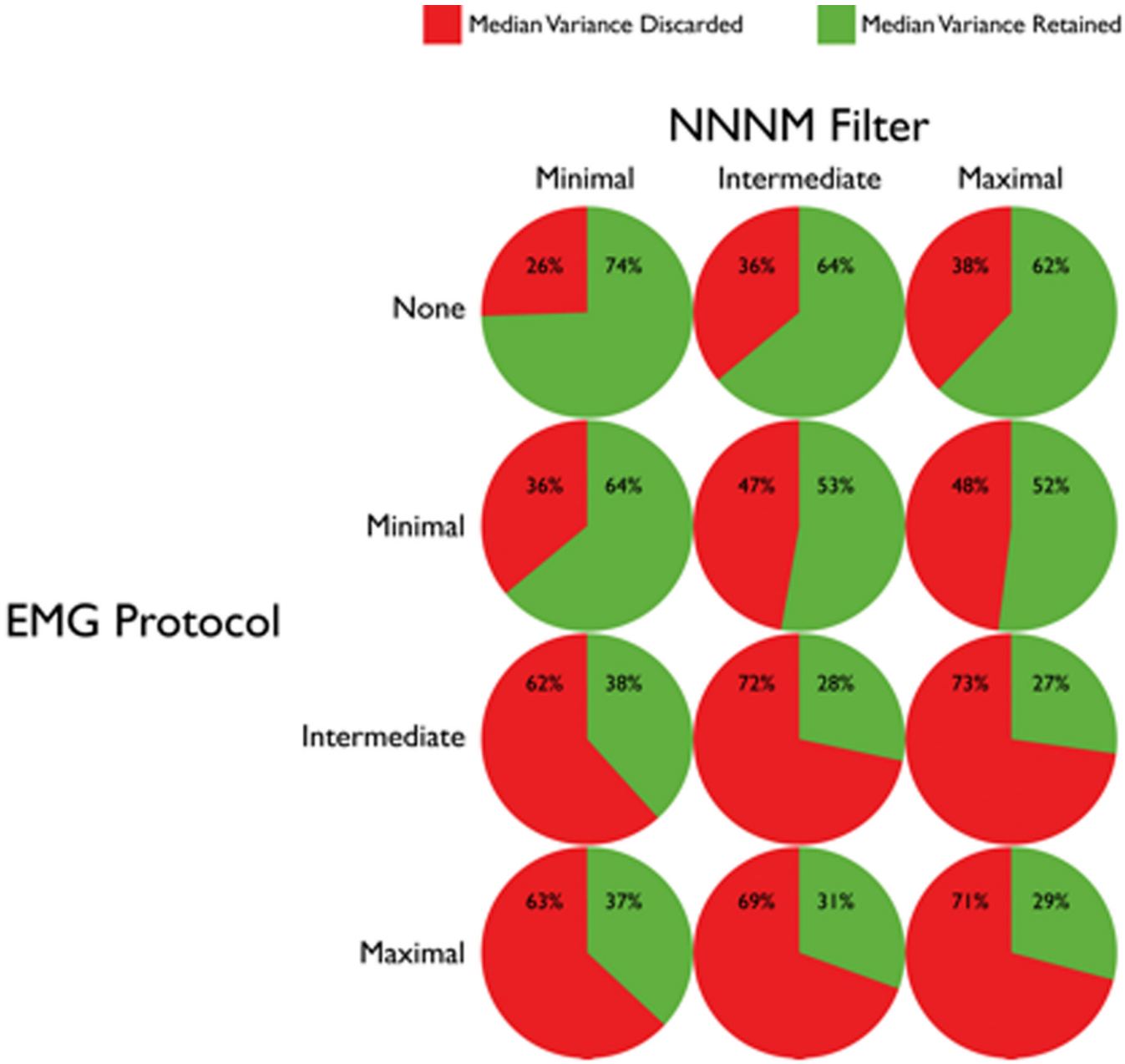

Figure 7. Median variance retained and discarded for each combination of NNNM filter and EMG correction protocol

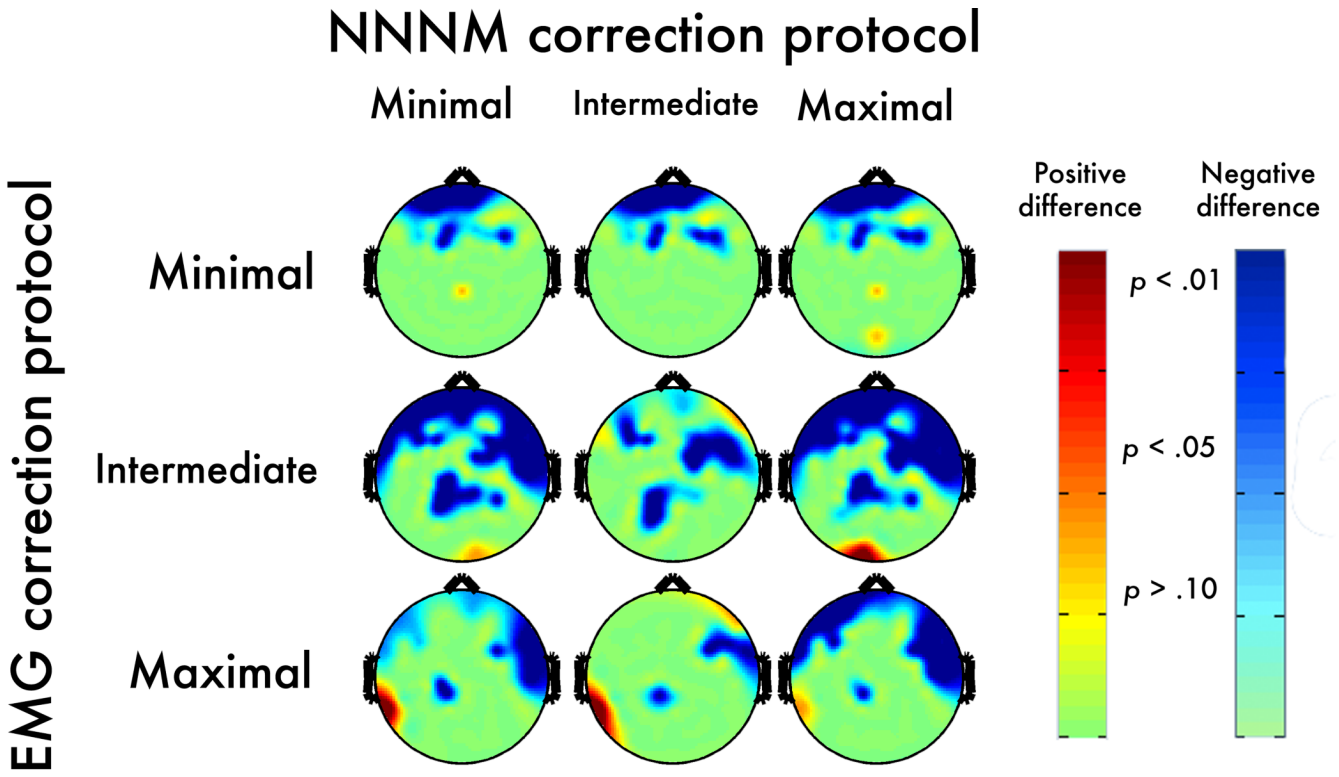

Figure 8. Neurogenic contrast: Post-correction error

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: $p < .05$; light-blue: $p < .10$; green: $p > .10$). Positive values, indicating correction-induced magnitude reduction of the neurogenic contrast, are shown in red (dark-red: $p < .05$; light-red: $p < .10$; yellow: $p > .10$).

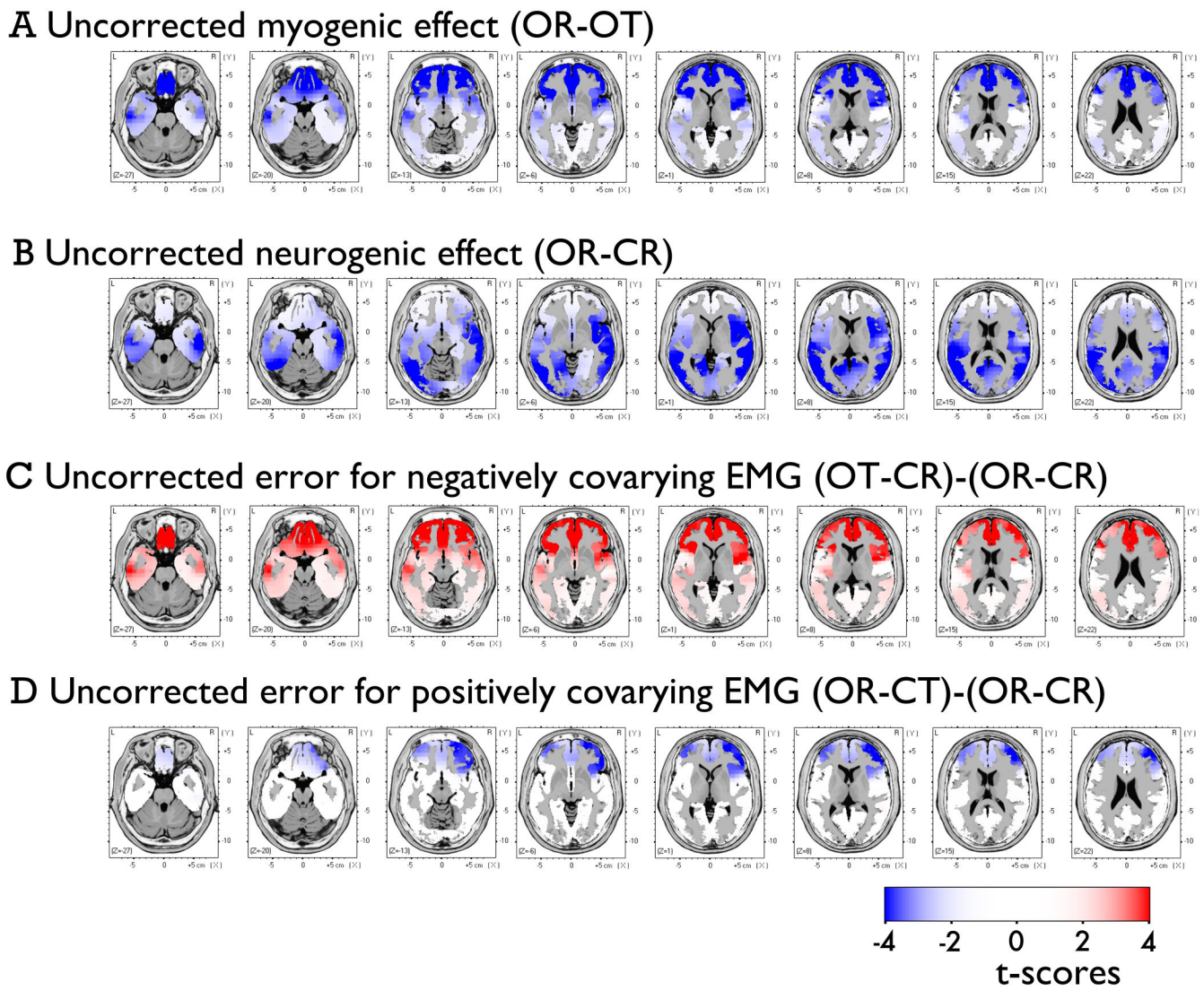

Figure 9. Contrasts of interest in the source-space after applying the Intermediate-NNNM filter

A) Myogenic contrast (OR-OT), B) Neurogenic contrast (OR-CR), C) Error induced by negatively-covarying artifact prior to EMG correction ([OT-CR] minus [OR-CR]), and D) Error induced by positively-covarying artifact ([OR-CT] minus [OR-CR]) prior to EMG correction.

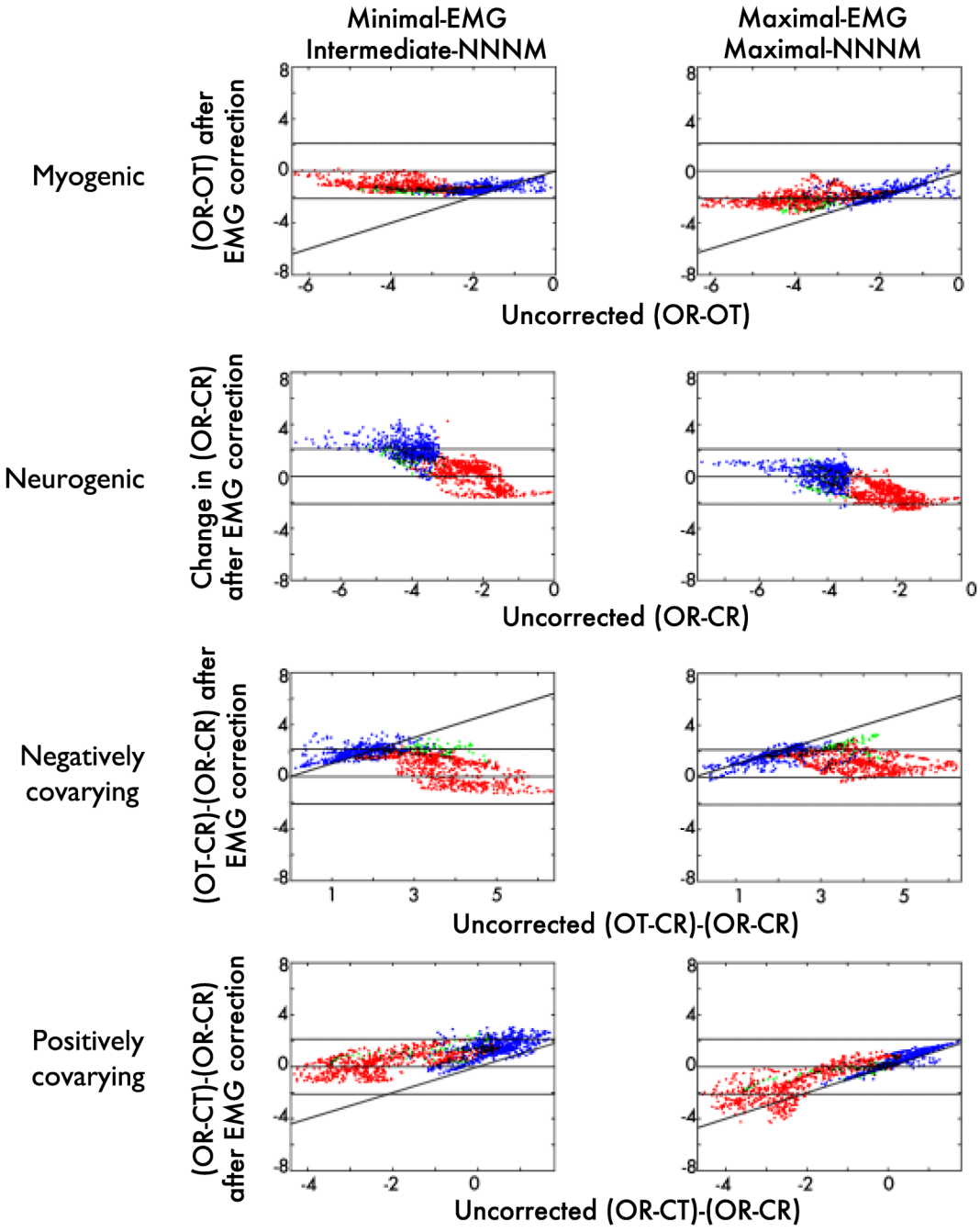

Figure 10. Source-space effects of EMG correction

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 ($p = .05$) 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 row B plots the uncorrected OR-CR contrast (*x*-axis) against the change in OR-CR after correction (i.e. correction induced error; *y*-axis).

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] Note: 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 Method).