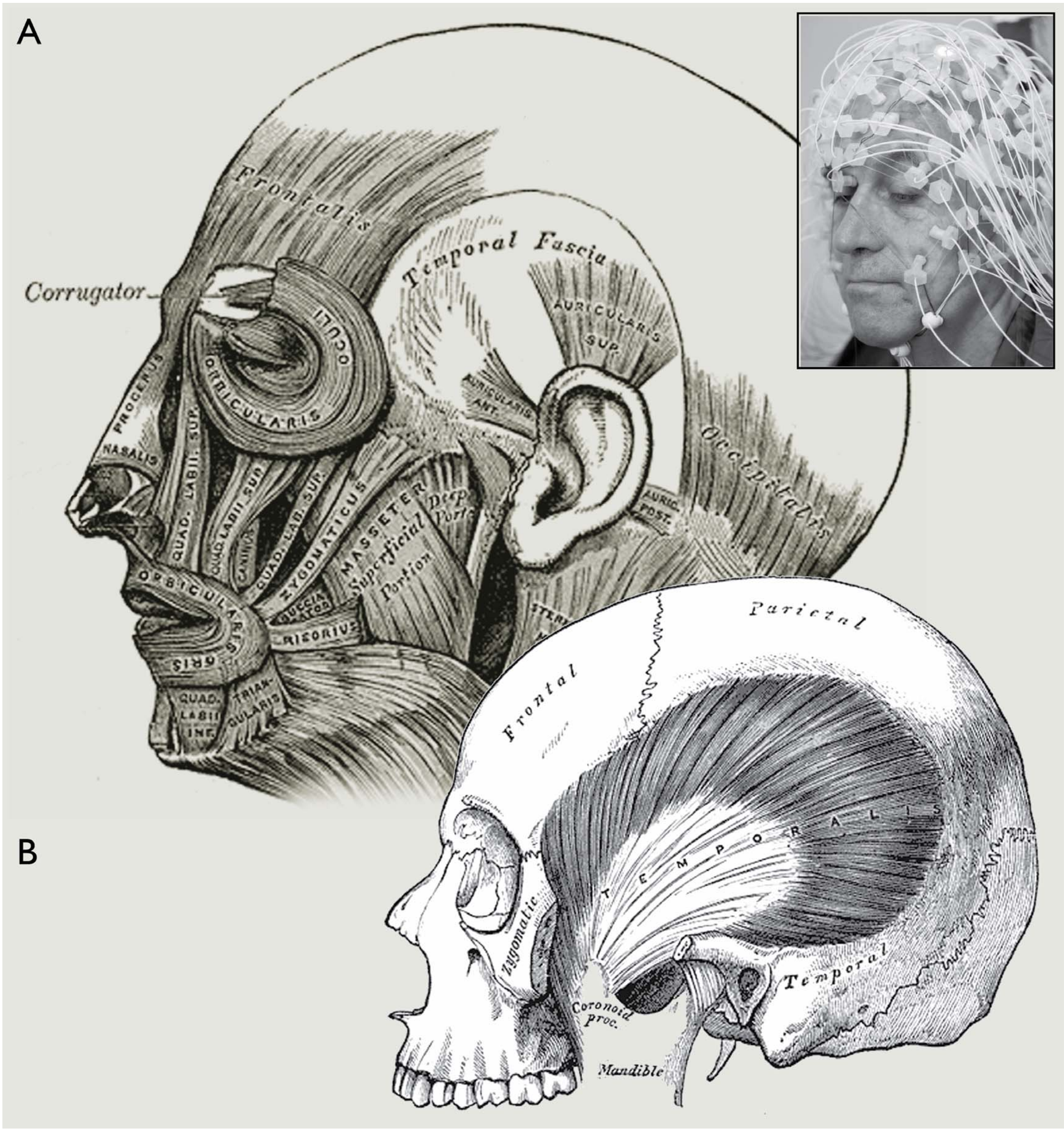
Supplementary Method and Results for

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

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

Contents

Cranial Musculature



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

Component Classification

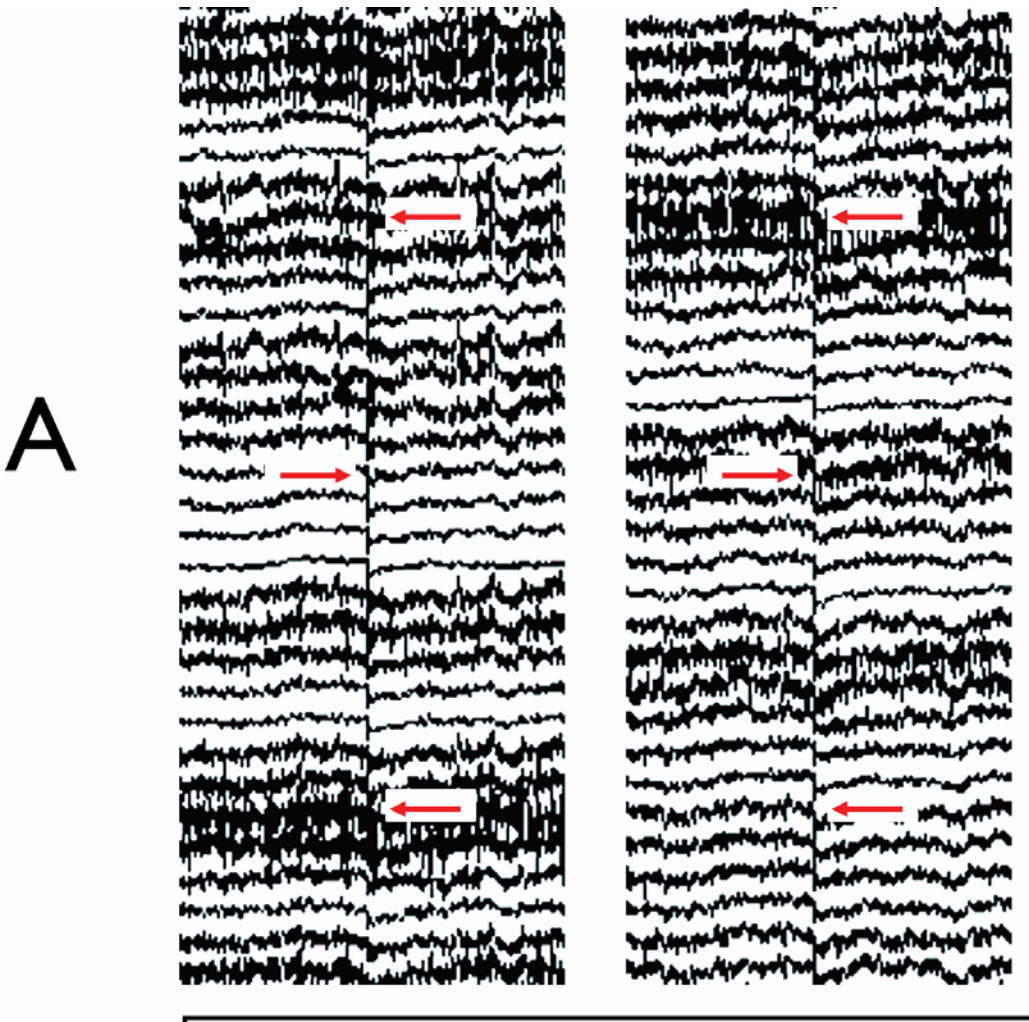
Components were classified by two independent, well-trained baccalaureate-level raters (DRWB and AMK) based on inspection of the time-series, power spectrum, and absolute topography. In general, our criteria were in close accord with classical descriptions (Gibbs & Gibbs, 1950). As detailed below, the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

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

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

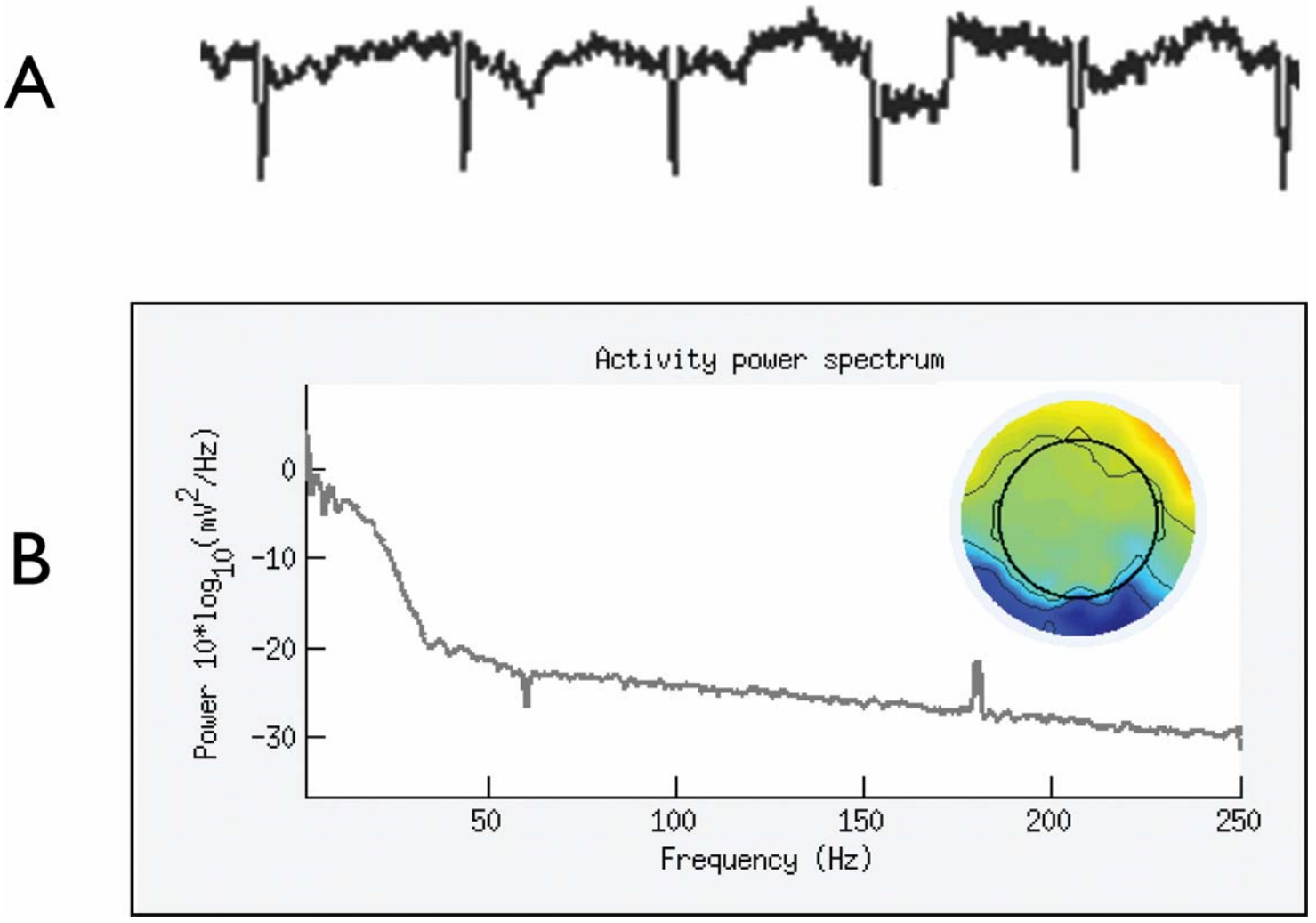
Reference and ground artifacts showed widespread, synchronous deflections in the raw time-series corresponding to periods of apparent "activation" in the component time-series, combined with a characteristically uniform topography (Supplementary Figure 2).

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



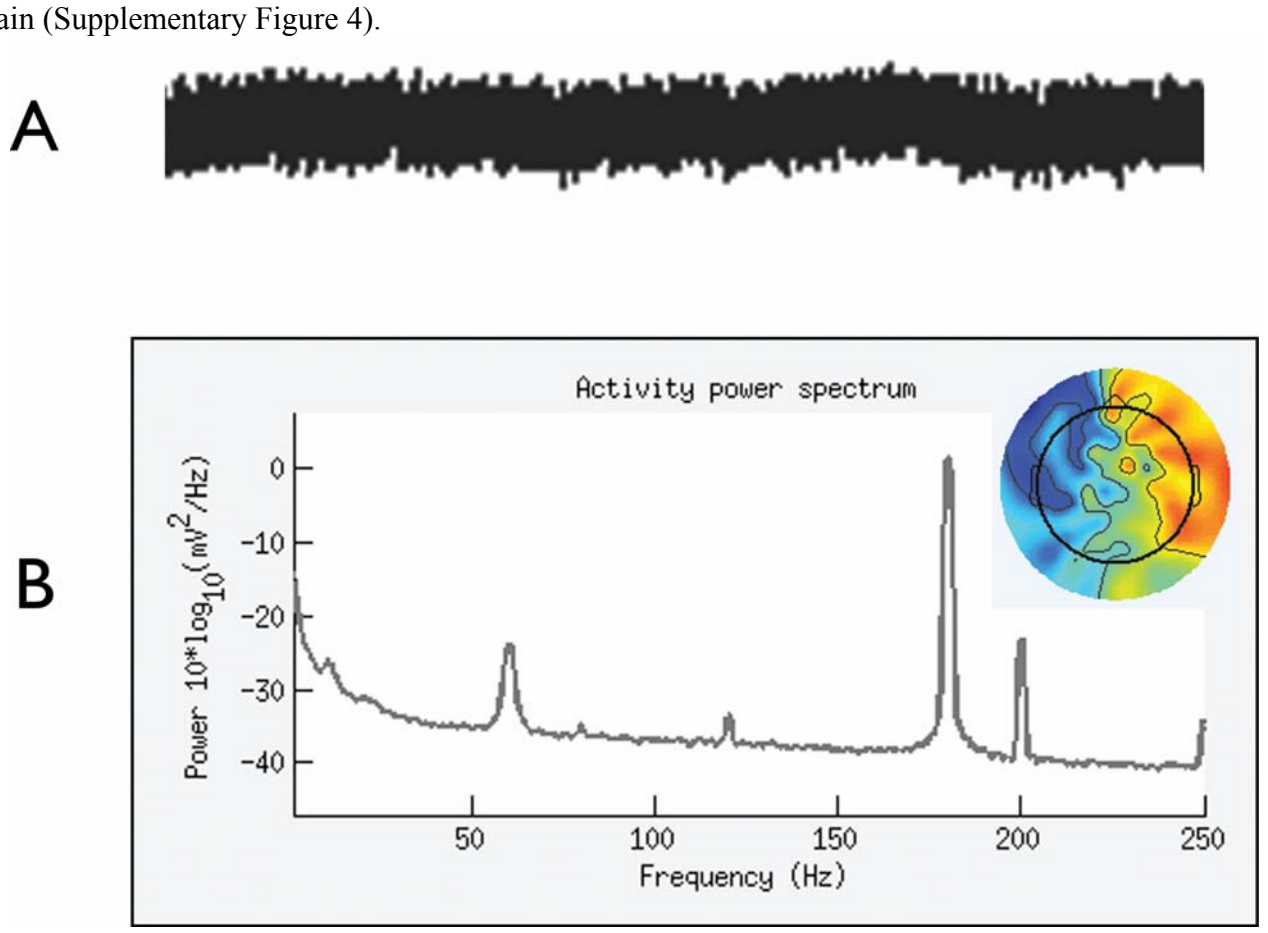
Supplementary Figure 2. *Reference Artifact.* **A)** Representative multi-channel slices of the raw time-series. Red arrows indicate location of transient reference artifacts. In both cases, all 128 channels (not displayed) were similarly affected. **B)** Back-projected and interpolated topographic map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

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



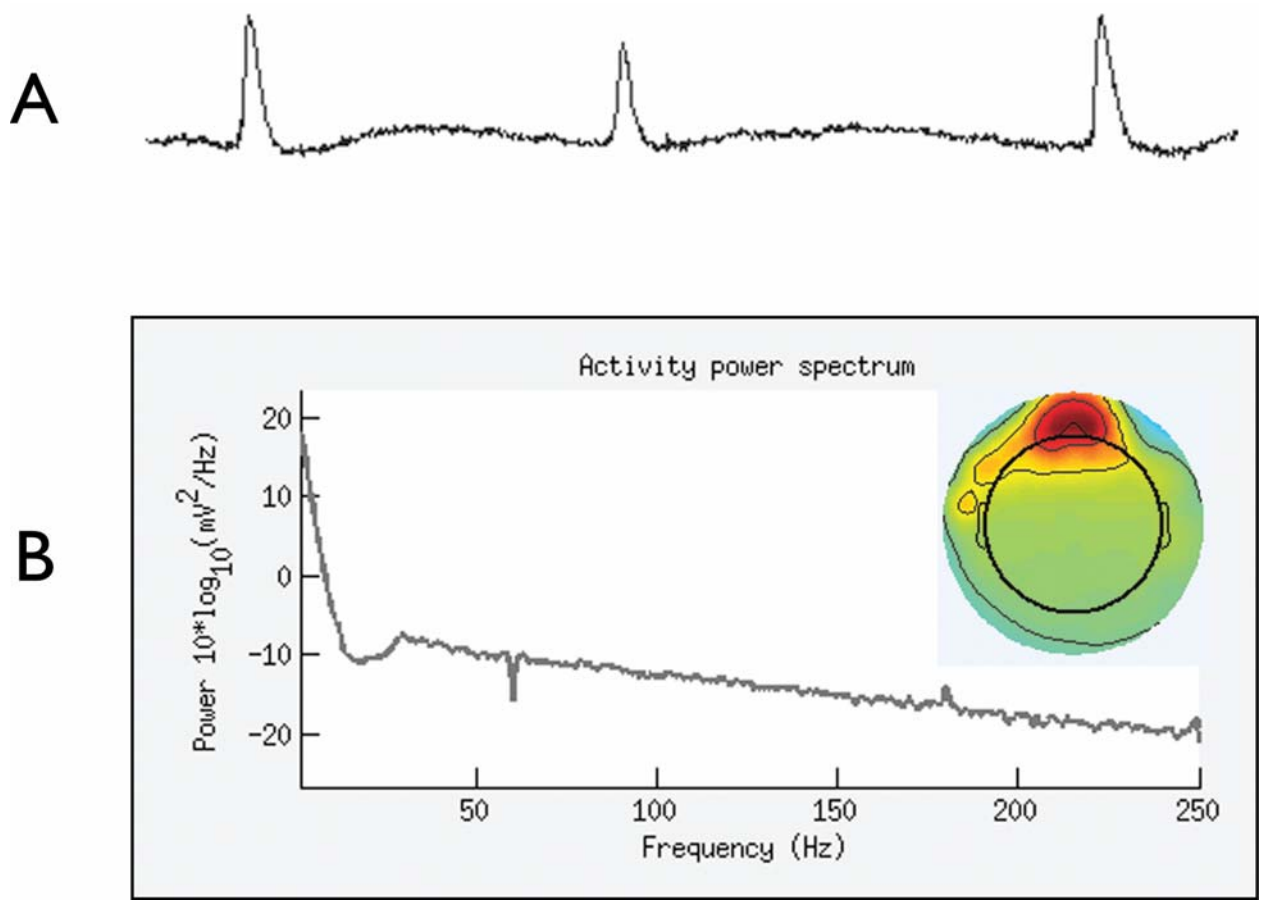
Supplementary Figure 3. *ECG Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

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

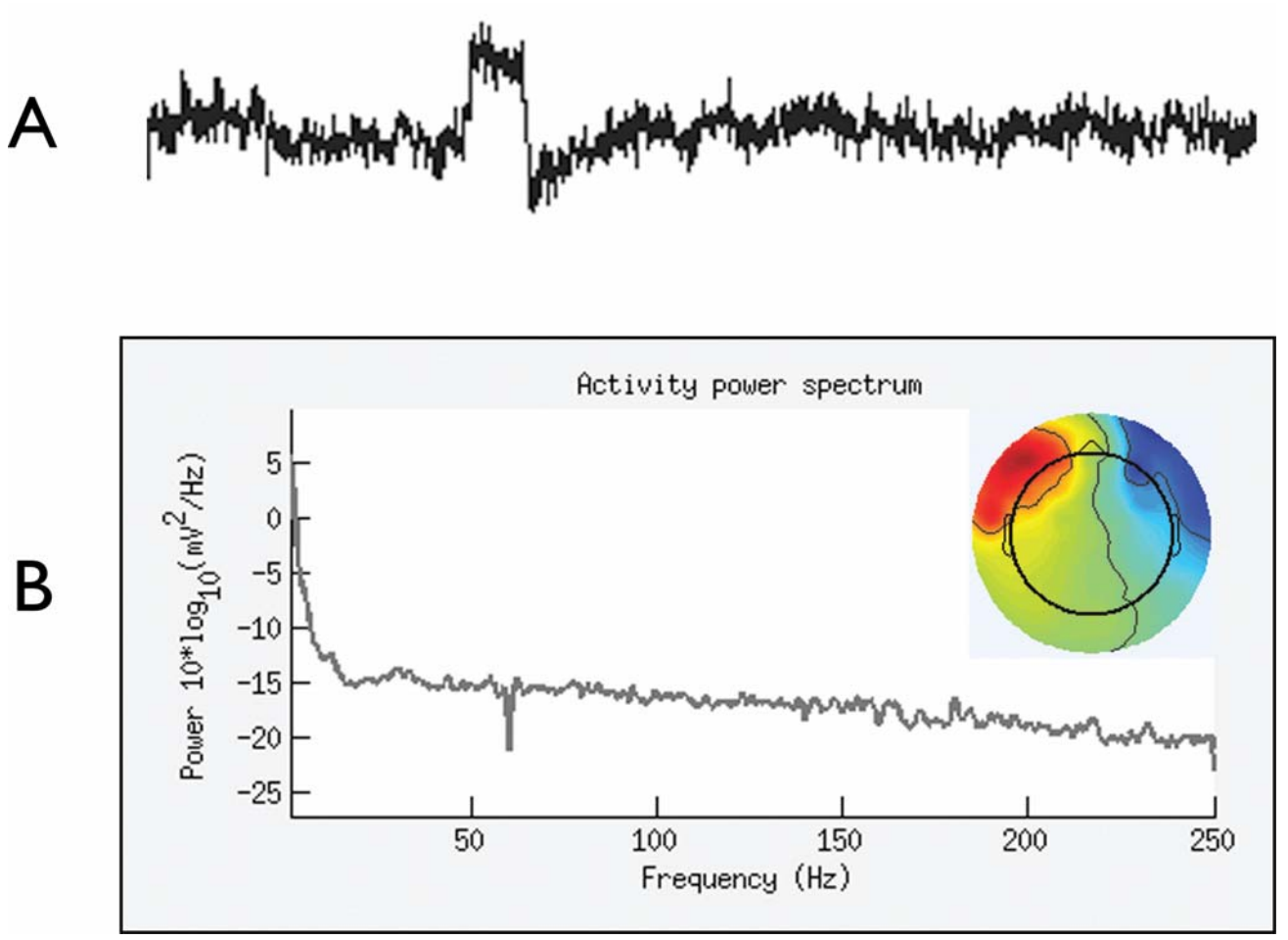


Supplementary Figure 4. *AC Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

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

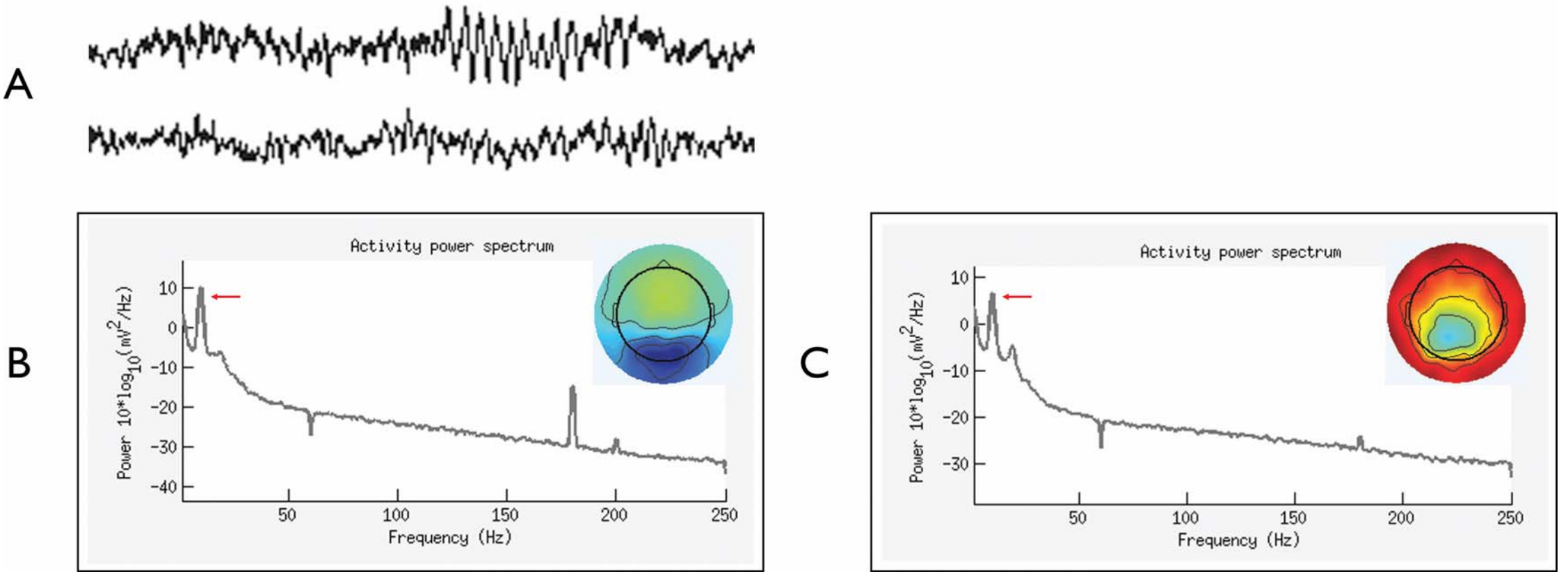


Supplementary Figure 5. *Blink Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

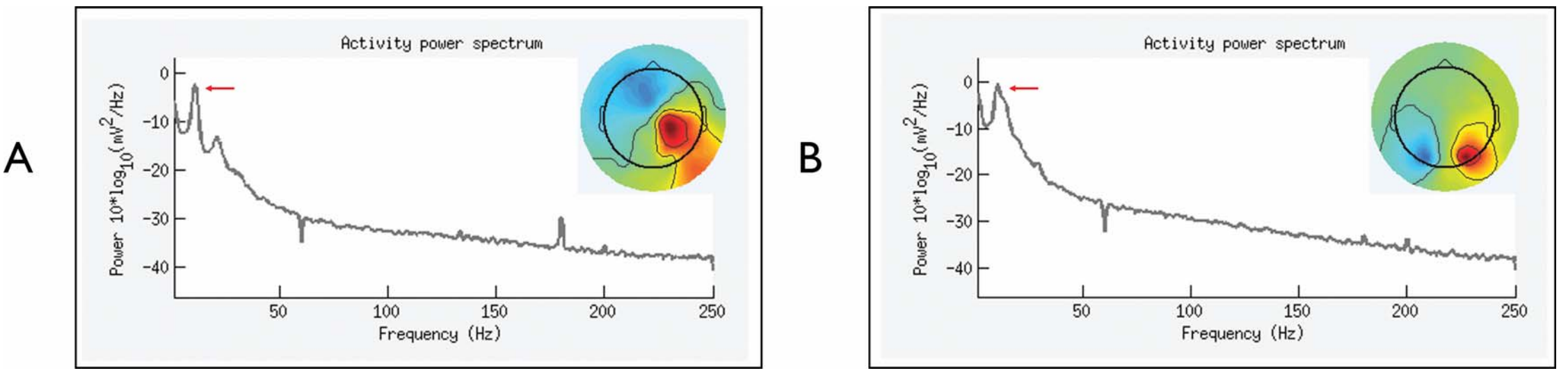


Supplementary Figure 6. *Saccadic Artifact.* **A)** Component time-series. **B)** Component power spectrum. Inset shows corresponding topographic map.

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

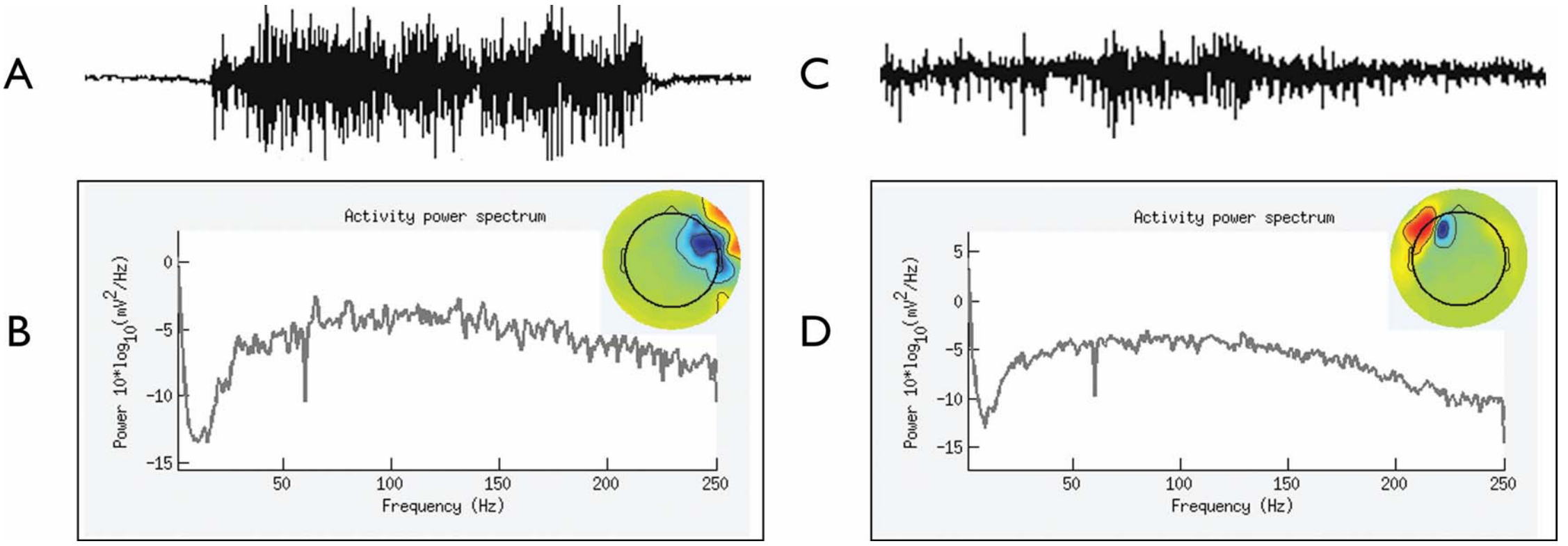


Supplementary Figure 7. *High-Variance Neurogenic Components*. **A)** Component time-series. **B) and C)** Component power spectra for the fourth and fifth components. Sixty-four components were extracted and ranked in descending order according to the amount of variance predicted. Red arrows indicate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



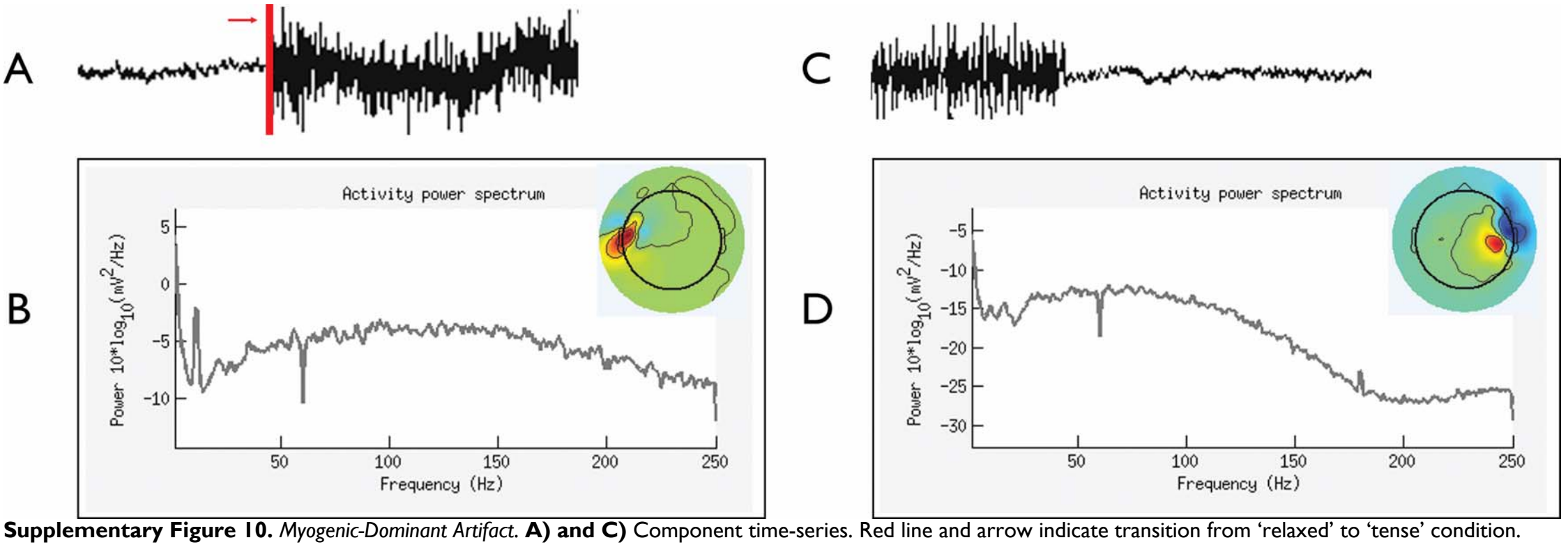
Supplementary Figure 8. *Low-Variance Neurogenic Components*. **A)** and **B)** Component power spectra for the sixteenth and eighteenth components. Red arrows indicate the location of the alpha peak (~10Hz). Insets show corresponding topographic maps.

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

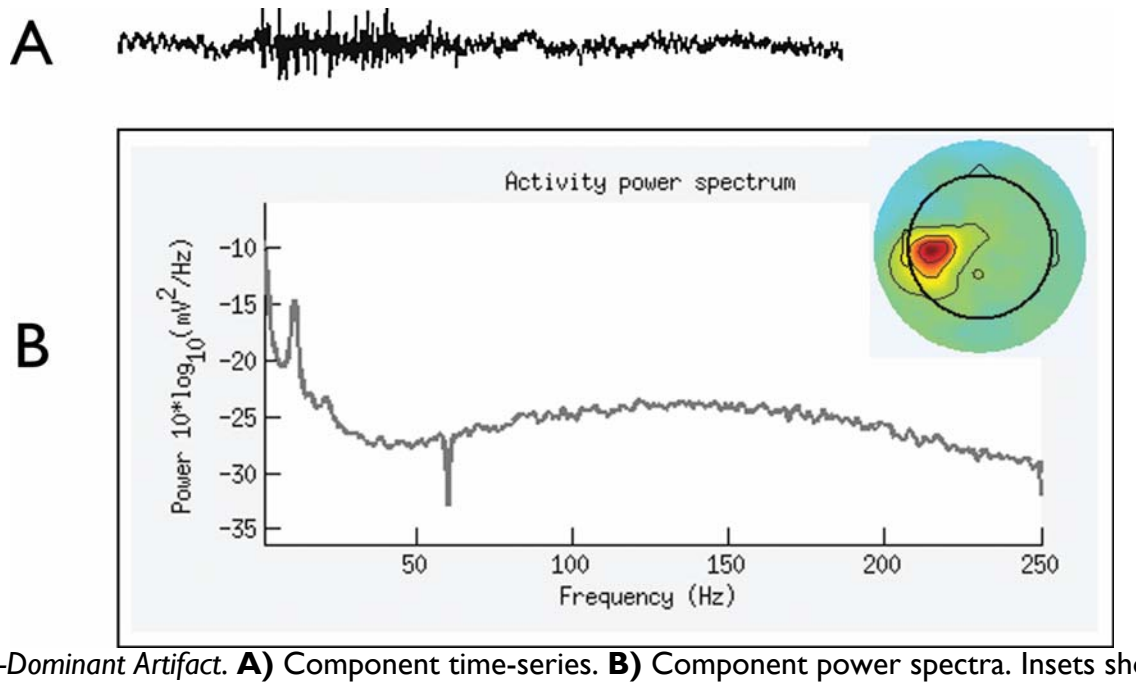


Supplementary Figure 9. *Myogenic Artifact.* **A) and C)** Component time-series. **B) and D)** Component power spectra. Insets show corresponding topographic maps.

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

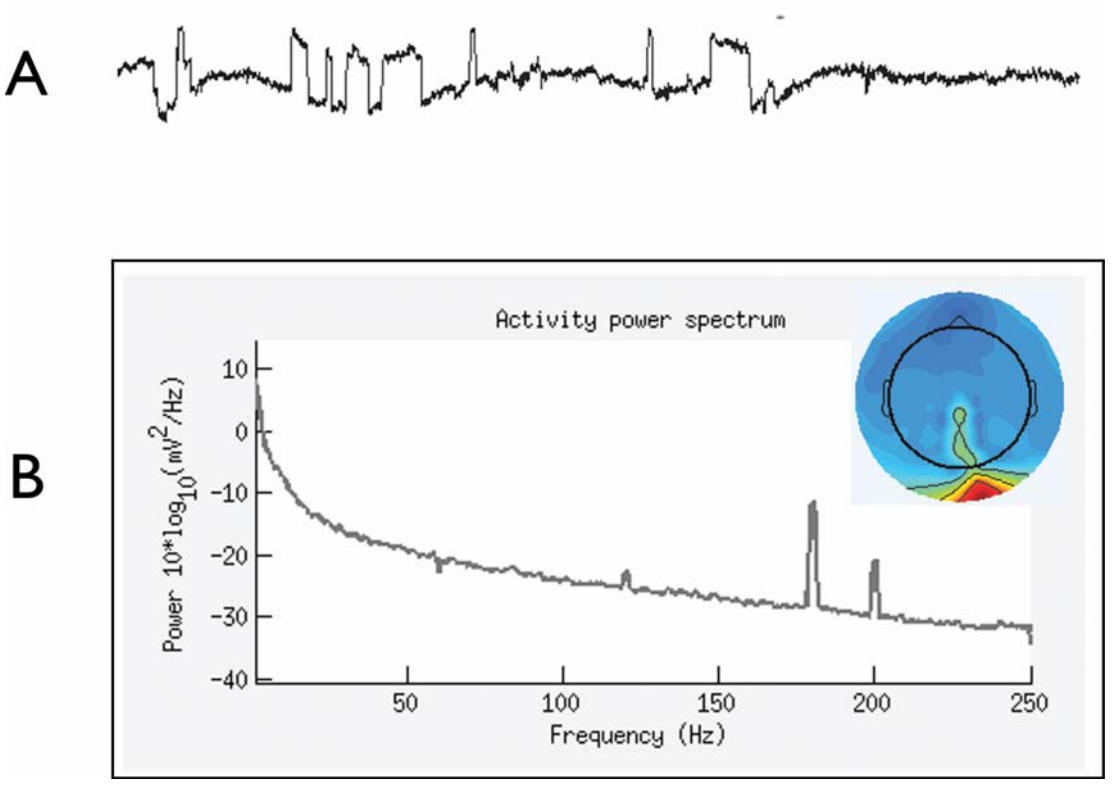


Supplementary Figure 10. *Myogenic-Dominant Artifact.* **A) and C)** Component time-series. Red line and arrow indicate transition from 'relaxed' to 'tense' condition. **B) and D)** Component power spectra. Insets show corresponding topographic maps.



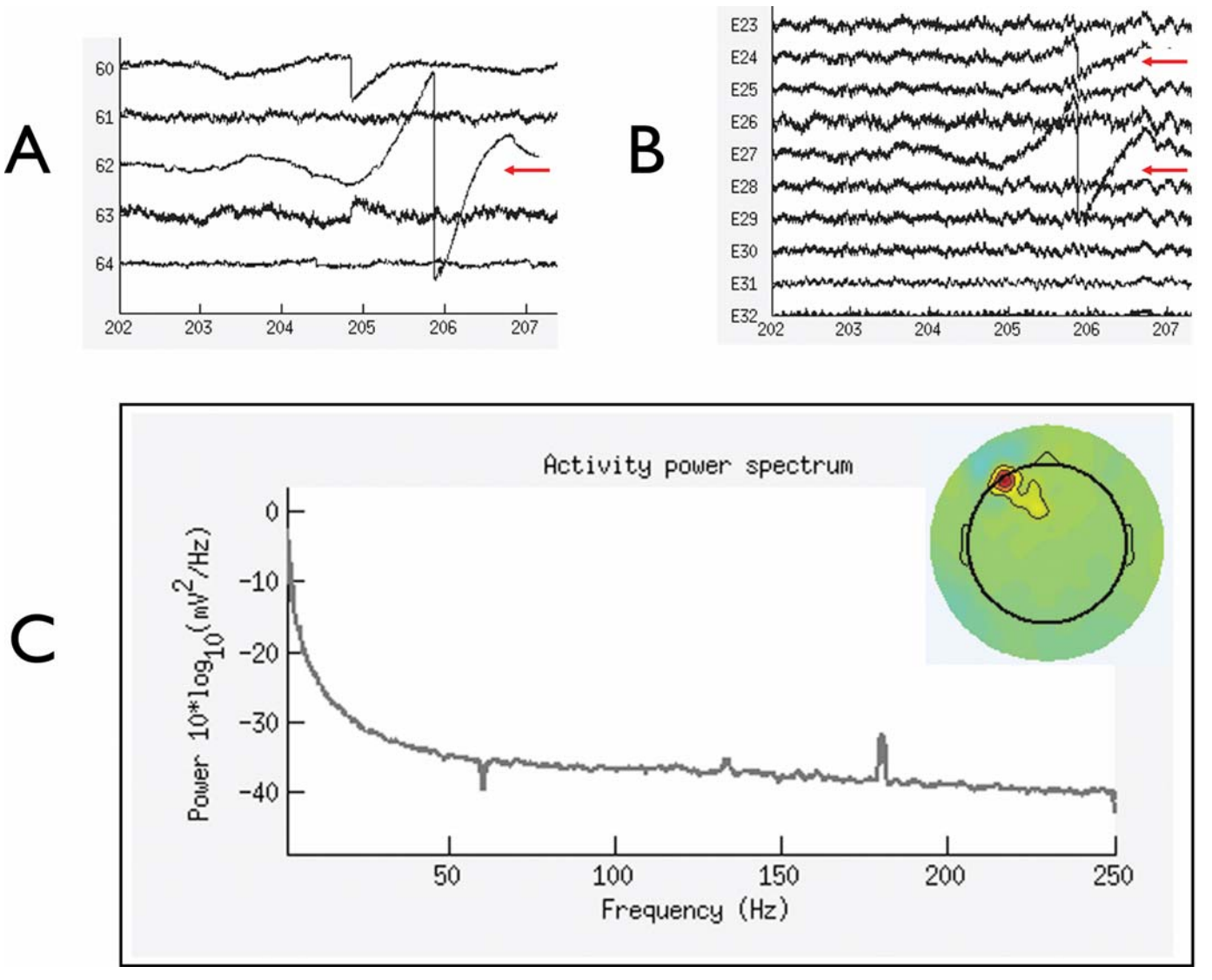
Supplementary Figure 11. *Neurogenic–Dominant Artifact.* **A)** Component time-series. **B)** Component power spectra. Insets show corresponding topographic maps.

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



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

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



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

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

Inter-Rater Reliability

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

Supplementary Tables for Scalp Analyses

Supplementary Table 1.
Sensitivity on the Scalp

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

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

Supplementary Table 2.
Specificity on the Scalp

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

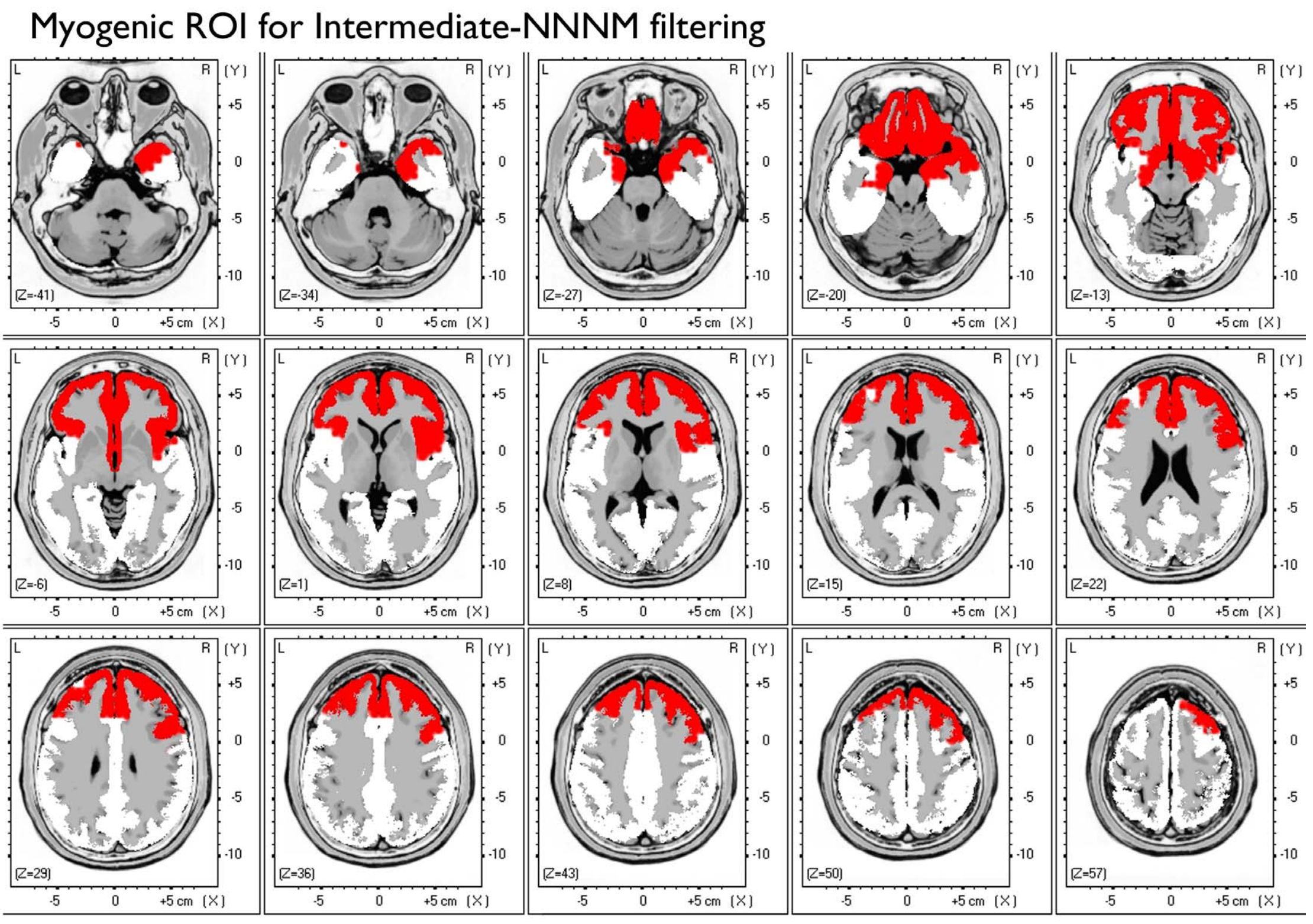
Note: [a] Corrected OR-CR vs. uncorrected OR-CR [b] Corrected OT-CR vs. uncorrected OR-CR. [c] Corrected OR-CT vs. uncorrected OR-CR. [d] Cells show the *t*-test.. [e]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data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

Supplementary Table 3.
Correction Artifact on the Scalp

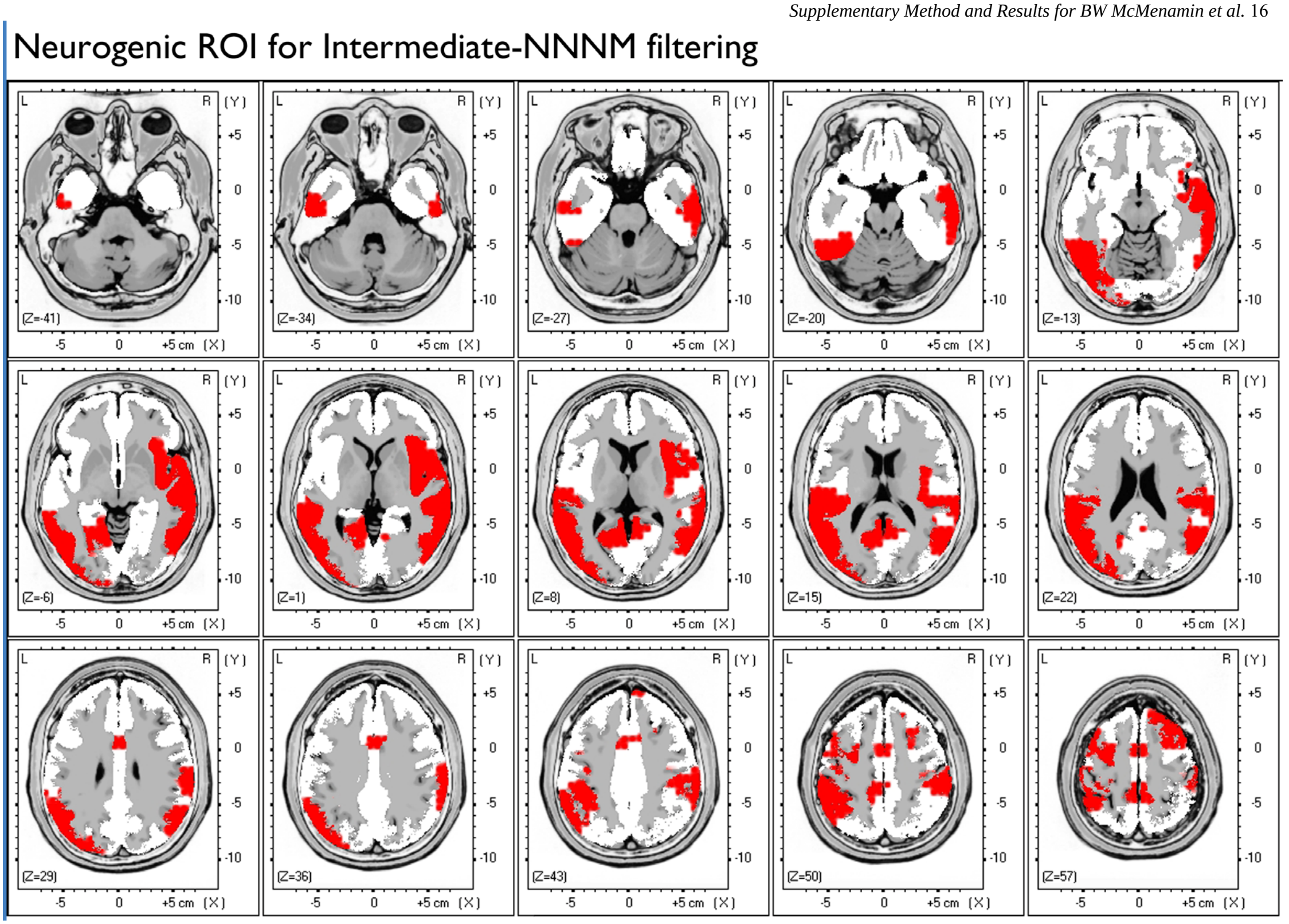
			ROI Median Contrast		ROI Peak Contrast	
	EMG Correction	NNNM Filtering	Myogenic[a]	Neurogenic[b]	Myogenic[a]	Neurogenic[b]
Corrected Contaminated vs. Uncorrected Analog[c]	None	Minimal	-1.25		-2.10†	
		Intermediate	-1.31		-2.06†	
		Maximal	-1.33		-2.12*	
	Minimal	Minimal	-1.29	-1.53	-2.08†	-2.41*
		Intermediate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

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

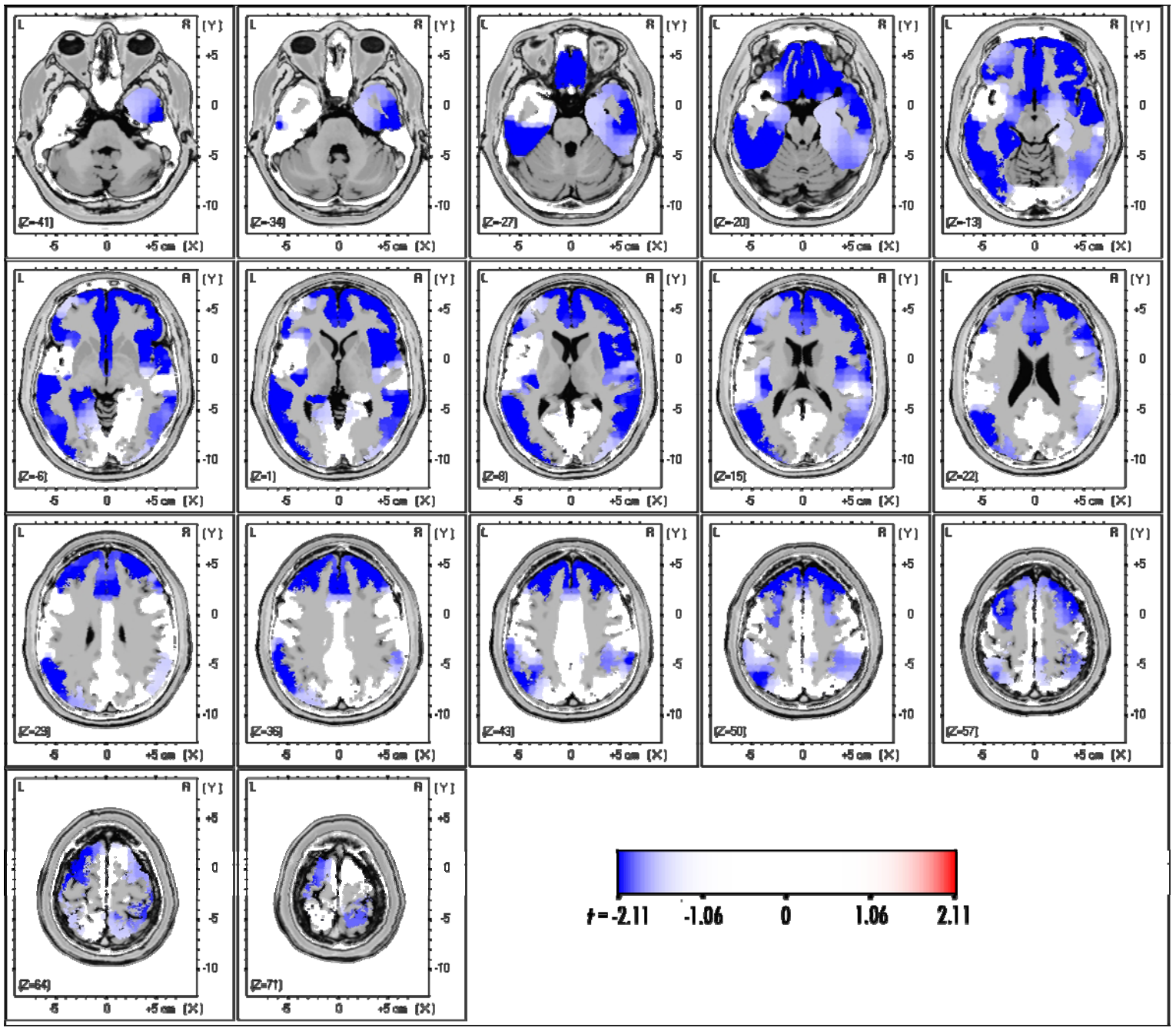
Supplementary Figures for Source-Modeling Analyses



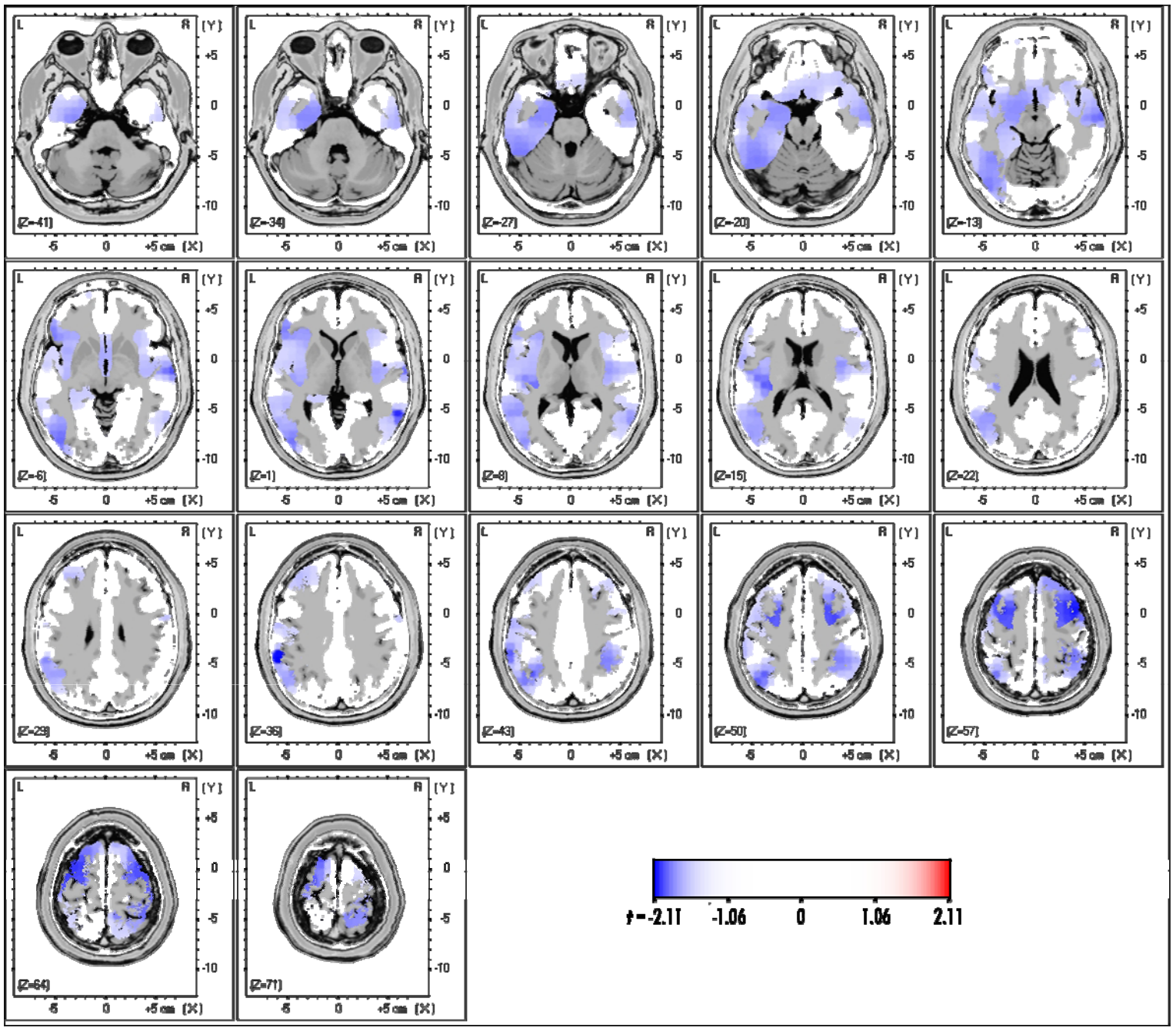
Supplementary Figure 14. *LORETA regions of interest (ROIs) for the myogenic contrast.*



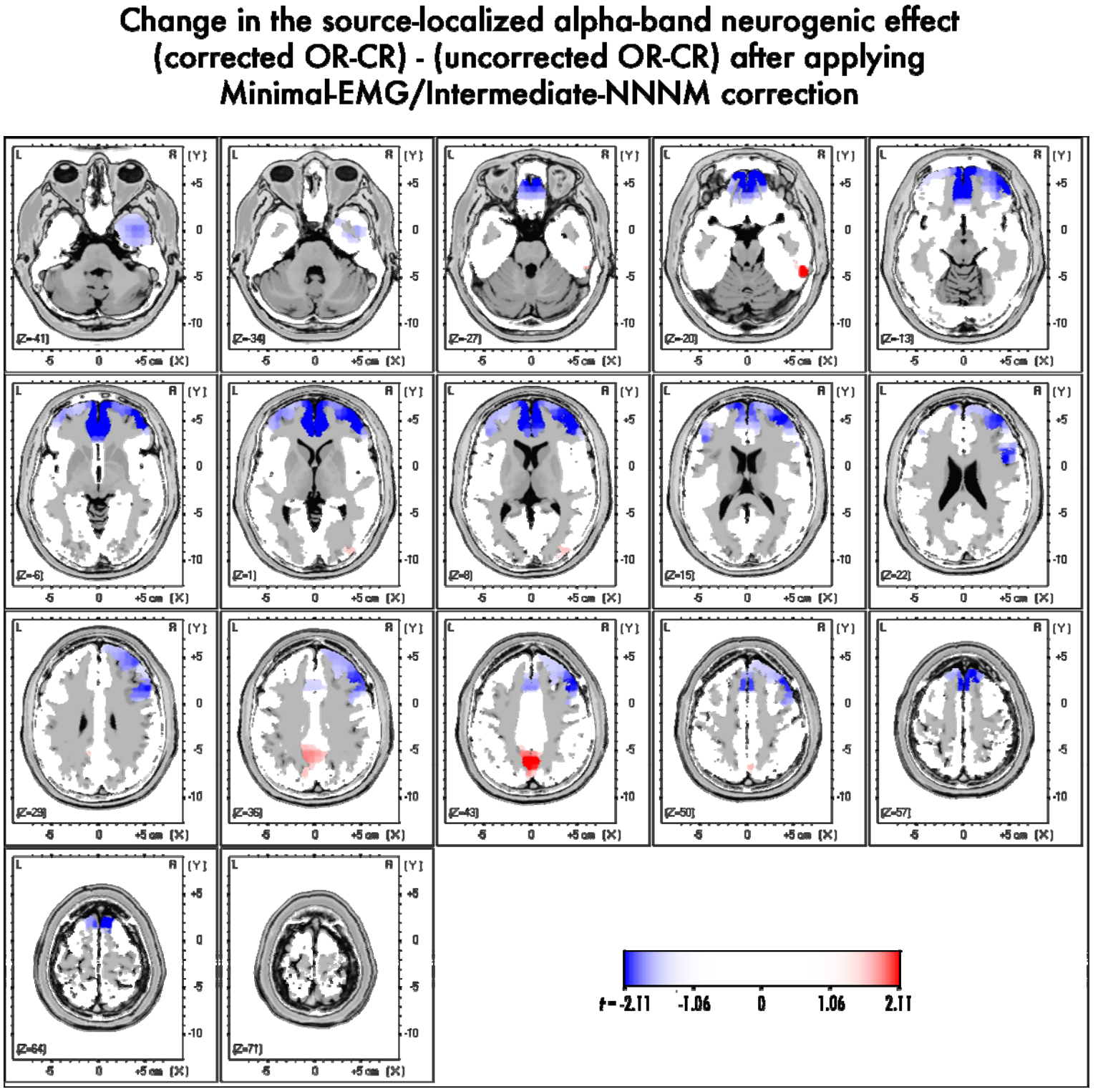
Supplementary Figure 15. *LORETA regions of interest (ROIs) for the neurogenic contrast.*



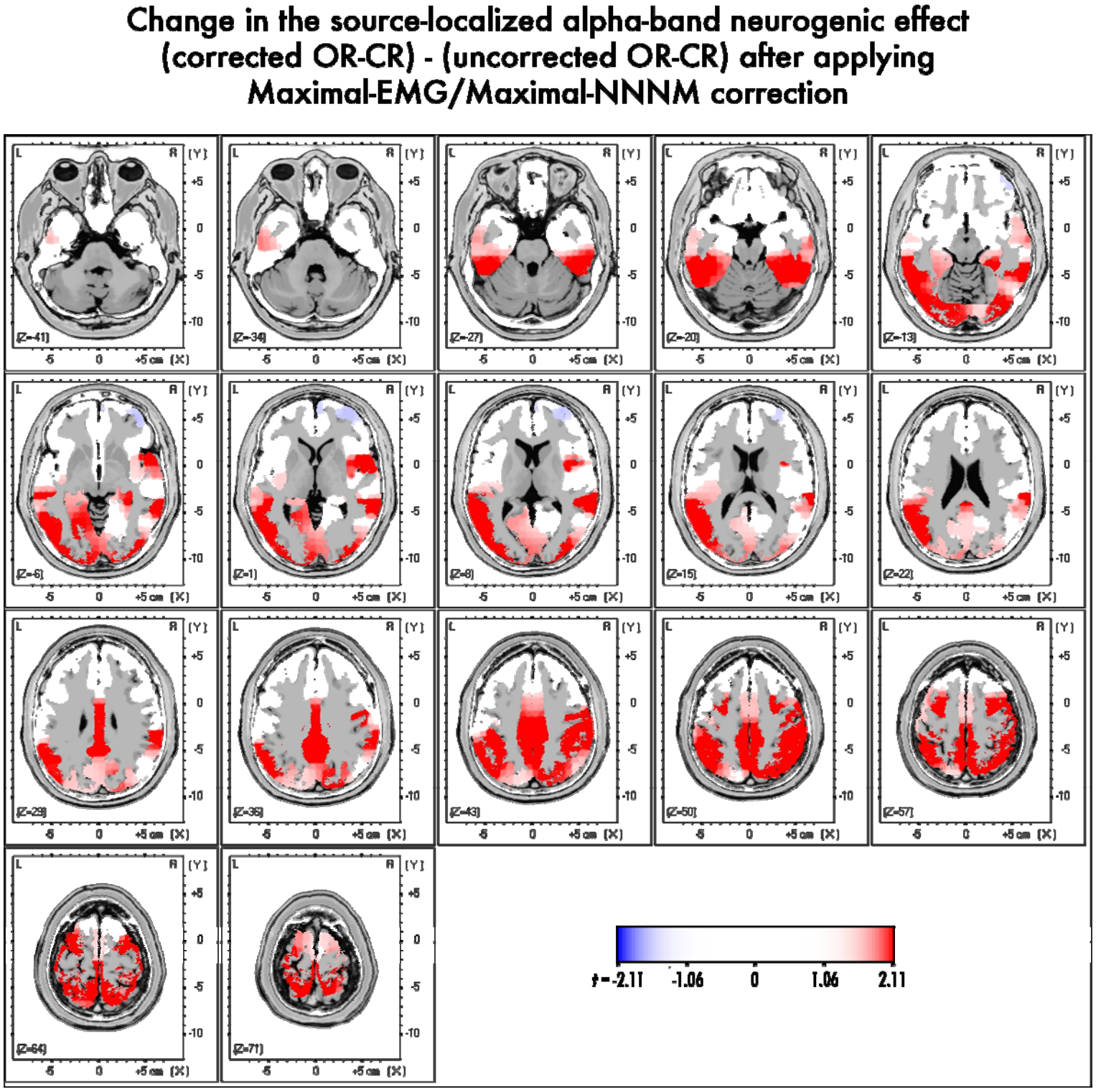
Supplementary Figure 16. *LORETA solution for the myogenic effect (OR-OT) after applying Minimal-EMG/Intermediate-NNNM correction*



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



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



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

Supplementary Tables for Source-Modeling Analyses

Supplementary Table 4.
Sensitivity in the Source-space

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

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

Supplementary Table 5.
Specificity in the Source-space

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

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

Supplementary Table 6.
Correction Artifact in the Source-Space

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

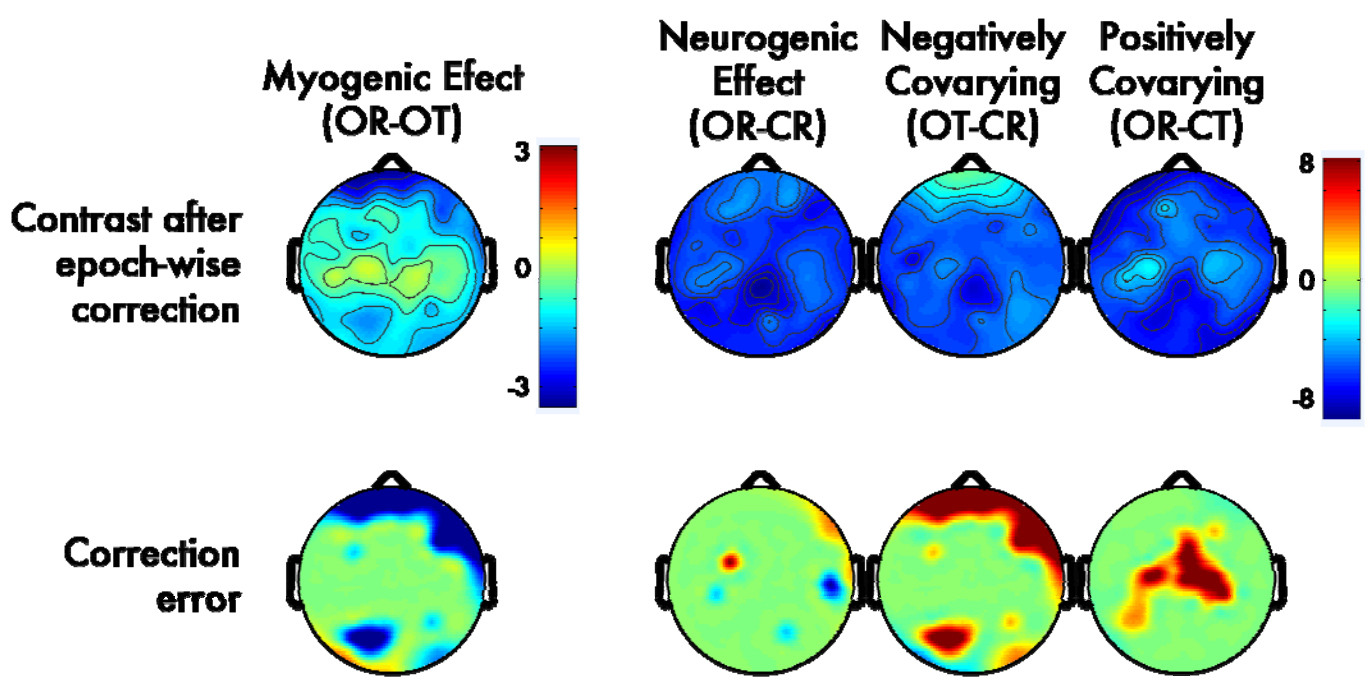
Note: [a] Corrected OR-OT contrast in the neurogenic ROI [b] Corrected OR-CR vs. uncorrected OR-CR in the myogenic ROI. [c]Cells show the *t*-test. [d]Cells show the TOST ε and equivalence region. In cases where the equivalence region was within +/- ε the corrected EMG-contaminated data was statistically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

Regression-Based EMG Correction

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

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

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



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

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

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

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

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

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

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

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

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

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

Post Hoc Model Order Estimation

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

Supplementary References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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