



Proposing Metrics for Benchmarking Novel EEG Technologies Towards Real-World Measurements

Anderson S. Oliveira^{1*}, Bryan R. Schlink¹, W. David Hairston², Peter König^{3,4}
and Daniel P. Ferris¹

¹ Human Neuromechanics Laboratory, School of Kinesiology, University of Michigan, Ann Arbor, MI, USA, ² Translational Neuroscience, Human Research and Engineering Directorate, U.S. Army Research Laboratory, Adelphi, MD, USA, ³ Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany, ⁴ Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Recent advances in electroencephalographic (EEG) acquisition allow for recordings using wet and dry sensors during whole-body motion. The large variety of commercially available EEG systems contrasts with the lack of established methods for objectively describing their performance during whole-body motion. Therefore, the aim of this study was to introduce methods for benchmarking the suitability of new EEG technologies for that context. Subjects performed an auditory oddball task using three different EEG systems (Biosemi wet—BSM, Cognionics Wet—Cwet, Conionics Dry—Cdry). Nine subjects performed the oddball task while seated and walking on a treadmill. We calculated EEG epoch rejection rate, pre-stimulus noise (PSN), signal-to-noise ratio (SNR) and EEG amplitude variance across the P300 event window (CV_{ERP}) from a subset of 12 channels common to all systems. We also calculated test-retest reliability and the subject's level of comfort while using each system. Our results showed that using the traditional 75 μ V rejection threshold BSM and Cwet epoch rejection rates are \sim 25% and \sim 47% in the seated and walking conditions respectively. However, this threshold rejects \sim 63% of epochs for Cdry in the seated condition and excludes 100% of epochs for the majority of subjects during walking. BSM showed predominantly no statistical differences between seated and walking condition for all metrics, whereas Cwet showed increases in PSN and CV_{ERP} , as well as reduced SNR in the walking condition. Data quality from Cdry in seated conditions were predominantly inferior in comparison to the wet systems. Test-retest reliability was mostly moderate/good for these variables, especially in seated conditions. In addition, subjects felt less discomfort and were motivated for longer recording periods while using wet EEG systems in comparison to the dry system. The proposed method was successful in identifying differences across systems that are mostly caused by motion-related artifacts and usability issues. We conclude that the extraction of the selected metrics from an auditory oddball paradigm may be used as a benchmark method for testing the performance of different EEG systems in mobile conditions. Moreover dry EEG systems may need substantial improvements to meet the quality standards of wet electrodes.

Keywords: EEG, event-related potentials, walking, locomotion, reliability, mobility, validation metrics

OPEN ACCESS

Edited by:

Mikhail Lebedev,
Duke University, USA

Reviewed by:

Tamer Demiralp,
Istanbul University, Turkey
Leon Y. Deouell,
The Hebrew University of Jerusalem,
Israel

*Correspondence:

Anderson S. Oliveira
oliveira_dkbr@hotmail.com

Received: 06 November 2015

Accepted: 14 April 2016

Published: 10 May 2016

Citation:

Oliveira AS, Schlink BR, Hairston WD,
König P and Ferris DP (2016)
Proposing Metrics for Benchmarking
Novel EEG Technologies Towards
Real-World Measurements.
Front. Hum. Neurosci. 10:188.
doi: 10.3389/fnhum.2016.00188

INTRODUCTION

Descriptions of brain activity using non-invasive electroencephalographic (EEG) have become an important topic in recent years. Special attention has been devoted to establish methodologies for successfully recording brain activity in real-world conditions (Casson et al., 2010; McDowell et al., 2013; Ries et al., 2014; Mihajlovic et al., 2015; Oliveira et al., 2016). The popular framework of embodied cognition requires investigating and modeling natural brain dynamics (Gramann et al., 2014). Along these lines, investigating sensorimotor coupling requires allowing a wide range of motor actions like complex movements, locomotion and navigation (Gwin et al., 2010; Ehinger et al., 2014). However, due to the high dynamic range and low signal-to-noise in these types of situations, acquiring reliable records of brain electrical activity in these motion conditions requires robust EEG acquisition systems beyond what has been conventionally used in more restricted laboratory or medical scenarios. Such achievement can contribute substantially to advance many research areas, such as psychology, motor rehabilitation, aging, robotics, and user-system interaction (Lance et al., 2012; McDowell et al., 2013; Kranczioch et al., 2014).

In the past few years, new EEG data acquisition approaches have been developed which leverage wireless data transmission and wet or dry electrodes as an attempt to increase the general usability of the systems for non-traditional use (Chi et al., 2010; Grozea et al., 2011; Zander et al., 2011; Hairston et al., 2014; Lopez-Gordo et al., 2014). While promising, an ongoing challenge is how to validate or prove the overall efficacy of these approaches, since there currently are no community-accepted or standard metrics for EEG signal quality as observed within real-world domains. In most cases, authors use the output from classic tasks such as steady-state visual evoked potential and oddball paradigms in stationary conditions as an attempt to benchmark the performance of these systems (Zander et al., 2011; Chi et al., 2012; Liao et al., 2012; Ries et al., 2014). All these investigations reported satisfactory performance of the dry EEG systems/electrodes in comparison to wet systems/electrodes, generally based on the performance of classifiers or the overall average waveform of an event-related potential (ERP). However some authors raised limitations related to timing (Hairston et al., 2014; Ries et al., 2014) and susceptibility to movement artifacts (Chi et al., 2012), which would be particularly problematic in real-world scenarios involving ambulation. Thus, these efforts do not properly encapsulate the challenges likely to be encountered using mobile EEG systems within their target application domain.

Proper benchmarking also requires the selection of metrics that can successfully underpin multiple potential limitations and not simply a comparison of classifier performance. Recent studies focusing on benchmarking EEG systems have been based on a wide range of metrics related to EEG spectral power and amplitude of specific ERP events (Zander et al., 2011; Chi et al., 2012; Ries et al., 2014; Fiedler et al., 2015). However when surveying these studies, a variety of specific

metrics are used, highlighting the difficulty of establishing solid methods and metrics for benchmarking. Moreover, studies related to whole-body motion have the addition of substantial movement artifacts (Gwin et al., 2010; Castermans et al., 2014; Reis et al., 2014; Oliveira et al., 2016). Therefore, effective benchmarking for EEG systems towards its usage in whole-body motion conditions is crucial for future advances in mobile brain measurements. In order for benchmarking metrics to be practically informative for real-world applications, they must involve the identification of appropriate metrics for defining the influence of motion on EEG parameters and the reliability of these metrics. The assessment of subject's comfort and motivation while using EEG systems is another important aspect of the benchmark.

The primary aim of the study described here was to introduce experimental protocols that may serve as benchmarks for assessing the quality of current EEG technologies devoted to mobile recording. Subjects performed an auditory oddball task while seated and during treadmill walking. We calculated basic variables such as absolute power, pre-stimulus noise (PSN), signal-to-noise ratio (SNR) and variance of ERP epochs (CV_{ERP}), as well as the level of similarity between ERP curves from seated *vs.* walking conditions. The secondary aim was to investigate the test-retest reliability of these variables. The results of this investigation have fundamental implications on both the usage and community-wide evaluation of mobile EEG systems for exploring brain activity from freely moving humans.

MATERIALS AND METHODS

Nine healthy volunteers (6 males and 3 females) between the ages of 21–36 years participated in the study. None had any history of major lower limb injury or known neurological or locomotor deficits. All study procedures were approved by the University of Michigan Internal Review Board and complied with the standards defined in the Declaration of Helsinki. All subjects provided informed, written consent before participating.

EEG Systems

All subjects were tested in three different sessions interspaced from 7–40 days. In each session they wore one of the following EEG systems in a random order: (1) Biosemi (248-channel, ActiveTwo, BioSemi, Amsterdam, Netherlands), abbreviated as BSM; (2) Cognionics (72-channel Mobile, Cognionics Inc, San Diego, CA, USA), abbreviated as Cwet; and (3) Cognionics (72-channel Dry, Cognionics Inc, San Diego, CA, USA), abbreviated as Cdry (**Figure 1**). The Biosemi system uses active electrodes, while both of the Cognionics systems use passive electrodes with active shielding. In each session one of the systems was mounted on the head of the subject following the manufacturers' recommendations. Subjects were seated in a comfortable, non-reclining chair without any head support or walked at 1.0 m/s on a treadmill while we recorded scalp EEG. For the Biosemi system, gel was used to lower the contact impedance of the electrodes and bring the electrode offset,

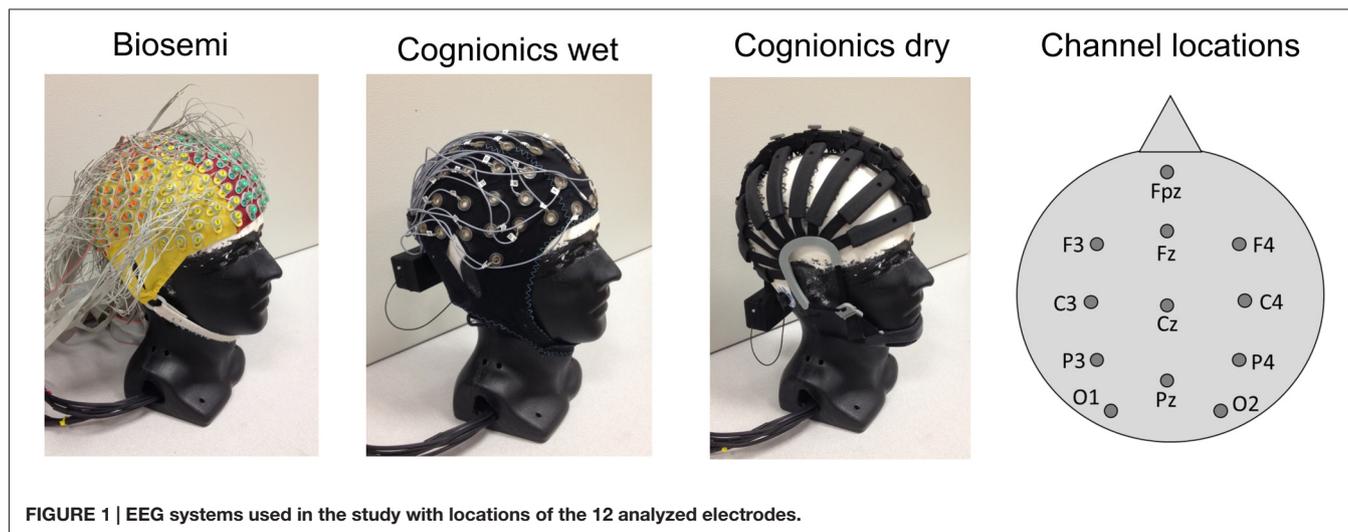


FIGURE 1 | EEG systems used in the study with locations of the 12 analyzed electrodes.

defined as the running average between the CMS electrode and each of the active electrodes, below 20 mV, as suggested by the manufacturer. For the Cognionics dry system, we made specific adjustments, including moving the electrode legs gently through the hair to touch the scalp and adding small amounts of water at the skin-electrode interface, to optimize the electrode contact with the scalp and bring the impedance for each individual electrode close to 100 k Ω . However, this threshold could not be achieved for all electrodes simultaneously in any data collection. For the wet Cognionics system, we again used gel to lower the contact impedance. We made sure that the impedance of each electrode was below 20 k Ω . Specifically for the Cdry system, after the cap was appropriately positioned on the subject's head according to the International 10–20 system, each row was lightly tightened so that the electrodes had a firm contact with the scalp without causing any discomfort. The electrodes were also maneuvered through the subject's hair to ensure they had direct contact with the scalp and were not resting on layers of hair. Subsequently, rows containing electrodes with high impedance were once more gently tightened in order to improve contact between electrodes and the scalp. The total time from cap application to the start of recording was 70 ± 19 min for Biosemi, 38 ± 10 min for Cwet and 28 ± 9 min for Cdry. In evaluating these numbers, please note that the Biosemi system had more than triple the number of electrodes than the two others. In order to assess the test-retest reliability, four subjects could be recruited to perform a second round of tests using the three EEG systems 30–60 days after finishing the first round of tests. This second round of tests were interspaced from 7 to 20 days. These intervals between sessions and for the test-retest intervals were within an acceptable range according to previous work investigating ERP reliability (Lew et al., 2007; Brunner et al., 2013).

Experimental Protocol

In each experimental session, subjects participated in auditory oddball discrimination and response tasks, once while seated

and once while walking. The proposed walking condition was meant to simulate a more realistic movement scenario when compared to the seated recordings. The order of these conditions was randomized across equipment and subjects. Subjects were seated in a padded, non-reclining, stationary chair without head support for the seated portions of the experiment. For the seated portions with the Biosemi system, the A/D box and electrode strands were placed behind their head on a small table so that their weight did not pull backwards on the subject. For the walking conditions, the A/D box was placed in a small backpack and the electrode strands were tethered to the backpack so that the subject had enough slack for a normal range of motion with their head. They were also tethered in such a way that they did not pull the subjects head backwards. The total weight of the backpack they carried was approximately 2.2 kg. There was no need for backpack setup for recordings using Cwet and Cdry systems as they were totally wireless. Subjects were also instructed not to hold onto the treadmill railing unless they needed to do so for their balance. The temperature of the room was set to 72° (F) throughout the experiment. In general, subjects were not visibly sweating and did not report sweating very much when asked. For each condition, participants listened to words presented using E-prime software (Psychology Software Tools, Version 2.0, Sharpsburg, PA, USA) from two speakers placed bilaterally 1.3 m high and 70 cm away from the subjects. Although hearing loss was not objectively assessed, all patients had no difficulty in hearing and responding to the sounds in both conditions. The stimuli were four different color names (blue, red, yellow and green), and one of these four words was selected to be a target (for instance, the word blue) while the other words (red, yellow and green) were non-target. Target and non-target stimuli appeared with a 0.25 and 0.75 probability respectively. Each word sound was presented for 400 ms with a 1420–1580 ms inter-stimulus interval. A fixation point was continuously displayed on a monitor placed at eye level about 1 m in front of the subjects while they were instructed to maintain a fixation on the center of the screen as

TABLE 1 | General specifications for each EEG system.

EEG system	Sampling rate (Hz)	Bandwidth (Hz)	Reference	# of EEG channels	Wireless transmission	Conductive mechanism
BSM	512	0–100	Averaged Mastoids	256	None	Gel
Cwet	500	0–131	Linked Mastoid	64	Bluetooth	Gel
Cdry	500	0–131	Linked Mastoid	64	Bluetooth	None

the stimuli were presented. Subjects were instructed to respond as fast and as accurately as possible by pressing a handheld button every time they heard the target word. If the word was a non-target, they were told not to press the button and wait for the next stimulus. Each participant completed four blocks of 160 trials per condition (seated or walking), and for each block, a different color name was randomly used as target stimulus.

Recording and Analysis

We recorded EEG for each system on a Windows 7 PC, separate from the E-prime computer, using the native software as provided by the system manufacturers (ActiView, Version 6.05, for BioSemi, and Cognonics Data Acquisition, Version 3.6, for Cognionics). **Table 1** shows EEG system specifications (number of channels, sampling rate, signal bandwidth, and reference used). All EEG data were processed in Matlab (The Mathworks, Natick, MA, USA) using EEGLAB and ERPLAB (Delorme and Makeig, 2004; Lopez-Calderon and Luck, 2014). Continuous EEG data from each recorded file were high-pass filtered (1 Hz) and with additional notch filtering using Cleanline.¹ The use of 1 Hz high-pass filtering might influence the ERP outcomes (Tanner et al., 2015); however for our methodological investigation it became necessary to use this filter for both sitting and walking conditions as described in previous literature (Gramann et al., 2010; Gwin et al., 2010; Ehinger et al., 2014). All datasets from the same condition were concatenated, generating a continuous dataset. Specific frame sequences containing large artifacts resulting from lost packets during wireless telemetry and muscle activity (EMG) were removed from the continuous EEG data. Channels with large fluctuations were marked for rejection and interpolated using EEGLAB. For BSM, we interpolated 1.2 ± 0.4 and 1.4 ± 0.6 channels for seated and walking condition respectively. For Cwet, we interpolated 1.3 ± 0.5 and 1.8 ± 0.5 channels for seated and walking condition respectively. For Cdry, we interpolated 2.1 ± 0.7 and 4.1 ± 1.6 channels for seated and walking condition respectively. Subsequently, independent component analysis was performed on these merged datasets for identifying and removing eye blinks (Ries et al., 2014). In addition, 1–3 components exhibiting large fluctuations in phase with the subject gait cycle were also removed from the walking datasets (Onikura et al., 2015). We identified these components by evaluating the frequency content, which usually has a peak around 1.2 Hz corresponding to the walking frequency. Additionally, we

inspected the temporal properties of the IC activation in order to confirm that there was clear periodic oscillation around 1.2 Hz.

The artifact-free datasets were downsized to 12 channels (Fpz, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, O1 and O2) in order to include samples from each of the primary regions of the cap (De Vos et al., 2014). Preliminary analyses showed these to be sufficiently representative of the primary regions, while also providing common positions across all three systems. Subsequently, these datasets were epoched from -0.3 s to 0.8 s window surrounding the onset of target and non-target stimuli. Non-target stimuli followed by button presses and target stimuli not followed by button presses (i.e., false positives and false negatives, respectively) were rejected. We also rejected epochs based on EEG amplitude thresholds, which were noted for comparison later in the “Results” Section. We evaluated all datasets from seated and walking conditions for each EEG system across a range of amplitude thresholds from 75 to $400 \mu\text{V}$. This analysis generated a range of rejection rates used for defining the optimal threshold for further analysis (see “Results” Section for more details). In order to perform fair comparison between seated and walking EEG datasets, we rejected the same number of epochs from both conditions for each system, based on the condition with the most rejected epochs.

From the remaining datasets containing single trials that passed the threshold analysis, we calculated the frequency power (*spectopo* function from EEGLAB, 256 point FFT, 128 point window) for each epoch from each channel separately. Subsequently, we averaged the data from all epochs for each channel and then averaged all channels. This measure, therefore, contains not only the evoked but also the induced power. We calculated the average absolute power across all 12 electrode locations for the, theta (5–8 Hz), alpha (9–13 Hz), beta (13–30 Hz) and gamma bands (30–80 Hz) from both seated and walking conditions. Subsequently we calculated a ratio between EEG absolute power walking/seated (Ratio W/S). Ratios W/S above 1 would indicate that EEG power during walking was higher in comparison to seated. In addition, PSN, defined as the root-mean square of the period from -300 to 0 ms, was calculated for each single trial. The single trial target and non-target datasets were subsequently averaged using ERPLAB and from these averaged curves we calculated the SNR—defined as the ratio of the peak amplitude for P300 and the averaged PSN (De Vos et al., 2014). In addition, we calculated the coefficient of variation—defined as the ratio of the standard

¹<https://www.nitrc.org/projects/cleanline/>

deviation to the mean—for trial i ($CV(i)$) in a window of 300–500 ms:

$$CV(i) = \frac{\sqrt{\text{var}_t(EEG(i, t))}}{\text{mean}_t(EEG(i, t))}$$

with i as the index of epochs and t as the interval of 300–500 ms. This window contains the P300 event for target stimuli (De Vos et al., 2014). These values were then averaged over epochs (CV_{ERP}):

$$CV_{ERP} = \frac{\sum_i CV(i)}{n_i}$$

with n_i as the number of epochs. This computation was performed for each channel and each subject in both seated and walking conditions for each system.

In order to quantitatively compare the outcomes of an oddball test during seated *vs.* walking, we calculated the scalar product of target averaged ERPs curves between each subject's seated and walking conditions for each system separately. Scalar products of normalized vectors yield a score ranging between -1 and $+1$ representing the overall similarity of a signal (d'Avella et al., 2003) and are equivalent to a correlation coefficient. Thus, scalar products $0.8 > r \geq 0.6$ would suggest only moderately similar ERPs recorded for these two conditions, and $r \geq 0.8$ would suggest highly similar curves for the two conditions. Fisher's transformation was applied on the individual scalar products values in order to normalize across the distribution for further averaging and back-transformation to the correlation domain.

Post-Experiment Survey

Immediately after finishing the first session using each of the three headsets, we asked subjects to fill a survey containing the following questions: (1) How well did the cap fit to your head? (2) Did you feel discomfort during the experiment? (3) Did you feel discomfort because the electrodes pinched your head? (4) When did the EEG cap/electrodes start to be uncomfortable? (5) When, if at all, did you lose motivation to continue the experiment? and (6) When, if at all, did you have troubles to stay focused during the experiment? Answers were based on a 5-point Likert scale. No survey was conducted to the four subjects when performing a second round of tests with each system.

Statistical Analysis

Initially, we performed analysis on EEG epochs based on cumulative distribution function of the single epoch amplitudes, in order to define the amplitude range in which the majority of epochs were found. Subsequently, we applied epoch rejection tool available on ERPLAB based on the range between the minimum and maximum amplitude (200 ms window, 100 ms overlap) in the seated and walking EEG datasets using thresholds varying from 75 to 400 μV . This analysis helped in defining a single threshold for rejecting epochs throughout the subsequent results. After defining the appropriate threshold, we used a 2-way analyses of variance (ANOVA; systems \times EEG

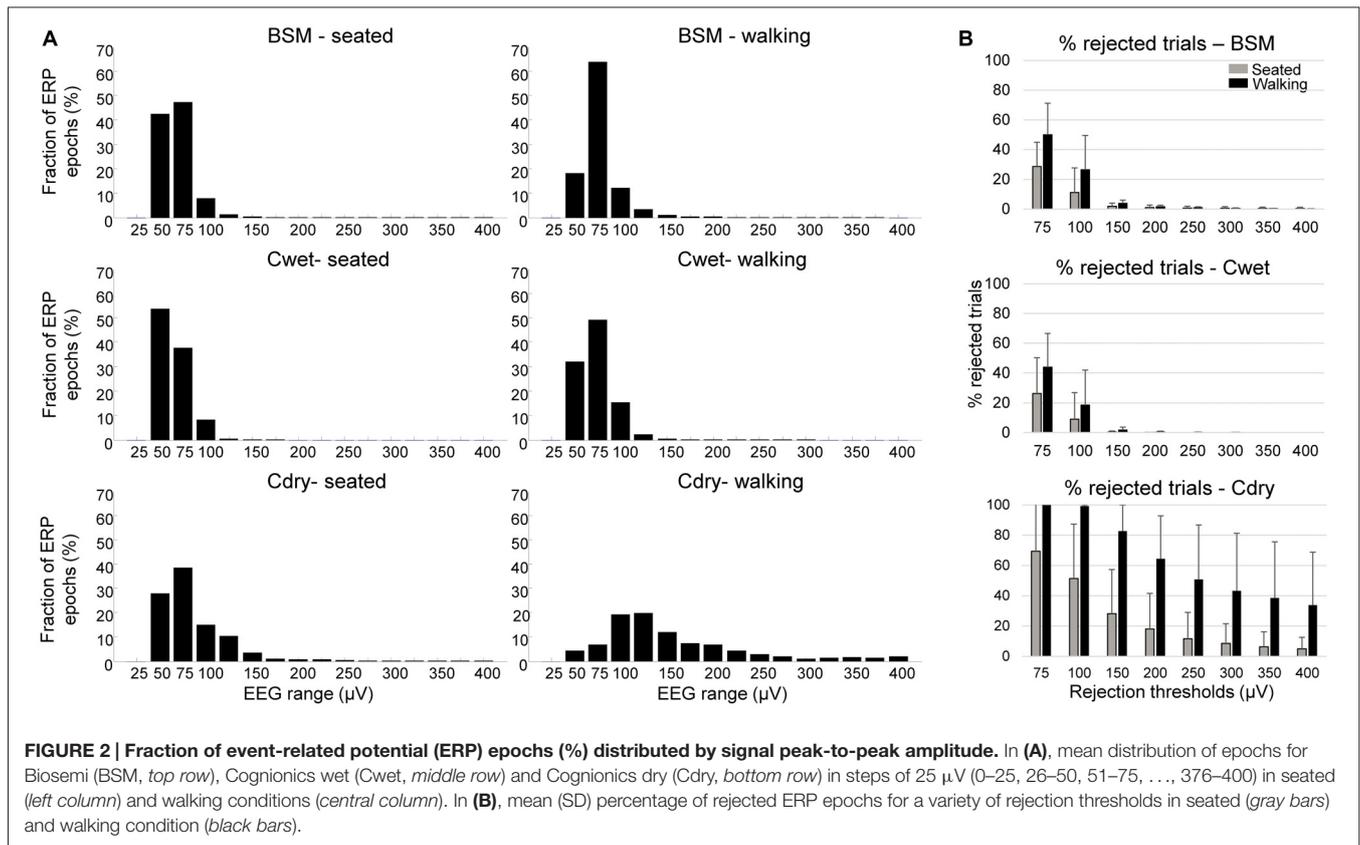
frequency band) to assess the effects of the EEG systems and EEG frequency band on the Ratio W/S. We also assessed the effects of EEG systems (three systems), condition (seated \times walking) and EEG channel (i.e., the 12 electrode locations) on the SNR and CV_{ERP} . Due to the exceedingly high rejection rate for the Cdry-walking condition a complete 3-way ANOVA covering all combinations was problematic. Therefore we used two types of 2-way ANOVAs: (1) 2-way ANOVA (3 EEG systems \times 12 electrode locations) for the seated condition; and (2) 2-way ANOVA (2 conditions \times 12 electrode locations) for Biosemi and Cwet separately. The data from Cdry walking was removed because there was not substantial data available from eight out of nine subjects after thresholding. In case of significant main effects of EEG system and/or EEG channel we performed *post hoc* pairwise *t*-tests with LSD multiple comparison correction.

Log transformation was applied to the variables exhibiting non-normal/skewed distribution. The alpha level of significance was set at $p < 0.05$, and Cohen's d effect size was calculated for all variables ($0.2 < d < 0.5$ = small effect, $0.5 < d < 0.8$ = medium effect, $d > 0.8$ = large effect). Test-retest reliability was calculated by intraclass correlation coefficient (ICC) method using a two-way mixed effect model set for absolute agreement, as in Brunner et al. (2013) for EEG absolute power in different bands, PSN, SNR, CV_{ERP} and scalar products between seated and walking ERP curves. We used 1-way ANOVA in order to assess the differences among systems for each of the questions in the post-experiment survey, followed by *post hoc* pairwise *t*-tests with LSD multiple comparison corrections.

RESULTS

Rejection Rate

The average rate at which data must be rejected due to high-amplitude swings, noise, or other artifacts has an impact on the overall usability of the recording. The majority of the ERP epochs showed peak-to-peak amplitudes ranging from 25 to 75 μV for all three EEG systems in seated conditions, as well as for BSM and Cwet in walking conditions (Figure 2A). In the case of Cdry during walking, the majority of the epochs showed peak-to-peak amplitudes ranging from 75 to 125 μV , and $\sim 32\%$ of all epochs showed peak-to-peak amplitudes above 125 μV . Varying the amplitude of cutoff thresholds showed that a traditional threshold at 75 μV may remove $\sim 30\%$ of the seated epochs for BSM and Cwet (Figure 2B), and this same threshold may remove $\sim 68\%$ of the seated epochs for Cdry. There were a higher number of removed epochs for the walking condition regardless the selected threshold. The 75 μV threshold removed 40–50% of the walking epochs for BSM and Cwet, whereas all epochs were marked for removal for Cdry in eight out of nine subjects, and the remaining subject had 94% of all epochs removed. Meanwhile, thresholds above 100 μV selected only a marginal number of walking ERP epochs for rejection for both BSM and Cwet ($< 4\%$ in average). On the other hand, thresholds at 175 μV and below marked more than 80% of all walking ERP epochs for rejection for



Cdry. These differences highlight the challenge associated with recording enough ERP epochs within lower amplitude ranges using Cdry.

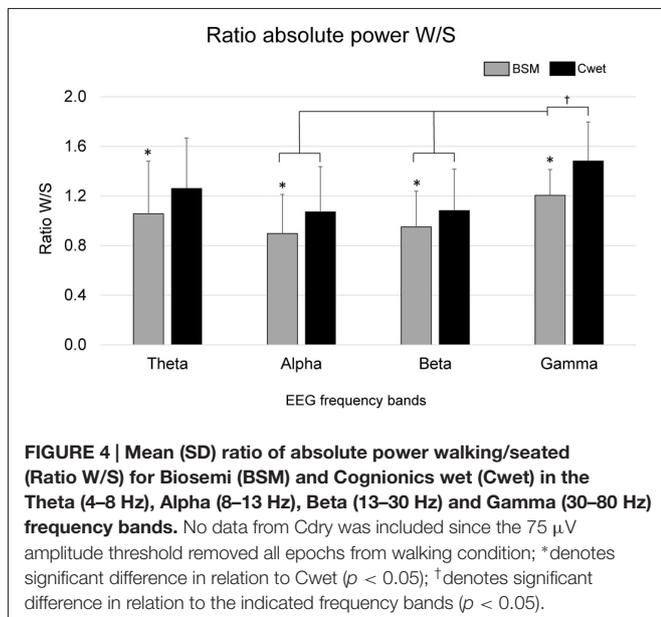
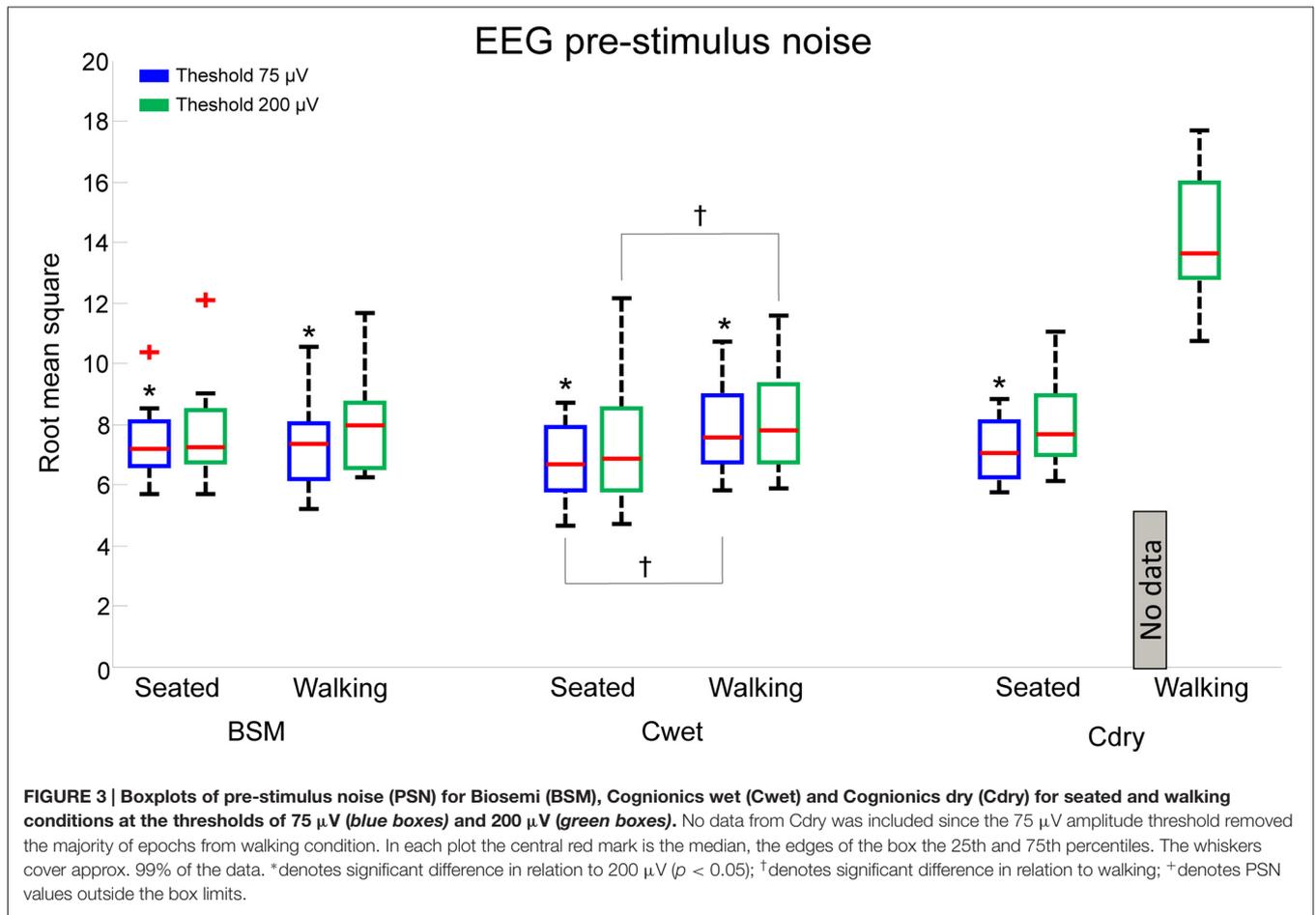
Pre-Stimulus Noise in Different Thresholds

In order to describe the effects of condition (seated vs. walking) and different thresholds on a parameter related to EEG data quality, we calculated single epoch PSN by using two thresholds (75 μV and 200 μV ; Figure 3). We did not include data from Cdry during walking for the 75 μV threshold as there was no data available from eight out of nine subjects. There were no interaction effects between electrode locations and any other effect (systems or conditions), therefore we averaged the 12 electrode location values for each subjects for further display in Figure 3. The separate ANOVAs for each system revealed a main effect of threshold on PSN results. This specific analysis did not aim at comparing performance across systems, but rather targeted the performance of a given system in two conditions (seated vs. walking) and two thresholds (75 vs. 200 μV) in order to define whether the threshold is a critical factor in determining the quality of the acquired EEG epochs. There was a significantly lower PSN for the 75 μV threshold in comparison to 200 μV for BSM ($F_{(1,64)} = 5.62, p < 0.05$), Cwet ($F_{(1,96)} = 6.01, p < 0.05$) and Cdry ($F_{(1,96)} = 34.3, p < 0.001$). In addition, there was a main effect of condition for Cwet ($F_{(1,96)} = 12.2, p < 0.001$, effect size = 0.7), in which lower PSN was found for the seated condition in comparison to walking. These results

suggested that there were no main effects of condition for the PSN recorded from BSM for both thresholds of 75 μV and 200 μV . Therefore this system performs equally in seated and motion conditions. Cdry showed substantial influence of motion, which made inviable the use of data from 75 μV threshold in walking conditions, and PSN from 200 μV threshold was substantially higher in comparison to those from wet systems. Based on these results, we fixed our threshold at 75 μV for the additional analysis presented in the following results, and there were no results from Cdry in the walking condition, since this threshold removed all epochs for eight out of nine subjects.

Spectral Analysis

In order to ascertain differences in sensitivity to spectral power in seated and walking conditions, we calculated the Ratio W/S. Cdry has been excluded from this analysis due to the lack of data from walking condition. A 2-way ANOVA (2 EEG systems vs. 5 EEG frequency bands) was used and a main effect of EEG system was found ($F_{(1,64)} = 6.73, p < 0.05$). The Ratio W/S for BSM was significantly lower in comparison to Cwet across all frequency bands (Figure 4). In addition, there was a main effect of EEG frequency band on the Ratio W/S ($F_{(3,64)} = 4.64, p < 0.01$). The ratio W/S from the delta band was significantly higher than theta and alpha bands ($p < 0.01$, effect sizes between 0.5 and 0.7), and the ratio from gamma band was significantly higher than alpha band ($p < 0.01$, effect size = 0.75) and beta band ($p < 0.01$, effect

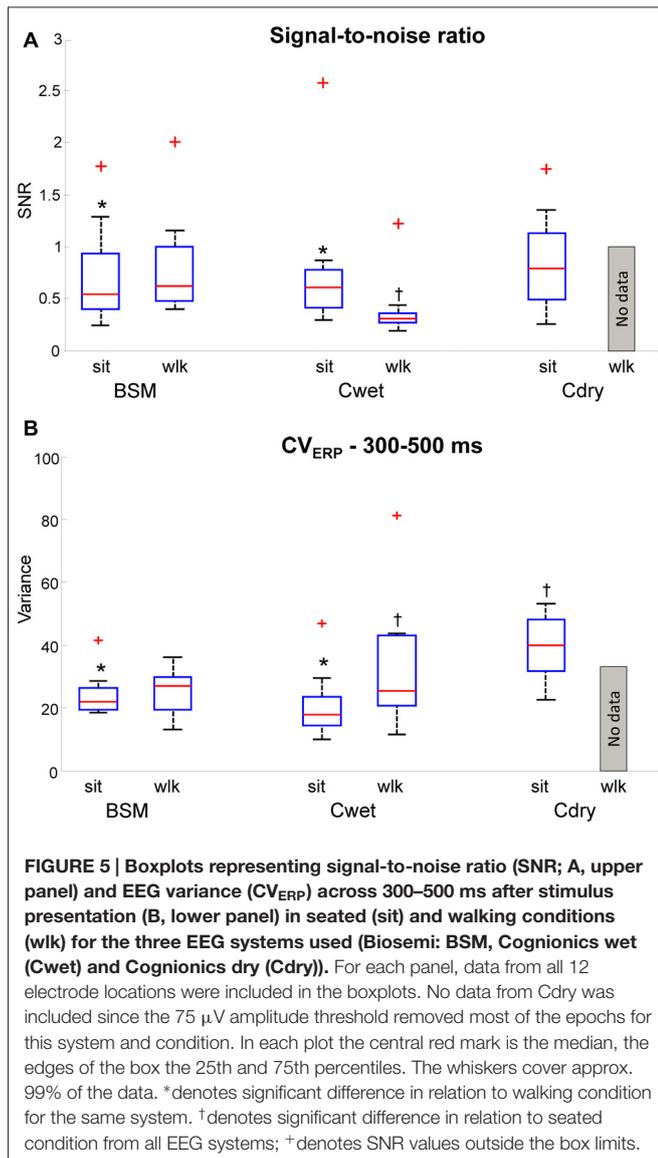


size = 0.82). The result from the ratio W/S show that BSM can provide ratios closer to 1, indicating better quality of results from walking recordings in comparison to Cwet. Moreover, we found

consistent differences across EEG frequency bands between BSM and Cwet.

Signal-to-Noise Ratio

The SNR is a variable widely used to describe the quality of EEG epochs with respect to the background noise inherent to EEG acquisition. We calculated SNR for each of the 12 channels for all EEG systems in the seated condition and from BSM and Cwet during walking. We investigated the effects of electrode location and condition for BSM and Cwet separately, using 2-way ANOVAs. In addition, we investigated the effects of electrode location and EEG system for the seated and walking conditions separately. For the seated condition we computed differences between BSM *vs.* Cwet *vs.* Cdry, whereas for walking we just computed differences between BSM *vs.* Cwet. There was no interaction between electrode location and condition, as well as no interaction between electrode location and systems ($p > 0.05$). There was interaction between EEG systems and condition for the SNR ($F_{(1,192)} = 8.70, p < 0.01$), showing that Cwet in walking had significantly lower values in comparison to BSM in the seated condition (Figure 5A). In addition, there was a significant main effect of EEG system ($F_{(2,192)} = 11.94, p < 0.001$). *Post hoc t*-tests with LSD multiple comparison correction revealed that Cwet had the highest SNR in the seated



condition, followed by BSM. We also found a significant main effect of condition ($F_{(1,96)} = 21.98, p < 0.001$), in which the SNR in the seated condition was significantly higher than walking for Cwet. The results from SNR indicated that all EEG systems have similar performances in the seated condition. However recording EEG data during walking reduces the quality of the signals for Cwet, whereas BSM can actually show increased SNR.

EEG Amplitude Variance

The variable CV_{ERP} was calculated in order to quantify the variability inherent to a window containing the P300 peak for individual epochs. Higher variances mean potential changes in the amplitude pattern that can influence the ERP results. For the purpose of this analysis we assume that CV_{ERP} might be higher during walking due to motion artifacts. Yet, high-quality systems must allow recordings

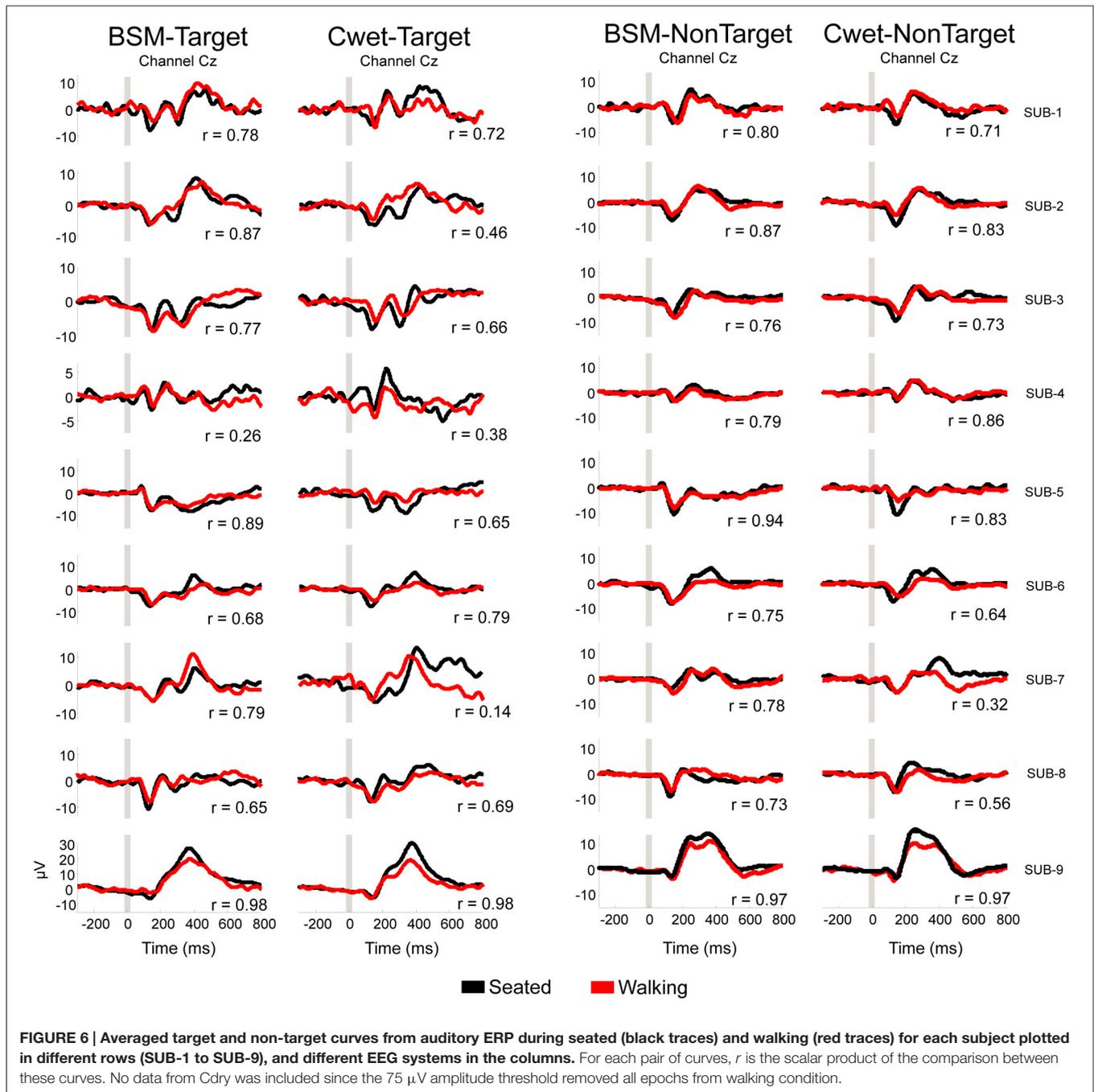
with minimal differences between seated and walking conditions. The statistical design was the same as applied for SNR. The results revealed interaction between EEG system and condition ($F_{(1,192)} = 55.78, p < 0.001$), showing that Cdry in seated condition had significantly higher CV_{ERP} in comparison to other EEG systems in the same condition (**Figure 5B**, data from all channels combined in the boxplots). In addition, we also found a significant main effect of condition ($F_{(1,96)} = 42.21, p < 0.001$), in which CV_{ERP} in seated condition was significantly higher than walking for BSM. No significant main effect of electrode location was found ($p > 0.05$). Our results from CV_{ERP} demonstrated that, similarly to the SNR, BSM can maintain the same level of quality on the data recorded during walking, whereas Cwet showed increases in variance in walking conditions. Moreover, Cdry delivered data from seated condition that was more variable than all datasets from wet systems.

Similarities Seated vs. Walking

Scalar products provide an objective measurement for the overall similarity of signals without regard to potential differences in scaling, and a means to assess any distortion observed between seated and walking ERP curves. The comparison between parameters such as P300 peak extracted from seated and walking conditions show similar results using traditional ERP methods (Gramann et al., 2010; Debener et al., 2012). However, it is possible that ERP epochs may be affected by motion artifacts during recordings while walking. The scalar products were included in this analysis in order to provide quantitative comparison of the entire ERP waveforms between conditions. Illustrations of target ERP curves from the sitting and walking conditions for both wet EEG systems are shown in **Figure 6** for the target and non-target stimuli from the Cz electrode. In general, BSM and Cwet presented predominantly similar patterns in the two conditions (5 out of 9 subjects with $r > 0.75$). Regarding all analyzed channels, the grand average across all channels (including target and non-target stimuli) for BSM was $r = 0.75 \pm 0.18$. There were 186 scalar products ($>80\%$ of the total from 9 subjects \times 12 electrode locations \times target/non-target stimuli) showing moderate or high similarity ($r > 0.6$). However, channels located posteriorly in the head (O1, O2) show reduced similarity (**Figure 7A**). For Cwet, the grand average was $r = 0.62 \pm 0.25$ and also follows the same trend of reduced similarities for channels located posteriorly in the head (**Figure 7B**). There were 134 scalar products ($>60\%$ of total) showing moderate or high similarity. The results from scalar products confirmed the superior performance of BSM and lower variability of observed event related potentials in comparison to Cwet. Moreover, both systems demonstrated limitations in delivering similar ERP curves from EEG electrodes located posteriorly in the head.

Test-Retest Reliability

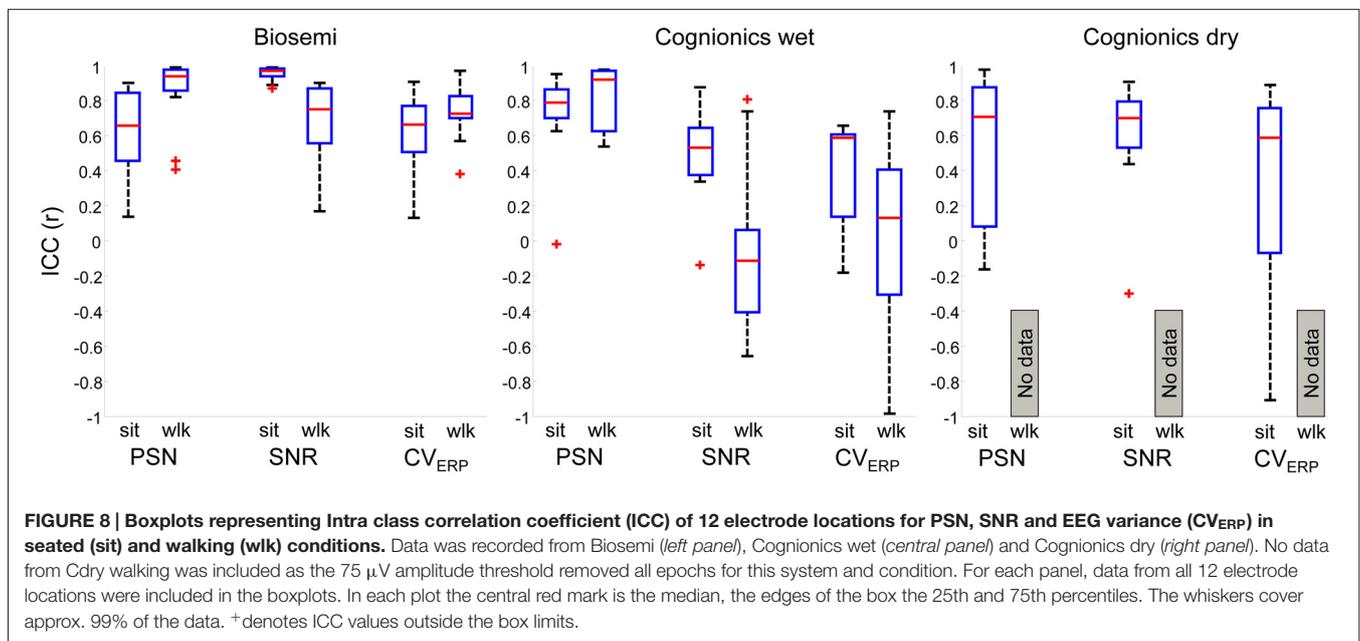
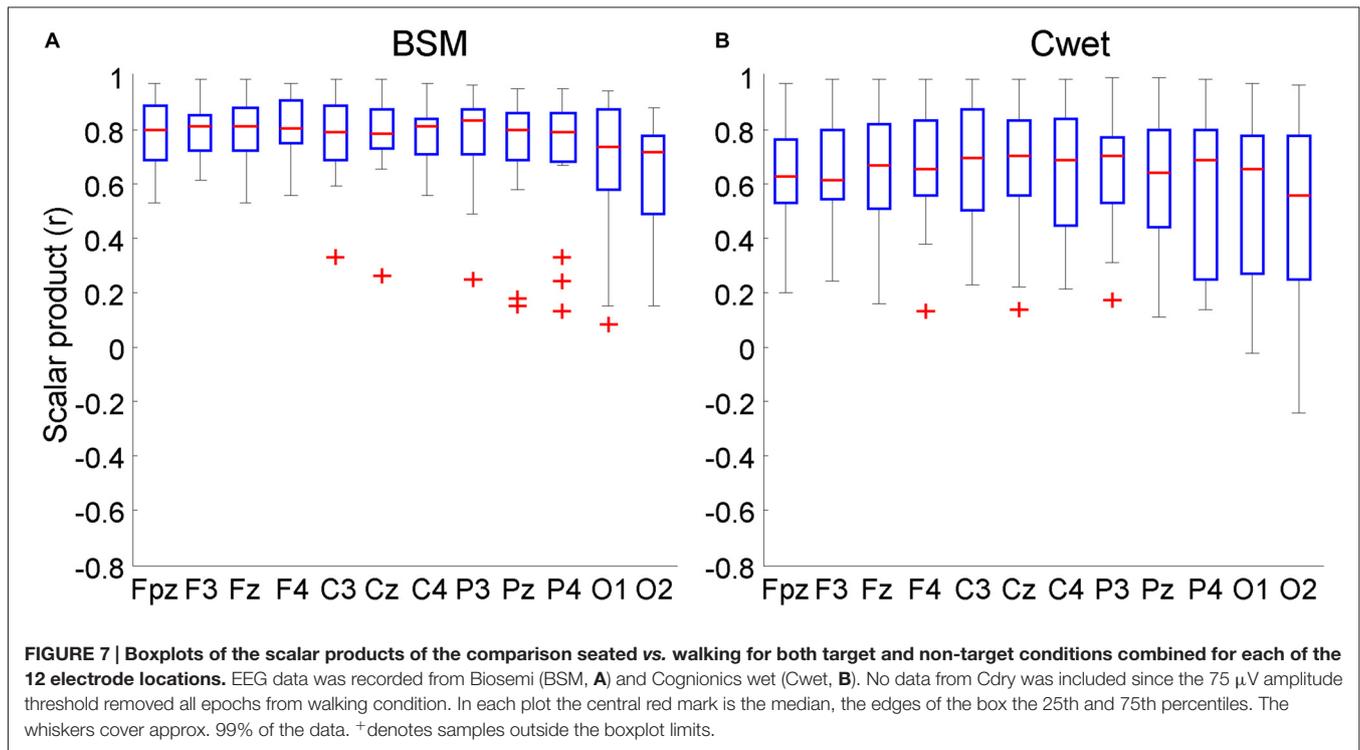
In order to assess overall consistency of the different metrics above across the recorded signal, test-retest reliability was



calculated by ICC. That is, we compared the characteristic derived parameters PSN and CV_{ERP} in different conditions between the first test and the re-test. This does not imply a direct comparison of an ERP waveform. We have found different results depending on the EEG system. BSM showed good median reliability ($\text{ICC} > 0.70$) for all variables across all 12 electrode locations. For PSN and CV_{ERP} the ICC was slightly higher for walking in comparison to the seated condition, whereas the ICC was higher for SNR in the seated condition (Figure 8). The ICC for BSM presented lower variability in comparison to the

two other systems. Therefore, data collected using this system can be considered better reproducible in seated and walking conditions.

The results for Cwet were less consistent. The PSN showed good reliability for both seated and walking conditions, however for both SNR and CV_{ERP} the ICC was reduced. The ICC for SNR was moderate for the seated condition, but for walking it was poor ($\text{ICC} < 0.4$). The same was found for CV_{ERP} . Moreover, the ICC for walking showed large variability for SNR and CV_{ERP} . Thus, the



use of Cwet can provide reproducible results depending on the variable, and also if data is recorded in stationary conditions.

Cdry showed moderate to good median reliability in all three variables in the seated condition, although the variability was high for PSN and CV_{ERP} . The SNR was the most stable variable while showing low variability. We could not evaluate the performance of this system during walking as

the data could not be used for the calculation of such variables.

The ICC calculated from the EEG absolute power was predominantly moderate/good for all EEG systems in seated and walking conditions (Table 2). Exception was found only for the theta band for Cdry seated condition ($ICC < 0.5$). These results suggest that the absolute power is also a reproducible variable from EEG recordings in seated and walking conditions.

TABLE 2 | Intra class correlation coefficient (ICC) of EEG absolute power in seated (sit) and walking (wlk) conditions for each EEG frequency band.

	Intra-class correlation coefficient (ICC) – Absolute power			
	Theta	Alpha	Beta	Gamma
<i>Seated</i>				
BSM	0.63	0.99	0.99	0.49
Cwet	0.76	0.84	0.93	0.68
Cdry	0.47	0.84	0.86	0.82
<i>Walking</i>				
BSM	0.96	0.98	0.98	0.98
Cwet	0.93	0.99	0.99	0.95

Data was recorded from three EEG systems (Biosemi: BSM, Cognionics wet (Cwet) and Cognionics dry (Cdry)). No data from Cdry was included in the walking condition since the 75 μ V amplitude threshold removed all epochs for this system and condition.

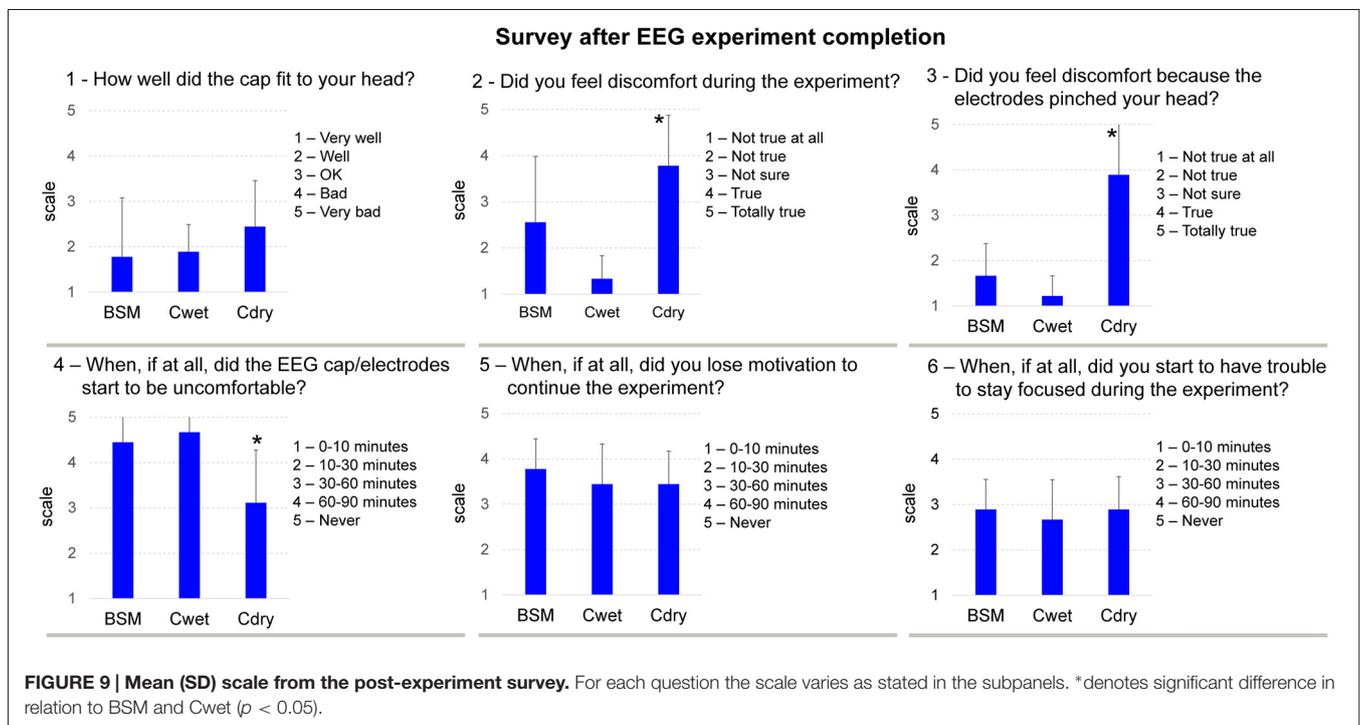
Post-Experiment Survey

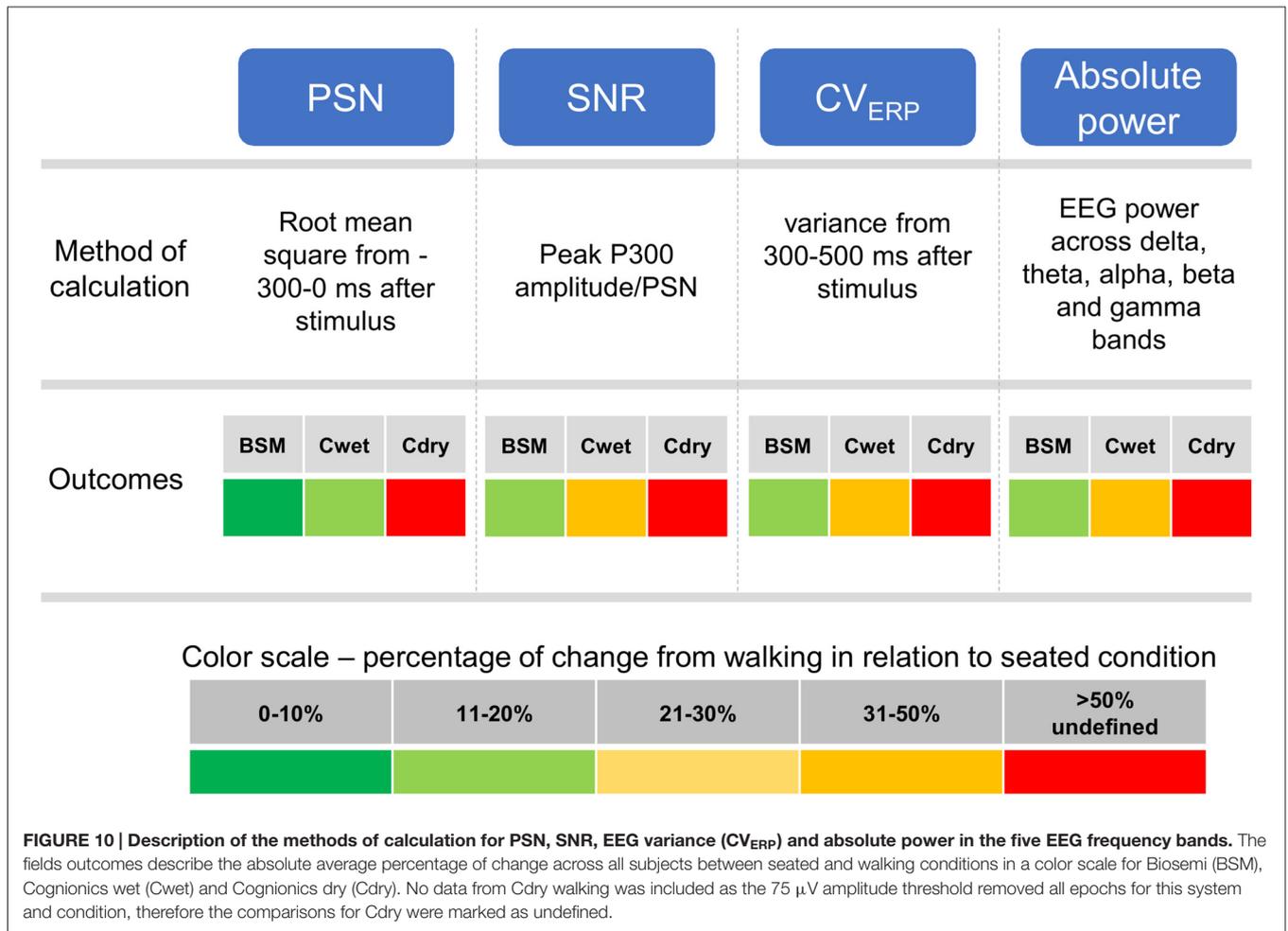
After completing the experiment, most of the subjects reported that the three different EEG systems fit their heads well/very well for the wet systems and well/OK for the dry system (Figure 9, question 1), with no significant differences across systems. For questions 2 and 3, subjects reported significantly higher discomfort using Cdry ($F_{(2,24)} = 11.62$, $p < 0.001$), and this discomfort was caused by the electrodes pinching their heads during the experiment ($F_{(2,24)} = 26.81$, $p < 0.001$) in comparison to the other two wet systems. Moreover, in question 4, subjects reported that they started feeling discomfort using Cdry approximately 30–60 min through the experiment, which was a significantly shorter time in comparison to the other two EEG systems ($F_{(2,24)} = 5.63$, $p < 0.05$). No significant

differences were found in relation to the time subjects started to lose motivation (question 5), but there was a trend to extended motivation while using BSM. In addition, no significant differences across EEG systems were found in relation to the time subjects started to lose focus (question 6).

DISCUSSION

Scalp EEG is currently the main tool for studying brain activity in real-world conditions. There is an increasing number of EEG systems suitable for recordings during whole-body motion, but there is an overall lack of community agreed-upon benchmarks for defining the performance and usability of such EEG systems. In this study, we proposed a standard auditory oddball task and basic metrics that were effective in underpinning differences in the performance and usability of three EEG systems. We first found that EEG epochs from the dry EEG system had substantially higher peak-to-peak amplitudes for both seated and walking conditions. Furthermore, data from the walking condition was mostly unusable due to high rejection rate when using a typical maximum threshold value. Secondly, calculations of PSN, SNR, CV_{ERP} and scalar products revealed differences across systems and/or across conditions (seated vs. walking). These variables showed predominantly moderate/good reliability, especially for the seated condition. Finally, subjects felt less discomfort and were motivated for longer periods while using wet EEG systems in comparison to the dry system. These results combined suggest that the proposed methodology and subsequent analysis were successful in identifying differences across systems that are mostly caused by motion-related artifacts and usability issues.





A short description of the computation of proposed variables and a brief summary of the results from the comparison between seated and walking for each EEG system is shown in **Figure 10**.

Previous investigations have applied a threshold of 75 μ V for cognitive tasks (De Vos et al., 2014; Ries et al., 2014). However, we found that different EEG systems may have widely different numbers of rejected epochs at such specific amplitude thresholds. Although BSM and Cwet had acceptable numbers of rejected epochs (about 15–40% rejection rate), the higher rejection rate for Cdry indicates an inherent source of noise, especially during walking. The use of the usual threshold determined that it was not possible to keep data from Cdry during walking throughout the entire investigation. There was an unexpectedly high rejection rate for both wet systems in the seated condition (~30%), which may be related to the use of 12 electrodes from different head locations containing artifacts not removed by the ICA cleaning. For both wet systems, there was an expected increase in the rejection rate in walking conditions, and both systems showed similar performance. Thus, we suggest that the use of standard amplitude thresholds for defining the rejection rate was successful in discriminating the performance of EEG systems. We speculate that the differences

between seated and walking conditions for Cdry can be explained by the higher susceptibility for changes in contact impedance of the dry electrodes from head motion caused by walking dynamics. Furthermore, a noted difficulty in establishing optimal levels of impedance across all rows of electrodes during subject preparation before the experiment can also be directly related to these high rejection rates. We based our ERP results on the averaging of a similar number of epochs across systems and conditions, and in some subjects we rejected ~50% of all trials' target epochs due to the high rejection rate for Cdry. Previous studies have recommended the average of at least 70 trials for optimal results concerning P300 (Kiesel et al., 2008). Our results follow this recommendation, as the rejection of 50% of target epochs would still result in averaging 80 epochs.

We calculated PSN as a benchmark metric since it can quantify the signal quality (De Vos et al., 2014). Ideally this amplitude fluctuation should be robust across seated and walking conditions, since the motion-related artifacts and the stimulus presentation were not coupled. Although three EEG systems showed similar PSN in the seated condition, only BSM did not show altered values while comparing seated vs. walking. A recent investigation did not find differences in PSN recorded in seated and walking conditions using an adapted Emotiv headset

(De Vos et al., 2014). However, there were differences in the methods used in this investigation and our results. De Vos et al. (2014) reported the PSN only for channel Pz, as representative of all other selected channels, whereas we used all 12 electrode locations in order to assure a robust representation of the EEG data across the scalp. Nonetheless, the use of PSN was effective in discriminating the performance of EEG systems when comparing seated vs. walking conditions.

The ratio W/S showed that different EEG frequency bands can have distinct influences of motion-related artifacts for both wet systems. In addition, the greater ratio W/S for Cwet suggests that power spectrum for this system is more influenced in walking conditions, when compared to BSM. The ratio W/S has the advantage of excluding normalization issues, since the ratio is calculated within the seated and walking EEG power for the same system. Previous investigations showed predominantly similar power spectrum for EEG recordings from wet and dry EEG electrodes in static (Zander et al., 2011; Chi et al., 2012) and dynamic conditions (Lin et al., 2014). However, these studies were conducted using different types of electrodes in simultaneous recordings. Our results for benchmarking focused on comparing entire systems, which can explain potential differences to previous research. As suggested in the “Results” Section, the ratio W/S from specific EEG epochs should ideally be as close to 1 as possible, if there are no major changes in continuous electrocortical activity between seated and walking conditions. Previous studies have shown reduction in alpha and beta power during motor activity (Tzagarakis et al., 2015), which can explain the ratios W/S below 1 for Biosemi for these two frequency bands. However, ratios from Cwet were consistently above 1, suggesting that higher power during walking recordings in the alpha and beta bands may be due to movement artifacts. Although the difference in the ratio W/S is not substantial across systems, this metric was effective in determining changes in the EEG frequency content related to motion.

The SNR was greater for BSM in comparison to all other EEG systems, and the CV_{ERP} for BSM was lower in comparison to all other EEG systems. The SNR is a traditional metric for quantifying the quality of the target ERP event in relation to the baseline noise. EEG is well-known for not presenting optimal SNR (Reis et al., 2014), and ERP events such as the P300 might be highly influenced by changes in EEG amplitude. It is noteworthy to mention that a few subjects have not presented clear P300 events (see subjects 4 and 5 in **Figure 6**) as previously reported elsewhere (Picton, 1992; Polikar et al., 2007). These subjects did not present clear P300 events regardless of the EEG system or condition, highlighting the importance of standardized benchmarking procedures for comparing the performance of EEG technologies. The use of SNR as a benchmark metric showed that motion did not negatively influence the SNR for BSM, whereas it was reduced for Cwet. The reduction in SNR for Cwet during walking can be explained by the greater PSN found in this condition. On the other hand, SNR can be also influenced by the amplitude of the P300 event. We calculated CV_{ERP} in order to quantify the influence of motion on the selected ERP epochs, even though all epochs passed the EEG

amplitude rejection method. CV_{ERP} showed the increase in trial-by-trial variance in walking conditions for both wet systems, especially for Cwet. Our results suggest that the use of SNR alone as a metric for benchmarking may generate questionable results, but the combination of SNR and CV_{ERP} can establish a robust method for evaluating the variability across individual epochs for different EEG systems. This suggestion can be supported by the results from Cdry while seated, as we found a high SNR, but there was also a high CV_{ERP} in the P300 window. The Cdry system showed to be very sensitive to motion, and minimal head motion can add artifacts to the acquired EEG data. Moreover, there might be other sources of artifacts related to dry electrodes not studied in our work. There is an inherent higher variability in the epochs for Cdry, as measured by the CV_{ERP} . Therefore it is not possible to assure that higher SNR is directly related to higher overall signal quality.

Previous studies comparing event-locked EEG data from stationary and dynamic conditions have used metrics such as peak P300 and PSN (Gramann et al., 2010; Debener et al., 2012) calculated from averaged ERP epochs. Our study introduced a comparison of the ERP curves from seated and walking conditions by means of scalar products. This methodology is equivalent to computing the correlation over time and allows for a direct mathematical comparison of the averaged epochs from both conditions, complementing the results already explored in the literature. Previous studies applied correlation analysis to investigate how similar power spectral curves from wet and dry electrodes were shaped, showing that both types of electrodes have similar spectral distribution (Zander et al., 2011; De Sanctis et al., 2014). Ideally, target and non-target curves from sitting and walking should be as similar as possible in a controlled laboratory-based setup, since no distraction other than the treadmill movement and its inherent noise are added while walking. This hypothesis is confirmed by the fact that more than 80% of all comparisons for BSM showed scalar products above 0.6 and out of these 40% being highly similar. Similar analyses for Cwet, showed poorer performance when compared to BSM, suggesting changes in the ERP curve shapes related to whole-body motion. Furthermore, channels located in the frontal and central areas of the head presented higher similarity to the target curves in comparison to those located in the posterior area (**Figures 7A,B**). These results suggest that the evaluation of the entire ERP curve can be used for establishing overall signal quality between different recording conditions.

In addition to identifying appropriate metrics for benchmarking EEG systems, we tested the potential reliability of such metrics. Overall, a general statement of which are the best metrics in terms of reliability is difficult to achieve, as different systems demonstrated different performances in each variable. Nonetheless, BSM continued to show moderate/good ICCs for all variables in both seated and walking conditions, whereas Cwet showed reduction in the reliability for SNR and CV_{ERP} , especially in walking conditions. Previous investigations have described good test-retest reliability of evoked EEG responses from transcranial magnetic stimulation (Lioumis et al., 2009), and auditory responses using ICC (Brunner et al., 2013) and Pearson correlation (McFadden et al., 2014). All these previous

investigations were conducted in seated conditions, and to our knowledge, the present investigation is the first to report test-retest reliability for EEG experiments in dynamic conditions. Although there was a limited sample size for test-retest reliability, our robust results suggest that wet systems deliver reliable results between sessions in seated conditions, but reliability is reduced in walking conditions.

Our results from the post-experiment survey showed that subjects felt discomfort while using Cdry, and this discomfort was caused by the electrode pressing against the scalp and pinching their heads. In fact, five subjects reported headaches after the experimental session, which can be related to the pressure of the dry electrodes onto the head. EEG experiments have shown that headache and discomfort can cause a reduction in cognitive performance (Evers et al., 1997; Lorenz and Bromm, 1997). These results are essential for EEG system developers because discomfort can limit cognitive performance and change the expected electrocortical signals in real-world conditions. The comfort rating of the Cdry in our experiment does not corroborate previous investigations that reported good comfort rates for dry EEG technologies (Hairston, 2012; Chen et al., 2014; Fiedler et al., 2015). However, these investigations were assessing comfort using different scales, from different number of electrodes and EEG sensors. The cap placement and preparation for the wet systems demanded more time in comparison to the dry system. Nonetheless, subjects did not report problems regarding the long procedure. Further studies and standardization of methods for accessing comfort are needed in order to establish optimal methodologies for benchmarking the usability of EEG systems. Additionally, our study has a limitation concerning number of subjects, especially for the test-retest reliability. We found robust ICC especially for the wet systems, which suggests that the proposed metrics are

reliable. However future studies aiming at developing these benchmarking testing protocols should consider using larger subject populations to gather more robust information about the different methodologies.

In summary, we propose the auditory oddball paradigm in combination with the calculations of epoch rejection rate, PSN, SNR, CV_{ERP} and scalar products can be used as a benchmark method for testing the performance of different EEG systems. Furthermore, we observed that a walking vs. seated paradigm elicits strong differences in these metrics. The use of wet systems has demonstrated the potential of acquiring EEG data with comparable quality in seated and walking conditions. The investigated metrics were successful in underpinning which systems were suitable to deliver comparable EEG data in whole-body motion, and the robustness of such results were attested by the assessment of test-retest reliability. Finally, EEG systems based on dry electrodes may need substantial improvement to meet the quality standards of wet electrodes. At the same time, improvements on the electrode construction and materials is needed, since comfort and performance may be compromised.

AUTHOR CONTRIBUTIONS

ASO, BRS, WDH, PK and DPF designed the experiment. ASO and BRS performed the experiments. ASO, BRS, WDH, PK and DPF analyzed and interpreted the data. ASO, BRS, and DPF drafted the manuscript and all authors approved the final version.

FUNDING

This work was supported in part by the US Army Research Laboratory ARL-CTA (W911NF-09-1-0139 and W911NF-10-2-0022).

REFERENCES

- Brunner, J. F., Hansen, T. I., Olsen, A., Skandsen, T., Håberg, A., and Kropotov, J. (2013). Long-term test-retest reliability of the P3 NoGo wave and two independent components decomposed from the P3 NoGo wave in a visual Go/NoGo task. *Int. J. Psychophysiol.* 89, 106–114. doi: 10.1016/j.ijpsycho.2013.06.005
- Casson, A., Yates, D., Smith, S., Duncan, J., and Rodriguez-Villegas, E. (2010). Wearable electroencephalography. What is it, why is it needed and what does it entail? *IEEE Eng. Med. Biol. Mag.* 29, 44–56. doi: 10.1109/MEMB.2010.936545
- Castermans, T., Duvinage, M., Cheron, G., and Dutoit, T. (2014). About the cortical origin of the low-delta and high-gamma rhythms observed in EEG signals during treadmill walking. *Neurosci. Lett.* 561, 166–170. doi: 10.1016/j.neulet.2013.12.059
- Chen, Y.-H., Op de Beeck, M., Vanderheyden, L., Carrette, E., Mihajlović, V., Vanstreels, K., et al. (2014). Soft, comfortable polymer dry electrodes for high quality ECG and EEG recording. *Sensors* 14, 23758–23780. doi: 10.3390/s141223758
- Chi, Y. M., Jung, T., and Cauwenberghs, G. (2010). Dry-contact and noncontact biopotential electrodes: methodological review. *IEEE Rev. Biomed. Eng.* 3, 106–119. doi: 10.1109/RBME.2010.2084078
- Chi, Y. M., Wang, Y. T., Wang, Y., Maier, C., Jung, T. P., and Cauwenberghs, G. (2012). Dry and noncontact EEG sensors for mobile brain-computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* 20, 228–235. doi: 10.1109/TNSRE.2011.2174652
- d'Avella, A., Saltiel, P., and Bizzi, E. (2003). Combinations of muscle synergies in the construction of a natural motor behavior. *Nat. Neurosci.* 6, 300–308. doi: 10.1038/nn1010
- De Sanctis, P., Butler, J. S., Malcolm, B. R., and Foxe, J. J. (2014). Recalibration of inhibitory control systems during walking-related dual-task interference: a mobile brain-body imaging (MOBI) study. *Neuroimage* 94, 55–64. doi: 10.1016/j.neuroimage.2014.03.016
- De Vos, M., Gandras, K., and Debener, S. (2014). Towards a truly mobile auditory brain-computer interface: exploring the P300 to take away. *Int. J. Psychophysiol.* 91, 46–53. doi: 10.1016/j.ijpsycho.2013.08.010
- Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small and wireless EEG for a walk? *Psychophysiology* 49, 1617–1621. doi: 10.1111/j.1469-8986.2012.01471.x
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Ehinger, B., Fischer, P., Gert, A., Klauhold, L., Weber, F., Pipa, G., et al. (2014). Kinesthetic and vestibular information modulate alpha activity during spatial navigation: a mobile EEG study. *Front. Hum. Neurosci.* 8:71. doi: 10.3389/fnhum.2014.00071
- Evers, S., Bauer, B., Suhr, B., Husstedt, I. W., and Grotemeyer, K. H. (1997). Cognitive processing in primary headache: a study on event-related potentials. *Neurology* 48, 108–113. doi: 10.1212/wnl.48.1.108
- Fiedler, P., Pedrosa, P., Griebel, S., Fonseca, C., Vaz, F., Supriyanto, E., et al. (2015). Novel multipin electrode cap system for dry electroencephalography. *Brain Topogr.* 28, 647–656. doi: 10.1007/s10548-015-0435-5

- Gramann, K., Ferris, D. P., Gwin, J., and Makeig, S. (2014). Imaging natural cognition in action. *Int. J. Psychophysiol.* 91, 22–29. doi: 10.1016/j.ijpsycho.2013.09.003
- Gramann, K., Gwin, J. T., Bigdely-Shamlo, N., Ferris, D. P., and Makeig, S. (2010). Visual evoked responses during standing and walking. *Front. Hum. Neurosci.* 4:202. doi: 10.3389/fnhum.2010.00202
- Grozea, C., Voinescu, C. D., and Fazli, S. (2011). Bristle-sensors–low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. *J. Neural Eng.* 8:025008. doi: 10.1088/1741-2560/8/2/025008
- Gwin, J. T., Gramann, K., Makeig, S., and Ferris, D. P. (2010). Removal of movement artifact from high-density EEG recorded during walking and running. *J. Neurophysiol.* 103, 3526–3534. doi: 10.1152/jn.00105.2010
- Hairston, W. D. (2012). *Accounting for Timing Drift and Variability in Contemporary Electroencephalography (EEG) systems. Report, ARL-TR-5945.* Aberdeen Proving Ground, MD: US Army Research Laboratory.
- Hairston, W. D., Whitaker, K. W., Ries, A. J., Vettel, J. M., Bradford, J. C., Kerick, S. E., et al. (2014). Usability of four commercially-oriented EEG systems. *J. Neural Eng.* 11:046018. doi: 10.1088/1741-2560/11/4/046018
- Kiesel, A., Miller, J., Jolicœur, P., and Brisson, B. (2008). Measurement of ERP latency differences: a comparison of single-participant and jackknife-based scoring methods. *Psychophysiology* 45, 250–274. doi: 10.1111/j.1469-8986.2007.00618.x
- Kranczoch, C., Zich, C., Schierholz, I., and Sterr, A. (2014). Mobile EEG and its potential to promote the theory and application of imagery-based motor rehabilitation. *Int. J. Psychophysiol.* 91, 10–15. doi: 10.1016/j.ijpsycho.2013.10.004
- Lance, B. J., Kerick, S. E., Ries, A. J., Oie, K. S., and McDowell, K. (2012). Brain-computer interface technologies in the coming decades. *Proc. IEEE* 100, 1585–1599. doi: 10.1109/jproc.2012.2184830
- Lew, H. L., Gray, M., and Poole, J. H. (2007). Temporal stability of auditory event-related potentials in healthy individuals and patients with traumatic brain injury. *J. Clin. Neurophysiol.* 24, 392–397. doi: 10.1097/wnp.0b013e31814a56e3
- Liao, L. D., Chen, C. Y., Wang, I. J., Chen, S. F., Li, S. Y., Chen, B. W., et al. (2012). Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors. *J. Neuroeng. Rehabil.* 9:5. doi: 10.1186/1743-0003-9-5
- Lin, Y.-P., Wang, Y., Wei, C.-S., and Jung, T.-P. (2014). Assessing the quality of steady-state visual-evoked potentials for moving humans using a mobile electroencephalogram headset. *Front. Hum. Neurosci.* 8:182. doi: 10.3389/fnhum.2014.00182
- Lioumis, P., Kikić, D., Savolainen, P., Mäkelä, J. P., and Kähkönen, S. (2009). Reproducibility of TMS - Evoked EEG responses. *Hum. Brain Mapp.* 30, 1387–1396. doi: 10.1002/hbm.20608
- Lopez-Calderon, J., and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* 8:213. doi: 10.3389/fnhum.2014.00213
- Lopez-Gordo, M. A., Sanchez-Morillo, D., and Pelayo Valle, F. (2014). Dry EEG electrodes. *Sensors* 14, 12847–12870. doi: 10.3390/s140712847
- Lorenz, J., and Bromm, B. (1997). Event-related potential correlates of interference between cognitive performance and tonic experimental pain. *Psychophysiology* 34, 436–445. doi: 10.1111/j.1469-8986.1997.tb02387.x
- McDowell, K., Lin, C.-T., Oie, K. S., Jung, T.-P., Gordon, S., Whitaker, K. W., et al. (2013). Real-world neuroimaging technologies. *IEEE Access* 1, 131–149. doi: 10.1109/access.2013.2260791
- McFadden, K. L., Steinmetz, S. E., Carroll, A. M., Simon, S. T., Wallace, A., and Rojas, D. C. (2014). Test-retest reliability of the 40 Hz EEG auditory steady-state response. *PLoS One* 9:e85748. doi: 10.1371/journal.pone.0085748
- Mihajlovic, V., Grundlehner, B., Vullers, R., and Penders, J. (2015). Wearable, wireless EEG solutions in daily life applications: what are we missing? *IEEE J. Biomed. Health Inform.* 19, 6–21. doi: 10.1109/JBHI.2014.2328317
- Oliveira, A. S., Schlink, B. R., Hairston, W. D., König, P., and Ferris, D. P. (2016). Induction and separation of motion artifacts in EEG data using a mobile phantom head device. *J. Neural Eng.* 13:036014. doi: 10.1088/1741-2560/13/3/036014
- Onikura, K., Katayama, Y., and Iramina, K. (2015). Evaluation of a method of removing head movement artifact from EEG by independent component analysis and filtering. *Adv. Biomed. Eng.* 4, 67–72. doi: 10.14326/abe.4.67
- Picton, T. W. (1992). The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456–479. doi: 10.1097/00004691-199210000-00002
- Polikar, R., Topalis, A., Green, D., Kounios, J., and Clark, C. M. (2007). Comparative multiresolution wavelet analysis of ERP spectral bands using an ensemble of classifiers approach for early diagnosis of Alzheimer's disease. *Comput. Biol. Med.* 37, 542–558. doi: 10.1016/j.compbiomed.2006.08.012
- Reis, P. M. R., Hebenstreit, F., Gabsteiger, F., von Tscharnner, V., and Lochmann, M. (2014). Methodological aspects of EEG and body dynamics measurements during motion. *Front. Hum. Neurosci.* 8:156. doi: 10.3389/fnhum.2014.00156
- Ries, A. J., Touryan, J., Vettel, J., McDowell, K., and Hairston, W. D. (2014). A comparison of electroencephalography signals acquired from conventional and mobile systems. *J. Neurosci. Neuroeng.* 3, 10–20. doi: 10.1166/jnsne.2014.1092
- Tanner, D., Morgan-Short, K., and Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology* 52, 997–1009. doi: 10.1111/psyp.12437
- Tzagarakis, C., West, S., and Pellizzer, G. (2015). Brain oscillatory activity during motor preparation: effect of directional uncertainty on beta, but not alpha, frequency band. *Front. Neurosci.* 9:246. doi: 10.3389/fnins.2015.00246
- Zander, T. O., Lehne, M., Ihme, K., Jatzew, S., Correia, J., Kothe, C., et al. (2011). A dry EEG-system for scientific research and brain-computer interfaces. *Front. Neurosci.* 5:53. doi: 10.3389/fnins.2011.00053

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Oliveira, Schlink, Hairston, König and Ferris. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.