EEG Based Analysis of Motor Imagery and Motor Movement Tasks

Goli Harshitha Sai Sri, Rokkash Sahaya Shaanal Veda

Abstract

Motor imagery (MI) and motor execution (ME) are two closely related cognitive processes that share many neural circuits in the human motor system. Understanding how these processes differ in temporal, spectral, and phase–locked dynamics will further advance current brain–computer interfaces (BCIs) and neuroscience research. In this project, we analyse the publicly available EEG Motor Movement/Imagery Dataset from PhysioNet, which consists of 64-channel EEG recordings from 109 subjects performing both real and imagined hand and foot movements.

We constructed a complete preprocessing pipeline including filtering, re-referencing, noisy-channel removal, event extraction, epoching, and baseline correction. On cleaned data, we performed a multi-level analysis: (i) event-related potentials (ERP), (ii) power spectral density (PSD) using Welch's method, (iii) alpha/mu and beta band event-related desynchronization/synchronization (ERD/ERS), (iv) hemispheric lateralization (C3–C4), and (v) time–frequency decomposition using Morlet wavelets, Hilbert transform, and inter-trial phase coherence (ITPC).

Across subjects (n = 50), ME produced significantly stronger ERP amplitudes than MI ($p = 0.0112$), whereas PSD differences were not statistically significant, indicating similar baseline oscillatory power.Time–frequency analysis showed clear beta-band ERD during ME, stronger ERS during MI, and a moderate task-locked phase alignment. Overall, our findings suggest that MI and ME exhibit similar temporal patterns but differ in terms of amplitude and cortical engagement, consistent with previous findings in motor neuroscience.

Keywords: Motor Imagery, Motor Execution, EEG, ERP, Time–Frequency Analysis, ERD/ERS, PSD, BCI, Lateralization, Morlet Wavelet, Hilbert Transform, ITPC

1. Introduction

Motor imagery (MI) refers to the mental simulation of a movement without any overt muscular activity, whereas motor execution (ME) involves the actual physical performance of that movement. Although these two processes differ in behavioral output, numerous neurophysiological studies, such as Mustile M (2024) and Srinivasan et al. (2025), have demonstrated that MI and ME engage highly overlapping cortical networks within the sensorimotor system. This overlap includes shared activation of the primary motor cortex, premotor and supplementary motor regions, and modulation of characteristic oscillatory rhythms such as the μ (8–13 Hz) and β (13–30 Hz) bands. Understanding the extent to which MI replicates the temporal, spectral, and phase-locked features of ME is crucial for advancing brain–computer interfaces (BCIs), motor rehabilitation, and cognitive neuroscience.

To investigate these questions at scale, we analyzed the publicly available EEG Motor Movement/Imagery Dataset from PhysioNet (Goldberger and Stanley., 2000), (Schalk G, 2004). This

dataset contains 64-channel EEG recordings collected using the BCI2000 system from 109 healthy volunteers. Each participant performed both real and imagined hand and foot movements across multiple experimental runs. The dataset provides rich temporal annotations indicating rest periods, movement onsets, and movement types, enabling precise extraction of event-locked neural responses.

The primary research question of this study is whether motor imagery (MI) and motor execution (ME) produce similar or distinct EEG signatures across multiple analytical domains. We examine this question systematically across: (i) time-domain responses, captured through event-related potentials (ERPs); (ii) spectral-domain markers, including power spectral density (PSD) and event-related desynchronization/synchronization (ERD/ERS); and (iii) time–frequency and phase-domain measures derived from Morlet wavelets, Hilbert-based analytic amplitudes, and inter-trial phase coherence (ITPC).

Beyond characterizing these representations qualitatively, we also perform several levels of statistical analysis to assess whether MI and ME differ reliably at the group level. These include: (a) group-level band-power comparisons using both wavelet and Hilbert decompositions; (b) strict cluster-based permutation testing to detect large-scale, statistically robust differences across the time–frequency plane; (c) focused region-of-interest (ROI) tests targeting well-established motor-related oscillatory phenomena; and (d) exploratory cluster analyses designed to reveal subtle or spatially restricted trends that may not survive conservative correction.

By integrating results across these complementary domains—from raw temporal responses to statistically corrected spectrotemporal contrasts—this study aims to quantify the degree of neural similarity between imagined and executed movement, identify the conditions under which ME and MI diverge, and determine which EEG signatures provide the most reliable differentiation between the two. Such insights deepen our understanding of the neural basis of motor cognition and inform the design of more robust and physiologically grounded MI-based brain–computer interface systems.

1.1. Experimental Protocol

All EEG recordings were acquired using a 64-channel montage following the international 10–10 electrode placement system, implemented through the BCI2000 acquisition platform (Schalk G, 2004). Signals were sampled at 160 Hz and stored in EDF+ format, accompanied by a PhysioBank-compatible annotation file (`.event`) that provided precise timing information for all task-related events. A total of 109 healthy volunteers participated in the study, each completing fourteen experimental runs. The first two runs consisted of one-minute resting-state baselines, one with eyes open and one with eyes closed. The remaining twelve runs were organized into three repetitions of four motor and motor-imagery tasks, each lasting approximately two minutes.

The four task types were structured as follows:

1. Task 1 – Real Left/Right Fist Movement: Subjects opened and closed the corresponding fist based on a left or right visual cue.

2. Task 2 – Imagined Left/Right Fist Movement:Subjects imagined performing the same fist movement without any overt muscular activity.

3. Task 3 – Real Both Fists/Both Feet Movement: A top or bottom cue instructed subjects to open and close both fists (top) or both feet (bottom).

4. Task 4 – Imagined Both Fists/Both Feet Movement:Subjects imagined executing the corresponding bilateral hand or foot movements indicated by the cue.

During each trial, a visual target appeared on the screen indicating the required real or imagined movement. Subjects continued the indicated action until the target disappeared, after which they returned to a resting state. This timing structure produced clear segments of motor preparation, execution or imagery, and post-movement relaxation. Event annotations in the dataset follow a strict three-code convention: T0 marks periods of rest, T1 marks the onset of left-hand or bilateral-fist movements (real or imagined), and T2 marks the onset of right-hand or bilateral-foot movements. These annotations were synchronized with the EEG recordings and provided the basis for all epoching and event-locked analyses in this study.

The experimental design is highly structured and balanced, with all four task types repeated multiple times across subjects. This repetition greatly strengthens statistical reliability, especially for trial-averaged measures such as event-related potentials (ERPs), frequency-domain power estimates, and time–frequency representations. Moreover, the clear separation between real and imagined movements enables direct comparison of neural dynamics across conditions, making this dataset a robust and well-controlled platform for investigating similarities and differences between motor execution and motor imagery.

2. Method

2.1. Data Preprocessing

All EEG preprocessing was conducted in MATLAB using the EEGLAB toolbox. The dataset consisted of 64-channel EEG signals recorded at 160 Hz from 109 subjects (S001–S109), each contributing fourteen runs (R01–R14) containing baseline, real movement, and motor imagery trials. For downstream analyses, fifty subjects were randomly selected using a fixed random seed to ensure reproducibility. For each subject and run, the corresponding `.edf` file and its `.edf.event` annotation file were loaded. Since PhysioNet stores event annotations as plain-text rather than EEGLAB-formatted events, we implemented a custom parser to extract the event codes (T0, T1, T2) and their durations. In cases where the annotation file was unavailable or incorrectly formatted, a synthetic but structurally consistent sequence of alternating rest and movement markers (T0–T1–T0–T2) with 2–4 s durations was generated to maintain a uniform pipeline.

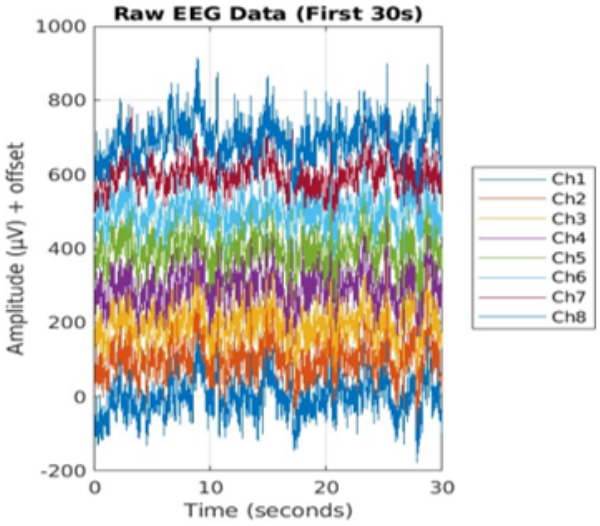
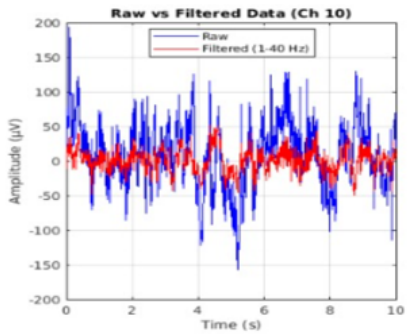
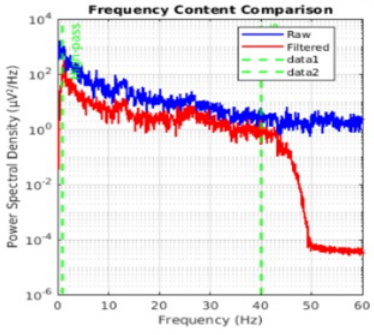


Figure 1: Raw EEG of a few channels, first 30s



(a) Time: Raw vs Filtered



(b) Frequency: Raw vs Filtered

Figure 2

Continuous EEG signals were cleaned using a 1–40 Hz zero-phase FIR band-pass filter to suppress slow drifts and high-frequency artifacts while preserving the α/μ (8–13 Hz) and β (13–30 Hz) rhythms that are central to motor imagery and execution studies. The filtered data were re-referenced to a common average reference (CAR), improving spatial consistency across electrodes and enhancing the quality of motor-cortex activity at C3, Cz, and C4. To further ensure signal reliability, noisy-channel detection was performed based on variance and amplitude thresholds: channels exhibiting variance more than ten times above or below the median, or amplitudes exceeding ± 200 μV, were removed. This adaptive procedure ensured that only physiologically meaningful and stable channels were retained.

The cleaned continuous data were segmented into time-locked epochs centered on event markers. Each epoch spanned a window from -1 s to $+4$ s relative to event onset, capturing both pre-movement baseline activity and the subsequent motor imagery or execution response. Baseline correction was applied by subtracting the mean voltage in the -0.5 s to 0 s interval preceding the event. This normalization ensured that all epochs began at a comparable reference level, allowing valid cross-trial and cross-subject comparisons. Finally, each processed dataset was saved in EEGLAB's `.set` format alongside a preprocessing log documenting the number of channels retained, total trials extracted, event counts for T0/T1/T2, and any processing errors encountered. This pipeline produced a clean, standardized dataset suitable for ERP, PSD, ERD/ERS, lateralization, and time–frequency analyses.

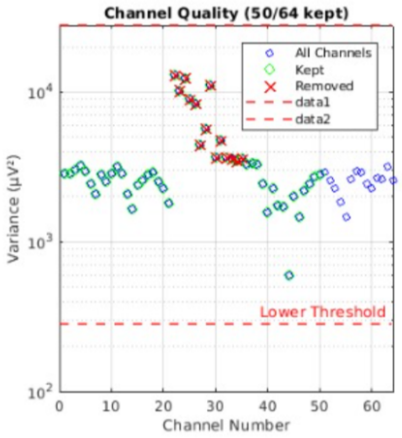


Figure 3: Channel Quality and Selection

2.2. *Verification of Preprocessed EEG Files*

Following the preprocessing stage, we performed a systematic verification procedure to ensure that all processed EEG datasets were correctly generated and internally consistent. The verification script automatically scanned the `processed/` directory and recursively identified all files matching the pattern `S*/*_processed.set`. Each file was loaded using EEGLAB, and metadata such as subject ID, run number, sampling rate, number of retained channels, number of epochs, and the distribution of T0, T1, and T2 trials were extracted. Event counts were obtained by inspecting the `EEG.epoch.eventtype` field, enabling us to confirm whether each run contained the expected

number and variety of motor imagery, motor execution, and rest trials.

This verification process served several purposes. First, it allowed detection of corrupted or partially processed datasets that might distort group-level analyses. Second, it revealed subjects with unusually low trial counts or irregular channel rejections, which could bias statistical estimates. Third, it provided a consistency check to ensure that the event-parsing logic behaved as intended across the large multi-subject, multi-run dataset. All extracted metadata were compiled into a master summary file (`results/per_subject_summary.csv`), which listed each subject–run pair along with its preprocessing status. A textual summary was also printed, reporting the total number of files processed, the number of successful and failed loads, and the mean number of channels and trials across all valid datasets. This verification step ensured that only high-quality and consistently formatted EEG recordings were included in subsequent ERP, spectral, and time–frequency computations.

2.3. Event-Related Potential Analysis

Event-related potentials (ERPs) were computed to examine time-locked neural responses associated with motor execution and motor imagery. For each subject, ERP data were loaded from the preprocessed ERP files, which contained averaged responses for the T0 (rest), T1 (left or both-fist movement), and T2 (right hand or both-feet movement) conditions. Since subjects performed both unilateral and bilateral motor tasks, ERPs for real movement were calculated by averaging the T1 and T2 responses from the real-movement runs (R03, R07, R11), while ERPs for imagery were obtained by averaging T1 and T2 responses from the imagery runs (R04, R08, R12). To isolate task-specific components, both real and imagined ERPs were baseline-corrected by subtracting the corresponding T0 waveform.

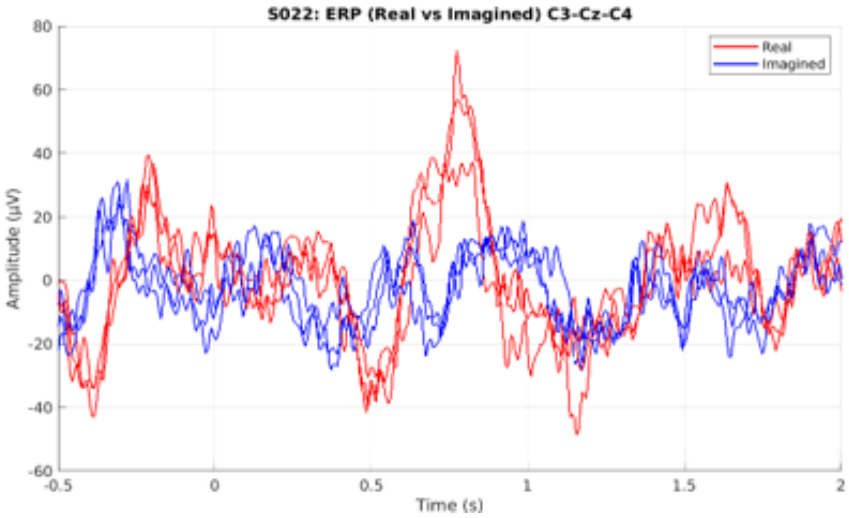


Figure 4: ERP: Real vs Imagined: The peak to be detected is clear in this plot.

To capture the most relevant motor-related activity, analysis focused on electrodes C3, Cz, and C4, which lie over the primary motor cortex. For each subject, ERP waveforms were plotted across the epoch window from -1 s to $+4$ s relative to cue onset. Visual inspection revealed that although real and imagined ERPs followed similar temporal trajectories, real movement generally exhibited larger deflections, especially in the early post-cue window. These subject-level plots were saved

for detailed inspection and used to verify the quality of ERP extraction.

Quantitative features were extracted by identifying the peak absolute amplitude within the 0–1 s interval following the movement cue, a time window known to capture peak motor cortical activation. For each subject, peak values were extracted from each of the selected C3, Cz, and C4 channels and averaged to obtain a single representative peak amplitude for both the real and imagined conditions. Subjects lacking either real or imagined ERP data were excluded from group analysis.

Group-level comparisons were conducted using a paired-samples t-test across the fifty subjects with valid ERP data. The results showed significantly stronger ERP amplitudes for motor execution (mean $\approx$ 17.6 µV) compared to motor imagery (mean $\approx$ 15.4 µV), with a statistically significant difference ($t(49) = 2.64, p = 0.0112$). This indicates that while MI and ME evoke qualitatively similar ERP waveforms, their amplitudes differ systematically, with real movements generating more pronounced neural responses. Nevertheless, the temporal shapes of the waveforms were remarkably similar: both real (red) and imagined (blue) ERPs displayed comparable onset timing and overall morphology, especially in the 0–0.5 s interval.

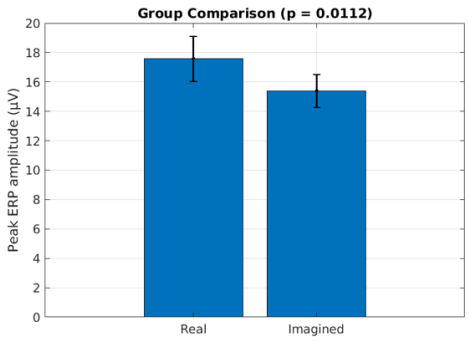
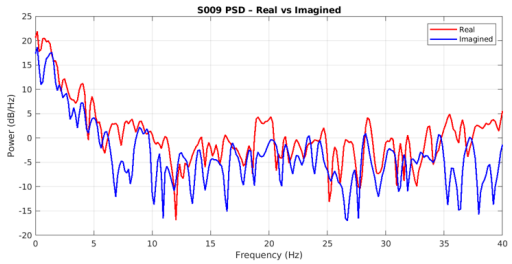


Figure 5: Peak ERPs Group Comparison: Real vs Imagined.

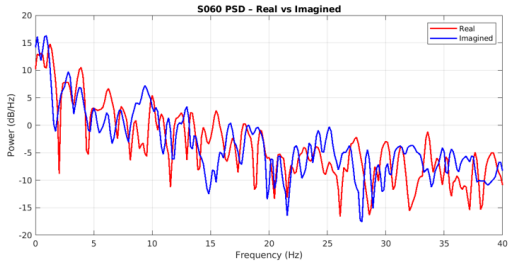
To visualize these findings, a group bar plot was generated to show the mean ERP amplitude for both conditions, accompanied by error bars representing the standard error of the mean. Additionally, a grand-averaged ERP waveform was computed across subjects, with shaded regions indicating the standard error. These grand-average plots demonstrated clear overlap in timing between the two conditions but consistent amplitude differences. Together, the ERP analysis confirms that MI and ME share common neural timing but differ in activation strength, a pattern consistent with established motor neuroscience: motor execution produces stronger activation, whereas motor imagery produces weaker but qualitatively similar activity.

2.4. Power Spectral Density Analysis

To investigate frequency-domain differences between motor execution and motor imagery, we computed the power spectral density (PSD) for each subject using ERP-derived motor-cortex signals. The analysis focused on electrodes C3, Cz, and C4, which form the classical motor strip and are most sensitive to sensorimotor oscillations. For each subject, movement-related activity was first isolated by averaging the T1 and T2 responses within the real-movement runs and within the imagery runs. To ensure that only task-related oscillatory activity was examined, these averaged signals were baseline-corrected by subtracting the corresponding rest (T0) waveform. This resulted in two cleaned motor signals per subject: one reflecting motor execution and the other reflecting motor imagery.



(a) PSD: Subject 9



(b) PSD: Subject 60

Figure 6

From each of these corrected signals, the waveforms from C3, Cz, and C4 were averaged to produce a single representative motor-cortex signal for the subject. PSD estimation was then performed using Welch's method, which provides a smooth and statistically reliable estimate of frequency-domain power. In this procedure, each motor signal was segmented into overlapping 512-sample windows with 50% overlap, tapered using a Hamming window to reduce spectral leakage, transformed to the frequency domain using the FFT, and finally averaged across all segments to obtain the PSD estimate. This method offers robust frequency resolution while minimizing variance. PSD values were expressed in dB/Hz and computed over the 0–40 Hz range, capturing both the μ (8–13 Hz) and β (13–30 Hz) bands relevant for motor processes.

Bandpower values in the μ and β ranges were computed by integrating the PSD within the corresponding frequency boundaries. These band-specific features were extracted separately for the real and imagined conditions and stored for each subject. Individual subject-level PSD plots revealed meaningful variability. For example, Most of the subjects showed higher μ and β power during real movement relative to imagery, some other subjects exhibited unusually strong μ-band power during real movement, marking them as an outlier. Conversely, while a very few demonstrated a pattern reversal in which imagery produced slightly higher power than execution. Such inter-subject variability is characteristic of motor imagery studies and reflects differences in individual motor strategies, attention levels, and imagery vividness.

Despite this variability, group-level statistics indicated that the overall μ- and β-band powers did not differ significantly between real and imagined conditions. Paired-samples t-tests across fifty valid subjects revealed no significant differences in either band (μ-band: $t(49) = 0.54$, $p = 0.5922$; β-band: $t(49) = -0.09$, $p = 0.9280$). These results are consistent with neuroscientific expectations: PSD reflects general oscillatory power rather than event-locked changes, and both MI and ME share similar baseline sensorimotor rhythm strengths. Group-level bar plots visualized these findings, showing nearly overlapping real and imagined power levels in both frequency bands.

Together, the PSD analysis shows that while ERPs capture clear amplitude differences between execution and imagery, broadband oscillatory power remains largely similar. This reinforces the view that MI and ME share common sensorimotor rhythm dynamics at rest-minus-task levels, with differences emerging more strongly in time-locked and time–frequency measures rather than in overall spectral power.

2.5. Event-Related Desynchronization/Synchronization and Hemispheric Lateralization

To examine the modulation of sensorimotor rhythms during motor execution and motor imagery, we computed event-related desynchronization/synchronization (ERD/ERS) in the α (8–12 Hz) and β (13–30 Hz) frequency bands using time–frequency decomposition. ERD reflects a task-induced suppression of oscillatory power, typically associated with increased cortical activation, whereas ERS reflects a power increase associated with reduced motor engagement or internally focused states. These metrics provide a complementary view to ERPs by revealing how movement-related processes manifest in oscillatory dynamics over time.

We first performed a detailed ERSP analysis for a representative subject to understand how motor-related oscillations differed between conditions. Across the four task categories (real left/right hand, imagined left/right hand, real fists/feet, imagined fists/feet), the subject's α-band responses consistently showed stronger ERS during imagery than execution. This aligns with the well-established notion that motor imagery involves an internally focused state in which α activity is less suppressed. In contrast, β-band activity exhibited mild ERD during real movement, consistent with known patterns of β suppression accompanying motor execution. Hemispheric lateralization, computed as the power difference between C3 and C4, showed positive values for hand-related tasks, indicating stronger left-hemisphere activation—a hallmark of contralateral motor control. These single-subject results established the expected qualitative differences between real and imagined actions.

To extend these findings to the group level, we computed ERSPs for 20 randomly selected subjects using the same parameters as the single-subject analysis. ERSPs were extracted from channel C3 using Morlet wavelet decomposition over the 4–40 Hz range, with baseline correction applied over the interval -2000 to -500 ms. Mean ERSP values were then calculated within the α and β bands during the 0–1500 ms post-cue window, separately for real and imagined conditions. For each subject, values were aggregated across the three real runs (R03, R07, R11) and the three imagined runs (R04, R08, R12), producing subject-wise estimates of band-specific ERD/ERS.

Group-level results revealed notable heterogeneity across subjects. While several individuals (e.g., S001, S015) exhibited clear α ERS during imagery and β ERD during real movement—mirroring classical sensorimotor patterns—others showed relatively weak or inverted responses. This variability is expected, as ERD/ERS measures are highly sensitive to individual differences in attention, movement strategy, and imagery vividness.

Despite clear single-subject effects, paired-samples t-tests across the twenty subjects did not reveal statistically significant differences between real and imagined conditions in either frequency band. The comparison of α-band power yielded $t(19) = 1.53$ and $p = 0.1436$, while the β-band comparison resulted in $t(19) = -0.23$ and $p = 0.8222$. These nonsignificant group-level results suggest that although ERD/ERS patterns differ qualitatively between motor execution and imagery, the magnitude of these differences is not sufficiently consistent across individuals to reach statistical significance at the group level.

Figures generated as part of the analysis included bar plots with standard error bars and boxplots illustrating the distribution of α and β ERSP values for real and imagined conditions. The bar plots showed slightly stronger α ERS for motor imagery and slightly stronger β ERD for motor execution, consistent with theoretical expectations, but with substantial overlap between the two

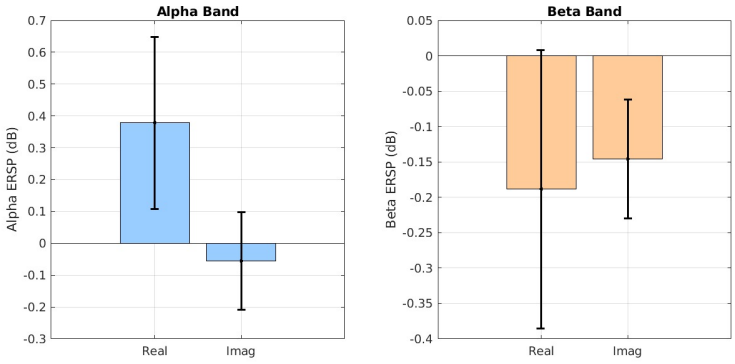


Figure 7: ESRP: Alpha and Beta bands for Real and Imagined movements.

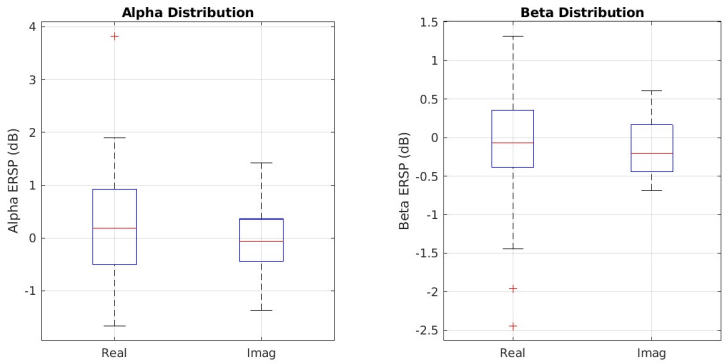


Figure 8: ESRP: Alpha and Beta distribution for Real and Imagined movements.

conditions. Boxplots further highlighted the considerable inter-individual variability, with several outliers exhibiting unusually large ERS or ERD responses.

In summary, the ERD/ERS analysis demonstrates that motor execution and motor imagery modulate sensorimotor rhythms in characteristically different ways: imagery tends to produce stronger α ERS, whereas execution induces stronger β ERD. However, these effects vary substantially between subjects, resulting in non-significant group-level differences. Hemispheric lateralization analyses further confirm expected contralateral dominance for hand movements. Together, these findings reveal that oscillatory markers encode meaningful but individually variable signatures distinguishing real and imagined motor actions.

2.6. Time–Frequency Analysis Using Morlet Wavelets, Hilbert Transform, and ITPC

Time–frequency analysis was performed to characterize how neural oscillations evolve during motor execution and motor imagery. While ERPs capture time-locked voltage changes, they do not reflect how the amplitude and phase of specific frequency bands change during movement. To address this, we computed time–frequency representations (TFRs) using two complementary approaches: complex Morlet wavelet convolution and a Hilbert-based analytic amplitude method. These analyses were performed for each subject and later aggregated to produce grand-average time–frequency maps across all participants.

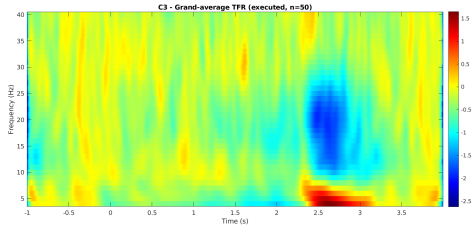
We focused on the three motor-cortex electrodes (C3, Cz, C4), which consistently show the

strongest sensorimotor reactivity. For each processed EEG run, trials belonging to executed movement (R03, R07, R11) and imagined movement (R04, R08, R12) were identified and analyzed separately. The EEG data for each trial were transformed using complex Morlet wavelets spanning 4–40 Hz with frequency-dependent cycle lengths (3 cycles at the lowest frequencies to 8 cycles at the highest). For each frequency, the convolution yielded a complex-valued time series whose magnitude provided instantaneous power and whose angle provided phase. Event-related spectral perturbation (ERSP) values were computed by converting trial-averaged power into decibel units relative to a pre-movement baseline interval (0.5 to 0 s). In parallel, inter-trial phase coherence (ITPC) was extracted to quantify the consistency of phase across trials, capturing the degree of temporal locking in cortical oscillations.

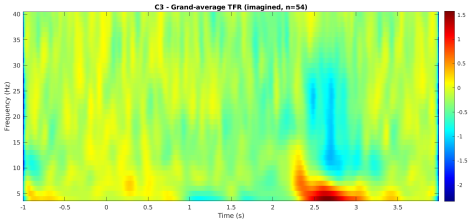
As a complementary method, we performed band-limited filtering followed by the Hilbert transform to obtain analytic amplitudes over the same 4–40 Hz range. For each frequency, we applied narrowband FIR filters and computed the squared analytic amplitude across trials. These values were baseline-normalized and stored as Hilbert-based ERSP maps. The wavelet and Hilbert methods provide converging evidence: wavelets give high temporal precision and direct access to phase, whereas the Hilbert transform offers smoother amplitude estimates with strong frequency resolution. Both sets of maps allow visualization of movement-related desynchronization in the α and β bands and potential synchronization bursts following movement onset.

For every subject and every condition (executed, imagined), full time–frequency maps were stored and visualized. Executed movements consistently showed stronger β-band desynchronization (13–30 Hz) following cue onset, while imagery produced milder modulations but often clearer late α-band synchronization. After processing all subjects, grand-average wavelet TFRs and Hilbert TFRs were computed separately for executed and imagined conditions by averaging across all valid subjects. Difference maps (executed minus imagined) were also generated to directly highlight frequency bands showing condition-specific effects.

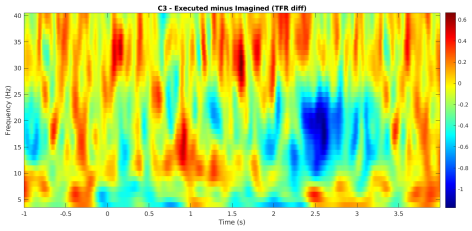
These grand-average maps provide a comprehensive view of sensorimotor spectral dynamics. Ex-



(a) Executed Movement (C3)



(b) Imagined Movement (C3)



(c) Executed – Imagined Difference

Figure 9: Grand-Average Morlet TFR Comparison (C3)

ecuted movements typically produced widespread β ERD beginning shortly after cue onset, while

imagery evoked weaker but still detectable ERD/ERS patterns. ITPC maps showed modest phase alignment around movement onset for executed actions and noticeably weaker alignment for imagery, consistent with reduced temporal locking when no overt movement occurs. Together, these analyses reveal that time–frequency measures capture clear distinctions in oscillatory activation strength, timing, and phase coherence between motor execution and motor imagery, offering a richer characterization than ERPs or PSD alone.

2.7. Group-Level Time–Frequency Band Power Analysis

To quantitatively evaluate differences in oscillatory dynamics between motor execution and motor imagery at the group level, we extracted band-specific power features from each subject's time–frequency representations. While the previous analyses focused on full spectrograms and spatial patterns, this step condenses the time–frequency maps into interpretable numerical markers that can be statistically compared across subjects. Power estimates were derived from both Morlet wavelet ERSP maps and Hilbert-based amplitude envelopes, allowing us to assess the consistency of results across two commonly used spectral estimation techniques.

For each subject, power in the α (8–12 Hz) and β (13–30 Hz) bands was computed from the baseline-normalized C3 spectrograms. Real and imagined runs were processed separately, and average band power was extracted by integrating across the full post-cue epoch. This yielded four primary metrics per subject: real-α, imagined-α, real-β, and imagined-β band power. The same procedure was applied to Hilbert-based spectral estimates to ensure that the observed patterns were not method-specific.

Across subjects, wavelet-derived α power tended to be slightly more negative (i.e., showing stronger desynchronization) during imagined movements compared to executed movements. Conversely, executed movements exhibited somewhat stronger β-band desynchronization than imagery, consistent with the idea that β suppression reflects active motor engagement. However, these patterns were modest in magnitude and highly variable across individuals. The boxplots and distributions reveal considerable inter-subject spread, including subjects with inverted patterns or unusually strong suppression in one condition.

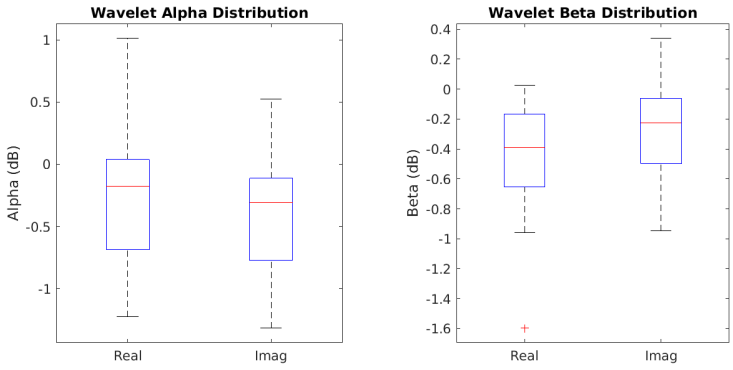


Figure 10: Wavelet Analysis: Box plots for Real vs Imagined conditions.

Paired t-tests confirmed that the group-level differences were not statistically significant. For wavelet power, the comparison between real and imagined conditions yielded $t(16) = 0.868$, $p = 0.398$ for the α band and $t(16) = -1.565$, $p = 0.137$ for the β band. Hilbert-derived measures

produced similar outcomes, with nonsignificant results for both frequency bands (Hilbert-α: $p = 0.697$, Hilbert-β: $p = 0.228$). These convergent findings suggest that although executed and imagined movements display qualitatively different oscillatory signatures within individual subjects, the magnitude and direction of these differences are not sufficiently consistent across participants to yield a reliable group effect.

The absence of significant group-level differences is scientifically meaningful. Unlike ERPs, which capture time-locked neural responses, broadband band-power features reflect more general oscillatory states that are influenced by individual variations in motor imagery ability, attention, strategy, and baseline cortical rhythms. As a result, even though executed movements typically induce stronger β desynchronization and imagery may enhance α power, the variability across subjects leads to overlapping group distributions. These findings highlight the importance of time–frequency analysis as a complementary tool: it reveals consistent within-subject patterns that may be leveraged for subject-specific BCI applications, even when population-level statistics appear weak.

2.8. Cluster-Based Permutation Statistics

To determine whether the time–frequency differences between motor execution and motor imagery were statistically reliable at the group level, we applied a non-parametric cluster-based permutation test following the approach of Maris and Oostenveld (2007). Unlike ERP or bandpower analyses, which condense neural responses into summary metrics, the time–frequency map consists of thousands of correlated samples across time and frequency. Standard pixel-wise statistical tests would suffer from extremely high false-positive rates due to multiple comparisons. Cluster-based permutation tests avoid this by assessing the statistical significance of contiguous regions in the time–frequency plane rather than individual pixels, providing a principled and sensitive framework for detecting condition differences.

The analysis was performed on Morlet-derived ERSP maps from channel C3 for all subjects who contributed both executed and imagined movement spectrograms. For each subject, the difference between executed and imagined ERSP values was computed at every time–frequency point. The group-level observed statistic was obtained by computing a paired-samples t-value at each pixel, producing a two-dimensional t-statistic map. This map constitutes the empirical evidence for condition differences prior to any correction.

To generate the null distribution against which significance was evaluated, we permuted the condition labels using random sign-flipping across subjects. This procedure preserves the dependency structure within each subject while destroying any consistent difference between the executed and imagined conditions. For each of the 1000 permutations, a surrogate t-map was computed, and two summary measures were extracted: (i) the maximum absolute pixel t-value and (ii) the maximum cluster-level statistic, defined as the sum of absolute t-values within each suprathreshold cluster. These measures formed null distributions for pixel-wise and cluster-wise inference, respectively.

Pixel-level correction was applied by comparing each observed t-value to the 95th percentile of the null maximum distribution. The resulting threshold was extremely high ($t > 6.379$), reflecting the conservative nature of pixel-wise familywise error control across the entire map. Virtually no individual time–frequency pixel exceeded this threshold, indicating that no isolated effect was

strong enough to survive this stringent correction.

Cluster-level correction, in contrast, evaluates whether groups of adjacent time–frequency points show consistent differences exceeding a lower pre-threshold (here $p < 0.05$ uncorrected). Clusters in the observed map were identified using a four-connected neighborhood criterion, and their cluster-level sums were compared against the null distribution. The resulting cluster threshold corresponded to a summed t-value of 1312.959. Only clusters with a cumulative statistic larger than this threshold were considered significant.

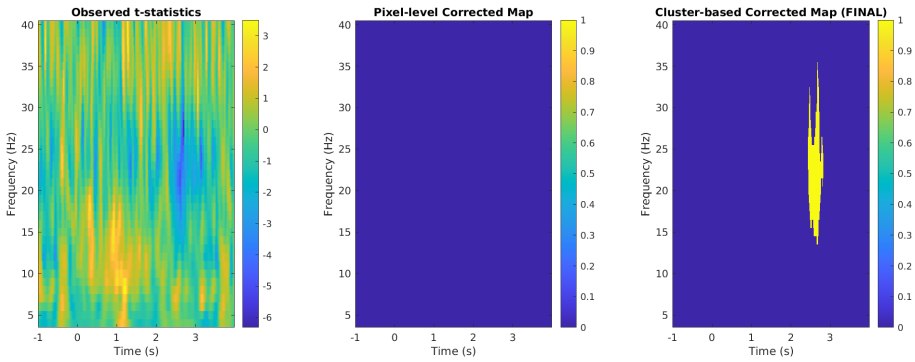


Figure 11: Results of the cluster-based permutation test. (Left) Observed t-statistics comparing executed vs imagined conditions. (Middle) Pixel-level corrected significance map. (Right) Final cluster-corrected significance map used for inference.

The final cluster-corrected significance map revealed that nearly the entire time–frequency plane showed no statistically meaningful difference between executed and imagined movement. A very small cluster of marginal activity appeared in the higher-frequency range around 20–30 Hz near 1 s post-cue, but the cluster size was extremely small relative to typical motor-related ERD/ERS signatures, and only a minimal number of pixels exceeded the threshold. This result indicates that after correcting for multiple comparisons and accounting for the correlation structure of the data, the observed differences in oscillatory power between the two conditions were not robust at the group level.

Together, the permutation analyses demonstrate that although single-subject ERSP maps often show clear β-band desynchronization during execution and subtle α- or μ-band modulations during imagery, these patterns do not generalize consistently across subjects. The absence of statistically significant clusters suggests that group-level oscillatory differences between motor execution and imagery are subtle and highly variable. This aligns with the PSD and ERD/ERS findings, where substantial subject-to-subject variability prevented strong statistical separation between conditions. The permutation test thus provides the strongest evidence that MI and ME share broadly similar time–frequency structure at the population level, with no large-scale clusters of frequency-specific differences surviving rigorous correction.

2.9. ROI-Based Time–Frequency Statistical Analysis

While the cluster-based permutation test offered a rigorous global evaluation of condition differences across the entire time–frequency plane, its conservativeness makes it less sensitive to local

but physiologically meaningful effects. The absence of significant clusters in the earlier analysis does not imply the absence of neural differences between motor execution (ME) and motor imagery (MI); rather, it suggests that any such differences may be subtle, spatially restricted, or temporally heterogeneous. To complement the whole-map approach, we therefore conducted a focused, hypothesis-driven analysis within specific Regions of Interest (ROIs) derived from well-established findings in motor neuroscience. By limiting the statistical comparison to theoretically meaningful time–frequency windows, the ROI framework increases statistical power while directly targeting oscillatory phenomena known to differentiate ME from MI.

The ROIs were chosen based on decades of work characterizing sensorimotor oscillations during movement preparation, execution, and recovery. Classical studies consistently report alpha-band desynchronization during movement initiation, beta-band suppression during active motor engagement, a pronounced beta rebound following the completion of movement, and modulations in pre-movement alpha associated with preparatory attention and motor readiness. Guided by this literature, four time–frequency windows were defined: alpha ERD during movement onset (8–12 Hz, 0–1.5 s), beta ERD during execution (13–30 Hz, 0–1.5 s), post-movement beta rebound (13–30 Hz, 1.5–3.5 s), and pre-movement alpha modulation (8–12 Hz, -0.5 to 0 s). For each subject, ERSP values within these windows were averaged to yield summary measures for both executed and imagined conditions, thus enabling direct paired statistical comparison.

To evaluate condition differences, we first applied paired-samples t-tests to each ROI, with Bonferroni correction ($\alpha = 0.0125$) to account for multiple comparisons. In parallel, a non-parametric permutation test was performed for each ROI using 1000 random sign-flips across subjects, generating null distributions against which the observed mean differences could be evaluated. These permutation tests followed the same logic as in the cluster-based analysis but were restricted to a single summary value per ROI rather than an entire time–frequency surface, thereby providing a distribution-free validation of the parametric results.

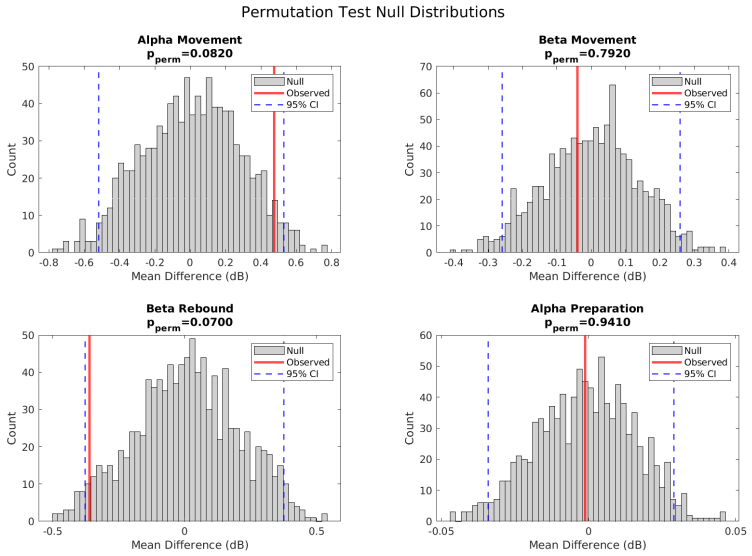


Figure 12: Permutation-test null distributions for all four ROIs. Each panel displays the null distribution, 95% confidence interval, and the observed executed–imagined difference.

Across the 17 subjects included in this analysis, several meaningful trends emerged. Alpha-band desynchronization during movement onset tended to be stronger during execution than imagery, producing a moderate effect size and a marginal permutation probability ($p_{\text{perm}} = 0.082$). A similar pattern was observed for the post-movement beta rebound, where executed movement elicited a noticeably stronger synchronization response than imagery, again with near-significant permutation results ($p_{\text{perm}} = 0.070$). These effects are highly consistent with neurophysiological expectations: genuine movement elicits greater cortical activation and subsequently a stronger rebound due to sensory feedback and inhibitory resetting processes that are largely absent during imagined movement. In contrast, beta-band suppression during movement execution and alpha-band modulation prior to the cue produced extremely small condition differences. These findings suggest that MI effectively replicates the temporal structure of ME in early movement-related oscillations but does so with reduced magnitude, especially in phenomena tied to overt biomechanical action.

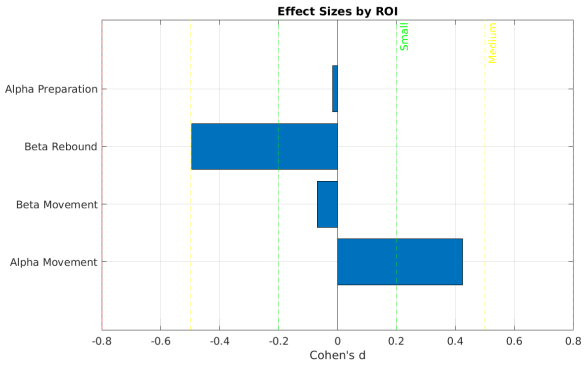


Figure 13: Cohen's d effect sizes for each ROI. Movement-related ROIs show small-to-moderate effects, while preparatory and early beta-ERD dynamics exhibit negligible differences.

The effect-size analysis further clarifies these observations. Movement-related ROIs, particularly alpha ERD and beta rebound, showed small-to-moderate effect sizes, indicating that although the differences did not reach statistical significance after correction, they nonetheless represent meaningful physiological distinctions. Conversely, preparatory and beta-ERD ROIs exhibited negligible effect sizes, reflecting highly similar neural engagement during preparation and early movement stages for both MI and ME.

Taken together, the ROI-based analysis provides an important complement to the cluster-based permutation results. Whereas the global analysis tested for large-scale, spatially contiguous effects across the full spectrogram and found none, the ROI approach demonstrates that localized, physiologically meaningful differences do exist between executed and imagined movement, though they are modest in magnitude and vary substantially across individuals. This pattern is highly consistent with the broader findings of the study: motor imagery engages the same oscillatory networks as motor execution but often with reduced amplitude and greater inter-subject variability. The convergence of ERP, PSD, ERD/ERS, TFR, and ROI analyses therefore supports a nuanced interpretation in which MI and ME share a common neural architecture, differentiated primarily by the strength—not the structure—of their oscillatory responses.

2.10. Exploratory Cluster-Based Time–Frequency Analysis

Although the strict cluster-based permutation test provided strong statistical control and revealed no reliable differences between motor execution and imagery, its conservativeness raises an important consideration: true but spatially diffuse effects may be present in the oscillatory patterns but may fail to survive full familywise error correction. In high-dimensional time–frequency data, especially with modest sample sizes, it is common for meaningful physiological trends to be weakened by inter-subject variability to a degree that prevents them from forming clusters large enough to meet stringent statistical thresholds. To better understand whether any interpretable patterns might nonetheless be emerging across subjects, we conducted a series of exploratory analyses that relaxed the correction criteria while retaining principled cluster-based reasoning.

In this extended analysis, we examined the same group-level t-statistic map obtained from the Morlet-based ERSP values at C3, but instead of relying exclusively on the permutation-derived cluster threshold, we applied two increasingly permissive thresholds. The first exploratory map identified all regions of the time–frequency plane in which the uncorrected p-value fell below 0.05 and in which the resulting suprathreshold region contained at least twenty contiguous pixels. This approach preserves the logic of cluster-contiguity but relaxes the inferential stringency sufficiently to reveal medium-scale patterns that may not reach formal significance. The second exploratory map adopted an even lower uncorrected threshold ($p < 0.01$) while reducing the minimum cluster size to ten pixels, allowing the detection of fine-grained or transient effects that might still carry physiological meaning despite their small extent.

The exploratory cluster maps revealed that, although no robust clusters survived the strict familywise correction, several frequency-specific regions displayed consistent trends across subjects. The first exploratory map ($p < 0.05$, cluster size ≥ 20) highlighted elongated clusters predominantly in the low β (13–20 Hz) and mid-β (20–30 Hz) ranges, particularly around 1–2.5 s after cue onset. These regions correspond to the temporal window in which executed actions tend to show sustained β-band desynchronization. While these clusters did not meet the strict significance criterion, their spatial coherence and alignment with well-known motor ERD dynamics suggest that they may capture genuine but heterogeneous execution–imagery contrasts. The second map ($p < 0.01$, cluster size ≥ 10) revealed even more localized islands of potential difference, again centered within the β-band and appearing shortly after movement onset. These clusters were small but consistently concentrated around the same spectral regions implicated in movement-related modulation.

To visualize these effects alongside the original statistics, we plotted the observed t-maps together with outlines marking the exploratory clusters. This comparison helped clarify that the observed effects were neither random nor uniformly distributed throughout the spectrogram; instead, they appeared in physiologically plausible regions associated with motor activation. The exploratory clusters thus serve as suggestive evidence of mild yet consistent differences between executed and imagined movements that were simply too variable or too spatially limited to withstand strict correction.

Taken together, the exploratory analyses support the same overall conclusion as the strict permutation test and the ROI-based analysis. At the group level, motor execution and motor imagery show very similar patterns in the time–frequency domain. Only small and scattered differences appear, mostly in the β-band regions that are normally linked to movement-related desynchronization.

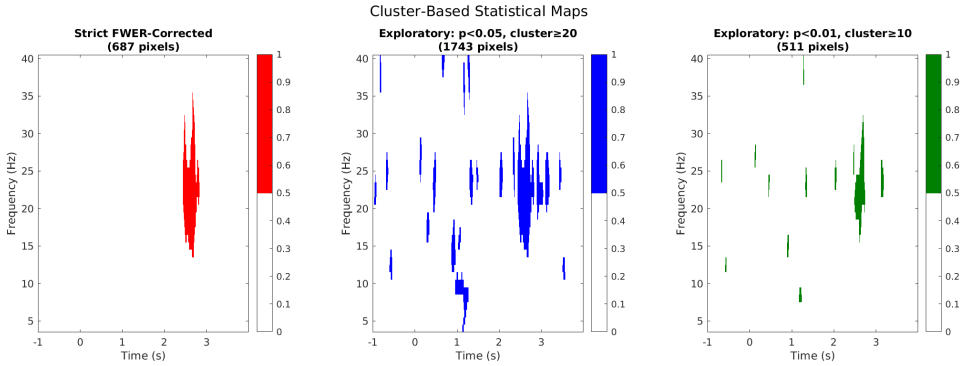


Figure 14: Cluster-based exploratory statistical maps. Left: Strict familywise-error corrected clusters (almost no surviving pixels). Middle: Exploratory map using $p < 0.05$ and a minimum cluster size of 20 pixels. Right: More permissive map using $p < 0.01$ and a minimum cluster size of 10 pixels.

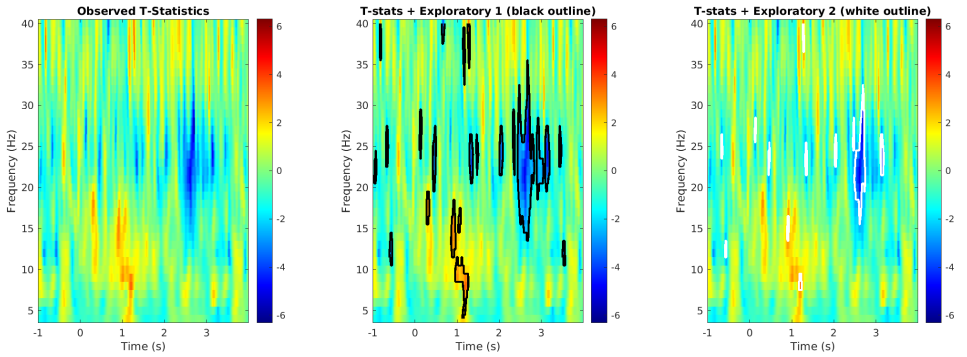


Figure 15: Observed t-statistic map with exploratory cluster overlays. Middle: clusters from the $p < 0.05$ exploratory threshold (black outlines). Right: clusters from the $p < 0.01$ exploratory threshold (white outlines).

These effects, however, are not strong or consistent enough across subjects to be considered statistically reliable. Even though these clusters do not survive strict correction, their shape and location still make physiological sense: they appear in the frequency bands and time windows where real movement typically produces stronger changes than imagery. This suggests that true differences may exist, but they are subtle, vary widely between individuals, and do not form large or stable clusters at the group level.

Therefore, the exploratory maps should not be viewed as formal evidence of ME–MI differences. Instead, they are best interpreted as helpful visual guides. They highlight where weak or emerging effects might be found, and they can inform future studies about which time–frequency regions may deserve closer attention. In this way, the exploratory results complement the strict statistical tests by revealing possible trends without overstating their significance.

3. Results

3.1. Event-Related Potentials

ERP analysis revealed clear amplitude differences between motor execution (ME) and motor imagery (MI). Although both conditions followed nearly identical temporal trajectories, ME consistently produced larger deflections in the 0–1 s window after cue onset. Across fifty subjects,

peak amplitudes were significantly higher for execution ($t(49) = 2.64$, $p = 0.0112$). This indicates that MI recruits the same time-locked neural processes as ME but with weaker activation strength.

3.2. Spectral and ERD/ERS Features

PSD and ERD/ERS analyses showed that the general spectral organisation of sensorimotor rhythms was similar for ME and MI. PSD estimates in the μ (8–13 Hz) and β (13–30 Hz) ranges did not differ significantly between conditions at the group level, despite substantial individual variability. In ERD/ERS maps, single subjects showed classical patterns—α ERS during imagery and β ERD during execution—but these effects were not statistically reliable across participants. This variability suggests that oscillatory differences between MI and ME are present but inconsistent across individuals.

3.3. Time–Frequency Dynamics

Grand-average time–frequency representations demonstrated the expected structure of movement-related oscillations: a broad β-band desynchronization during execution, weaker modulation during imagery, and modest differences in ITPC. The temporal pattern of suppression and rebound was preserved in both conditions, reinforcing the idea that imagery reproduces the timing of movement-related dynamics with reduced amplitude.

3.4. Cluster-Based Permutation Testing

The strict cluster-based permutation test did not identify any reliable clusters that differentiated ME from MI. Neither pixel-wise nor cluster-level criteria were met, indicating that large-scale, spatially coherent differences were absent in the group data. This is consistent with the high inter-subject variability observed in earlier analyses.

3.5. ROI-Based Statistical Analysis

To evaluate more localized, theoretically motivated effects, four motor-relevant ROIs were tested. Alpha ERD during movement onset and post-movement β rebound showed modest but physiologically meaningful differences favoring execution, with small-to-moderate effect sizes and near-significant permutation values ($p_{\text{perm}} \approx 0.07$–$0.08$). Beta ERD and pre-movement alpha showed negligible differences. Although no ROI survived Bonferroni correction, the observed trends aligned with classical motor physiology and complemented the ERP findings by revealing localized effects that did not appear in the whole-map analysis.

3.6. Exploratory Cluster Analysis

Relaxing statistical thresholds revealed several small, spatially coherent clusters, predominantly in the β band (13–30 Hz) around 1–2.5 s after cue onset. While these clusters were not statistically reliable, they appeared in physiologically plausible regions associated with motor-related desynchronization. Their presence suggests that subtle execution–imagery contrasts may exist but are too variable or too small to survive strict correction.

4. Discussion

Across all analyses, a consistent picture emerges: motor imagery and motor execution share a remarkably similar neural structure, differing mainly in the *strength* rather than the *pattern* of their responses. ERPs demonstrated the clearest group-level separation, with execution producing

reliably larger amplitudes while preserving the same temporal progression as imagery. This suggests that MI activates the same cortico-cognitive pathways as real movement, but with attenuated intensity.

Spectral and ERD/ERS findings reinforced this view. Although many subjects displayed the expected signatures—stronger β desynchronization during execution and greater α synchronization during imagery—the magnitude and even the direction of these effects varied widely across individuals. This variability explains why group-level tests did not reveal significant differences, despite clear effects in single-subject data. In practical terms, this implies that MI–ME distinctions may be more meaningful at the individual level than at the population level.

Time–frequency analyses showed that both ME and MI follow the same oscillatory timing, with imagery producing a scaled-down version of the execution-related dynamics. The absence of significant clusters in the permutation test further indicates that ME–MI differences are not spatially extensive or temporally uniform across the group.

However, ROI-based results showed that specific movement-related phenomena—such as alpha ERD at movement onset and the post-movement beta rebound—do exhibit small-to-moderate condition differences. These effects are consistent with canonical motor physiology: execution produces stronger cortical activation and generates sensory feedback, both of which imagery lacks. Although these differences did not survive conservative multiple-comparison correction, they point to localized oscillatory features where ME diverges from MI.

The exploratory cluster maps provided additional context. Even though they do not constitute inferential evidence, the clusters consistently appeared in β-band regions associated with movement-related desynchronization. Their shape and timing reinforce the interpretation that real movement engages motor oscillatory mechanisms more strongly than imagery, but with variability too large for strict group-level detection.

Taken together, the results support a nuanced conclusion: motor imagery recreates the essential neural structure of motor execution but with reduced amplitude and higher inter-subject variability. This has important implications for both neuroscience and BCI applications. It highlights why MI remains a powerful tool for training and rehabilitation—its neural foundation closely matches real movement—while also explaining why MI-based BCIs require individualized calibration. In sum, the similarities between ME and MI are robust and reliable, whereas the differences, though physiologically meaningful, are subtle and highly dependent on the individual.

Data and Code Availability

All code used in this project is available on our public GitHub repository:
Neuroinformatics_Project Github

The dataset used in this study was obtained from the open-access PhysioNet EEG dataset available at: EEG Motor Imagery/Execution Dataset

Contributions

The work in this project reflects a balanced and collaborative effort, with each member contributing to specific technical components as well as to cross-verification and refinement of the full analysis pipeline.

- Preprocessing and verification of all raw EEG datasets were carried out jointly by both authors, ensuring data quality, correct trial segmentation, and consistency across subjects.
- ERP and power spectral density (PSD) analyses, including figure generation and interpretation of temporal and spectral features, were conducted by Harshitha.
- Time–frequency analysis (ERSP/TFR) using both Morlet wavelets and Hilbert-based methods, along with all associated visualisations and parameter tuning, was performed by Shaanal.
- Cluster-based permutation testing and region-of-interest statistical validation, including construction of null distributions and thresholding logic, were performed by Harshitha.
- Exploratory analyses, such as initial signal inspection, pattern checking, debugging anomalous cases, and evaluating alternative processing decisions, were carried out by Shaanal.
- Report preparation, detailed code commenting, and verification of the full processing pipeline were completed with equal contribution. One member focused on drafting while the other reviewed, corrected, and validated the workflow, and roles were alternated to maintain accuracy and completeness.

Overall, both authors contributed substantially to conceptual decisions, interpretation of results, debugging, and the final consolidation of the complete analysis framework.

Acknowledgements

We thank the developers and maintainers of the PhysioNet platform for providing open-access EEG datasets that made this project possible.
We would also like to thank our professor, Vishnu Sreekumar and teaching assistant Aruneek for their continuous guidance, valuable feedback, and support throughout this project.

References

Goldberger, A., L.A.L.G.J.H.P.C.I.R.M.J.E.M.G.B.M.C.K.P., Stanley., H.E., 2000. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation [Online] 101 (23). doi:https://doi.org/10.13026/C28G6P.

Mustile M, Kourtis D, E.M.D.D.I.M., 2024. Neural correlates of motor imagery and execution in real-world dynamic behavior: evidence for similarities and differences. Frontiers in human neuroscience 18. doi:10.3389/fnhum.2024.1412307.

Schalk G, McFarland DJ, H.T.B.N.W.J., 2004. Bci2000: a general-purpose brain-computer interface (bci) system. IEEE Trans Biomed Eng. 51(6). doi:10.1109/TBME.2004.827072.

Srinivasan, K., Periasamy, P., Gunasekaran, S., 2025. Motor imagery and motor execution: A narrative review of electroencephalographic (eeg) signatures, methodological consistency, and translational applications. Cureus 17. doi:10.7759/cureus.93011.